Riccati-based solution to the optimal control of linear evolution equations with finite memory

RICCATI-BASED SOLUTION TO THE OPTIMAL CONTROL OF LINEAR EVOLUTION EQUATIONS WITH FINITE MEMORY

PAOLO ACQUISTAPACE AND FRANCESCA BUCCI

Abstract. In this article we study the optimal control problem with quadratic functionals for a linear Volterra integro-differential equation in Hilbert spaces. With the finite history seen as an (additional) initial datum for the evolution, following the variational approach utilized in the study of the linear-quadratic problem for memoryless infinite dimensional systems, we attain a closed-loop form of the unique optimal control via certain operators that are shown to solve a coupled system of quadratic differential equations. This result provides a first extension to the partial differential equations realm of the Riccati-based theory recently devised by L. Pandolfi in a finite dimensional context.

1. Introduction

Given a linear control system $y'(t) = Ay(t) + Bu(t)$, $t \in [0,T)$, in a Hilbert space H, existence and uniqueness for the Cauchy problems associated with the (differential) Riccati equations

$$P'(t) + A^*P(t) + P(t)A - P(t)BB^*P(t) + Q = 0 \qquad t \in [0,T)$$

play a central role in the study of the related optimal control problems with quadratic functionals on a finite time interval. (We recall that the operator Q occurs in the functional in connection with the observed state.) Indeed, the Riccati equation corresponding to the minimization problem is expected to yield the – hopefully unique – operator $P = P(t)$ which enters the feedback representation of the optimal control, thereby allowing its synthesis.

The question of well-posedness of Riccati equations is particularly interesting in the context of partial differential equations (PDE), where major technical difficulties stem from the presence of unbounded control operators B – these are brought about, in particular, by the modeling of boundary inputs. The analysis become even more challenging in the case of hyperbolic or composite dynamics: the actual meaning of the gain operator $B^*P(t)$ that occurs in the quadratic term of the Riccati equation (or lack thereof) is the central issue that must be tackled, even for the only purpose of existence of a solution to the Riccati equation.

Forty years of research on this subject have brought about distinct functional-analytic frameworks that mirror parabolic PDE, hyperbolic PDE, and also certain systems of hyperbolic-parabolic PDE, along with respective Riccati theories. The reader is referred to the Lecture Notes [27] for a brief overview and significant illustrations on the subject, and to the monographs by Bensoussan *et al.* [2] and by Lasiecka and Triggiani [28] for an ample treatment of the Riccati theories pertaining

2020 *Mathematics Subject Classification.* 49N10, 35R09; 93C23, 49N35.

Key words and phrases. integro-differential equation, linear-quadratic problem, evolution equations with memory, closed-loop optimal control, Riccati equation.

to the finite time horizon problem until 2000. A literature review on the major contributions and latest achievements in the study of the linear-quadratic (LQ) problem for coupled systems of hyperbolic-parabolic PDEs – now spanning more than two decades – is found in the recent [1] and [38] (in the deterministic and stochastic cases, respectively).

In this paper we are instead interested in the LQ problem for integro-differential equations. These arise, as is widely known, in the modeling of certain diffusion processes and other phenomena that exhibit hereditary effects such as viscoelasticity; see for instance the monograph [37] by Renardy *et al.* The class of evolutions we consider is described by problem (1.1), which is introduced in full detail in Section 1.1 below. It is a reasonably simple model equation, which reduces to the archetypical differential system in the absence of the integral term, with the control operator B here assumed bounded. We believe the setting and the chosen approach should also serve as a baseline for further developments.

When it comes to the LQ problem for integro-differential PDE, the picture of the existing literature is not as complete as in the memoryless case. On one hand, the work by Cannarsa *et al.* [6] – as first, in 2013 – deals with the Bolza problem for semilinear evolution equations, with the dynamics displaying an *infinite* memory; the analysis encompasses both parabolic and hyperbolic equations with nonlocal terms (the respective results and methods of proof differ[1], though). Thus, [6] deals with a more general problem than the LQ one; results for the existence of an *open-loop* solution to the optimization problem are established, which apply to relevant physical evolutions.

The optimal control problem with quadratic functionals for a class of semilinear parabolic equations is studied in the recent [7]. Interestingly, this work explores and establishes necessary and/or sufficient second-order optimality conditions.

On the other hand, our framework and goals are more specific: the present study focuses on the finite horizon LQ problem in the presence of *finite* memory; we aim at its *closed-loop* optimal solution and more precisely, at a synthesis of the unique optimal control by way of solving a corresponding Riccati (or Riccati-like) equation. Thus, we must go back in time a bit. The 1996 work [35] by Pritchard and You addresses a similar problem (more precisely, our integro-differential model can be subsumed under the one considered therein), in the same Hilbert space setting. A semi-causal representation formula for the optimal control is established, with the feedback operator depending on an another operator which is shown to solve a Fredholm integral equation. In the authors' words, the said equation "... plays a role similar to that of the operator Riccati equation".

The question as to whether a Riccati-based theory is actually viable was addressed (and answered affirmatively) just recently by L. Pandolfi [33], with a study restricted to an uncomplicated setting, namely, considering the integro-differential model (1.1) in a finite dimensional space $H = \mathbb{R}^n$ and neglecting the generator A (governing the free dynamics in the absence of memory). A feedback representation of the optimal control is established, and in addition the operators involved in it are shown to solve a coupled system of three quadratic (matrix) equations.

[1]Just to give a glimpse of some of the methods employed therein, we recall that in order to tackle a second order in time problem the authors appeal to the celebrated *history approach* introduced by Dafermos in 1970, which allows to utilize semigroup theory for the mathematical analysis of an equivalent coupled system satisfied by a suitable augmented variable.

In this work we fully extend – and in a sense enhance, see our Theorem 1.7 – the results obtained in [33] to the controlled integro-differential system (1.1) in a true infinite dimensional context. In order to do so,

- we adopt the Volterra equations perspective of [33], along with the consideration of the history as a component of the state (shared by [6] as well), while
- we perform (and adapt) the plan carried out in the study of the LQ problem for *memoryless* control systems in infinite dimensional spaces, whose line of argument can be summarized (in broad terms) as follows: (i) a convex optimization argument brings about the existence of a unique optimal pair; (ii) an operator $P(t)$ is introduced, defined in terms of the optimal evolution; this operator enters the quadratic form which yields the optimal cost as well as the causal representation of the optimal control; (iii) $P(t)$ is shown to solve the differential Riccati equation; (iv) if uniqueness holds true as well, the said solution $P(t)$ renders effective the optimal synthesis.

The formulation of the optimal control problem, our assumptions and main results, i.e. Theorems 1.6 and 1.7, are made explicit in the next Section 1.1. An expanded outline of the paper found at the end of this section will provide guide and insight into the sequence of proofs.

In concluding this introduction we provide a (minimal, for obvious reasons) bibliographical selection of general textbooks as well as articles with specific focuses, the latter ones still within control theory for linear models.

An explicit account of the well-posedness and regularity results devised for abstract linear equations in Banach spaces is beyond the scope of this article. We point out and give credit to a few pioneering works from the late sixties on, such as [21], [14] and [15]. So, with regard to the broad topic of integro-differential equations the reader is referred to the monographs (listed in cronological order, along with the above-mentioned [37]) by Gripenberg *et al.* [24] (in finite dimensional spaces), Prüss [36] and Pandolfi [34]; [34] is especially valuable for the historical insight and up-to-date references, besides the discussion of modeling and analytical aspects. (In this connection we note that even a higher order PDE such as the Moore-Gibson-Thompson equation[2] has been related – with different aims and in a different fashion – to wave equations with memory; see [16], [4], [5].)

Now, moving on to various control-theoretic properties – distinct from quadratic optimal control – that have been explored and established in the case of evolution equations with (infinite, more often than finite) memory, we recall that these include controllability ([26], [25], [9, 10], [19, 20]), reachability ([29], [22]), unique continuation ([17]); observability and inverse problems via Carleman estimates ([8], [31, 32]); stability and uniform decay rates ([23], [11], [12]). The mathematical tools utilized include: purely PDE methods, semigroup theory, harmonic analysis. (The variety of notable advances in the study of the long-time behaviour of solutions to semilinear and nonlinear equations with memory is inevitably left out.)

We note in particular that a line of investigation which has been followed specifically in the study of controllability is the reduction of the integro-differential equation to a system comprising a PDE and an ordinary differential equation; see [9, 10].

[2]The Moore-Gibson-Thompson (or Stokes-Moore-Gibson-Thompson) equation is a widely studied third order in time PDE arising from the linearization of a quasilinear model for the propagation of ultrasound waves.

Whether this method could be pursued successfully in order to study the LQ problem and devise a Riccati-based theory, as well, is a question that is left open here.

Finally, for context and pertinent background about the optimal control of *stochastic* Volterra equations, see for instance the recent [3] along with its references.

1.1. **The integro-differential model, associated minimization problem.** Let H and U be two separable Hilbert spaces; U is the *control* space. Given $T>0$, we consider a linear Volterra integro-differential equation in the space H, and the corresponding Cauchy problem

$$\begin{cases} w'(t) = Aw(t) + \displaystyle\int_0^t K(t-s)w(s)\,ds + Bu(t)\,, & t\in(0,T) \\ w(0) = w_0 \in H\,, \end{cases} \tag{1.1}$$

under the following basic Assumptions on the operators A, B and K which appear in (1.1). These operators describe the uncontrolled (or free) dynamics (A and K) and the action of control functions (B), respectively; the "memory kernel" K enters the convolution term which specifically accounts for a past history of the component w of the state variable. The actual *state space* will be described next; see (1.6).

Assumptions 1.1 (**Basic Assumptions**). *Let H, U be separable complex Hilbert spaces.*

- *The closed linear operator $A\colon \mathcal{D}(A)\subset H\to H$ is the infinitesimal generator of a strongly continuous semigroup $\{e^{tA}\}_{t\ge 0}$ on H;*
- *$K\in L^2(0,T;\mathbb{R})$;*
- *$B\in\mathcal{L}(U,H)$.*

Remarks 1.2. Our analysis here focuses on a simple pattern for the integro-differential model. In particular, (i) the hypothesis on the control operator B covers the case of partial differential equations (PDE) systems subject to *distributed* control. The study of an integro-differential model in the presence of an *unbounded* control operator B – which is naturally brought about by boundary or point control actions – is left to subsequent work. (ii) On the other hand, the assumption that the memory kernel is real valued can be relaxed to $K\in L^2(0,T;\mathcal{L}(H))$, with the computations carried out still valid, provided the operator $K(\cdot)$ commutes with the semigroup $e^{\cdot A}$. (iii) The absence of (the realization of) a differential operator within the convolution term in (1.1) is a restriction which is rendered milder in view of MacCamy's trick; see e.g. [34, Sections 3.3 and 4.2].

To the state equation (1.1) we associate the following quadratic functional over the preassigned time interval $[0,T]$:

$$J(u) = \int_0^T \big(\langle Qw(t),w(t)\rangle_H + \|u(t)\|_U^2\big)\,dt\,, \tag{1.2}$$

where the operator Q simply satisfies

$$Q\in\mathcal{L}(H)\,,\quad Q=Q^*\ge 0\,.$$

In Section 2 we will derive a representation formula for the solution to the state equation in (1.1), as well as to the solutions to the family of Cauchy problems obtained taking as initial time $\tau\in(0,T)$ (in place of $\tau=0$), that is (1.4) below.

The optimal control problem is formulated in the usual (classical) way.

Problem 1.3 (**The optimal control problem**)**.** *Given $w_0 \in H$, seek a control function $u \in L^2(0,T;U)$ which minimizes the cost functional* (1.2)*, where $w(\cdot)$ is the solution to* (1.1) *corresponding to the control function $u(\cdot)$ (and with initial datum w_0.*

Remark 1.4. The solutions to (1.1) are meant in a *mild* sense; see (2.1).

If $\tau \in (0,T)$ is given, and having set

$$\xi(\cdot) = w(\cdot)\big|_{[0,\tau]}\,, \quad \xi_0 = w(\tau^+)\,, \quad X_0 = \begin{pmatrix} \xi_0 \\ \xi(\cdot) \end{pmatrix}, \tag{1.3}$$

we introduce the family of Cauchy problems

$$\begin{cases} w'(t) = Aw(t) + \displaystyle\int_\tau^t K(t-s)w(s)\,ds + \int_0^\tau K(t-s)\xi(s)\,ds + Bu(t)\,,\ t \in (\tau,T) \\ w(\tau^+) = \xi_0 \end{cases} \tag{1.4}$$

and the associated cost functional

$$J_\tau(u,X_0) = \int_\tau^T \left(\langle Qw(t),w(t)\rangle_H + \|u(t)\|_U^2\right) dt\,. \tag{1.5}$$

We will set

$$Y_\tau := H \times L^2(0,\tau;H)\,; \tag{1.6}$$

Y_τ is the *state space* of our integro-differential problem (1.4)-(1.5).

Remark 1.5. It is interesting to note the analogy with the state space pertaining to retarded differential equations, that is $H \times L^2(-h,0;H)$ in the case of a delay $h > 0$. The reader is referred to [2, Part II, Chapter 4] for a discussion about the "product space approach" to the analysis of differential systems with delays in the state, control, and observation.

1.2. **Main results.** Our main results establish the synthesis of the optimal control for our problem (1.1)-(1.2), as certain operators P_i, $i = 1,2,3$, which occur in its feedback representation – as well as in the quadratic form that actualizes the optimal cost – are entries of an operator matrix P which is shown to be the unique solution to a quadratic Riccati-type equation.

Theorem 1.6. *With reference to the optimal control problem* (1.4)-(1.5)*, under the Assumptions 1.1, the following statements are valid for each $\tau \in (0,T)$.*

S1. *For each $X_0 \in Y_\tau$ there exists a unique optimal pair $(\hat u(\cdot,\tau;X_0), \hat w(\cdot,\tau;X_0))$ which satisfies*

$$\hat u(\cdot,\tau;X_0) \in C([\tau,T],U)\,, \quad \hat w(\cdot,\tau;X_0) \in C([\tau,T],H)\,.$$

S2. *Given $t \in [\tau,T]$, the linear bounded operator $\Phi(t,\tau)\colon Y_\tau \longrightarrow Y_t$ defined by*

$$\Phi(t,\tau)X_0 := \begin{pmatrix} \hat w(t,\tau;X_0) \\ \hat y(\cdot) \end{pmatrix} \ \text{where } \hat y(\cdot) = \begin{cases} \xi(\cdot) & \text{in } [0,\tau] \\ \hat w(\cdot,\tau,X_0) & \text{in } [\tau,t] \end{cases} \tag{1.7}$$

is an evolution operator, namely, it satisfies

$$\Phi(t,t) = I\,, \qquad \Phi(t,\tau) = \Phi(t,\tau_1)\Phi(\tau_1,\tau) \quad \text{for } \tau \le \tau_1 \le t \le T.$$

S3. *There exist three bounded operators, denoted by $P_0(\tau)$, $P_1(\tau,s)$, $P_2(\tau,s,q)$ – defined in terms of the optimal evolution and of the data of the problem (see the expressions* (4.8) *and* (4.12)*) –, such that the optimal cost is given by*

$$\begin{aligned} J_\tau(\hat{u},X_0) &= \big\langle P_0(\tau)w_0,w_0\big\rangle_H + 2Re\int_0^\tau \big\langle P_1(\tau,s)\xi(s),w_0\big\rangle_H\,ds \\ &\quad + \int_0^\tau\!\!\int_0^\tau \big\langle P_2(\tau,s,q)\xi(s),\xi(q)\big\rangle_H\,ds\,dq \equiv \langle P(\tau)X_0,X_0\rangle_{Y_\tau}\,. \end{aligned} \tag{1.8}$$

$P_0(\tau)$ and $P_2(\tau,s,q)$ are non-negative self-adjoint operators in the respective functional spaces H and $L^2(0,\tau;H)$.

S4. *The optimal control admits the following feedback representation*

$$\hat{u}(t,\tau;X_0) = -B^*P_0(t)\hat{w}(t;\tau,X_0) - \int_0^t B^*P_1(t,s)\hat{y}(s)\,ds\,,\ \tau\le t\le T\,, \tag{1.9}$$

with $\hat{y}(\cdot)$ given by (1.7)*.*

S5. *The operators $P_0(t)$, $P_1(t,s)$, $P_2(t,s,q)$ – as above in* **S3.** *– satisfy the following coupled system of equations, for every $t\in[0,T)$, $s,q\in[0,t]$, and for any $x,y\in\mathcal{D}(A)$:*

$$\begin{cases} \frac{d}{dt}\langle P_0(t)x,y\rangle_H + \langle P_0(t)x,Ay\rangle_H + \langle Ax,P_0(t)y\rangle_H + \langle Qx,y\rangle_H \\ \qquad -\langle B^*P_0(t)x,B^*P_0(t)y\rangle_U + \langle P_1(t,t)x,y\rangle_H + \langle x,P_1(t,t)y\rangle_H = 0 \\ \frac{\partial}{\partial t}\langle P_1(t,s)x,y\rangle_H + \langle P_1(t,s)x,Ay\rangle_H + \langle K(t-s)x,P_0(t)y\rangle_H \\ \qquad +\langle P_2(t,s,t)x,y\rangle_H - \langle B^*P_1(t,s)x,B^*P_0(t)y\rangle_U = 0 \\ \frac{\partial}{\partial t}\langle P_2(t,s,q)x,y\rangle_H + \langle P_1(t,s)x,K(t-q)y\rangle_H + \langle K(t-s)x,P_1(t,q)y\rangle_H \\ \qquad -\langle B^*P_1(t,s)x,B^*P_1(t,q)y\rangle_U = 0 \end{cases} \tag{1.10}$$

with final conditions

$$P_0(T)=0\,,\ P_1(T,s)=0\,,P_2(T,s,q)=0\,.$$

The coupled system of (four) equations satisfied by the operators $P_0(t)$, $P_1(t,s)$, $P_1(t,q)^*$ and $P_2(t,s,q)$ – with the one for $P_1(t,q)^*$ derived later in the paper, see (5.10) – can be shown to be equivalent to a single equation satisfied by the matrix operator

$$P(t) := \begin{pmatrix} P_0(t) & P_1(t,\cdot) \\ P_1(t,:)^* & P_2(t,\cdot,:) \end{pmatrix} \tag{1.11}$$

in the product space $H\times L^2(0,t;H)$ (the symbols $\cdot$ and $:$ stand for the hidden variables s and q, respectively). Such a unified form of the equation, that is (1.12) below,

- renders more explicit its Riccati-type nature,
- reduces to a standard Riccati equation in the absence of memory.

The unbounded operator coefficients involve Dirac delta distributions[3].

[3]More precisely, the Dirac delta is considered here as a linear functional acting on the space of continuous functions with values in a Banach space Y as follows: if $U\in C([a,b],Y)$, and $t\in[a,b]$, then $\delta_t(U(\cdot))=U(t)$.

