

This article was downloaded by:[Technical University of Eindhoven]  
[Technical University of Eindhoven]

On: 22 May 2007

Access Details: [subscription number 777257291]

Publisher: Taylor & Francis

Informa Ltd Registered in England and Wales Registered Number: 1072954

Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## International Journal of Control

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t713393989>

### Linear quadratic regulator problem with positive controls

W. P. M. H. Heemels; S. J. L. Van Eijndhoven; A. A. Stoorvogel

To cite this Article: Heemels, W. P. M. H., Van Eijndhoven, S. J. L. and Stoorvogel, A. A., 'Linear quadratic regulator problem with positive controls', International Journal of Control, 70:4, 551 - 578

To link to this article: DOI: 10.1080/002071798222208

URL: <http://dx.doi.org/10.1080/002071798222208>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article maybe used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

© Taylor and Francis 2007

## Linear quadratic regulator problem with positive controls

W. P. M. H. HEEMELS<sup>†</sup>, S. J. L. VAN EIJDHOVEN<sup>‡</sup> and  
A. A. STOOORVOGEL<sup>‡</sup>

In this paper, the Linear Quadratic Regulator Problem with a positivity constraint on the admissible control set is addressed. Necessary and sufficient conditions for optimality are presented in terms of inner products, projections on closed convex sets, Pontryagin's maximum principle and dynamic programming. The main results are concerned with smoothness of the optimal control and the value function. The maximum principle will be extended to the infinite horizon case. Based on these analytical methods, we propose a numerical algorithm for the computation of the optimal controls for the finite and infinite horizon problem. The numerical methods will be justified by convergence properties between the finite and infinite horizon case on one side and discretized optimal controls and the true optimal control on the other.

### 1. Introduction

In the literature, the Linear Quadratic Regulator Problem has been solved by the use of Riccati equations (Anderson and Moore 1990). In this paper the same problem will be treated with the additional constraint that the control function can only take values in a closed convex polyhedral cone, like the non-negative orthant in a Euclidean space. More precisely, we will consider the regular case, where the weight on the control in the integrand of the cost functions is positive definite.

In many real-life problems, the influence we have on the system can be used only in one direction. One could, for instance, think of regulating the temperature of a room: the heating element can only put energy into the room, but it cannot extract energy. In the process industry or in mechanical systems, the flow of a certain fluid or gas is regulated by one-way valves. Other examples arise in the control of electrical networks with diode elements and in economic systems, where quantities like investments and taxes are always non-negative.

A basic issue to be solved before computing or analysing optimal controls is related to the existence and uniqueness of optimal controls. The existence part leads to a verification whether the set of admissible controls has been chosen properly. In section 2 we discuss this problem. The existence of solutions to the optimal control problem is studied by Pachter (1980). The form of the conditions in Pachter (1980) is only explicit in the case where the integrand in the cost functional is independent of the trajectory  $x$  (i.e.  $Q = 0$  in terms of Pachter (1980) and  $C^T D = 0$ ,  $C^T C = 0$  in terms of section 2 below). This special case is very restrictive and, in our paper, these

---

Received 16 January 1998.

<sup>†</sup> Eindhoven University of Technology, Department of Electrical Engineering, Measurement and Control Systems, Section: Measurement and Control, PO Box 513, 5600 MB Eindhoven, The Netherlands. E-mail: w.p.m.h.heemels@ele.tue.nl; Fax: +31 40 2434582.

<sup>‡</sup> Eindhoven University of Technology, Department of Mathematics and Computing Science, PO Box 513, 5600 MB Eindhoven. Fax: +31 40 2442489.

assumptions are not made. After suitable transformation of our optimal control problem we can rely on a result from functional analysis (cf. Theorem 5 below) to prove the existence and uniqueness of optimal controls. Moreover, a characterization of optimal controls in terms of inner products is obtained as a side product. For the infinite horizon case the existence of an admissible control is not clear at all. By admissible control we mean a non-negative control function that keeps the cost functional finite and the state trajectory square integrable. Necessary and sufficient conditions for the existence of admissible controls for all initial states can be found in Heemels (1998). If we assume that these conditions are satisfied, the existence and uniqueness of the optimal controls for the infinite horizon case are established.