Theorem 1.7 (Well-posedness for the Riccati equation)**.** *With reference to the optimal control problem* (1.4)-(1.5)*, under the Assumptions 1.1, let* $P_0(t)$*,* $P_1(t,\cdot)$ *and* $P_2(t,\cdot,:)$ *as from the statement S3. of Theorem 1.6. Then, the (matrix) operator* $P(t)$ *defined by* (1.11) *is the unique solution of the following quadratic equation with unbounded coefficients:*

$$\begin{aligned}\frac{d}{dt}P(t)+P(t)(\mathcal{A}+\mathcal{K}_1(t)+\mathcal{D}_{1,t})+(\mathcal{A}^*+\mathcal{K}_2(t)+\mathcal{D}_{2,t})P(t)&\\ -P(t)\mathcal{B}\mathcal{B}^*P(t)+\mathcal{Q}=0\,,&\end{aligned}\tag{1.12}$$

where we set

$$\mathcal{A}:=\begin{pmatrix}A&0\\0&0\end{pmatrix},\qquad \mathcal{B}:=\begin{pmatrix}B&0\\0&0\end{pmatrix},\qquad \mathcal{Q}:=\begin{pmatrix}Q&0\\0&0\end{pmatrix},$$

$$\mathcal{K}_1(t):=\begin{pmatrix}0&K(t-\cdot)\\0&0\end{pmatrix},\qquad \mathcal{K}_2(t):=\begin{pmatrix}0&0\\K(t-:)&0\end{pmatrix},$$

$$\mathcal{D}_{1,t}:=\begin{pmatrix}0&0\\\delta_t(\cdot)&0\end{pmatrix},\qquad \mathcal{D}_{2,t}:=\begin{pmatrix}0&\delta_t(:)\\0&0\end{pmatrix},$$

and $\delta_t(\cdot)$ *denotes the Dirac delta distribution: to wit,* $\delta_t f(\cdot)=f(t)$*.*

Remark 1.8. An analogous appearance of Dirac delta distributions is found, e.g., in [18, Chapter VI, Section 7, p. 447], where a different integro-differential problem – in the absence of control actions, though – is studied. The state space is $X\times L^1(0,\infty;X)$ (X being a Banach space) therein, and the Dirac delta is used to represent the integral part (of the right member of the equation) in the resolvent operator; the mild solution of the problem is given by the first component of such operator. Similarly, the optimal control $\hat{u}(t,\tau;X_0)$ of our problem, in feedback form, is given by the first component of the vector $-\mathcal{B}^*P(t)\Phi(t,\tau)X_0$, where $\mathcal{B}$ is the matrix operator defined in the statement of Theorem 1.7; see Proposition 4.5.

1.3. **Notation.** We note that various operators that occur in the article – such as the ones in (2.3) – are strongly continuous in appropriate spaces; this means that they are continous when evaluated on a given element (say, x) of the said space. However, we will omit x for the sake of simplicity (with a few exceptions). For instance, by a slight abuse of notation continuity of

$$F(t,\tau)=e^{(t-\tau)A}-\int_\tau^t R(t-s)e^{(s-\tau)A}\,ds$$

is to be interpreted as continuity of

$$t,\tau\longmapsto F(t,\tau)x=e^{(t-\tau)A}x-\int_\tau^t R(t-s)e^{(s-\tau)A}x\,ds\,.$$

The concise notation ∂_τ (∂_t, etc.) in place of $\frac{\partial}{\partial\tau}$ ($\frac{\partial}{\partial t}$, etc., respectively), will be adopted throughout.

1.4. **An overview of the paper.** To a large extent, the proof of Theorem 1.6 retraces the principal steps of the proofs carried out in the study of the LQ problem for memoryless control systems; see [28]. However, the solution formula corresponding to the integro-differential initial value problem naturally accounts for the more involved computations *at any step* of the line of investigation. Also, a difference – and technically challenging element – in comparison with the memoryless case stands in the fact that the optimal cost operator is not readily identified in a first formula which relates the optimal control to the optimal state, as it follows from the optimality condition.

Furthermore, recasting the system of quadratic equations satisfied by the optimal cost operators as a single Riccati-type matrix equation in the state space as well as proving uniqueness (these are the contents of Theorem 1.7) demand that nontrivial analytical matters are addressed and overcome.

In the following section, i.e. in Section 2, we derive an explicit representation for the solutions to the integro-differential problem in terms of the initial state X_0 (which actually comprises the component $w(\cdot)$ at an initial time and the memory up to it) and the control function $u(\cdot)$. The said representation, achieved by solving a certain Volterra equation of the second kind, generalizes the well-known *input-to-state formula* for (memoryless) linear control systems.

In Section 3 we prove the first two statements (namely, S1. and S2.) of Theorem 1.6. We not only infer existence of the unique optimal pair, but also pinpoint the transition properties fulfilled by both the optimal state and the optimal control.

In Section 4 we deal with the statements S3. and S4. of Theorem 1.6. We find readily that the representation of the optimal cost as a quadratic form involves three operators P_i, $i = 0, 1, 2$. Attaining certain alternative expressions for the said operators – as it is pursued in Lemma 4.4 – is a critical (and nontrivial) step in our analysis, since it enables us to establish a sought *closed-loop* form of the optimal control; see Proposition 4.5.

Section 5 is entirely devoted to the proof of statement S5., namely, to derive the coupled system of three differential equations satisfied by the operators P_i, $i = 0, 1, 2$.

Section 6 focuses on the proof of Theorem 1.7. Its existence and uniqueness parts rely upon two instrumental results (Proposition 6.1 and Lemma 6.2) which resolve respective technical points.

In Appendix A we gather several analytical results which are primarily utilized in the proofs of Lemma 4.4 and of the fundamental assertion S5. of Theorem 1.6, as well as for the question of uniqueness within Theorem 1.7.

2. Preliminaries. A representation formula for the solutions

The starting point for a study of the optimal control problem with quadratic functionals for (memoryless) linear differential systems of general form $y' = Ay + Bu$ is a representation formula for the *mild* solutions corresponding to an initial state $y_0 := y(\tau)$ and a control function $u(\cdot)$, that is

$$y(t) = e^{(t-\tau)A}y_0 + \int_\tau^t e^{(t-s)A}Bu(s)\, ds\,, \quad t \in [\tau, T)\,;$$

see [28]. Then, as it is well known, the analyses of boundary control systems governed by PDE split apart, depending on the distinct regularity properties of the (so called) *input-to-state map*

$$L\colon L^2(\tau, T; U) \ni u(\cdot) \longrightarrow (Lu(\cdot))(t) := \int_\tau^t e^{(t-s)A}Bu(s)\, ds$$

that occurs in the said formula, as well as of its adjoint, in accordance with a parabolic or hyperbolic character of the free dynamics.

In the present context, the following definition appears natural.

Definition 2.1. *We say that a function $w(t)$ is a mild solution to the Cauchy problem* (1.4) *corresponding to an initial state X_0 – that subsumes the value ξ_0 at time $\tau \in (0,T)$ and the past history $\xi(\cdot)$ on $[0,\tau)$ – and a control function $u(\cdot) \in L^2(\tau,T;U)$, if it belongs to $L^2(\tau,T;H)$ and satisfies the integral equation*

$$\begin{aligned} w(t,\tau,X_0) \equiv w(t) &= e^{(t-\tau)A}\xi_0 + \int_\tau^t e^{(t-s)A} \int_\tau^s K(s-\sigma)w(\sigma)\,d\sigma\,ds \\ &+ \int_\tau^t e^{(t-s)A} \int_0^\tau K(s-\sigma)\xi(\sigma)\,d\sigma\,ds + \int_\tau^t e^{(t-s)A} Bu(s)\,ds \end{aligned} \tag{2.1}$$

a.e. on $[\tau,T]$.

Lemma 2.2. *Under the Assumptions 1.1, let $w = w(\cdot)$ be a mild solution to the Cauchy problem* (1.4) *according with the Definition 2.1. Then $w \in C([\tau,T],H)$.*

Proof. It suffices to note that $w(\cdot) \in L^2(\tau,T;H)$ and $B \in \mathcal{L}(U,H)$. Then all the integrals in the right hand side of (2.1) yield continuous functions on $[\tau,T]$. $\square$

The theory of linear Volterra equations allows to achieve a representation formula for the mild solutions to the family of Cauchy problems (1.4) (depending on the parameter τ).

Proposition 2.3. *For any X_0 as in* (1.3) *and $u \in L^2(\tau,T;U)$, the (controlled) integro-differential problem* (1.4) *admits a unique mild solution $w = w(t)$, given by*

$$w(t) = w(t;\tau,X_0) = F(t,\tau)\xi_0 + \int_0^\tau M(t,\sigma,\tau)\xi(\sigma)\,d\sigma + \int_\tau^t F(t,s)Bu(s)\,ds,\; t \in [\tau,T]\,, \tag{2.2}$$

where

$$F(t,\tau) := e^{(t-\tau)A} - \int_\tau^t R(t-s)e^{(s-\tau)A}\,ds\,, \qquad \tau \le t, \tag{2.3a}$$

$$M(t,\sigma,\tau) := G(t,\sigma,\tau) - \int_\tau^t R(t-s)G(s,\sigma,\tau)\,ds\,, \qquad \sigma \le \tau \le t\,, \tag{2.3b}$$

$$G(t,\sigma,\tau) := \mu(t-\sigma) - e^{(t-\tau)A}\mu(\tau-\sigma)\,, \qquad \sigma \le \tau \le t, \tag{2.3c}$$

$$\mu(t) := \int_0^t e^{(t-s)A}K(s)\,ds\,, \tag{2.3d}$$

while $R(t)$ is the unique solution to the Volterra equation (of the second kind)

$$R(t) - \int_0^t \mu(t-s)R(s)\,ds = -\mu(t)\,, \quad t > 0, \tag{2.4}$$

explicitly given by

$$R(t) = -\mu(t) - \int_0^t \mu(t-s)\mu(s)\,ds - \int_0^t \mu(t-s)\int_0^\sigma \mu(\sigma-s)\mu(s)\,ds\,d\sigma - \dots\,. \tag{2.5}$$

Proof. We limit ourselves to a brief outline of the argument; the full details are omitted. With (2.1) as a starting point, let $w = w(t)$ be an L^2-in time solution. The key is to attain a reformulation of the original (2.1) as a simple Volterra equation of the form

$$w(t) - \int_\tau^t \mu(t-\sigma)w(\sigma)\,d\sigma = \mathcal{F}(t)\,, \tag{2.6}$$

with μ as in (2.3d) and $\mathcal{F}(t)$ depending on the initial datum, the past history and the control. More precisely, we have

$$\mathcal{F}(t) := e^{(t-\tau)A}\xi_0 + \int_0^\tau G(t,\sigma,r)\xi(\sigma)\,d\sigma + \int_\tau^t e^{(t-s)A}Bu(s)\,ds\,. \tag{2.7}$$

The conclusion follows recalling that the solution to the Volterra equation (2.6) is given by

$$w(t) = \mathcal{F}(t) - \int_\tau^t R(t-s)\mathcal{F}(s)\,ds\,, \tag{2.8}$$

where $R(t)$ is the resolvent kernel of (2.6), namely, the unique solution to the integral equation (2.4); see e.g. [13, Chapter 5]. A simple verification confirms that the function in (2.2) is indeed the unique mild solution of (2.1). □

3. The unique optimal pair. Transition properties, regularity, statements S1. and S2. of Theorem 1.6

We begin this section by proving that, given $\tau \in (0,T)$, every solution $w(t;\tau,X_0)$ to the integro-differential problem (1.4) satisfies a transition property, just like in the memoryless case.

Next, we show that the cost functional (1.5) is a quadratic form in the space $Y_\tau = H \times L^2(0,\tau;H)$. This brings about the existence of a unique optimal control $\hat{u}(t,\tau,X_0)$, along with a first (pointwise in time) representation of $\hat{u}(\cdot)$ in dependence on the component $\hat{w}(\cdot)$ of the optimal state. The latter is a straightforward outcome of the optimality condition.

A further analysis allows then to prove that the optimal control inherits a transition property from the optimal state, as well.

3.1. Transition property for the state variable.

Let $w(t) := w(t;\tau,X_0)$, $t \in [\tau,T]$, be the mild solution to the integro-differential equation (1.4) corresponding to an initial state X_0 and a control function $u(\cdot)$. For $\tau_1 \in (\tau,t)$, define

$$X_1 = \begin{pmatrix} y_0 \\ y(\cdot) \end{pmatrix}, \quad \text{with} \quad y_0 = w(\tau_1^+)\,, \quad y(\cdot) = \begin{cases} \xi(\cdot) & \text{in } [0,\tau] \\ w(\cdot,\tau,X_0) & \text{in } (\tau,\tau_1]. \end{cases} \tag{3.1}$$

Then, the following result holds true.

Proposition 3.1. *The following transition property*

$$w(t;\tau,X_0) = w(t;\tau_1,X_1) \qquad \forall t \in (\tau_1,T)$$

is valid.

Proof. The function $w_1(t) := w(t;\tau_1,X_1)$ solves the initial value problem

$$\begin{cases} w_1'(t) = Aw_1(t) + \displaystyle\int_{\tau_1}^t K(t-s)w_1(s)\,ds + \int_0^{\tau_1} K(t-s)y(s)\,ds + Bu(t) \\ \qquad\;\; = Aw_1(t) + Bu(t) + \displaystyle\int_{\tau_1}^t K(t-s)w_1(s)\,ds + \int_0^\tau K(t-s)\xi(s)\,ds \\ \qquad\qquad + \displaystyle\int_\tau^{\tau_1} K(t-s)w(s)\,ds \\ w_1(\tau_1^+) = y_0 = w(\tau_1^+) \end{cases} \tag{3.2}$$

(in fact its mild form). Thus the function $z(t)=w_1(t)-w(t)$ is a mild solution to

$$\begin{cases} z'(t)=Az(t)+\displaystyle\int_{\tau_1}^t K(t-s)w_1(s)\,ds-\int_\tau^t K(t-s)w(s)\,ds+\int_\tau^{\tau_1} K(t-s)w(s)\,ds \\ \qquad = Az(t)+\displaystyle\int_{\tau_1}^t K(t-s)z(s)\,ds \\ z(\tau_1^+)=0\,; \end{cases}$$

then

$$z(t)=\int_{\tau_1}^t e^{(t-\sigma)A}\int_{\tau_1}^\sigma K(\sigma-s)z(s)\,ds\,d\sigma=\int_{\tau_1}^t\Big[\int_s^t e^{(t-\sigma)A}K(\sigma-s)\,d\sigma\Big]\,z(s)\,ds\,.$$

Therefore, since by semigroup theory $\|e^{rA}\|_{\mathcal{L}(H)}\le C\,e^{\omega r}$ for some $C\ge 1$, $\omega\in\mathbb{R}$ and any $r\ge 0$, we get

$$\|z(t)\|_H\le C\,e^{\omega T}\|K(\cdot)\|_{L^1(0,T;\mathbb{R})}\int_{\tau_1}^t\|z(s)\|_H\,ds\,,\quad t\in[\tau_1,T]\,,$$

which implies $\|z(t)\|_H\equiv 0$ on $[\tau_1,T]$ by the Gronwall Lemma. The conclusion $w_1\equiv w$ on $[\tau_1,T]$ follows.

□

3.2. **The optimal pair. Proof of the statement S1.** Recall the representation formula (2.2) for the solution to the Cauchy problem (1.4). Just like in the study of memoryless control systems, it is useful to introduce the operator $L_\tau\colon L^2(\tau,T;U)\longrightarrow L^2(\tau,T;H)$ defined as follows,

$$[L_\tau u(\cdot)](t):=\int_\tau^t F(t,\sigma)Bu(\sigma)\,d\sigma\,,\qquad t\in[\tau,T]\,,\tag{3.3}$$

along with its adjoint $L_\tau^*\colon L^2(\tau,T;H)\longrightarrow L^2(\tau,T;U)$ that is deduced readily:

$$[L_\tau^* g(\cdot)](\sigma):=\int_\sigma^T B^*F(t,\sigma)^*g(t)\,dt\,,\qquad \sigma\in[\tau,T]\,.\tag{3.4}$$

We will use the abbreviated notations $[L_\tau u](t)$ and even the neat $L_\tau u(t)$, in place of $[L_\tau u(\cdot)](t)$, etc. Then, (2.2) reads as

$$w(t;\tau,X_0)\equiv w(t)=F(t,\tau)\xi_0+\int_0^\tau M(t,\sigma,\tau)\xi(\sigma)\,d\sigma+L_\tau u(t)\,.\tag{3.5}$$

By inserting the expression (3.5) of $w(t)$ in the cost functional (1.5), we obtain readily

$$J_\tau(u,X_0)=\langle\mathcal{M}_\tau X_0,X_0\rangle_{Y_\tau}+2\mathrm{Re}\,\langle N_\tau X_0,u\rangle_{L^2(\tau,T;U)}+\langle\Lambda_\tau u,u\rangle_{L^2(\tau,T;U)}\,,\tag{3.6}$$

where

$$
\begin{aligned}
\langle \mathcal{M}_\tau X_0, X_0\rangle_{Y_\tau} &:= \int_\tau^T \langle QE(t,\tau)X_0, E(t,\tau)X_0\rangle_H\, dt = \int_\tau^T \langle QF(t,\tau)\xi_0, F(t,\tau)\xi_0\rangle_H\, dt \\
&\quad + 2\mathrm{Re} \int_\tau^T \left\langle QF(t,\tau)\xi_0, \int_0^\tau M(t,\sigma,\tau)\xi(\sigma)\, d\sigma \right\rangle_H dt \\
&\quad + \int_\tau^T \left\langle \int_0^\tau M(t,\sigma,\tau)\xi(\sigma)\, d\sigma, \int_0^\tau M(t,q,\tau)\xi(q)\, dq \right\rangle_H dt \\
N_\tau X_0 &:= \big[L_\tau^* QE(\cdot,\tau)X_0\big](\cdot) \\
\Lambda_\tau &:= I + L_\tau^* Q L_\tau\,,
\end{aligned}
\tag{3.7}
$$

having set

$$
E(t,\tau)X_0 := F(t,\tau)\xi_0 + \int_0^\tau M(t,\sigma,\tau)\xi(\sigma)\, d\sigma \tag{3.8}
$$

for the sake of brevity (although this abbreviated notation will seldom occur) and where I denotes the identity operator on $L^2(\tau,T;U)$.

A pretty standard argument is invoked now: notice that from the assumption $Q \geq 0$ it follows $\Lambda_\tau \geq I$; namely, the cost functional is coercive in the space $\mathcal{U}_\tau = L^2(\tau,T;U)$ of admissible controls, and hence there exists a unique optimal control minimizing the cost (1.5). The optimality condition

$$
\Lambda_\tau \hat{u} + N_\tau X_0 = 0
$$

yields, on the one side,

$$
\hat{u} = -\Lambda_\tau^{-1} N_\tau X_0\,. \tag{3.9}
$$

On the other side, recalling (3.7) and rewriting explicitly Λ_τ, we see that

$$
\hat{u} + L_\tau^* Q L_\tau \hat{u} + L_\tau^* QE(\cdot,\tau)X_0 = 0\,,
$$

that is

$$
\hat{u} = -L_\tau^* Q\big[E(\cdot,\tau)X_0 + L_\tau \hat{u}\big] = -L_\tau^* Q\hat{w}\,.
$$

This is nothing but a first representation of the optimal control in terms of the component $\hat{w}(\cdot)$ of the optimal state:

$$
\hat{u}(t;\tau,X_0) = -[L_\tau^* Q\hat{w}(\cdot,\tau;X_0)](t) = -\int_t^T B^* F(\sigma,t)^* Q\hat{w}(\sigma,\tau;X_0)\, d\sigma\,. \tag{3.10}
$$

We note that, as $\hat{w}(\cdot)$ is a continuous function in view of Lemma 2.2, formula (3.10) establishes that the optimal control is continuous in time as well. Thus, the statement S1. is proved.