In addition to papers like Pachter (1980) where only existence of optimal controls is studied, we characterize the optimal controls in various equivalent forms. Moreover, smoothness properties of the optimal control and value function, and convergence results between the finite and infinite horizon are stated. In Pachter (1980) no attention is paid to the infinite horizon case nor to the problem of how to compute or approximate the optimal control in both the finite and infinite horizon problem. All these issues will be considered in the current paper.

Essentially, there are two classical approaches in optimal control theory. The first approach is the maximum principle, initiated by Pontryagin *et al.* (1962). The original maximum principle has been used and extended by many others (Macki and Strauss 1982, Feichtinger and Hartl 1986, Lee and Markus 1967). To get a complete treatise of dealing with constrained finite horizon optimal control problems the application of the maximum principle is incorporated. However, for the infinite horizon case a similar result is non-trivial. The convergence results between the finite and infinite horizon problem that will be established, can be exploited to derive, under rather mild conditions, a maximum principle on an infinite horizon for the 'positive Linear Quadratic Regulator problem.'

The second approach, dynamic programming, was originally conceived by Bellman (1967) as a fruitful numerical method to compute optimal controls in discrete time processes. Later, people realized that the same ideas can be used for optimal control problems in continuous time. For continuous time problems, dynamic programming leads to a partial differential equation, the so-called Hamilton–Jacobi–Bellman (HJB) equation, which has the value function among its solutions (Fleming and Rishel 1975). Traditionally, one needed assumptions on the smoothness of the value function to apply this theory. However, in many problems, the value function does not behave smoothly, which causes both analytical and numerical problems. Recently, the notion of viscosity solutions has been introduced, which generalizes the concept of a solution to the HJB equation (Fleming and Soner 1993). In this paper, it will be shown that the value function in our problem is continuously differentiable and satisfies the HJB equation in the classical sense. This simplifies dynamic programming in both analytical and numerical aspects and allows us to obtain a direct relationship with the results obtained via the maximum principle. It will become clear that both dynamic programming and the maximum principle yield necessary and sufficient conditions for optimality.

The maximum principle results in a two-point boundary value problem. When this problem is solved, the optimal open-loop control for one specific initial state has been found. By implementing such an open-loop control, we only use *a priori* knowledge of the initial condition and there is no adaptation possible for disturbances acting on the system such as measurement noise, unmodelled dynamics

of the plant, etc. In contrast to the maximum principle, the solution of the HJB equation gives the optimal state-feedback controller. If disturbances are active, deviations occur from the optimal path. The feedback controller uses the current state (and time) to determine its best current control value and hence can adapt to such deviations. This is the main advantage of feedback control over open-loop control and explains why dynamic programming is used more often for control than the maximum principle.

However, the computation of the optimal controls with the techniques mentioned above are not explicit in the sense that they do not lead to a suitable form for simple numerical implementation. Hence, the need arises for an easily implementable approximation method. Our aim is not to give a complete treatise of solving the finite horizon problem numerically, but briefly state a possible approximation method based on discrete dynamic programming, as initiated by Bellman. However, the analysis of convergence of the approximations to the exact optimal control is crucial and justifies the proposed algorithm. For more details on implementation aspects, we refer to more specialized books like Kushner and Dupuis (1992). We discretize our optimization problem in both time and state. Similar techniques can be found, for instance in Kirk (1970) and Kushner and Dupuis (1992) with some illustrative examples. In later sections, we will see how this method can be used to approximate also the infinite horizon optimal feedback. This method is justified by the convergence results between the finite and infinite horizon problems.

The organization of the paper is as follows. In section 2 the problem is formulated for the finite horizon case. Section 3 contains a motivating example of a pendulum to show that there is no obvious connection between the unconstrained and constrained Linear Quadratic Regulator problem. In section 4, we recall the concept of projections on closed convex sets in Hilbert spaces. The next section, section 5, contains, under a regularity condition, the existence and uniqueness of optimal controls with various characterizations of the optimal control. In section 6, a recursive scheme to approximate the optimal control will be described with convergence results of the approximations to the optimal control. The infinite horizon case and its connection with the finite horizon case will be the object of study in section 7. Finally, the conclusions are stated.