3.3. **Transition property for the optimal pair. Proof of the statement S2.** In order to infer that the transition property fulfilled by the component $\hat{w}$ of the optimal state is inherited by the optimal control $\hat{u}$, we follow an argument which is pretty standard in the case of memoryless control systems.

Given $\tau_1 > \tau$, and with X_1 as in (3.1) (y_0 and $y(\cdot)$ are defined therein), we associate to the equation (1.4) with initial time τ_1 and initial state X_1 the cost functional $J_{\tau_1}(u,X_1)$. With $\hat{w}(\cdot,\tau;X_0)$ the first component of the optimal state of the original control problem, assuming for the time being that $\hat{w}(\cdot;\tau,X_0)$ restricted to $[\tau_1, T]$ is the first component of the optimal state for J_{τ_1} as well, then it follows from (3.10) that

$$\begin{aligned}\hat{u}(t;\tau_1,X_1) &= -\int_t^T B^*F(\sigma,t)^*Q\hat{w}(\sigma;\tau_1,X_1)\,d\sigma \\ &= -\int_t^T B^*F(\sigma,t)^*Q\hat{w}(\sigma;\tau,X_0)\,d\sigma = \hat{u}(\cdot;\tau,X_0)\,, \qquad \tau_1 < t \le T\,;\end{aligned}$$

to wit, the optimal control satisfies a transition property as well. In the following Lemma we prove that indeed the claimed condition holds true.

Lemma 3.2. *Let $(\hat{u}(\cdot;\tau,X_0),(\hat{w}(\cdot;\tau,X_0),\xi(\cdot)))$ be the optimal pair of problem* (1.4)-(1.5). *Then $(\hat{u}|_{[\tau_1,T]},(\hat{w}|_{[\tau_1,T]},y(\cdot)))$ is the optimal pair of the minimization problem with functional $J_{\tau_1}(u,X_1)$, where $y(\cdot)$ is given by* (3.1).

Proof. Consider the optimal control problem (1.4)-(1.5). By definition,

$$J_\tau(\hat{u},X_0) \le J_\tau(u,X_0) \qquad \forall u \in L^2(\tau,T;U)\,.$$

Given $u \in L^2(\tau_1,T;U)$, let w be the function which corresponds to the input u and to the initial state X_1 defined by (3.1): then w satisfies (the same Cauchy problem as (3.2))

$$\begin{cases} w'(t) - \displaystyle\int_{\tau_1}^t K(t-s)w(s)\,ds = Aw(t) + Bu(t) + \int_0^{\tau_1} K(t-s)y(s)\,ds \\ w_1(\tau_1) = \hat{w}(\tau_1)\,. \end{cases} \tag{3.11}$$

Let us introduce

$$\overline{u}(\cdot) = \begin{cases} \hat{u}(\cdot) & \text{in } [\tau,\tau_1] \\ u(\cdot) & \text{in } [\tau_1,T] \end{cases}, \qquad \overline{w}(\cdot) = \begin{cases} \hat{w}(\cdot) & \text{in } [\tau,\tau_1] \\ w(\cdot) & \text{in } [\tau_1,T] \end{cases}.$$

If $t \in [\tau,\tau_1]$, then $\overline{w}(\cdot)$ is such that

$$\begin{aligned}\overline{w}'(t) - \int_\tau^t K(t-s)\overline{w}(s)\,ds - A\overline{w}(t) &= \hat{w}'(t) - \int_\tau^t K(t-s)\hat{w}(s)\,ds - A\hat{w}(t) \\ &= B\hat{u}(t) + \int_0^\tau K(t-s)\xi(s)\,ds\,.\end{aligned} \tag{3.12}$$

When $t \in [\tau_1,T]$, one has

$$\begin{aligned}&\overline{w}'(t) - \int_\tau^t K(t-s)\overline{w}(s)\,ds - A\overline{w}(t) \\ &\quad = w'(t) - \int_\tau^{\tau_1} K(t-s)\hat{w}(s)\,ds - \int_{\tau_1}^t K(t-s)w(s)\,ds - Aw(t) \\ &\quad = -\int_\tau^{\tau_1} K(t-s)\hat{w}(s)\,ds + Bu(t) + \int_0^{\tau_1} K(t-s)y(s)\,ds \\ &\quad = Bu(t) + \int_0^\tau K(t-s)\xi(s)\,ds\,,\end{aligned} \tag{3.13}$$

instead. In view of (3.13) and (3.12), we find that $\overline{w}$ satisfies

$$\overline{w}'(t) - \int_\tau^t K(t-s)\overline{w}(s)\,ds - A\overline{w}(t) = B\overline{u}(t) + \int_0^\tau K(t-s)\xi(s)\,ds \quad \forall t \in [\tau, T]\,, \tag{3.14}$$

which means that $\overline{w}(\cdot)$ is the first component of the state corresponding to the control $\overline{u}(\cdot)$, with initial state X_0, in $[\tau, T]$. Therefore we have

$$J_\tau(\hat{u}, X_0) \le J_\tau(\overline{u}, X_0)\,,$$

which reads as

$$\int_\tau^T \big[\langle Q\hat{w}(t), \hat{w}(t)\rangle_H + \|\hat{u}(t)\|_U^2\big]\,dt \le \int_\tau^T \big[\langle Q\overline{w}(t), \overline{w}(t)\rangle_H + \|\overline{u}(t)\|_U^2\big]\,dt\,.$$

Deleting the integrals between τ and τ_1 in both sides, as $\hat{w}(\cdot)$ and $\overline{w}(\cdot)$ coincide on $[\tau, \tau_1]$, we obtain

$$\int_{\tau_1}^T \big[\langle Q\hat{w}(t), \hat{w}(t)\rangle_H + \|\hat{u}(t)\|_U^2\big]\,dt \le \int_{\tau_1}^T \big[\langle Qw(t), w(t)\rangle_H + \|u(t)\|_U^2\big]\,dt\,.$$

that is

$$J_{\tau_1}(\hat{u}|_{[\tau_1,T]}, X_1) \le J_{\tau_1}(u, X_1)\,.$$

Since $u(\cdot)$ was an arbitrarily chosen admissible control in $[\tau_1, T]$, it follows that $\hat{u}|_{[\tau_1,T]}$ is the optimal cost and $\hat{w}|_{[\tau_1,T]}$ is the first component of the optimal state for the minimization problem in the time interval $[\tau_1, T]$, whose functional is $J_{\tau_1}(u, X_1)$. This concludes the proof. □

In light of the above Lemma, the statement S2. of Theorem 1.6 is established.

Proposition 3.3. *The transition property*

$$\begin{cases} \hat{u}(t;\tau,X_0) = \hat{u}(t;\tau_1,X_1) \\ \hat{w}(t;\tau,X_0) = \hat{w}(t;\tau_1,X_1) \end{cases} \qquad \forall t \in (\tau_1, T)$$

holds true.

Now we introduce the (evolution) operator $\Phi(t,\tau)$ by setting, for $t > \tau$,

$$\Phi(t,\tau)X_0 := \begin{pmatrix} \hat{w}(t;\tau,X_0) \\ \hat{y}(\cdot) \end{pmatrix} \tag{3.15}$$

with $\hat{y}(\cdot)$ as in (1.7). It is seen immediately that $\Phi(t,\tau)\colon Y_\tau \longrightarrow Y_t$ and that it is a linear map, owing to (3.9) combined with (3.5). And notably, in view of Proposition 3.3, the transition property

$$\Phi(t,\tau) = \Phi(t,\tau_1)\,\Phi(\tau_1,\tau) \qquad \forall \tau \in (\tau, t)$$

holds true. Indeed, let $\tau < \tau_1 < t$. Set

$$X_0 = \begin{pmatrix} \xi_0 \\ \xi(\cdot) \end{pmatrix}\,, \quad X_1 = \Phi(\tau_1,\tau)X_0 = \begin{pmatrix} \hat{w}(\tau_1,\tau;X_0) \\ \hat{y}(\cdot) \end{pmatrix}$$

where $\hat{y}(\cdot)$ is defined by

$$\hat{y}(\cdot) = \begin{cases} \xi(\cdot) & \text{in } [0,\tau] \\ \hat{w}(\cdot,\tau,X_0) & \text{in } (\tau,\tau_1]. \end{cases} \tag{3.16}$$

Then, we have

$$\Phi(t,\tau_1)X_1 = \begin{pmatrix} \hat{w}(t,\tau_1;X_1) \\ z(\cdot) \end{pmatrix}\,, \quad \text{where} \quad z(\cdot) = \begin{cases} \hat{y}(\cdot) & \text{in } [0,\tau_1] \\ \hat{w}(\cdot,\tau_1;X_1) & \text{in } [\tau_1,t] \end{cases}$$

that is

$$z(\cdot) = \begin{cases} \xi & \text{in } [0,\tau] \\ \hat{w}(\cdot,\tau;X_0) & \text{in } [\tau,\tau_1] \\ \hat{w}(\cdot,\tau_1;X_1) & \text{in } [\tau_1,t]. \end{cases}$$

Thus, since in view of Proposition 3.3 $\hat{w}(\cdot,\tau_1;X_1) = \hat{w}(\cdot,\tau;X_0)$ in $[\tau_1,T]$, we infer

$$\Phi(t,\tau)X_0 = \begin{pmatrix} \hat{w}(t,\tau;X_0) \\ X(\cdot) \end{pmatrix}\,, \quad \text{where} \quad X(\cdot) = \begin{cases} \xi & \text{in } [0,\tau] \\ \hat{w}(\cdot,\tau;X_0) & \text{in } [\tau,t], \end{cases}$$

i.e. $X(\cdot) \equiv z$; namely,

$$\Phi(t,\tau)X_0 = \Phi(t,\tau_1)X_1 = \Phi(t,\tau_1)\,\Phi(\tau_1,\tau)X_0\,, \qquad \forall \tau_1 \in (\tau,t)\,.$$

Furthermore, as shown above, $\hat{u}(t,\tau;X_0) = \hat{u}(t,\tau_1;\Phi(\tau_1,\tau)X_0)$ holds as well. The proof of the statement S2. is concluded.

4. An ensemble of optimal cost operators, the feedback formula. Statements S3. and S4. of Theorem 1.6

While the existence of a unique optimal control for the optimization problem (1.4)-(1.5) follows by a standard argument, we aim at providing a representation of the said optimal control in *feedback* form (*open- vs closed-loop* control). In order to achieve the intended goal, we will establish a first representation formula, that is (4.3) below. This can be done rather easily, starting from the formula (3.10) that connects the optimal control $\hat{u}$ to $\hat{w}$ (a consequence of the optimality condition), and next taking advantage of the transition property satisfied by the optimal pair. This analysis is carried out in Section 4.1.

To single out within the said representation certain operators ($P_0(\tau)$ and $P_1(t,\tau)$) that also occur in the quadratic form that yields the optimal cost, derived in section 4.2, further computations are necessitated. This is eventually explored and achieved in section 4.3.

4.1. A first representation of the optimal control in terms of the optimal state.

The relation (3.9) between the optimal control and the initial state, following from the optimality condition, can be rendered more explicit, via a representation formula that will play a crucial role in the sequel. Indeed, we have

$$\begin{aligned} \hat{u}(t) &= -\big[\Lambda_\tau^{-1}N_\tau X_0\big](t) = -\Lambda_\tau^{-1}L_\tau^* Q\Big[F(\cdot,\tau)\xi_0 + \int_0^\tau M(\cdot,\sigma,\tau)\xi(\sigma)\,d\sigma\Big](t) \\ &= -\left[\begin{pmatrix} \Lambda_\tau^{-1}L_\tau^* QF(\cdot,\tau) \\ \Lambda_\tau^{-1}L_\tau^* QM(\cdot,:,\tau) \end{pmatrix} \begin{pmatrix} \xi_0 \\ \xi(:) \end{pmatrix}\right](t)\,. \end{aligned}$$

Thus, by introducing the notation

$$\Psi_1(t,\tau) := -\Big[\Lambda_\tau^{-1}L_\tau^* QF(\cdot,\tau)\Big](t)\,, \tag{4.1a}$$

$$\Psi_2(t,\sigma,\tau) := -\Big[\Lambda_\tau^{-1}L_\tau^* QM(\cdot,\sigma,\tau)\Big](t)\,, \tag{4.1b}$$

we finally attain

$$\hat{u}(t) = \Psi_1(t,\tau)\xi_0 + \int_0^\tau \Psi_2(t,\sigma,\tau)\xi(\sigma)\,d\sigma\,. \tag{4.2}$$

Lemma 4.1. *Let* $(\hat{u}(\cdot,\tau;X_0),(\hat{w}(\cdot,\tau;X_0),\xi(\cdot)))$ *– $(\hat{u},(\hat{w},\xi))$, in short – be the optimal pair for the minimization problem* (1.4)-(1.5)*, with initial state X_0 (as in* (1.3)*). Then,*

$$\begin{aligned}\hat{u}(t,\tau;X_0) = &-\int_\tau^T B^*F(\sigma,t)^*QZ_1(\sigma,t)\hat{w}(t,\tau;X_0)\,d\sigma\\ &-\int_\tau^T B^*F(\sigma,t)^*Q\int_0^t Z_2(\sigma,s,t)\hat{y}(s)\,ds\,d\sigma\,,\end{aligned} \tag{4.3}$$

where $\hat{y}(\cdot)$ is given by (1.7)*, and having set for $\tau\le s\le t$*

$$Z_1(t,\tau) := F(t,\tau) + \int_\tau^t F(t,\sigma)B\Psi_1(\sigma,\tau)\,d\sigma\,, \tag{4.4a}$$

$$Z_2(t,s,\tau) := M(t,s,\tau) + \int_\tau^t F(t,\sigma)B\Psi_2(\sigma,s,\tau)\,d\sigma \tag{4.4b}$$

(with Ψ_i, $i=1,2$, as in (4.1)*).*

Proof. We rewrite the representation formula (2.2) for the mild solutions to the integro-differential problem, with $\hat{w}$ and $\hat{u}$ in place of w and u, respectively:

$$\hat{w}(t,\tau;X_0) = F(t,\tau)\xi_0 + \int_0^\tau M(t,\sigma,\tau)\xi(\sigma)\,d\sigma + \int_\tau^t F(t,\sigma)B\hat{u}(\sigma)\,d\sigma\,.$$

Next, taking into account the expression (4.2) of the optimal control $\hat{u}$, we obtain

$$\begin{aligned}\hat{w}(t,\tau;X_0) = F(t,\tau)\xi_0 &+ \int_0^\tau M(t,\sigma,\tau)\xi(\sigma)\,d\sigma\\ &+\int_\tau^t F(t,\sigma)B\Big[\Psi_1(\sigma,\tau)\xi_0 + \int_0^\tau \Psi_2(\sigma,s,\tau)\xi(s)\,ds\Big]\,d\sigma\,,\end{aligned}$$

which becomes

$$\hat{w}(t,\tau;X_0) = Z_1(t,\tau)\xi_0 + \int_0^\tau Z_2(t,s,\tau)\xi(s)\,ds\,, \tag{4.5}$$

by making use of the novel functions $Z_1(t,\tau)$ and $Z_2(t,s,\tau)$ defined in (4.4). (Just note the similar structure of the representations (4.2) and (4.5).)

Thus, recalling the expression (3.10) of the optimal control $\hat{u}(t;\tau,X_0)$ in terms of the component $\hat{w}(\cdot,\tau;X_0)$ of the optimal state (that follows from the optimality condition), we arrive at

$$\begin{aligned}\hat{u}(t,\tau;X_0) &= -\int_t^T B^*F(\sigma,t)^*Q\hat{w}(\sigma,\tau;X_0)\,d\sigma\\ &= -\int_t^T B^*F(\sigma,t)^*Q\hat{w}(\sigma,t;\Phi(t,\tau)X_0)\,d\sigma\\ &= -\int_t^T B^*F(\sigma,t)^*QZ_1(\sigma,t)\hat{w}(t,\tau;X_0)\,d\sigma\\ &\qquad -\int_t^T B^*F(\sigma,t)^*Q\int_0^t Z_2(\sigma,s,t)\hat{y}(s)\,ds\,d\sigma\,,\end{aligned}$$

which establishes (4.3). We just observe that in the second equality we have used the transition property fulfilled by the optimal state, while in the third one we appealed to the representation (4.5) of the component $\hat{w}(\sigma,t;\Phi(t,\tau)X_0)$ of the optimal state in terms of the initial state $\Phi(t,\tau)X_0$, which reads as

$$\begin{aligned}\hat{w}(\sigma,t;\Phi(t,\tau)X_0) &= Z_1(\sigma,t)\hat{w}(t,\tau;X_0)+\int_0^t Z_2(\sigma,s,t)\hat{y}(s)\,ds\\ &= Z_1(\sigma,t)\hat{w}(t,\tau;X_0)+\int_0^\tau Z_2(\sigma,s,t)\xi(s)\,ds+\int_\tau^t Z_2(\sigma,s,t)\hat{w}(s,\tau;X_0)\,ds\,.\end{aligned}$$

□

4.2. **The optimal cost operators. Proof of statement S3.** We now use the formulas (4.5) and (4.2) – with Ψ_i, Z_i, $i=1,2$, defined in (4.1) and (4.4), respectively – to compute the optimal value of the cost functional (1.5):

$$\begin{aligned}J_\tau(u,X_0) &= \int_\tau^T \big[\langle Q\hat{w}(\sigma,\tau;X_0),\hat{w}(\sigma,\tau;X_0\rangle_H+\|\hat{u}(\sigma,\tau;X_0)\|_U^2\big]\,d\sigma\\ &= \int_\tau^T \Big\{\Big\|Q^{1/2}\Big[Z_1(\sigma,\tau)\xi_0+\int_0^\tau Z_2(\sigma,s,\tau)\xi(s)\,ds\Big]\Big\|_H^2\\ &\qquad +\Big\|\Psi_1(\sigma,\tau)\xi_0+\int_0^\tau \Psi_2(\sigma,s,\tau)\xi(s)\,ds\Big\|_U^2\Big\}\,d\sigma\,.\end{aligned} \tag{4.6}$$

We replicate the formula, with $\Phi(\tau_1,\tau)X_0$ and $\hat{y}(\cdot)$ as in (3.16) in place of X_0 and $\xi(\cdot)$, respectively, to find

$$\begin{aligned}J_{\tau_1}(\hat{u},\Phi(\tau_1,\tau)X_0) &= \int_{\tau_1}^T \Big\{\Big\|Q^{1/2}\Big[Z_1(\sigma,\tau_1)\hat{w}(\tau_1,\tau;X_0)+\int_0^{\tau_1} Z_2(\sigma,s,\tau_1)\hat{y}(s)\,ds\Big]\Big\|_H^2\\ &\qquad +\Big\|\Psi_1(\sigma,\tau_1)\hat{w}(\tau_1,\tau;X_0)+\int_0^{\tau_1} \Psi_2(\sigma,s,\tau_1)\hat{y}(s)\,ds\Big\|_U^2\Big\}\,d\sigma\,.\end{aligned} \tag{4.7}$$