## 2. Problem formulation

We consider a linear system with input or control  $u: [t, T] \rightarrow \mathbb{R}^m$ , state  $x: [t, T] \rightarrow \mathbb{R}^n$  and output  $z: [t, T] \rightarrow \mathbb{R}^p$ , given by

$$\dot{x}(s) = Ax(s) + Bu(s) \quad (1)$$

$$z(s) = Cx(s) + Du(s) \quad (2)$$

where  $s$  denotes the time,  $A$ ,  $B$ ,  $C$  and  $D$  are matrices of appropriate dimensions and  $T$  is a fixed end time. We consider inputs in the Lebesgue space of square integrable, measurable functions on  $[t, T]$  taking values in  $\mathbb{R}^m$ , denoted by  $L_2[t, T]^m$ . For every input function  $u \in L_2[t, T]^m$  and initial condition  $(t, x_0)$ , i.e.  $x(t) = x_0$ , the solution of (1) is an absolutely continuous state trajectory, denoted by  $x_{t, x_0, u}$ . The

corresponding output can be written as

$$z_{t,x_0,u}(s) = \underbrace{C e^{A(s-t)} x_0}_{(\mathcal{M}_{t,T}x_0)(s)} + \underbrace{\int_t^s C e^{A(s-\tau)} Bu(\tau) d\tau + Du(s)}_{(\mathcal{L}_{t,T}u)(s)} \tag{3}$$

for  $s \in [t, T]$   $\mathcal{M}_{t,T}$  is a bounded linear operator from  $\mathbb{R}^n$  to  $L_2[t, T]$  and  $\mathcal{L}_{t,T}$  is a bounded linear operator from  $L_2[t, T]^m$  to  $L_2[t, T]^p$ .

The closed convex cone of positive functions in  $L_2[t, T]^p$  is defined by

$$P[t, T] := \{u \in L_2[t, T]^p \mid u(s) \in \Omega \text{ almost everywhere on } [t, T]\}$$

where the control restraint set  $\Omega$  equals  $\mathbb{R}_+^m := \{\mu \in \mathbb{R}^m \mid \mu_i \geq 0\}$ , the non-negative orthant in  $\mathbb{R}^m$ . In fact, we can take the control restraint set to be any closed convex polyhedral cone  $\tilde{\Omega} \in \mathbb{R}^m$ . Parametrizing  $\tilde{\Omega}$  as  $\tilde{\Omega} = F\Omega$ , where  $F$  is an  $\tilde{m} \times m$  matrix and taking  $(\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D}) = (A, BF, C, DF)$  translates the problem into an optimization problem with positive controls (as in Pachter 1980).

First, we consider the finite horizon case ( $T < \infty$ ). The infinite horizon version ( $T = \infty$ ) will be postponed to section 7.

**Problem 1:** *The objective is to determine for every initial condition  $(t, x_0) \in [0, T] \times \mathbb{R}^n$  a control input  $u \in P[t, T]$  an optimal control, such that*

$$\begin{aligned} J(t, x_0, u) &:= \|z_{t,x_0,u}\|_2^2 = \|\mathcal{M}_{t,T}x_0 + \mathcal{L}_{t,T}u\|_2^2 \\ &= \int_t^T z_{t,x_0,u}^T(s) z_{t,x_0,u}(s) ds \\ &= \int_t^T [x_{t,x_0,u}^T(s) C^T C x_{t,x_0,u}(s) + x_{t,x_0,u}^T(s) C^T D u(s) + u^T(s) D^T D u(s)] ds \end{aligned} \tag{4}$$

is minimal.

Along with this problem formulation, there are several questions to be answered. Does there exist an optimal control? And moreover, if it exists, is it unique and how is it characterized? And last but not least: can the optimal control be explicitly computed or is there a numerical method that approximates the exact optimal control? In the latter case, the method is acceptable if there are corresponding convergence results that justify the use of the method. Answering these questions is the main goal of this paper. Note that the only question considered in Pachter (1980) is the first one, which is concerned with existence of optimal controls.

In fact, this is a minimum-norm problem over a closed convex cone. The control functions minimizing the criterion in (4) for given initial conditions and horizon are called *optimal controls*. The optimal value of  $J$  for all considered initial conditions is described by the value function.

**Definition 1—Value function:** The value function  $V$  is a function from  $[0, T] \times \mathbb{R}^n$  to  $\mathbb{R}$  and is defined for every  $(t, x_0) \in [0, T] \times \mathbb{R}^n$  by

$$V(t, x_0) := \inf_{u \in P[t, T]} J(t, x_0, u) \tag{5}$$

□

As announced in the introduction, we will restrict ourselves to the regular problem. By this we mean that the matrix  $D$  is injective, i.e.  $D$  has full column rank. To motivate the regularity assumption, we consider a couple of alternatives.

In the singular case, the optimal controls may not exist in an  $L_2$ -setting. Look at the following simple example.

**Example 1:** Consider the system

$$\dot{x}(s) = u(s)$$

with criterion  $\int_0^T x^2(s) ds$  and initial condition  $x(0) = -1$ . The optimal control will be the Dirac pulse  $\delta$ , which results in an initial jump to 0 at time instant 0. The optimal costs are 0 in this case. However, the Dirac pulse is no  $L_2$ -function. In case we put a weight on the control, the Dirac pulse leads to an infinite value of the cost function. □

To circumvent the above problem without imposing weights on the control in the cost function, traditionally one restricted the controls to take values in compact sets. The resulting optimal controllers are non-smooth ‘bang-bang’ controllers attaining only values at the saturation borders of the restraint set.

**Example 2:** Looking at the optimal control problem above with restraint set  $\Omega := [0, 1]$  it is obvious that  $u(s) = 1, s \in [0, 1)$  and  $u(s) = 0, s \geq 1$  is the optimal control. Note that the optimal control is discontinuous. □

We are using the weighting of the control to guarantee boundedness and moreover obtain smoothness: the optimal control and the value function behave rather smoothly, as we will see. Physical implementation of bounded smooth controllers is preferred in most cases. Saturation characteristics and rate limiters often obscure the physical implementation of impulsive, bang-bang or non-smooth (discontinuous) controls.

### 3. Motivating example

A first thought may be that there exists a simple connection between the optimal static feedback corresponding to the unconstrained optimal control problem ( $\Omega = \mathbb{R}^m$ ) and the optimal control in the constrained problem ( $\Omega = \mathbb{R}_+^n$ ). It is well-known that the optimal feedback in the unconstrained case arises from the Algebraic Riccati Equation (Anderson and Moore 1990).

To show that such a simple connection is not easily established, we consider the pendulum as in figure 1. On a pendulum the gravitation  $f$  and the control force  $u$  act as a vertical force and horizontal force, respectively. The angle the pendulum makes with the vertical dashed line, is called  $\theta$ . Since it is only allowed to push against the pendulum from one side, the control is restricted to be non-negative. After proper scaling of the time and linearizing around the equilibrium  $\theta = \dot{\theta} = u = 0$  we obtain the equations

$$\dot{x}_1 = x_2$$

$$\dot{x}_2 = -x_1 - u$$

where  $x_1$  is the first-order approximation of  $\theta$  and  $x_2$  the first-order approximation of  $\dot{\theta}$ .