Computing the squares, we find

$$\begin{aligned}J_{\tau_1}(\hat{u},\Phi(\tau_1,\tau)X_0) = \int_{\tau_1}^T &\Big\{\Big\|Q^{1/2}Z_1(\sigma,\tau_1)\hat{w}(\tau_1,\tau;X_0)\Big\|_H^2\\ &+2\mathrm{Re}\Big\langle QZ_1(\sigma,\tau_1)\hat{w}(\tau_1,\tau;X_0),\int_0^{\tau_1} Z_2(\sigma,s,\tau_1)\hat{y}(s)\,ds\Big\rangle_H\\ &+\Big\|Q^{1/2}\int_0^{\tau_1} Z_2(\sigma,s,\tau_1)\hat{y}(s)\,ds\Big\|_H^2+\Big\|\Psi_1(\sigma,\tau_1)\hat{w}(\tau_1,\tau;X_0)\Big\|_U^2\\ &+2\mathrm{Re}\Big\langle \Psi_1(\sigma,\tau_1)\hat{w}(\tau_1,\tau;X_0),\int_0^{\tau_1} \Psi_2(\sigma,s,\tau_1)\hat{y}(s)\,ds\Big\rangle_U\\ &+\Big\|\int_0^{\tau_1} \Psi_2(\sigma,s,\tau_1)\hat{y}(s)\,ds\Big\|_U^2\Big\}\,d\sigma\,.\end{aligned}$$

Suitably rearranging the summands and setting

$$P_0(t)=\int_t^T \big[\Psi_1(p,t)^*\Psi_1(p,t)+Z_1(p,t)^*QZ_1(p,t)\big]\,dp\,, \tag{4.8a}$$

$$P_1(t,s)=\int_t^T \big[\Psi_1(p,t)^*\Psi_2(p,s,t)+Z_1(p,t)^*QZ_2(p,s,t)\big]\,dp \tag{4.8b}$$

$$P_2(t,s,q)=\int_t^T \big[\Psi_2(p,q,t)^*\Psi_2(p,s,t)+Z_2(p,q,t)^*QZ_2(p,s,t)\big]\,dp\,, \tag{4.8c}$$

we obtain

$$\begin{aligned}J_{\tau_1}(\hat u,\Phi(\tau_1,\tau)X_0)&=\big\langle P_0(\tau_1)\hat w(\tau_1,\tau;X_0),\hat w(\tau_1,\tau;X_0)\big\rangle_H\\&\quad+2\mathrm{Re}\int_0^{\tau_1}\big\langle P_1(\tau_1,s)\hat y(s),\hat w(\tau_1,\tau;X_0)\big\rangle_H\,ds\\&\quad+\int_0^{\tau_1}\int_0^{\tau_1}\big\langle P_2(\tau_1,s,q)\hat y(s),\hat y(q)\big\rangle\,ds\,dq\,,\end{aligned}$$

which is rewritten as

$$J_{\tau_1}(\hat u,\Phi(\tau_1,\tau)X_0)=\big\langle P(\tau_1)\Phi(\tau_1,\tau)X_0,\Phi(\tau_1,\tau)X_0\big\rangle_{Y_{\tau_1}}\,, \tag{4.9}$$

with $P(\cdot)$ the operator defined by (1.12).

If we set now $\tau_1=\tau$ in (4.9), we finally establish the sought representation (1.8) of the optimal cost as a quadratic form on the state space Y_τ.

Remark 4.2. The matrix operator $P(\cdot)$ is the 'want-to-be' Riccati operator of the optimal control problem, namely, a candidate solution to an appropriate Riccati equation in a dense subspace of the space Y_τ. This constitutes a first part of the statement of Theorem 1.7. It will be proved rigorously in Section 6, on the basis of the assertion S5. of Theorem 1.6 and in the light of Proposition 6.1.

We note that the properties

$$P_0(t)=P_0(t)^*\ge 0,\quad P_2(t,s,q)=P_2(t,q,s)=P_2(t,s,q)^*\ge 0\,. \tag{4.10}$$

are intrinsic to the respective definitions (4.8a) and (4.8c) of the operators P_0 and P_2.

It is worth emphasizing at the outset and explicitly the basic regularity properties of the optimal cost operators P_i ($i=1,2,3$).

Proposition 4.3. *The operators $P_0(t)$, $P_1(t,s)$, $P_2(t,s,q)$ defined in* (4.8) *possess the following regularity:*

- *for every $t\in[0,T]$ and $s,q\in[0,t]$, $P_0(t)$, $P_1(t,s)$, $P_2(t,s,q)$ belong to $\mathcal{L}(H)$, with respective norms bounded by some constant c;*
- *$P_0(t)$, $P_1(t,s)$, $P_2(t,s,q)$ are strongly continuous functions (with respect to their variables).*

Proof. The assertions follow given the respective definitions of $P_0(t)$, $P_1(t,s)$ and $P_2(t,s,q)$, after a careful analysis of the strong continuity properties of the operators Ψ_i, $i=1,2$ (see (4.1)), Z_i, $i=1,2$ (see (4.4), F, M (and G, R, μ; see (2.3a), (2.3b), (2.3c), (2.3d), (2.4)). We omit the details. □

4.3. **The feedback representation of the optimal control. Proof of statement S4.** In the previous subsections we derived

- a first (pointwise in time) representation of the optimal control in terms of the optimal state, that is (4.3);
- the representation of the optimal cost $J_\tau(\hat{u})$ as the quadratic form (1.8) in the space Y_τ.

However, differently from the case of the LQ problem for memoryless equations, one cannot single out readily the presence of the operator $P(\cdot) \in \mathcal{L}(Y_\tau)$ associated with this quadratic form within the formula (4.3). To disclose the said presence, the following result is critical.

Lemma 4.4 (Key Lemma)**.** *With the functions $\Psi_1(\sigma,t)$ and $\Psi_2(\sigma,s,t)$ defined in* (4.1)*, and $Z_1(\sigma,t)$ and $Z_2(\sigma,s,t)$ defined in* (4.4)*, the following identities hold true:*

$$\begin{aligned}
\int_t^T F(\sigma,t)^* Q Z_1(\sigma,t)\,d\sigma &= \int_t^T \big[Z_1(\sigma,t)^* Q Z_1(\sigma,t) + \Psi_1(\sigma,t)^*\Psi_1(\sigma,t)\big]\,d\sigma\,,\\
\int_t^T F(\sigma,t)^* Q Z_2(\sigma,s,t)\,d\sigma &= \int_t^T \big[Z_1(\sigma,t)^* Q Z_2(\sigma,s,t) + \Psi_1(\sigma,t)^*\Psi_2(\sigma,s,t)\big]\,d\sigma\,.\\
\int_t^T M(p,q,t)^* Q Z_2(p,s,t)\,dp &= \int_t^T \big[Z_2(p,q,t)^* Q Z_2(p,s,t) + \Psi_2(p,q,t)^*\Psi_2(p,s,t)\big]\,dp\,.
\end{aligned} \tag{4.11}$$

As a consequence, the optimal cost operators P_i, $i=1,2,3$ in (4.8) *admit the following respective representations, as well:*

$$P_0(t) = \int_t^T F(\sigma,t)^* Q Z_1(\sigma,t)\,d\sigma \tag{4.12a}$$

$$P_1(t,s) = \int_t^T F(\sigma,t)^* Q Z_2(\sigma,s,t)\,d\sigma \tag{4.12b}$$

$$P_2(t,s,q) = \int_t^T M(p,q,t)^* Q Z_2(p,s,t)\,dp\,. \tag{4.12c}$$

Proof. (i) In order to establish the first one of the identities (4.11), we subtract its right hand side from the left hand side, to get

$$\begin{aligned}
&\int_t^T \big[F(\sigma,t)^* - Z_1(\sigma,t)^*\big] Q Z_1(\sigma,t)\,d\sigma - \int_t^T \Psi_1(\sigma,t)^*\Psi_1(\sigma,t)\,d\sigma\\
&\quad = -\int_t^T \Big[\int_t^\sigma \Psi_1(q,t)^* B^* F(\sigma,q)^*\,dq\Big] Q Z_1(\sigma,t)\,d\sigma - \int_t^T \Psi_1(\sigma,t)^*\Psi_1(\sigma,t)\,d\sigma\\
&\quad = -\int_t^T \int_q^T \Psi_1(q,t)^* B^* F(\sigma,q)^* Q Z_1(\sigma,t)\,d\sigma dq - \int_t^T \Psi_1(q,t)^*\Psi_1(q,t)\,dq\\
&\quad = -\int_t^T \Psi_1(q,t)^* \Big[\big[L_t^* Q Z_1(\cdot,t)\big](q) + \Psi_1(q,t)\Big]\,dq\,,
\end{aligned}$$

where we made use of the definitions of Z_1 (see (4.4)) as well as the one of L_τ^* (see (3.4)).

With the last expression as a starting point we substitute once more the expression of Z_1 and move on with the computations, to find

$$\begin{aligned}
&\int_t^T \big[F(\sigma,t)^* - Z_1(\sigma,t)^*\big] Q Z_1(\sigma,t)\, d\sigma - \int_t^T \Psi_1(\sigma,t)^*\Psi_1(\sigma,t)\, d\sigma \\
&\quad = -\int_t^T \Psi_1(q,t)^* \big[L_t^* Q[F(\cdot,t) + L_t\Psi_1(\cdot,t)](q) + \Psi_1(q,t)\big]\, dq \\
&\quad = -\int_t^T \Psi_1(q,t)^* \big[\big(L_t^* Q F(\cdot,t)\big)(q) + [\Lambda_t \Psi_1(\cdot,t)](q)\big]\, dq \\
&\quad = -\int_t^T \Psi_1(q,t)^* \big[\big(L_t^* Q F(\cdot,t)\big)(q) - \big(L_t^* Q F(\cdot,t)\big)(q)\big]\, dq \equiv 0\,,
\end{aligned}$$

as desired. We note that in the last but one equality we recalled $L_t^* Q L_t + I =: \Lambda_t$, while in the last equality we utilized once again the definition of Ψ_1 in (4.1).

(ii) To prove the second one of the identities (4.11), we proceed in an analogous way, *mutatis mutandis*. A first series of passages leads to

$$\begin{aligned}
&\int_t^T \Big\{\big[F(\sigma,t)^* - Z_1(\sigma,t)^*\big] Q Z_2(\sigma,s,t) - \Psi_1(\sigma,t)^*\Psi_2(\sigma,s,t)\Big\}\, d\sigma \\
&\quad = -\int_t^T \Big\{\Big[\int_t^\sigma \Psi_1(q,t)^* B^* F(\sigma,q)^*\, dq\Big] Q Z_2(\sigma,s,t) + \Psi_1(\sigma,t)^*\Psi_2(\sigma,s,t)\Big\}\, d\sigma \\
&\quad = -\int_t^T \int_q^T \Psi_1(q,t)^* B^* F(\sigma,q)^* Q Z_2(\sigma,s,t)\, d\sigma dq - \int_t^T \Psi_1(q,t)^*\Psi_2(q,s,t)\, dq \\
&\quad = -\int_t^T \Psi_1(q,t)^* \Big[\big(L_t^* Q Z_2(\cdot,s,t)\big)(q) + \Psi_2(q,s,t)\Big]\, dq\,.
\end{aligned}$$

Similarly as before, we utilize $\Lambda_t := L_t^* Q L_t + I$ and the definitions of Ψ_2 and Z_2, to find

$$\begin{aligned}
&\int_t^T \big[F(\sigma,t)^* - Z_1(\sigma,t)^*\big] Q Z_2(\sigma,s,t)\, d\sigma - \int_t^T \Psi_1(\sigma,t)^*\Psi_2(\sigma,s,t)\, d\sigma \\
&\quad = -\int_t^T \Psi_1(q,t)^* \big[L_t^* Q[M(\cdot,s,t) + L_t\Psi_2(\cdot,s,t)](q) + \Psi_2(q,s,t)\big]\, dq \\
&\quad = -\int_t^T \Psi_1(q,t)^* \big[\big(L_t^* Q M(\cdot,s,t)\big)(q) + [\Lambda_t \Psi_2(\cdot,s,t)](q)\big]\, dq \\
&\quad = -\int_t^T \Psi_1(q,t)^* \big[\big(L_t^* Q M(\cdot,s,t)\big)(q) - \big(L_t^* Q M(\cdot,s,t)\big)(q)\big]\, dq \equiv 0\,.
\end{aligned}$$

(iii) Once again, we take the difference

$$\begin{aligned}
&\int_t^T \big[Z_2(p,q,t)^* Q Z_2(p,s,t) + \Psi_2(p,q,t)^*\Psi_2(p,s,t)\big]\, dp - \int_t^T M(p,q,t)^* Q Z_2(p,s,t)\, dp \\
&\quad = \int_t^T \Big[\int_t^p \Psi_2(\sigma,q,t)^* B^* F(p,\sigma)^* Q Z_2(p,s,t)\, d\sigma + \Psi_2(p,q,t)^*\Psi_2(p,s,t)\Big]\, dp \\
&\quad = \int_t^T \int_\sigma^T \Psi_2(\sigma,q,t)^* B^* F(p,\sigma)^* Q Z_2(p,s,t)\, dp\, d\sigma + \int_t^T \Psi_2(p,q,t)^*\Psi_2(p,s,t)\, dp\,.
\end{aligned}$$

Recall the definition of L_τ^* and move on with the computations to find

$$\begin{aligned}
&\int_t^T \big[Z_2(p,q,t)^* Q Z_2(p,s,t) + \Psi_2(p,q,t)^*\Psi_2(p,s,t)\big]\,dp - \int_t^T M(p,q,t)^* Q Z_2(p,s,t)\,dp \\
&\quad = \int_t^T \Psi_2(\sigma,q,t)^* \big[L_t^* Q Z_2(\cdot,s,t)\big](\sigma)\,d\sigma + \int_t^T \Psi_2(p,q,t)^*\Psi_2(p,s,t)\,dp \\
&\quad = \int_t^T \Psi_2(\sigma,q,t)^* \Big[L_t^* Q\big[M(\cdot,s,t) + L_t\Psi_2(\cdot,s,t)\big](\sigma) + \Psi_2(\sigma,s,t)\Big]\,d\sigma \\
&\quad = \int_t^T \Psi_2(\sigma,q,t)^* \Big[\big[L_t^* Q M(\cdot,s,t)\big](\sigma) + \big[\Lambda_t \Psi_2(\cdot,s,t)\big](\sigma)\Big]\,d\sigma \\
&\quad = \int_t^T \Psi_2(\sigma,q,t)^* \Big[\big[L_t^* Q M(\cdot,s,t)\big](\sigma) - \big[L_t^* Q M(\cdot,s,t)\big](\sigma)\Big]\,d\sigma \equiv 0\,,
\end{aligned}$$

as expected. (In the last two equalities, we used again $L_\tau^* Q L_\tau + I = \Lambda_t$ first and the representation of Ψ_2 in (4.1b) next.)

(iv) Thus, the formulae (4.12) follow combining the attained identities (4.11) with the original representations in (4.8). □

The reformulation (4.12) of the optimal cost operators allows for a derivation of the sought-after feedback representation of the optimal control.

Proposition 4.5. *Let $(\hat u(\cdot,\tau;X_0),(\hat w(\cdot,\tau;X_0),\xi(\cdot)))$ – $(\hat u,(\hat w,\xi))$, in short – be the optimal pair for the minimization problem* (1.4)-(1.5)*, with initial state X_0 (as in* (1.3)*). Then, the optimal control $\hat u$ admits the feedback representation* (1.9)*, that is*

$$\hat u(t,\tau;X_0) = -B^* P_0(t)\hat w(t;\tau,X_0) - \int_0^t B^* P_1(t,s)\hat y(s)\,ds\,, \qquad \tau \le t \le T\,,$$

with $\hat y(\cdot)$ given by (1.7)*.*

Proof. The validity of the feedback formula (1.9) now follows readily from the former representation (4.3) of the optimal control, in the light of the first two identities in (4.11). □

5. The coupled system of quadratic equations satisfied by the optimal cost operators. Statement S5. of Theorem 1.6

This section is entirely devoted to the proof of the crucial assertion **S5.** of Theorem 1.6, namely of the fundamental fact that the operators $P_0(t)$, $P_1(t,s)$, $P_2(t,s,q)$ – arisen as the 'building blocks' of the quadratic form defining the optimal cost (1.8), eventually shown to be given by (4.12), and with the former two of them entering the feedback formula (1.9) – solve a system of three coupled partial differential equations, that is (1.10).

We note that the system (1.10) cannot be recast immediately as a quadratic *equation* in the augmented space Y_τ, since the scalar products in the second and third equations of (1.10) involve only elements of $\mathcal{D}(A) \subset H$ rather than also functions with values in H. This subtle issue will be addressed later in the proof of Theorem 1.7.

Let $P_0(t)$, $P_1(t,s)$ and $P_2(t,s,q)$ be the operators defined in (4.8). The basic regularity of P_i, $i = 0,1,2$, is stated in Proposition 4.3. We will make use of the equivalent representations (4.12) obtained in Lemma 4.4 throughout. We additionally note that combining the definitions (4.4) of $Z_2(p,s,\tau)$ and (4.1) of $\Psi_2(\sigma,s,\tau)$, the following equivalent representation for $P_1(\tau,s)$ holds true,

$$P_1(\tau,s)=\int_\tau^T F(\sigma,\tau)^*Q\big[M(\sigma,s,\tau)-[L_\tau\Lambda_\tau^{-1}L_\tau^*QM(\cdot,s,\tau)](\sigma)\,\big]\,d\sigma\,, \tag{5.1}$$

which in turn yields, using (A.2) in Appendix A,

$$\begin{aligned} P_1(\tau,\tau)&=\int_\tau^T F(\sigma,\tau)^*Q\big[[I-L_\tau\Lambda_\tau^{-1}L_\tau^*]QM(\cdot,\tau,\tau)\big](\sigma)\,d\sigma\\ &=-\int_\tau^T F(\sigma,\tau)^*Q\big[[I-L_\tau\Lambda_\tau^{-1}L_\tau^*]QR(\cdot-\tau)\big](\sigma)\,d\sigma\,. \end{aligned} \tag{5.2}$$

This information will be essential later in the computations that follow.

Proof of the statement S5. of Theorem 1.6.

i) Equation satisfied by P_0. We start from $P_0(\tau)$ as in (4.12a): according to (1.10), we need to differentiate the quantity $\langle P_0(t)x,y\rangle_H$ for given $x,y\in D(A)$. We recall preliminarly that in view of Propositions A.5 and A.10 the operators $F(\sigma,\tau)$ and $Z_1(\sigma,\tau)$ can be differentiated with respect to τ, when acting on elements of $\mathcal{D}(A)$. In addition, and in particular, (A.6) of Propositions A.5 tells us that there exists $\partial_\tau\langle F(\sigma,\tau)^*x,y\rangle_H$ for $x\in H$ and $y\in\mathcal{D}(A)$. For the sake of simplicity we will neglect x and y throughout.