Our objective is to minimize

$$\int_0^\infty \{x_1^2(s) + x_2^2(s) + u^2(s)\} ds \tag{6}$$

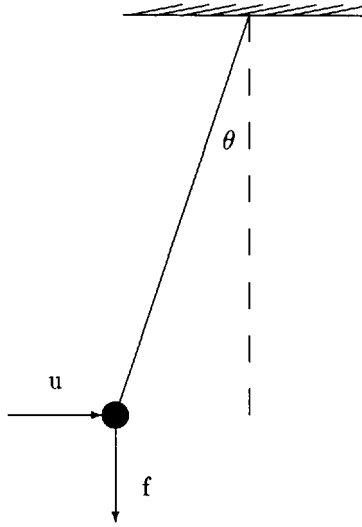
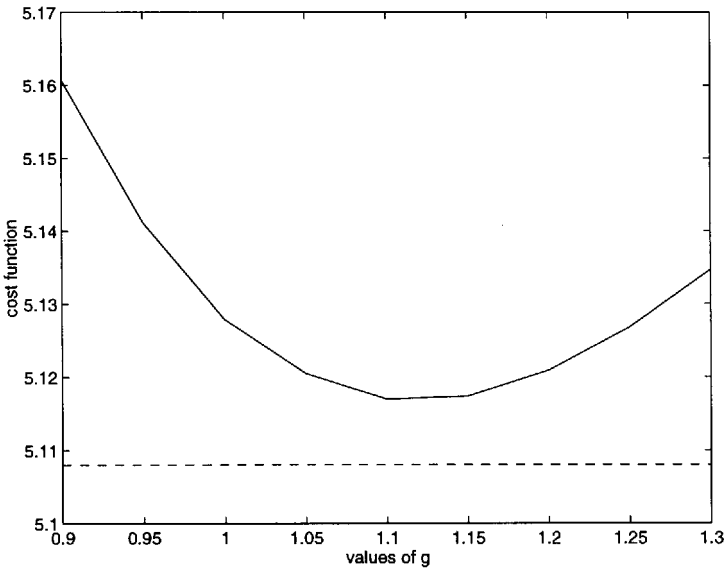


Figure 1. Pendulum.

with initial state  $(1,0)^T$  on time instant  $t = 0$ . Without the positivity constraint the optimal solution would be given by static state-feedback  $u(x) = Kx$ . In the case of this single input system a good guess of the optimal positive control could be  $u_g(x) = g \max(0, Kx)$  for some suitable chosen gain  $g \geq 0$ . In particular,  $u_1$  seems a good candidate. For different values of  $g$ , we computed the expression (6) by simulating the system with corresponding feedback  $u_g$  for initial state  $(1,0)^T$ . The results are plotted in figure 2.

Indeed, we observe  $u_1$  is not optimal, because  $u_{1.13}$  performs better. In section 7

Figure 2. Cost function for various values of the gain  $g$ .

an approximation of the optimal state feedback is computed. The cost function for this controller equals 5.108, the level indicated by the dashed line in the picture. Our observation is that this controller is even better and indicates that for this example no trivial connection exists between the constrained and unconstrained problem.

**4. Projections on closed convex sets**

As mentioned before, the problem we consider is a minimum norm problem over a closed convex set. This motivates the introduction of a generalization of orthogonal projections on closed subspaces in Hilbert spaces. Consider the following fundamental theorem in Hilbert space theory concerning the minimum distance to a closed convex set. For a proof, see chapter 3 of Luenberger (1969).

**Theorem 1—Minimum distance to a convex set:** *Let  $x$  be a vector in a Hilbert space  $H$  with inner product  $(\cdot | \cdot)$  and let  $K$  be a closed convex subset of  $H$ . Then there is exactly one  $k_0 \in K$  such that  $\|x - k_0\| \leq \|x - k\|$  for all  $k \in K$ . Furthermore, a necessary and sufficient condition that  $k_0$  is the unique minimizing vector is that  $(x - k_0 | k - k_0) \leq 0$  for all  $k \in K$ .*

**Definition 2:** Let  $K$  be a closed convex set of the Hilbert space  $H$ . We introduce the projection  $P_K$  onto  $K$  as the operator that assigns to each vector  $x$  in  $H$ , the vector contained in  $K$  that is closest to  $x$  in the norm induced by the inner product. Formally,

$$P_K x = k_0 \in K \iff \|x - k_0\| \leq \|x - k\| \quad \forall k \in K \tag{7}$$

for  $x \in H$ . □

This concept can be found for instance in Kirk (1970). Theorem 1 justifies this definition and gives an equivalent characterization of  $P_K$

$$P_K x = k_0 \in K \iff (x - k_0 | k - k_0) \leq 0 \quad \forall k \in K \tag{8}$$

A set  $K$  in a linear vector space is a cone, if  $x \in K$  implies  $\alpha x \in K$  for all  $\alpha \geq 0$ . In case  $K$  is such a closed convex cone, we can take  $k = \frac{1}{2}k_0$  and  $k = 2k_0$  in the above equation to observe that  $(x - k_0 | k_0) = 0$  and hence

$$(x - k_0 | k) \leq 0 \quad \forall k \in K \tag{9}$$

To keep the paper self-contained, the following two lemmas, describing further properties that can also be found in Hiriart-Urruty and Lemaréchal (1993), are included.