Using also Lemma A.1, the derivative with respect to τ is computed to be initially

$$\begin{aligned} P_0'(\tau)&=\frac{d}{d\tau}\int_\tau^T F(\sigma,\tau)^*QZ_1(\sigma,\tau)\,d\sigma\\ &=-Q+\int_\tau^T \partial_\tau F(\sigma,\tau)^*QZ_1(\sigma,\tau)\,d\sigma+\int_\tau^T F(\sigma,\tau)^*Q\partial_\tau Z_1(\sigma,\tau)\,d\sigma\\ &=-Q+\int_\tau^T \big[-A^*F(\sigma,\tau)^*+R(\sigma-\tau)^*\big]QZ_1(\sigma,\tau)\,d\sigma\\ &\qquad+\int_\tau^T F(\sigma,\tau)^*Q\big[\partial_\tau F(\sigma,\tau)-F(\sigma,\tau)B\Psi_1(\tau,\tau)\big]\,d\sigma\\ &\qquad\qquad+\int_\tau^T F(\sigma,\tau)^*Q\big[L_\tau\partial_\tau\Psi_1(\cdot,\tau)\big](\sigma)\,d\sigma\\ &=-Q-A^*P_0(\tau)+\int_\tau^T R(\sigma-\tau)^*QZ_1(\sigma,\tau)\,d\sigma\\ &\quad+\int_\tau^T F(\sigma,\tau)^*Q\big[-F(\sigma,\tau)A+R(\sigma-\tau)\big]\,d\sigma\\ &\quad-\int_\tau^T F(\sigma,\tau)^*QF(\sigma,\tau)B\Psi_1(\tau,\tau)\,d\sigma+\int_\tau^T F(\sigma,\tau)^*Q\big[L_\tau\partial_\tau\Psi_1(\cdot,\tau)\big](\sigma)\,d\sigma\,. \end{aligned}$$

So far, we used simply the formula in (A.6) for $\partial_\tau F(\sigma,\tau)^*$, as found in Proposition A.5, and singled out the term $-A^*P_0(\tau)$. We recall that the term $-A^*P_0(\tau)$ means in fact $\langle P_0(t)x,Ay\rangle_H$.

It turns out to be useful to replace $F(\sigma,\tau)$, that occurs in the fourth summand (in the right hand side), with its expression following from the definition of Z_1 in (4.4); at the same time, we use the formula (A.13) for $\partial_\tau\Psi_1(\cdot,\tau)$ established in

Proposition A.8 to rewrite the last summand. Thus, we obtain

$$
\begin{aligned}
P_0'(\tau) = -Q - A^*P_0(\tau) + \int_\tau^T R(\sigma-\tau)^* Q Z_1(\sigma,\tau)\big]\, d\sigma \\
&+ \underbrace{\int_\tau^T F(\sigma,\tau)^* Q\Big[-Z_1(\sigma,\tau)A}_{-P_0(\tau)A} + \big[L_\tau \Psi_1(\cdot,\tau)A\big](\sigma) + R(\sigma-\tau)\Big]\, d\sigma \\
&- \int_\tau^T F(\sigma,\tau)^* Q F(\sigma,\tau) B \Psi_1(\tau,\tau)\, d\sigma \\
&- \int_\tau^T F(\sigma,\tau)^* Q \big[L_\tau \Lambda_\tau^{-1} L_\tau^* Q F_\tau(\cdot,\tau)\big](\sigma)\, d\sigma \\
&+ \int_\tau^T F(\sigma,\tau)^* Q \big[L_\tau \Lambda_\tau^{-1} L_\tau^* Q F(\cdot,\tau) B \Psi_1(\tau,\tau)\big](\sigma)\, d\sigma\,.
\end{aligned}
$$

We move on with the computations substituting the expression (4.1a) of $\Psi_1(t,\tau)$ to find

$$
\begin{aligned}
P_0'(\tau) = -Q - A^*P_0(\tau) - P_0(\tau)A + \int_\tau^T R(\sigma-\tau)^* Q Z_1(\sigma,\tau)\, d\sigma \\
&- \int_\tau^T F(\sigma,\tau)^* Q \big[L_\tau \Lambda_\tau^{-1} L_\tau^* Q F(\cdot,\tau)A\big](\sigma)\, d\sigma + \int_\tau^T F(\sigma,\tau)^* Q R(\sigma-\tau)\, d\sigma \\
&- \int_\tau^T F(\sigma,\tau)^* Q \big[I - L_\tau \Lambda_\tau^{-1} L_\tau^* Q F(\cdot,\tau) B \Psi_1(\tau,\tau)\big](\sigma)\, d\sigma) \\
&- \int_\tau^T F(\sigma,\tau)^* Q \big[L_\tau \Lambda_\tau^{-1} L_\tau^* Q[-F(\cdot,\tau)A + R(\cdot-\tau)]\big](\sigma)\, d\sigma\,.
\end{aligned}
$$

We note first that the fifth summand in the right hand side cancels with a portion of the eighth summand, then recall that $R(\sigma-\tau) = -M(\sigma,\tau,\tau)$; thus, we get

$$
\begin{aligned}
P_0'(\tau) = -Q - A^*P_0(\tau) - P_0(\tau)A \\
&- \int_\tau^T M(\sigma,\tau,\tau)^* Q Z_1(\sigma,\tau)\, d\sigma - \int_\tau^T F(\sigma,\tau)^* Q M(\sigma,\tau,\tau)\, d\sigma \\
&+ \int_\tau^T F(\sigma,\tau)^* Q \big[I - L_\tau \Lambda_\tau^{-1} L_\tau^* Q F(\cdot,\tau)\big](\sigma) B \big[\Lambda_\tau^{-1} L_\tau^* Q F(\cdot,\tau)\big](\tau)\, d\sigma \\
&+ \int_\tau^T F(\sigma,\tau)^* Q \big[L_\tau \Lambda_\tau^{-1} L_\tau^* Q M(\cdot,\tau,\tau)]\big](\sigma)\, d\sigma\,.
\end{aligned}
$$

Recalling the representations (4.1b), (4.4) and (4.12b) for Ψ_2, Z_2 and P_1, respectively, we disclose the presence of $P_1(\tau,\tau)$ as the sum of the fifth and seventh summands. Therefore,

$$
\begin{aligned}
P_0'(\tau) &= -Q - A^*P_0(\tau) - P_0(\tau)A \\
&\quad - \int_\tau^T M(\sigma,\tau,\tau)^*Q\Big[F(\sigma,\tau) - \big[L_\tau\Lambda_\tau^{-1}L_\tau^*QF(\cdot,\tau)\big](\sigma)\Big]\,d\sigma - P_1(\tau,\tau) \\
&\quad + \int_\tau^T F(\sigma,\tau)^*Q\Big[\big[I - L_\tau\Lambda_\tau^{-1}L_\tau^*Q\big]F(\cdot,\tau)\big[B\Lambda_\tau^{-1}L_\tau^*QF(\cdot,\tau)\big](\tau)\Big](\sigma)\,d\sigma \\
&\quad + \int_\tau^T F(\sigma,\tau)^*Q\big[L_\tau\Lambda_\tau^{-1}L_\tau^*QM(\cdot,\tau,\tau)\big](\sigma)\,d\sigma\,,
\end{aligned}
$$

where the right hand side embeds the representation formula for $Z_1(t,\tau)$ in (A.9).

Taking into account that

$$
\begin{aligned}
&\int_\tau^T M(\sigma,\tau,\tau)^*Q\big[I - L_\tau\Lambda_\tau^{-1}L_\tau^*QF(\cdot,\tau)\big](\sigma)\,d\sigma \\
&\qquad\qquad = \int_\tau^T Z_2(\sigma,\tau,\tau)^*QF(\sigma,\tau)\,d\sigma = P_1(\tau,\tau)^*\,,
\end{aligned}
$$

we achieve

$$
\begin{aligned}
P_0'(\tau) &= -Q - A^*P_0(\tau) - P_0(\tau)A - P_1(\tau,\tau) - P_1(\tau,\tau)^* \\
&\quad \underbrace{+ \int_\tau^T F(\sigma,\tau)^*Q\Big[\big[I - L_\tau\Lambda_\tau^{-1}L_\tau^*\big]F(\cdot,\tau)B\big[\Lambda_\tau^{-1}L_\tau^*QF(\cdot,\tau)\big](\tau)\Big](\sigma)\,d\sigma}_{T(\tau)}\,.
\end{aligned}
$$

In order to complete this first part of the proof of the assertion S5. of Theorem 1.6, it remains to unveil that the term $T(\tau)$ coincides exactly with the quadratic term $P_0(\tau)BB^*P_0(\tau)$. We compute

$$
\begin{aligned}
&P_0(\tau)BB^*P_0(\tau) \\
&\quad = \int_\tau^T\int_\tau^T F(\sigma,\tau)^*Q\big[F(\sigma,\tau) + [L_\tau\Psi_1(\cdot,\tau)](\sigma)\big]BB^*\big[F(q,\tau)^*QF(q,\tau) \\
&\qquad\qquad\qquad + \Psi_1(q,\tau)^*[L_\tau^*QF(\cdot,\tau)](q)\big]\,dq\,d\sigma \\
&\quad = \int_\tau^T\int_\tau^T F(\sigma,\tau)^*Q\big[F(\sigma,\tau) - [L_\tau\Lambda_\tau^{-1}L_\tau^*QF(\cdot,\tau)](\sigma)\big]BB^*F(q,\tau)^*Q\big[F(q,\tau)- \\
&\qquad\qquad\qquad - [L_\tau\Lambda_\tau^{-1}L_\tau^*QF(\cdot,\tau)](q)\big]\,dq\,d\sigma \\
&\quad = \int_\tau^T F(\sigma,\tau)^*Q\big[[I - L_\tau\Lambda_\tau^{-1}L_\tau^*Q]F(\cdot,\tau)\big]B\Big[L_\tau^*Q\big[[I - L_\tau\Lambda_\tau^{-1}L_\tau^*Q]F(\cdot,\tau)\big](\tau)\Big](\sigma)\,d\sigma \\
&\quad = \int_\tau^T F(\sigma,\tau)^*Q\big[[I - L_\tau\Lambda_\tau^{-1}L_\tau^*Q]F(\cdot,\tau)\big]B\Big[\big[[L_\tau^*Q(\Lambda_\tau - I)\Lambda_\tau^{-1}L_\tau^*Q]F(\cdot,\tau)\big](\tau)\Big](\sigma)\,d\sigma \\
&\quad = \int_\tau^T F(\sigma,\tau)^*Q\big[[I - L_\tau\Lambda_\tau^{-1}L_\tau^*Q]F(\cdot,\tau)\big]B\Big[\big[\Lambda_\tau^{-1}L_\tau^*QF(\cdot,\tau)\big](\tau)\Big](\sigma)\,d\sigma \equiv T(\tau)\,,
\end{aligned}
$$

where in the last but one equality we have rewritten $L_\tau^*QL_\tau$ as $\Lambda_\tau - I$, and in the last one we have seen that

$$
L_\tau^*Q - (\Lambda_\tau - I)\Lambda_\tau^{-1}L_\tau^*Q = L_\tau^*Q - L_\tau^*Q + \Lambda_\tau^{-1}L_\tau^*Q = \Lambda_\tau^{-1}L_\tau^*Q\,.
$$

The argument is complete.

ii) Equation satisfied by P_1. Achieving the equation satisfied by the operator $P_1(\tau,s)$ is a bit trickier. Our starting point is the representation in (4.12b); we recall that the terms $F(\sigma,\tau)x$ and $Z_2(\sigma,s,\tau)x$ are differentiable with respect to τ, when $x\in\mathcal{D}(A)$, by Propositions A.5 and A.11. We compute the derivative (with respect to τ) of $P_1(\tau,s)$:

$$\begin{aligned}
\partial_\tau P_1(\tau,s) &= -F(\tau,\tau)^*QZ_2(\tau,s,\tau)+\int_\tau^T \partial_\tau F(\sigma,\tau)^*QZ_2(\sigma,s,\tau)\,d\sigma \\
&\quad+\int_\tau^T F(\sigma,\tau)^*Q\partial_\tau Z_2(\sigma,s,\tau)\,d\sigma \\
&=\int_\tau^T [-F(\sigma,\tau)A+R(\sigma-\tau)]^*QZ_2(\sigma,s,\tau)\,d\sigma \\
&\quad+\int_\tau^T F(\sigma,\tau)^*Q\big[\partial_\tau M(\sigma,s,\tau)-F(\sigma,\tau)B\Psi_2(\tau,s,\tau)+[L_\tau\partial_\tau\Psi_2(\cdot,s,\tau)](\sigma)\big]\,d\sigma\,,
\end{aligned}$$

where we used that $Z_2(\tau,s,\tau)=0$ (see (A.17) of Proposition A.11), and the representation (4.4) of $Z_2(\sigma,s,\tau)$.

Then, we see that

$$\begin{aligned}
\partial_\tau P_1(\tau,s) &= -A^*P_1(\tau,s)+\int_\tau^T R(\sigma-\tau)^*QZ_2(\sigma,s,\tau)\,d\sigma \\
&\quad-\int_\tau^T F(\sigma,\tau)^*QF(\sigma,\tau)K(\tau-s)\,d\sigma+\int_\tau^T F(\sigma,\tau)^*QF(\sigma,\tau)B\big[\Lambda_\tau^{-1}L_\tau^*QM(\cdot,s,\tau)\big](\tau)\,d\sigma \\
&\quad-\int_\tau^T F(\sigma,\tau)^*QL_\tau\Big[\big[\Lambda_\tau^{-1}L_\tau^*Q\partial_\tau M(\cdot,s,\tau)\big]-\Lambda_\tau^{-1}L_\tau^*QF(\cdot,\tau)B\Psi_2(\tau,s,\tau)\Big](\sigma)\,d\sigma\,;
\end{aligned}$$

it is important to emphasize as before that $A^*P_1(\tau,s)$ here means $\langle P_1(\tau,s)x,Ay\rangle_H$. Thus we have

$$\partial_\tau P_1(\tau,s)=\sum_{i=1}^6 S_i\,, \tag{5.3}$$

where it is immediately seen that

$$S_1=-A^*P_1(\tau,s)\,, \tag{5.4a}$$

$$S_2:=\int_\tau^T R(\sigma-\tau)^*QZ_2(\sigma,s,\tau)\,d\sigma=-P_2(\tau,s,\tau) \tag{5.4b}$$

(the latter equality follows recalling $M(\sigma,\tau,\tau)=-R(\sigma-\tau)$ from (A.2) in the Appendix A).

As for the summand S_3, we get

$$\begin{aligned}
S_3 &= -\int_\tau^T F(\sigma,\tau)^*QF(\sigma,\tau)\,d\sigma\,K(\tau-s) \\
&= -\int_\tau^T F(\sigma,\tau)^*Q\big\{Z_1(\sigma,\tau)-[L_\tau\Psi_1(\cdot,\tau)](\sigma)\big\}\,d\sigma\,K(\tau-s) \\
&= -P_0(\tau)K(\tau-s)+\int_\tau^T F(\sigma,\tau)^*Q\Big[L_\tau\big[-\Lambda_\tau^{-1}L_\tau^*QF(\cdot,\tau)\big]\Big](\sigma)\,d\sigma\,K(\tau-s) \\
&=: S_{31}+S_{32}\,.
\end{aligned} \tag{5.5}$$

Thus, we note that

$$\begin{aligned} S_5 &:= -\int_\tau^T F(\sigma,\tau)^* Q L_\tau \big[\Lambda_\tau^{-1} L_\tau^* Q \partial_\tau M(\cdot,s,\tau)\big](\sigma)\,d\sigma \\ &= \int_\tau^T F(\sigma,\tau)^* Q L_\tau \big[\Lambda_\tau^{-1} L_\tau^* Q F(\cdot,\tau) K(\tau-s)\big](\sigma)\,d\sigma \equiv -S_{32}\,, \end{aligned}$$

so that

$$S_{32} + S_5 = S_{32} - S_{32} = 0\,. \tag{5.6}$$

It remains to consider $S_4 + S_6$, where

$$\begin{aligned} S_4 &= \int_\tau^T F(\sigma,\tau)^* Q F(\sigma,\tau) B \big[\Lambda_\tau^{-1} L_\tau^* Q M(\cdot,s,\tau)\big](\tau)\,d\sigma \\ S_6 &= \int_\tau^T F(\sigma,\tau)^* Q \big[L_\tau \Lambda_\tau^{-1} L_\tau^* Q F(\cdot,\tau) B \Psi_2(\tau,s,\tau)\big](\sigma)\,d\sigma \\ &= \int_\tau^T F(\sigma,\tau)^* Q L_\tau \Lambda_\tau^{-1} \big[L_\tau^* Q F(\cdot,\tau) B[-\Lambda_\tau^{-1} L_\tau^* Q M(\cdot,s,\tau)](\tau)\big](\sigma)\,d\sigma \end{aligned}$$

so that

$$S_4 + S_6 = \Big[\int_\tau^T F(\sigma,\tau)^* Q \big[I - L_\tau \Lambda_\tau^{-1} L_\tau^* Q\big] F(\cdot,\tau)(\sigma)\,d\sigma\Big] B [\Lambda_\tau^{-1} L_\tau^* Q M(\cdot,s,\tau)](\tau)\,.$$

We prove now the following

Lemma 5.1.

$$S_4 + S_6 = P_0(\tau) B B^* P_1(\tau,s)\,. \tag{5.7}$$

Proof. To compute the term $P_0(\tau)BB^*P_1(\tau,s)$ we appeal once more to the representations (4.12a) and (4.12b) (of $P_0(\tau)$ and $P_1(\tau,s)$, respectively), this time making use of the formulas (A.9) for $Z_1(\sigma,t)$ and $Z_2(\sigma,s,t)$, respectively. This gives

$$\begin{aligned} P_0(\tau)BB^*P_1(\tau,s) &= \int_\tau^T F(\sigma,\tau)^* Q \big\{\big[I - L_\tau \Lambda_\tau^{-1} L_\tau^* Q\big] F(\cdot,\tau)\big\}(\sigma)\,d\sigma\, B \\ &\qquad \times B^* \int_\tau^T F(\sigma,\tau)^* Q \big\{\big[I - L_\tau \Lambda_\tau^{-1} L_\tau^* Q\big] M(\cdot,s,\tau)\big\}(q)\,dq \\ &= \int_\tau^T F(\sigma,\tau)^* Q \big\{\big[I - L_\tau \Lambda_\tau^{-1} L_\tau^* Q\big] F(\cdot,\tau)\big\}(\sigma)\,d\sigma\, B \\ &\qquad \times L_\tau^* Q \big\{\big[I - L_\tau \Lambda_\tau^{-1} L_\tau^* Q\big] M(\cdot,s,\tau)\big\}(\tau)\,, \end{aligned} \tag{5.8}$$

where the last equality is justified by the identification

$$\begin{aligned} &B^* \int_\tau^T F(\sigma,\tau)^* Q \big\{\big[I - L_\tau \Lambda_\tau^{-1} L_\tau^* Q\big] M(\cdot,s,\tau)\big\}(q)\,dq \\ &\qquad \equiv L_\tau^* Q \big\{\big[I - L_\tau \Lambda_\tau^{-1} L_\tau^* Q\big] M(\cdot,s,\tau)\big\}(\tau)\,, \end{aligned}$$

on the basis of the definition (3.4) of L_τ^*.

A key observation now is that the operator $L_\tau^* Q[I - L_\tau \Lambda_\tau^{-1} L_\tau^* Q]$ can be replaced by $\Lambda_\tau^{-1} L_\tau^* Q$: indeed,

$$\begin{aligned} L_\tau^* Q\big[I - L_\tau \Lambda_\tau^{-1} L_\tau^* Q\big] &= L_\tau^* Q - [L_\tau^* Q L_\tau]\Lambda_\tau^{-1} L_\tau^* Q = L_\tau^* Q - [\Lambda_\tau - I]\Lambda_\tau^{-1} L_\tau^* Q \\ &= \underbrace{L_\tau^* Q - L_\tau^* Q}_{\equiv 0} + \Lambda_\tau^{-1} L_\tau^* Q = \Lambda_\tau^{-1} L_\tau^* Q\,, \end{aligned} \tag{5.9}$$

which inserted in (5.8) shows

$$\begin{aligned} P_0(\tau)BB^* P_1(\tau,s) = \int_\tau^T F(\sigma,\tau)^* Q\big\{\big[I - L_\tau \Lambda_\tau^{-1} L_\tau^* Q\big]F(\cdot,\tau)\big\}(\sigma)\,d\sigma\, B \\ \times \big[\Lambda_\tau^{-1} L_\tau^* Q M(\cdot,s,\tau)\big](\tau)\,. \end{aligned}$$

But this is nothing but (5.7), which ends the proof of the lemma.

□

Now we are ready to return to (5.3), taking into account (5.4), (5.5), (5.6) and (5.7) of Lemma 5.1, to achieve the desired conclusion that $P_1(\tau,s)$ (more exactly, $\langle P_1(\tau,s)x,y\rangle_H$ with $x,y\in D(A)$) satisfies the equation

$$\partial_\tau P_1(\tau,s) = -A^* P_1(\tau,s) - P_2(\tau,s,\tau) - P_0(\tau)K(\tau-s) + P_0(\tau)BB^* P_1(\tau,s)\,,$$

just like in (1.10). Then, when acting on elements of $D(A)$,

$$\partial_\tau P_1(\tau,q)^* = -P_1(\tau,q)^* A - P_2(\tau,\tau,q) - K(\tau-q)P_0(\tau) + P_1(\tau,q)^* BB^* P_0(\tau) \tag{5.10}$$

as a consequence, since $P_0(\tau)$ and $P_2(t,s,q)$ are self-adjoint (see (4.10)) and we have $P_2(t,q,s) = P_2(t,s,q)$.

iii) Equation satisfied by P_2. On the basis of the representation (4.12c) of $P_2(\tau,s,q)$, we recall that in view of Propositions A.6 and A.11 the operators $M(p,q,\tau)$ and $Z_2(p,s,\tau)$ are differentiable with respect to τ (still when acting on elements of $\mathcal{D}(A)$). We begin the computation of the partial derivative $\partial_\tau P_2(\tau,s,q)$, obtaining first

$$\begin{aligned} \partial_\tau P_2(\tau,s,q) = \cancel{-M(\tau,q,\tau)^* Q Z_2(\tau,s,\tau)} + \int_\tau^T \partial_\tau M(p,q,\tau)^* Q Z_2(p,s,\tau)\,dp \\ + \int_\tau^T M(p,q,\tau)^* Q \partial_\tau Z_2(p,s,\tau)\,dp\,, \end{aligned}$$

where the first summand cancels, as $M(\tau,q,\tau) = 0$ (see Lemma A.1). Then we appeal to Proposition A.6 to find $\partial_\tau M(p,q,\tau) = -F(p,\tau)K(\tau-q)$, and to Proposition A.11 to compute

$$\begin{aligned} \partial_\tau Z_2(p,s,\tau) &= \partial_\tau M(p,s,\tau)x - F(p,\tau)B\Psi_2(\tau,s,\tau)x + \big[L_\tau \partial_\tau \Psi_2(\cdot,s,\tau)x\big](p) \\ &= \partial_\tau M(p,s,\tau) - F(p,\tau)B\Psi_2(\tau,s,\tau) + \int_\tau^p F(p,\sigma)B\Psi_2(\sigma,s,\tau)\,d\sigma \\ &= -F(p,\tau)K(\tau-s) - F(p,\tau)B\Psi_2(\tau,s,\tau) \\ &\quad + \int_\tau^p F(p,\sigma)B\big[\Lambda_\tau^{-1} L_\tau^* Q F(\cdot,\tau)K(\tau-s)\big](\sigma)\,d\sigma \\ &\quad + \int_\tau^p F(p,\sigma)B\big[\Lambda_\tau^{-1} L_\tau^* Q F(\cdot,\tau)B\Psi_2(\tau,s,\tau)\big](\sigma)\,d\sigma\,. \end{aligned}$$

Then

$$
\begin{aligned}
\partial_\tau P_2(\tau,s,q) &= -\int_\tau^T K(\tau-q)F(p,\tau)^*QM(p,s,\tau)\,dp \\
&- \int_\tau^T K(\tau-q)F(p,\tau)^*Q[L_\tau\Psi_2(\cdot,s,\tau)](p)\,dp \\
&- \int_\tau^T M(p,q,\tau)^*QF(p,\tau)K(\tau-s)\,dp - \int_\tau^T M(p,q,\tau)^*QF(p,\tau)B\Psi_2(\tau,s,\tau)\,dp \\
&+ \int_\tau^T M(p,q,\tau)^*Q\int_\tau^p F(p,\sigma)B\big[\Lambda_\tau^{-1}L_\tau^*QF(\cdot,\tau)K(\tau-s)\big](\sigma)\,d\sigma\,dp \\
&+ \int_\tau^T M(p,q,\tau)^*Q\int_\tau^p F(p,\sigma)B\big[\Lambda_\tau^{-1}L_\tau^*QF(\cdot,\tau)B\Psi_2(\tau,s,\tau)\big](\sigma)\,d\sigma\,dp \\
&= \sum_{i=1}^6 T_i\,.
\end{aligned}
$$

First we observe that

$$
\begin{aligned}
T_1+T_2 &= -K(\tau-q)\Big[\int_\tau^T F(p,\tau)^*Q\big[M(p,s,\tau)+[L_\tau\Psi_2(\cdot,s,\tau)](p)\big]\,dp\Big] \\
&= -K(\tau-q)P_1(\tau,s)\,;
\end{aligned}
\tag{5.11}
$$

then, we compute

$$
\begin{aligned}
T_3+T_5 &= -\int_\tau^T M(p,q,\tau)^*QF(p,\tau)\,dp\,K(\tau-s) \\
&\qquad + \int_\tau^T M(p,q,\tau)^*Q[L_\tau\Lambda_\tau^{-1}L_\tau^*QF(\cdot,\tau)](p)\,dp\,K(\tau-s) \\
&= -\Big[\int_\tau^T M(p,q,\tau)^*QF(p,\tau)\,dp - \int_\tau^T \Psi_2(p,q,\tau)^*[L_\tau^*QF(\cdot,\tau)](p)\,dp\Big]\,K(\tau-s) \\
&= -\int_\tau^T Z_2(p,q,\tau)^*QF(p,\tau)\,dp\,K(\tau-s)
\end{aligned}
$$

(where in the last two equalities we used the second and fourth of the (A.10), respectively), to find

$$
T_3+T_5 = -P_1(\tau,q)^*K(\tau-s)\,. \tag{5.12}
$$

It remains to pinpoint T_4+T_6. We have

$$
\begin{aligned}
T_4+T_6 &:= -\int_\tau^T M(p,q,\tau)^* Q F(p,\tau) B\Psi_2(\tau,s,\tau)\,dp \\
&\quad + \int_\tau^T M(p,q,\tau)^* Q\big[L_\tau\Lambda_\tau^{-1}L_\tau^* Q F(\cdot,\tau)B\Psi_2(\tau,s,\tau)\big](p)\,dp \\
&= \int_\tau^T M(p,q,\tau)^* Q\big[[I-L_\tau\Lambda_\tau^{-1}L_\tau^*Q]F(\cdot,\tau)B\Psi_2(\tau,s,\tau)\big](p)\,dp \\
&= -\int_\tau^T M(p,q,\tau)^* Q\big[[I-L_\tau\Lambda_\tau^{-1}L_\tau^*Q]F(\cdot,\tau)B[\Lambda_\tau^{-1}L_\tau^*QM(\cdot,s,\tau)](\tau)\big](p)\,dp \\
&= \int_\tau^T M(p,q,\tau)^* Q\Big[\big[I-L_\tau\Lambda_\tau^{-1}L_\tau^*Q\big]F(\cdot,\tau)B\big[L_\tau^*Q[I-L_\tau\Lambda_\tau^{-1}L_\tau^*Q]M(\cdot,s,\tau)\big](\tau)\Big](p)dp,
\end{aligned}
$$

where we used once again $\Psi_2(\tau,s,\tau)=-\big[\Lambda_\tau^{-1}L_\tau^*QM(\cdot,s,\tau)\big](\tau)$ first, while in the last equality we replaced $\Lambda_\tau^{-1}L_\tau^*Q$ by $L_\tau^*Q[I-L_\tau\Lambda_\tau^{-1}L_\tau^*Q]$ on the strenght of (5.9).

We insert the explicit meaning of the operator L_τ^* and carry on with the computations, to find

$$
\begin{aligned}
T_4+T_6 :=& \int_\tau^T M(p,q,\tau)^* Q\big[[I-L_\tau\Lambda_\tau^{-1}L_\tau^*Q\big]F(\cdot,\tau)(p)\big]\,dp \\
&\qquad \times BB^*\int_\tau^T F(\sigma,\tau)^*Q\big[[I-L_\tau\Lambda_\tau^{-1}L_\tau^*Q]M(\cdot,s,\tau)\big](\sigma)\,d\sigma \\
&= \int_\tau^T M(p,q,\tau)^*QF(p,\tau)\,dp - \int_\tau^T M(p,q,\tau)^*Q\big[L_\tau\Lambda_\tau^{-1}L_\tau^*QF(\cdot,\tau)\big](p)\,dp \\
&\qquad \times BB^*\int_\tau^T F(\sigma,\tau)^*Q\big[M(\sigma,s,\tau)+[L_\tau\Psi_2(\cdot,s,\tau)](\sigma)\big]\,d\sigma \\
&= \int_\tau^T Z_2(p,q,\tau)^*QF(p,\tau)\,dp\,BB^*\int_\tau^T F(\sigma,\tau)^*QZ_2(\sigma,s,\tau)(\sigma)\,d\sigma \\
&= P_1(\tau,q)^*BB^*P_1(\tau,s)\,.
\end{aligned}
\tag{5.13}
$$

Combining (5.13) with (5.11) and (5.12) we finally attain

$$
\partial_\tau P_2(\tau,s,q) = -K(\tau-q)P_1(\tau,s) - P_1(\tau,s)^*K(\tau-s) + P_1(\tau,q)^*BB^*P_1(\tau,s)\,,
$$

as desired. □

6. Proof of Theorem 1.7

1. (Existence). To infer the *existence* of (at least) an operator solution to the Riccati-type equation (1.12), it is natural to claim that it is the matrix operator $P(t)$ defined in (1.11) to solve (1.12) (in some dense subspace of $H\times L^2(0,t;H)$). This is indeed true. However, there is a subtle analytical gap that needs to be dealt with, beforehand. The task is fulfilled proving the following result in the first place.

Proposition 6.1. *A triple* (P_0, P_1, P_2) *is a solution to the coupled system* (1.10) *if and only if it solves the following one:*

$$\begin{cases} \frac{d}{dt}\langle P_0(t)x,y\rangle_H + \langle P_0(t)x, Ay\rangle_H + \langle Ax, P_0(t)y\rangle_H + \langle Qx,y\rangle_H \\ \qquad -\langle B^*P_0(t)x, B^*P_0(t)y\rangle_U + \langle P_1(t,t)x,y\rangle_H + \langle x, P_1(t,t)y\rangle_H = 0 \\ \frac{\partial}{\partial t}\langle P_1(t,\cdot)f(\cdot),y\rangle_{L^2(0,t;H)} + \langle P_1(t,\cdot)f(\cdot),Ay\rangle_{L^2(0,t;H)} \\ \quad +\langle K(t-\cdot)f(\cdot),P_0(t)y\rangle_{L^2(0,t;H)} + \langle P_2(t,\cdot,t)f(\cdot),y\rangle_{L^2(0,t;H)} \\ \quad -\langle B^*P_1(t,\cdot)f(\cdot), B^*P_0(t)y\rangle_{L^2(0,t;U)} = 0 \\ \frac{\partial}{\partial t}\langle P_2(t,\cdot,:)f(\cdot),g(:)\rangle_{L^2((0,t)^2;H)} + \langle P_1(t,\cdot)f(\cdot),K(t-:)g(:)\rangle_{L^2((0,t)^2;H)} \\ \quad +\langle K(t-\cdot)f(\cdot),P_1(t,:)g(:)\rangle_{L^2((0,t)^2;H)} \\ \quad -\langle B^*P_1(t,\cdot)f(\cdot), B^*P_1(t,:)g(:)\rangle_{L^2((0,t)^2;U)} = 0 \end{cases} \tag{6.1}$$

(for all $t\in[0,T]$*,* $y\in\mathcal{D}(A)$ *and* $f,g\in L^2(0,t;\mathcal{D}(A))$*).*

Proof. Assune that (P_0,P_1,P_2) solves the coupled system (6.1) (for all $t\in(0,T]$, $y\in H$ and $f,g\in L^2(0,t;H)$). Given $\tau\in(0,t)$ and $x\in\mathcal{D}(A)$, we choose

$$f(s)=\begin{cases} x & \text{in } [0,\tau] \\ 0 & \text{in } (\tau,t] \end{cases}$$

in the second of the equations in (6.1); the scalar products bring about integrals that are performed on the interval $[0,\tau]$, as a consequence. If we write these explicitly, we see that the partial derivative ∂_t can be moved inside any integral. Then, differentiating with respect to τ, we obtain

$$\begin{aligned} &\frac{\partial}{\partial t}\langle P_1(t,\tau)x,y\rangle_H + \langle P_1(t,\tau)x,Ay\rangle_H + \langle K(t-\tau)x,P_0(t)y\rangle_H \\ &\qquad + \langle P_2(t,\tau,t)x,y\rangle_H - \langle B^*P_1(t,\tau)x, B^*P_0(t)y\rangle_U = 0\,, \end{aligned}$$

which is nothing but the second equation of (1.10), as $\tau<t$ was (given, and yet) arbitrary.

Similarly, for any given $\tau,\sigma\in(0,t)$ and $x,y\in\mathcal{D}(A)$, we choose f as above and set

$$g(s)=\begin{cases} y & \text{in } [0,\sigma] \\ 0 & \text{in } (\sigma,t]. \end{cases}$$

We insert these functions in the third equation of (6.1) and write explicitly the integrals; once again the derivative ∂_t can be moved inside the integrals. Differentiate with respect to τ first, and next with respect to σ, to find

$$\begin{aligned} &\frac{\partial}{\partial t}\langle P_2(t,\tau,\sigma)x,y\rangle_H + \langle P_1(t,\tau)x, K(t-\sigma)y\rangle_H \\ &\qquad + \langle K(t-\tau)x, P_1(t,\sigma)y\rangle_H - \langle B^*P_1(t,\tau)x, B^*P_1(t,\sigma)y\rangle_U = 0\,, \end{aligned}$$

which is just the third equation in (1.10) (since both τ,σ are arbitrary).

Suppose now, conversely, that (P_0,P_1,P_2) solves the coupled system (1.10) (for any $t\in[0,T]$ and $x,y\in\mathcal{D}(A)$). With focus on the second equation, set $x=f(s)$

where $f \in C_0^0((0,T),\mathcal{D}(A))$ and $f \equiv 0$ in $[t,T]$, and integrate over $[0,t]$. Then,

$$\begin{aligned}\int_0^t \frac{\partial}{\partial t}\langle P_1(t,s)f(s),y\rangle_H\,ds &= \frac{d}{dt}\left\langle \int_0^t P_1(t,s)f(s)\,ds, y\right\rangle_H \\ &= \frac{d}{dt}\langle P_1(t,\cdot)f(\cdot),y\rangle_{L^2(0,t;H)}\,,\end{aligned}$$

where in the first equality we used that $f(t)=0$. This means that the second equation in (6.1) is valid for any $f \in C_0^0((0,t),\mathcal{D}(A))$ and $y \in H$.

If now $f \in L^2(0,t;\mathcal{D}(A))$, we extend it on the entire interval $[0,T]$ by setting it to 0 on $(t,T]$. Select a sequence $\{f_n\} \subset C_0^0((0,t),D(A))$ such that f_n converges to f in $L^2(0,t;\mathcal{D}(A))$. The second equation in (6.1) is satisfied for any f_n; letting $n \to \infty$, all the terms on the right hand side converge to the corresponding ones, with f in place of f_n. In addition, it is verified readily that the convergence is uniform with respect to $t \in [0,T]$. Thus, in the first hand side of the equation we also have

$$\begin{aligned}\lim_{n\to\infty} \frac{d}{dt}\langle P_1(t,\cdot)f_n(\cdot),y\rangle_{L^2(0,t;H)} &= -\langle P_1(t,\cdot)f(\cdot),Ay\rangle_{L^2(0,t;H)} \\ &- \langle K(t-\cdot)f(\cdot),P_0(t)y\rangle_{L^2(0,t;H)} - \langle P_2(t,s,t)f(\cdot),y\rangle_{L^2(0,t;H)} \\ &+ \langle B^*P_1(t,\cdot)f(\cdot),B^*P_0(t)y\rangle_{L^2(0,t;U)}\,,\end{aligned} \tag{6.2}$$

uniformly with respect to $t \in [0,T]$. Similarly,

$$\lim_{n\to\infty}\langle P_1(t,\cdot)f_n(\cdot),y\rangle_{L^2(0,t;H)} = \langle P_1(t,\cdot)f(\cdot),y\rangle_{L^2(0,t;H)} \tag{6.3}$$

uniformly with respect to $t \in [0,T]$.

In view of (6.3) and (6.2), we obtain that $t \longmapsto \langle P_1(t,\cdot)f_n(\cdot),y\rangle_{L^2(0,t;H)}$ is differentiable and that the second equation in (6.1) is satisfied (for any $f \in L^2(0,t;\mathcal{D}(A))$ and any $y \in H$).

A similar argument proves that the third equation in (6.1) is satisfied as well. Just set $x = f(s)$, $y = g(q)$, with $f,g \in C_0^0((0,T),\mathcal{D}(A))$ and $f \equiv g \equiv 0$ in $[t,T]$, and proceed (*mutatis mutandis*) as before. This concludes the proof of the proposition. □

In light of Proposition 6.1, the existence part of the proof of Theorem 1.7 now demands only a rewriting of the coupled system of four equations satisfied by the operators P_0, P_1, P_1^*, P_2 – to wit, system (6.1) complemented with the equation satisfied by P_1^* following (5.10) – as the unique matrix equation (1.12). As such, it is omitted.