**Lemma 1—Continuity of  $P_K$ :**  *$P_K$  is globally Lipschitz continuous. In particular, for all  $x, y \in H$  it holds that*

$$\|P_K x - P_K y\| \leq \|x - y\| \tag{10}$$

**Proof:** Using the characterization (8) for  $x$  with  $k = P_K y$  we arrive at  $(x - P_K x | P_K y - P_K x) \leq 0$ . Changing the roles of  $x$  and  $y$  we get  $(y - P_K y | P_K x - P_K y) \leq 0$ . Adding both inequalities gives  $(x - y + P_K y - P_K x | P_K y - P_K x) \leq 0$ , which leads to  $\|P_K x - P_K y\|^2 \leq (x - y | P_K x - P_K y)$ . Applying the Cauchy-Schwarz inequality to the right side and dividing by  $\|P_K x - P_K y\|$  completes the proof. □

**Lemma 2—Positive-homogeneity of  $P_K$ :** *Besides  $K$  being closed and convex, assume it is a cone. Then*



$$P_K(\alpha x) = \alpha P_K x \quad \forall \alpha \geq 0, x \in X$$

**Proof:** The proof follows immediately from (8).  $\square$

## 5. Optimal controls

In this section, we answer the first two questions raised in section 2, just after the problem formulation.

### 5.1. Existence and uniqueness

As mentioned before, to establish the existence of optimal controls the results from Pachter (1980) cannot be used, because the condition formulated there is only explicit in case of  $C^T C = 0$ ,  $C^T D = 0$ .

As discussed in section 2, a standing assumption in the remainder of the paper will be the full column rank (or injectivity) of  $D$ . In this regular case,  $\mathcal{L}_{t,T}$ , as defined in section 2, has a bounded left inverse from  $L_2[t, T]$  to  $L_2[t, T]^n$ . An operator  $\tilde{\mathcal{L}}_{t,T}$  is a bounded left inverse of  $\mathcal{L}_{t,T}$ , if  $\mathcal{L}_{t,T} u = z$  implies  $\tilde{\mathcal{L}}_{t,T} z = u$ . By manipulating (1)–(2) it follows that one of the bounded left inverses of  $\mathcal{L}_{t,T}$  is described by the following state–space representation

$$\begin{aligned} \dot{x}(s) &= (A - B(D^T D)^{-1} D^T C)x(s) + B(D^T D)^{-1} D^T z(s), & x(t) &= 0 \\ u(s) &= (D^T D)^{-1} D^T \{z(s) - Cx(s)\} \end{aligned}$$

We denote this particular left inverse by  $\tilde{\mathcal{L}}_{t,T}$ .

Using this bounded left inverse, we can reformulate our problem as the minimization of  $\|v - \mathcal{M}_{t,T} x_0\|$  over  $v \in \mathcal{L}_{t,T}(P[t, T])$ . However, this is exactly the setting of Theorem 1 with  $K = \mathcal{L}_{t,T}(P[t, T])$  and  $k_0 = \mathcal{L}_{t,T}(u^*)$ . The linearity of  $\mathcal{L}_{t,T}$  shows that  $\mathcal{L}_{t,T}(P[t, T])$  is a convex cone. Moreover, since  $K$  is the inverse image of a closed set under  $\mathcal{L}_{t,T}$ , it is closed as well. Hence, we have proved Theorem 2.

**Theorem 2—Existence and uniqueness of the optimal control:** *Let  $t, T \in \mathbb{R}$  with  $t < T$  and  $x_0 \in \mathbb{R}^n$ . There is a unique control  $u_{t,T,x_0} \in P[t, T]$  such that*

$$\|\mathcal{M}_{t,T} x_0 + \mathcal{L}_{t,T} u_{t,T,x_0}\|_2 \leq \|\mathcal{M}_{t,T} x_0 + \mathcal{L}_{t,T} u\|_2$$

for all  $u \in P[t, T]$ . A necessary and sufficient condition for  $u^* \in P[t, T]$  to be the unique minimizing control is that

$$(\mathcal{M}_{t,T} x_0 + \mathcal{L}_{t,T} u^* \mid \mathcal{L}_{t,T} u - \mathcal{L}_{t,T} u^*) \geq 0 \quad (11)$$

for all  $u \in P[t, T]$

The optimal control, the optimal trajectory and the optimal output with initial conditions  $(t, x_0)$  and final time  $T$  will be denoted by  $u_{t,T,x_0}$ ,  $x_{t,T,x_0}$  and  $z_{t,T,x_0}$ , respectively.