2. (Uniqueness). In order to show uniqueness, it is necessary to make some preliminary remarks. The operator $P(\tau)$ defined by (1.11) belongs to the space $\mathcal{L}(Y_\tau)$ for any $\tau \in (0,T]$, and hence $P(\cdot)$ belongs to the Banach space

$$Z := \big\{U(\cdot)\colon U(\tau)\in\mathcal{L}(Y_\tau),\ \tau\in(0,T]\big\}, \quad \|U\|_Z = \sup_{\tau\in(0,T]}\|U(\tau)\|_{\mathcal{L}(Y_\tau)} < \infty\,.$$

Of course, for $\tau < t$ we have $Y_t \subseteq Y_\tau$. On the other hand, given $g \in L^2(0,\tau;H)$, set

$$G(t) := \begin{cases} g(t) & \text{in } [0,\tau] \\ 0 & \text{in } (\tau,t]; \end{cases}$$

then, for $\tau < t$ the map $(y,g) \longmapsto (y,G)$ from Y_τ to the set

$$E_{\tau,t} = \big\{(y,h) \in Y_t = H \times L^2(0,t;H)\colon h(\cdot) \equiv 0 \text{ in } (\tau,t]\big\}$$

is an isometry. We may therefore identify Y_τ with $E_{\tau,t}$. Hence, if $U \in Z$ and $(y,g) \in Y_\tau$ we may define $\|U(t)(y,g)\|_{Y_\tau} = \|U(t)(y,G)\|_{Y_t}$, and we have

$$\|U(t)(y,g)\|_{Y_\tau} = \|U(t)(y,G)\|_{Y_t} \le \|U(t)\|_{\mathcal{L}(Y_t)}\|(y,G)\|_{Y_t} = \|U(t)\|_{\mathcal{L}(Y_t)}\|(y,g)\|_{Y_\tau}.$$

In other words, we may say that $U(t) \in \mathcal{L}(Y_\tau)$ for every $\tau < t$, with

$$\|U(t)\|_{\mathcal{L}(Y_\tau)} \le \|U(t)\|_{\mathcal{L}(Y_t)}, \qquad 0 < \tau \le t \le T\,. \tag{6.4}$$

We proceed by contradiction: let $U(\cdot) \in Z$ be another solution of (1.12) (besides $P(\tau)$ defined by (1.11)), and set $V := P - U$. Of course, we have as well

$$U(\tau) = \begin{pmatrix} U_0(\tau) & U_1(\tau,\cdot) \\ U_1(\tau,:)^* & U_2(\tau,\cdot,:) \end{pmatrix}, \qquad V(\tau) = \begin{pmatrix} V_0(\tau) & V_1(\tau,\cdot) \\ V_1(\tau,:)^* & V_2(\tau,\cdot,:) \end{pmatrix}, \quad \tau \in (0,T).$$

For given $\tau, t \in (0,T)$, with $\tau < t$, the difference operator $V(\cdot)$ satisfies in $[t,T)$ the differential equation

$$\begin{aligned} \frac{d}{dr}V(r) + V(r)(\mathcal{A} + \mathcal{K}_1(r) + \mathcal{D}_{1,r}) + (\mathcal{A}^* + \mathcal{K}_2(r) + \mathcal{D}_{2,r})V(r) \\ - V(r)\mathcal{B}\mathcal{B}^*P(r) - U(r)\mathcal{B}\mathcal{B}^*V(r) = 0, \end{aligned} \tag{6.5}$$

with $V(T) = P(T) - U(T) = 0$. We remark that the equation (6.5) is to be interpreted as the ones in (6.1): namely, the first term in the left hand side of (6.5) is

$$\frac{d}{dr}\langle V(r)X_0, X_1\rangle_{Y_t}$$

(with $X_0, X_1 \in Y_t$); the other terms are understood in a similar way.

We note that the operator $\mathcal{A}$ is the infinitesimal generator of a C_0-semigroup $\{e^{s\mathcal{A}}\}_{s\ge 0}$ in Y_τ (for any $\tau \in (0,T)$), explicity given by

$$e^{s\mathcal{A}} = \begin{pmatrix} e^{sA} & 0 \\ 0 & I \end{pmatrix};$$

the bound $\|e^{s\mathcal{A}}\|_{\mathcal{L}(Y_\tau)} \le C_T$ holds true for any $s \in [0,T]$ and $\tau \in (0,T)$.

Thus, the mild form of (6.5) in $[t,T]$ is

$$\begin{aligned} V(r) = \int_r^T e^{(p-r)\mathcal{A}^*}\Big[V(p)[\mathcal{K}_1(p) + \mathcal{D}_{1,p}] + [\mathcal{K}_2(p) + \mathcal{D}_{2,p}]V(p) \\ - V(p)\mathcal{B}\mathcal{B}^*P(p) - U(p)\mathcal{B}\mathcal{B}^*V(p)\Big]e^{(p-r)\mathcal{A}}\,dp \end{aligned} \tag{6.6}$$

(this mild form is understood as (6.5) did). We can now observe that in view of (6.4) it makes sense to estimate $\|V(r)\|_{\mathcal{L}(Y_t)}$ when $r \ge t$. In addition, it can be shown that for $r \ge t$ the map $r \longmapsto \|V(r)\|_{\mathcal{L}(Y_t)}$ is lower semi-continuous and hence it is a measurable function in $[t,T]$ (the proof is postponed; see Lemma 6.2 at the end of this section).

Starting from (6.6), we find

$$
\begin{aligned}
\|V(r)\|_{\mathcal{L}(Y_t)} \le \int_r^T & C_T^2\Big[2\|K\|_{L^2(0,T)}\|V(p)\|_{\mathcal{L}(Y_t)} + 2\|V(p)\|_{\mathcal{L}(Y_t)} \\
& + \big(\|P\|_Z + \|U\|_Z\big)\|BB^*\|_{\mathcal{L}(H)}\|V(p)\|_{\mathcal{L}(Y_t)}\Big]\,dp \\
\le C_1 & \int_r^T \|V(p)\|_{\mathcal{L}(Y_t)}\,dp\,,
\end{aligned}
$$

where C_1 is a suitable positive constant depending on T, $\|BB^*\|_{\mathcal{L}(H)}$, $\|K\|_{L^2(0,T)}$, $\|P\|_Z$, $\|U\|_Z$.

By the Gronwall Lemma it follows that

$$
\|V(r)\|_{\mathcal{L}(Y_t)} = 0
$$

for $r \in [t,T]$. This means, in particular, $\|V(t)\|_{\mathcal{L}(Y_t)} = 0$. Since $t > \tau$ was given and yet arbitrary, it follows that

$$
V_0(t) = 0, \quad V_1(t,s) = 0, \quad V_2(t,s,q) = 0 \quad \forall s,q \in [0,\tau],\ \forall t \in [\tau,T]\,.
$$

This means that the operators U and P coincide for any t,s,q such that $s,q \in [0,\tau]$ and $t \in [\tau,T]$ (with arbitrary $\tau \in (0,T)$).

Thus, this second part of the proof is complete once the following result is established.

Lemma 6.2. *For any $\tau \in (0,T)$, the map $t \longmapsto \|V(t)\|_{\mathcal{L}(Y_\tau)}$ is lower semicontinuous in $[\tau,T]$.*

Proof. Let $D_\tau := \mathcal{D}(A) \times L^2(0,\tau;\mathcal{D}(A))$. Clearly D_τ is dense in Y_τ. Recalling that $\|V(t)\|_{\mathcal{L}(Y_\tau)} \le \|P\|_Z + \|U\|_Z \le C$, and that $V(\cdot)X$ is differentiable at the point t when $X \in D_\tau$, a straightforward density argument shows that $V(\cdot)$ is strongly continuous at the point t, i.e.

$$
\lim_{s\to t} \|[V(s) - V(t)]X\|_{Y_\tau} = 0 \qquad \forall X \in Y_\tau.
$$

In particular,

$$
\|V(t)X\|_{Y_\tau} = \lim_{s\to t} \|V(s)X\|_{Y_\tau} \le \liminf_{s\to t} \|V(s)\|_{\mathcal{L}(Y_\tau)}\|X\|_{Y_\tau}, \qquad \forall X \in Y_\tau
$$

and consequently

$$
\|V(t)\|_{\mathcal{L}(Y_\tau)} \le \liminf_{s\to t} \|V(s)\|_{\mathcal{L}(Y_\tau)}.
$$

The proof is complete. □

Appendix A. Further analytical results

Here we collect a series of results which are instrumental for the proof of the feedback formula (1.9) and/or critically utilized in the derivation of the coupled system of (three) differential equations satisfied by the operators P_i, $i = 0,1,2$, initially identified via the representation (4.8), subsequently shown to be equivalent to (4.12).

A.1. **Instrumental results, I.** A first set of results specifically pertain to the operators

$$\begin{aligned} \mu(t) &= \int_0^t e^{(t-s)A} K(s)\, ds \\ F(t,\tau) &= e^{(t-\tau)A} - \int_\tau^t R(t-s) e^{(s-\tau)A}\, ds \\ G(t,\sigma,\tau) &= \mu(t-\sigma) - e^{(t-\tau)A}\mu(\tau-\sigma) \\ M(t,\sigma,\tau) &= G(t,\sigma,\tau) - \int_\tau^t R(t-s) G(s,\sigma,\tau)\, ds \end{aligned} \tag{A.1}$$

introduced at the very outset in Section 2 and whose respective definitions are recorded in this appendix for the reader's convenience; see (2.3) ($R(\cdot)$ is the solution of the Volterra equation (2.4)).

First, it is useful to list certain basic properties of the operators in (A.1) which follow immediately from the respective definitions.

Lemma A.1. *The operators $\mu(t)$, $F(t,\tau)$, $G(t,\sigma,\tau)$ and $M(t,\sigma,\tau)$ recalled above in* (A.1) *satisfy*

$$\begin{aligned} \mu(0) = 0\,, &\qquad F(\tau,\tau) = I\,; \\ G(\tau,\sigma,\tau) = 0\,, &\quad G(t,\tau,\tau) = \mu(t-\tau)\,; \\ M(\tau,\sigma,\tau) = 0\,, &\quad M(t,\tau,\tau) = -R(t-\tau)\,. \end{aligned} \tag{A.2}$$

Proof. The conditions grouped collectively in (A.2) are inferred via an easy verification. The two conditions in the first row are immediate, on the basis of (A.1). Then, we note that

$$G(\tau,\sigma,\tau) = \mu(\tau-\sigma) - \mu(\tau-\sigma) = 0\,;$$

then $M(\tau,\sigma,\tau) = G(\tau,\sigma,\tau) = 0$ as a consequence. We have instead

$$G(t,\tau,\tau) = \mu(t-\tau) - \mu(0) = \mu(t-\tau)\,,$$

and therefore,

$$\begin{aligned} M(t,\tau,\tau) = G(t,\tau,\tau) - \int_\tau^t R(t-s)G(s,\tau,\tau)\, ds &= \mu(t-\tau) - \int_\tau^t R(t-s)\mu(s-\tau)\, ds \\ &= \mu(t-\tau) - \int_0^{t-\tau} R(t-\tau-\lambda)\mu(\lambda)\, d\lambda = -R(t-\tau)\,, \end{aligned}$$

where in the last equality the Volterra equation (2.4) satisfied by the resolvent operator R is employed. □

We move on exploring the differentiability of the operator μ.

Proposition A.2. *Let $\mu(\cdot)$ be the operator in* (A.2). *If $x \in \mathcal{D}(A)$, then $\mu(\cdot)x$ is differentiable, and*

$$\mu'(t)x = K(t)x + \mu(t)Ax\,, \qquad x \in \mathcal{D}(A) \tag{A.3}$$

holds true.

Proof. The proof is straighforward. With $t \geq 0$ and $h \neq 0$ ($h > 0$, in the case $t = 0$, also in the sequel) we compute

$$\frac{\mu(t+h)-\mu(t)}{h} = \frac{1}{h}\int_t^{t+h} e^{(t+h-\lambda)A} K(\lambda)\, d\lambda + \int_0^t \frac{e^{(t+h-\lambda)A} - e^{(t-\lambda)A}}{h} K(\lambda)\, d\lambda\,,$$

which implies readily, for $x \in \mathcal{D}(A)$, that there exists

$$\begin{aligned}\lim_{h\to 0} \frac{\mu(t+h)-\mu(t)}{h} x &= K(t)x + \int_0^t Ae^{(t-\lambda)A} K(\lambda) x\, d\lambda \\ &= K(t)x + \int_0^t e^{(t-\lambda)A} K(\lambda) Ax\, d\lambda\,,\end{aligned}$$

that is nothing but (A.3). □

Remark A.3. It is worth noting that the property that K commutes with the operator A – and hence, with the semigroup e^{tA} as well – is essential for the proof of Proposition A.2. This should be kept in mind in case of extensions to non-scalar kernels K.

Proposition A.4. *The resolvent operator $R(\cdot)$ is differentiable on $\mathcal{D}(A)$, with*

$$R'(t)x = -K(t)x + R(t)Ax + \int_0^t K(t-\sigma)R(\sigma)\, d\sigma\,, \qquad x \in \mathcal{D}(A)\,. \tag{A.4}$$

Proof. Recall that $R(t)$ is expressed by the iterated formula (2.5), and take the convolution of μ and R, to find

$$\mu * R = R * \mu = -\mu * \mu(t) - \mu * \mu * \mu(t) - \dots\,,$$

which once read from the left to the right yields $R(t) + \mu(t) = R * \mu$, that is

$$R(t) = -\mu(t) + \int_0^t R(\sigma)\mu(t-\sigma)\, d\sigma\,.$$

The above applied to $x \in \mathcal{D}(A)$ establishes that $R(t)x$ is differentiable in the first place in view of Proposition A.2; then, a straightforward computation yields (A.4) by virtue of (A.3).

□

Proposition A.5. *Let $F(t,\tau)$ be the operator recalled in* (A.1)*, $\tau < t$. If $x \in \mathcal{D}(A)$, then there exist both $\partial_t F(t,\tau)x$ and $\partial_\tau F(t,\tau)x$, given by the following espressions:*

$$\begin{aligned}\partial_t F(t,\tau)x &= e^{(t-\tau)A}Ax - \int_\tau^t R'(t-s)e^{(s-\tau)A}x\, ds\,, \\ \partial_\tau F(t,\tau)x &= -e^{(t-\tau)A}Ax + R(t-\tau)x - \int_\tau^t R(t-s)e^{(s-\tau)A}Ax\, ds \\ &= -F(t,\tau)Ax + R(t-\tau)x\,.\end{aligned} \tag{A.5}$$

Moreover, if $x \in X$ and $y \in \mathcal{D}(A)$

$$\langle \partial_\tau F(t,\tau)^* x, y\rangle_H = -\langle F(t,\tau)^* x, Ay\rangle_H + \langle R(t-\tau)x, y\rangle_H. \tag{A.6}$$

Proof. The thesis follows readily from the definition of $F(t,\tau)$, in light of Proposition A.4. □

Proposition A.6. *If $x \in \mathcal{D}(A)$, then there exists*

$$\partial_\tau M(t,\sigma,\tau)x \tag{A.7}$$

Proof. Let $x \in \mathcal{D}(A)$ be given. In view of the definition of $M(t,\sigma,\tau)x$ in (A.1), we preliminarly examine $G(t,\sigma,\tau)x$. Because of Proposition A.2 and (A.3), we see that there exists

$$\begin{aligned}
\partial_\tau G(t,\sigma,\tau)x &= -\frac{\partial}{\partial\tau}\big[e^{(t-\tau)A}\mu(\tau-\sigma)x\big] \\
&= Ae^{(t-\tau)A}\mu(\tau-\sigma)x - e^{(t-\tau)A}\mu'(\tau-\sigma)x \\
&\underset{(A.3)}{=} e^{(t-\tau)A}A\mu(\tau-\sigma)x - e^{(t-\tau)A}\big[K(\tau-\sigma)x + \mu(\tau-\sigma)Ax\big] \\
&= -e^{(t-\tau)A}K(\tau-\sigma)x\,.
\end{aligned}$$

Going back once more to the definitions in (A.1), the above implies that – when acting on $\mathcal{D}(A)$ – there exists

$$\begin{aligned}
\partial_\tau M(t,\sigma,\tau) &= \partial_\tau G(t,\sigma,\tau) + R(t-\tau)\underbrace{G(\tau,\sigma,\tau)}_{\equiv 0} - \int_\tau^t R(t-s)\partial_\tau G(s,\sigma,\tau)\,ds \\
&= -e^{(t-\tau)A}K(\tau-\sigma) + \int_\tau^t R(t-s)e^{(s-\tau)A}K(\tau-\sigma)\,ds \\
&= -\Big[e^{(t-\tau)A} - \int_\tau^t R(t-s)e^{(s-\tau)A}\,ds\Big]\,K(\tau-\sigma) = -F(t,\tau)K(\tau-\sigma)\,,
\end{aligned}$$

which confirms (A.7). □

A.2. **Instrumental results, II.** A second set of results pertains to the couples of operators $\Psi_1(t,\tau)$ and $\Psi_2(t,\sigma,\tau)$, $Z_1(t,\tau)$ and $Z_2(t,s,\tau)$, identified earlier in the paper (see (4.1) and (4.4)). Their respective expressions are collectively recalled here for the reader's convenience:

$$\begin{aligned}
\Psi_1(t,\tau) &= -\Big[\Lambda_\tau^{-1}L_\tau^* QF(\cdot,\tau)\Big](t)\,, \\
\Psi_2(t,\sigma,\tau) &= -\Big[\Lambda_\tau^{-1}L_\tau^* QM(\cdot,\sigma,\tau)\Big](t)\,, \\
Z_1(t,\tau) &= F(t,\tau) + \int_\tau^t F(t,\sigma)B\Psi_1(\sigma,\tau)\,d\sigma\,, \\
Z_2(t,s,\tau) &= M(t,s,\tau) + \int_\tau^t F(t,\sigma)B\Psi_2(\sigma,s,\tau)\,d\sigma\,.
\end{aligned} \tag{A.8}$$

The representation formulas

$$\begin{aligned}
Z_1(t,\tau) &= \big\{\big[I - L_\tau\Lambda_\tau^{-1}L_\tau^* Q\big]F(\cdot,\tau)\big\}(t)\,, \\
Z_2(t,s,\tau) &= \big\{\big[I - L_\tau\Lambda_\tau^{-1}L_\tau^* Q\big]M(\cdot,s,\tau)\big\}(t)\,.
\end{aligned} \tag{A.9}$$

that follow inserting the expressions of Ψ_1 and Ψ_2 within the ones of Z_1 and Z_2, respectively, are especially useful for the derivation of the Riccati equation satisfied by P_2 (specifically in the proof of the statement S5. of Theorem 1.6, part (ii), Lemma 5.1).

A.2.1. *Adjoint operators.* We begin by computing the adjoints of the operators listed in (A.8), as they occur in the very definition (4.8) of the operators P_i, $i = 0, 1, 2$. The obtained expressions play a critical role in the derivation of the alternative (and neater) representations (4.12) of the P_i, $i = 0, 1, 2$; see Lemma 4.4.

Lemma A.7. *Let $\Psi_1(t,\tau)$, $\Psi_2(t,\sigma,\tau)$, $Z_1(t,\tau)$ and $Z_2(t,s,\tau)$ be the operators recalled in* (A.8)*. The respective adjoint operators are acting as follows:*

$$\begin{aligned} \Psi_1(\cdot,\tau)^* g &= F(\cdot,\tau)^* Q[L_\tau\Lambda_\tau^{-1}g](\cdot)\,, \\ \Psi_2(\cdot,\sigma,\tau)^* g &= M(\cdot,\sigma,\tau)^* Q[L_\tau\Lambda_\tau^{-1}g](\cdot) \\ Z_1(\cdot,\tau)^* g &= F(\cdot,\tau)^* g(\cdot) + \Psi_1(\cdot,\tau)^*[L_\tau^* g](\cdot) \\ Z_2(\cdot,s,\tau)^* g &= M(\cdot,s,\tau)^* g(\cdot) + \Psi_2(\cdot,s,\tau)^*[L_\tau^* g](\cdot) \end{aligned} \tag{A.10}$$

for every $g \in L^2(\tau,T;H)$.

Proof. (i) With $f, g \in L^2(\tau,T;H)$ we have

$$\begin{aligned} \int_\tau^T \big\langle \Psi_1(\sigma,\tau)f(\sigma), g(\sigma)\big\rangle_H\,d\sigma &= -\int_\tau^T \big\langle \big[\Lambda_\tau^{-1}L_\tau^* QF(\cdot,\tau)\big](\sigma)f(\sigma), g(\sigma)\big\rangle_H\,d\sigma \\ &= -\int_\tau^T \big\langle F(\sigma,\tau)f(\sigma), Q\big[L_\tau\Lambda_\tau^{-1}g\big](\sigma)\big\rangle_H\,d\sigma \\ &= -\int_\tau^T \big\langle f(\sigma), F(\sigma,\tau)^* Q\big[L_\tau\Lambda_\tau^{-1}g\big](\sigma)\big\rangle_H\,d\sigma\,, \end{aligned} \tag{A.11}$$

which establishes the first one of (A.10).

(ii) We just repeat the preceding computation with $M(\cdot,\sigma,\tau)$ in place of $F(\cdot,\tau)$, and the second one of (A.10) follows.

(iii) Once again, given $f, g \in L^2(0,T;H)$ we have

$$\begin{aligned} \int_\tau^T \big\langle Z_1(q,\tau)f(q), g(q)\big\rangle_H\,d\sigma &= -\int_\tau^T \Big\langle \Big[F(q,\tau) + \big[L_\tau\Psi_1(\cdot,\tau)\big](q)\Big]f(q), g(q)\Big\rangle_H\,dq \\ &= \int_\tau^T \big\langle f(q), F(q,\tau)^* g(q) + \Psi_1(q,\tau)^*\big[L_\tau^* g\big](q)\big\rangle_H\,dq \end{aligned} \tag{A.12}$$

which establishes the third one of the (A.10).

(iv) The verification of the last one of the identities (A.10) easily follows as in (iii), just replacing $F(\cdot,\tau)$ and $\Psi_1(\cdot,\tau)$ by $M(\cdot,s,\tau)$ and $\Psi_2(\cdot,s,\tau)$, respectively. □

A.2.2. *Differentiation of the operators.*

Proposition A.8. *Let $\Psi_1(t,\tau)$ as in* (A.8)*. If $x \in \mathcal{D}(A)$, then the derivative $\partial_\tau\Psi_1(t,\tau)x$ exists, and it is given by*

$$\partial_\tau\Psi_1(t,\tau)x = -\big[\Lambda_\tau^{-1}L_\tau^* QF_\tau(\cdot,\tau)x\big](t) + \big[\Lambda_\tau^{-1}L_\tau^* QF(\cdot,\tau)B\Psi_1(\tau,\tau)x\big](t)\,. \tag{A.13}$$

Proof. The proof's strategy is as follows: we consider the implicit representation of $\Psi_1(t,\tau)$, that is (since $\Lambda_\tau = I + L_\tau^* QL_\tau$)

$$\Psi_1(t,\tau) + \big[L_\tau^* QL_\tau\Psi_1(\cdot,\tau)\big](t) = -\big[L_\tau^* QF(\cdot,\tau)\big](t)\,,$$

which explicitly reads as

$$\begin{aligned}\Psi_1(t,\tau) + \int_t^T B^*F(p,t)^*Q\int_\tau^p F(p,\sigma)B\Psi_1(\sigma,\tau)\,d\sigma\,dp \\ = -\int_t^T B^*F(p,t)^*QF(p,\tau)\,dp\,.\end{aligned}$$

The Fubini-Tonelli Theorem yields

$$\begin{aligned}\Psi_1(t,\tau) + \Big[\int_\tau^t\int_t^T + \int_t^T\int_\sigma^T\Big] B^*F(p,t)^*QF(p,\sigma)B\Psi_1(\sigma,\tau)\,dp\,d\sigma \\ = -\int_t^T B^*F(p,t)^*QF(p,\tau)\,dp\,,\end{aligned}$$

that is

$$\begin{aligned}\Psi_1(t,\tau) &+ \int_\tau^t \Big[\int_t^T B^*F(p,t)^*QF(p,\sigma)B\,dp\Big]\Psi_1(\sigma,\tau)\,d\sigma \\ &\qquad + \int_t^T \Big[\int_\sigma^T B^*F(p,t)^*QF(p,\sigma)B\Big]\,dp\Psi_1(\sigma,\tau)\,d\sigma \\ &= -\int_t^T B^*F(p,t)^*QF(p,\tau)\,dp\,.\end{aligned} \tag{A.14}$$

Taking (A.14) as a starting point, we compute for $x \in \mathcal{D}(A)$ the incremental ratio of $\Phi_1(t,\cdot)x$ (with $h \neq 0$) to find the identity

$$\begin{aligned}&\frac{\Psi_1(t,\tau+h)x - \Psi_1(t,\tau)x}{h} \\ &\qquad - \frac{1}{h}\int_\tau^{\tau+h} \Big[\int_t^T B^*F(p,t)^*QF(p,\sigma)B\,dp\Big]\Psi_1(\sigma,\tau)x\,d\sigma \\ &\qquad + \int_\tau^t \Big[\int_t^T B^*F(p,t)^*QF(p,\sigma)B\,dp\Big]\frac{\Psi_1(\sigma,\tau+h)x - \Psi_1(\sigma,\tau)x}{h}\,d\sigma \\ &\qquad + \int_t^T \Big[\int_\sigma^T B^*F(p,t)^*QF(p,\sigma)B\Big]\,dp\frac{\Psi_1(\sigma,\tau+h)x - \Psi_1(\sigma,\tau)x}{h}\,d\sigma \\ &= -\int_t^T B^*F(p,t)^*Q\,\frac{F(p,\tau+h)x - F(p,\tau)x}{h}\,dp\,.\end{aligned}$$

Thus we proceed as before, but somewhat in the reverse direction: we

- use once again the Fubini-Tonelli Theorem, this time to merge the third and fourth summands in the left hand side,
- recognize that the sum of the first, third and fourth terms is nothing but $I + L_\tau^*QL_\tau =: \Lambda_\tau$ applied to $[\Psi_1(\sigma,\tau+h)x - \Psi_1(\sigma,\tau)x]/h$,
- move the second summand from the left to the right hand side,

to attain

$$\frac{\Psi_1(t,\tau+h)x-\Psi_1(t,\tau)x}{h} = \underbrace{\Lambda_\tau^{-1}\Big[-L_\tau^* Q\frac{F(\cdot,\tau+h)x-F(\cdot,\tau)x}{h}\Big](t)}_{T_1(t)x} + \underbrace{\Lambda_\tau^{-1}\Big[\frac{1}{h}\int_\tau^{\tau+h}\big[L_\tau^* QF(\cdot,\sigma)B\,\Psi_1(\sigma,\tau)x\big](t)\,d\sigma\Big]}_{T_2(t)x}.$$

Now the existence of the limit, as $h\to 0$, of the incremential ratio of $\Psi_1(t,\cdot)x$, along with the formula (A.15), follows observing that, since $x\in\mathcal{D}(A)$, we have

$$\begin{cases} T_1(t)x \to -\big[\Lambda_\tau^{-1}L_\tau^* Q\partial_\tau F(\cdot,\tau)x\big](t)\,, \text{ as } h\to 0 \quad \text{(in view of Proposition A.5)},\\ T_2(t)x \to \big[\Lambda_\tau^{-1}L_\tau^* QF(\cdot,\tau)B\Psi_1(\tau,\tau)x\big](t)\,, \text{ as } h\to 0. \end{cases}$$

This establishes (A.13), thus concluding the proof of the proposition.

□

Proposition A.9. *Let $\Psi_2(t,\sigma,\tau)x$ as in* (A.8)*. If $x\in\mathcal{D}(A)$, then there exists $\partial_\tau\Psi_2(t,\sigma,\tau)x$ and it is given by*

$$\partial_\tau\Psi_2(t,\sigma,\tau)x = -\big[\Lambda_\tau^{-1}L_\tau^* QM_\tau(\cdot,\sigma,\tau)x\big](t) + \big[\Lambda_\tau^{-1}L_\tau^* QF(\cdot,\tau)B\Psi_2(\tau,\sigma,\tau)x\big](t)\,. \tag{A.15}$$

Proof. The proof of the proposition can be carried out employing, *mutatis mutandis*, a similar path as the one utilized in the proof of Proposition A.8, with the implicit equation satisfied by $\Psi_2(t,\sigma,\tau)$ as a starting point, and the use of Proposition A.6 – in place of Proposition A.5 – at the end. The details are omitted.

□

A direct consequence of Propositions A.8 and A.9 are the following results, whose respective proofs are straightforward and hence are omitted.

Proposition A.10. *Let $Z_1(t,\tau)$ as in* (A.8)*. Then $Z_1(\tau,\tau)=F(\tau,\tau)=I_H$. If $x\in\mathcal{D}(A)$, there exists $\partial_\tau Z_1(t,\tau)x$ and it is given by*

$$\partial_\tau Z_1(t,\tau)x = \partial_\tau F(t,\tau)x - F(t,\tau)B\Psi_1(\tau,\tau)x + \big[L_\tau\partial_\tau\Psi_1(\cdot,\tau)x\big](t). \tag{A.16}$$

Proposition A.11. *Let $Z_2(t,s,\tau)$ as in* (A.8)*. Then, $Z_2(\tau,s,\tau)=M(\tau,s,\tau)=G(\tau,s,\tau)=0$. If $x\in\mathcal{D}(A)$, there exists $\partial_\tau Z_2(t,s,\tau)x$ and it is given by*

$$\partial_\tau Z_2(t,s,\tau)x = \partial_\tau M(t,s,\tau)x - F(t,\tau)B\Psi_2(\tau,s,\tau)x + \big[L_\tau\partial_\tau\Psi_2(\cdot,s,\tau)x\big](t). \tag{A.17}$$

Acknowledgements

We would like to thank the reviewer for their insightful comments and suggestions, which led to an improved version of the manuscript.

The research of F. Bucci has been performed in the framework of the MIUR-PRIN Grant 2020F3NCPX "Mathematics for industry 4.0 (Math4I4)". Bucci's research was also supported by the Università degli Studi di Firenze under the 2022 Project *Analisi e controllo di equazioni di evoluzione che descrivono propagazione ondosa o viscoelasticità*, of which she was responsible. F. Bucci is a member of the Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni (GNAMPA) of the Istituto Nazionale di Alta Matematica (INdAM).

References

[1] P. Acquistapace, F. Bucci, Uniqueness for Riccati equations with application to the optimal boundary control of composite systems of evolutionary partial differential equations, *Ann. Mat. Pura Appl.* **202** (2023), 1611-1642.

[2] A. Bensoussan, G. Da Prato, M. Delfour, S. Mitter, *Representation and Control of Infinite Dimensional Systems*, 2^{nd} edition, Birkhäuser, Boston, 2007.

[3] S. Bonaccorsi, F. Confortola, Optimal control for stochastic Volterra equations with multiplicative Lévy noise, *NoDEA Nonlinear Differential Equations Appl.* **27** (2020), no. 3, Paper No. 26, 26 pp.

[4] F. Bucci, L. Pandolfi, On the regularity of solutions to the Moore-Gibson-Thompson equation: a perspective via wave equations with memory, *J. Evol. Equ.* **20** (2020), no. 3, 837-867.

[5] F. Bucci, M. Eller, The Cauchy-Dirichlet problem for the Moore-Gibson-Thompson equation, *C. R. Math. Acad. Sci. Paris 359* (2021), 881-903.

[6] P. Cannarsa, H. Frankowska, E. M. Marchini, Optimal control for evolution equations with memory, *J. Evol. Equ.* **13** (2013), no. 1, 197-227.

[7] E. Casas, J. Yong, Optimal control of a parabolic equation with memory, *ESAIM Control Optim. Calc. Var.* **29** (2023), Paper No. 23, 16 pp.

[8] C. Cavaterra, A. Lorenzi, M. Yamamoto, A stability result via Carleman estimates for an inverse source problem related to a hyperbolic integro-differential equation, *Comput. Appl. Math.* **25** (2006), no. 2-3, 229-250.

[9] F.W. Chaves-Silva, L. Rosier, E. Zuazua, Null controllability of a system of viscoelasticity with a moving control, *J. Math. Pures Appl.* **101** (2014), no. 2, 198-222.

[10] F. W. Chaves-Silva, X. Zhang, E. Zuazua, Controllability of evolution equations with memory, *SIAM J. Control Optim.* **55** (2017), no. 4, 2437-2459.

[11] V.V. Chepyzhov, V. Pata, Some remarks on stability of semigroups arising from linear viscoelasticity, *Asymptot. Anal.* **46** (2006), no. 3-4, 251-273.

[12] M. Conti, S. Gatti, V. Pata, Uniform decay properties of linear Volterra integro-differential equations, *Math. Models Methods Appl. Sci.* **18** (2008), no. 1, 21-45.

[13] C. Corduneanu, *Integral Equations and Applications*, Cambridge University Press, Cambridge, 2010.

[14] C. M. Dafermos, An abstract Volterra equation with applications to linear viscoelasticity, *J. Differential Equations* **7** (1970), 554-569.

[15] G. Da Prato, M. Iannelli, Linear integro-differential equations in Banach spaces. *Rend. Sem. Mat. Univ. Padova* **62** (1980), 207-219.

[16] F. Dell'Oro, V. Pata, On the Moore-Gibson-Thompson equation and its relation to linear viscoelasticity, *Appl. Math. Optim.* **76** (2017), no. 3, 641-655.

[17] A. Doubova, E. Fernández-Cara, On the control of viscoelastic Jeffreys fluids. (English summary) *Systems Control Lett.* **61** (2012), no. 4, 573-579.

[18] K.-J. Engel, R. Nagel, *One-parameter semigroups for linear evolution equations*, *With contributions by S. Brendle, M. Campiti, T. Hahn, G. Metafune, G. Nickel, D. Pallara, C. Perazzoli, A. Rhandi, S. Romanelli and R. Schnaubelt*, Graduate Texts in Mathematics, 194, Springer-Verlag, New York, 2000. xxii+586 pp. ISBN: 0-387-98463-1

[19] E. Fernández-Cara, Q. Lü, E. Zuazua, Null controllability of linear heat and wave equations with nonlocal spatial terms, *SIAM J. Control Optim.* **54** (2016), no. 4, 2009-2019.

[20] E. Fernández-Cara, J. L. Machado, D.A. Souza, Non null controllability of Stokes equations with memory, *ESAIM Control Optim. Calc. Var.* **26** (2020), Paper No. 72, 18 pp.

[21] A. Friedman, M. Shinbrot, Volterra integral equations in Banach space, *Trans. Amer. Math. Soc.* **126** (1967), 131-179.

[22] P. Gamboa, V. Komornik, O. Vera, Partial reachability of a thermoelastic plate with memory, *Quart. Appl. Math.* **74** (2016), no. 2, 235-243.

[23] M. Grasselli, M. Squassina, Exponential stability and singular limit for a linear thermoelastic plate with memory effects, *Adv. Math. Sci. Appl.* **16** (2006), no. 1, 15-31.

[24] G. Gripenberg, S. O. Londen, O. Staffans, *Volterra integral and functional equations*, Encyclopedia of Mathematics and its Applications, 34. Cambridge University Press, Cambridge, 1990. xxii+701 pp. ISBN: 0-521-37289-5

[25] S. Guerrero, O. Y. Imanuvilov, Remarks on non controllability of the heat equation with memory, *ESAIM Control Optim. Calc. Var.* **19** (2013), no. 1, 288-300.

[26] S. Ivanov, L. Pandolfi, Heat equation with memory: lack of controllability to rest, *J. Math. Anal. Appl.* **355** (2009), no. 1, 1-11.

[27] I. Lasiecka, R. Triggiani, Differential and algebraic Riccati equations with application to boundary/point control problems: continuous theory and approximation theory. *Lecture Notes in Control and Information Sciences*, 164, Springer-Verlag, Berlin, 1991. xii+160 pp.

[28] I. Lasiecka, R. Triggiani, *Control Theory for Partial Differential Equations: Continuous and Approximation Theories.* I. Abstract Parabolic Systems; II. Abstract Hyperbolic-like Systems over a Finite Time Horizon, Encyclopedia Math. Appl., 74–75, Cambridge University Press, Cambridge, 2000.

[29] P. Loreti, D. Sforza, Reachability problems for a class of integro-differential equations, *J. Differential Equations 248* (2010), no. 7, 1711-1755.

[30] P. Loreti, D. Sforza, Inverse observability inequalities for integrodifferential equations in square domains, *Evol. Equ. Control Theory* **7** (2018), no. 1, 61-77.

[31] P. Loreti, D. Sforza, M. Yamamoto, Carleman estimates for integro-differential parabolic equations with singular memory kernels, *J. Elliptic Parabol. Equ.* **3** (2017), no. 1-2, 53-64.

[32] P. Loreti, D. Sforza, M. Yamamoto, Carleman estimate and application to an inverse source problem for a viscoelasticity model in anisotropic case, *Inverse Problems* **33** (2017), no. 12, 125014, 28 pp.

[33] L. Pandolfi, The quadratic regulator problem and the Riccati equation for a process governed by a linear Volterra integrodifferential equations, *IEEE Trans. Automat. Control* **63** (2018), no. 5, 1517-1522.

[34] L. Pandolfi, *Systems with persistent memory – controllability, stability, identification*, Interdisciplinary Applied Mathematics, 54. Springer, Cham, [2021], x+356 pp.

[35] A. J. Pritchard, Y. You, Causal feedback optimal control for Volterra integral equations, *SIAM J. Control Optim.* **34** (1996), 1874-1890.

[36] J. Prüss, *Evolutionary integral equations and applications*, [2012] reprint of the 1993 edition, Modern Birkhäuser Classics, Birkhäuser/Springer Basel AG, Basel, 1993. xxvi+366 pp. ISBN: 978-3-0348-0498-1

[37] M. Renardy, W. J. Hrusa, J.A. Nohel, *Mathematical problems in viscoelasticity*, Pitman Monographs and Surveys in Pure and Applied Mathematics, 35. Longman Scientific & Technical, Harlow; John Wiley & Sons, Inc., New York, 1987. x+273 pp. ISBN: 0-582-00320-2

[38] A. Tuffaha, The Stochastic Linear Quadratic Optimal Control Problem on Hilbert Spaces: The Case of Non-analytic Systems, *Appl. Math. Optim.* **87** (2023), Paper No. 58, 41 pp.

Paolo Acquistapace, Università di Pisa (*Ret.*), Dipartimento di Matematica, Largo Bruno Pontecorvo 5, 56127 Pisa, ITALY

Email address: `paolo.acquistapace(at)unipi.it`

Francesca Bucci, Università degli Studi di Firenze, Dipartimento di Matematica e Informatica, Via S. Marta 3, 50139 Firenze, ITALY

Email address: `francesca.bucci(at)unifi.it`