Considering the proof above, it is not hard to see that, in terms of projections, we can write

$$u_{t,T,x_0} = \tilde{\mathcal{L}}_{t,T} \mathcal{P}(-\mathcal{M}_{t,T} x_0) \quad (12)$$

where  $\mathcal{P}$  is the projection on the closed convex cone  $\mathcal{L}_{t,T}(P[t, T])$  in the Hilbert space  $L_2[t, T]$ .

From (12) and the positive-homogeneity of  $\mathcal{P}$ , it is clear that for  $\alpha \geq 0$

$$u_{t,T,\alpha x_0} = \alpha u_{t,T,x_0} \quad (13)$$

and

$$V(t, \alpha x_0) = \alpha^2 V(t, x_0) \tag{14}$$

### 5.2. Maximum principle

In spite of the fact that the results in this subsection are classical (see for example Pontryagin *et al.* 1962), they are stated here because in section 7 this theory is exploited to derive a maximum principle for the infinite horizon case and convergence results between finite and infinite horizon optimal controls. Moreover, in a complete overview of available techniques to tackle constrained optimal control problems, the maximum principle cannot be absent.

Consider the Hamiltonian

$$H(x, \mu, \phi) = x^T A^T \phi + \mu^T B^T \phi + x^T C^T Cx + 2x^T C^T D\mu + \mu^T D^T D\mu$$

for  $(x, \mu, \phi) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ .

The adjoint or costate equation and the corresponding terminal conditions read

$$\dot{\phi} = - \frac{\partial H(x, \mu, \phi)}{\partial x} = - A^T \phi - 2C^T Cx - 2C^T D\mu = - A^T \phi - 2C^T z, \quad \phi(T) = 0 \tag{15}$$

where  $z = Cx + D\mu$  and the column vector valued function  $\partial H / \partial x$  denotes the partial derivative of  $H$  with respect to  $x$ . Pontryagin's maximum principle states that the optimal control  $u_{t,T,x_0}$ , satisfies for all  $s \in [t, T]$

$$u_{t,T,x_0}(s) \in \arg \min_{\mu \in \mathbb{R}^m_+} H(x_{t,T,x_0}(s), \mu, \phi_{t,T,x_0}(s)) \tag{16}$$

where  $\phi_{t,T,x_0}$  is the solution to the adjoint equation (15) with  $z$  equal to the optimal output  $z_{t,T,x_0}$ .

We introduce  $\|x\|_{D^T D} := x^T D^T D x$  as the norm induced by the inner product  $(x|y)_{D^T D} = x^T D^T D y$  for  $x, y \in \mathbb{R}^n$  in the Hilbert space  $\mathbb{R}^n$ . Furthermore, in this Hilbert space  $P_\Omega$  denotes the projection on  $\Omega := \mathbb{R}^m_+$ .

**Lemma 3:** *Let the initial time  $t$ , the final time  $T > 0$  and initial state  $x_0 \in \mathbb{R}^n$  be fixed. The optimal control  $u_{t,T,x_0}$  satisfies*

$$u_{t,T,x_0}(s) = P_\Omega(-\frac{1}{2}(D^T D)^{-1} \{B^T \phi_{t,T,x_0}(s) + 2D^T Cx_{t,T,x_0}(s)\}) \tag{17}$$

for all  $s \in [t, T]$  where the functions  $x_{t,T,x_0}$  and  $\phi_{t,T,x_0}$  are given by

$$\left. \begin{aligned} \dot{x}_{t,T,x_0} &= Ax_{t,T,x_0} + Bu_{t,T,x_0} & x(t) &= x_0 \\ \dot{\phi}_{t,T,x_0} &= -A^T \phi_{t,T,x_0} - 2C^T Cx_{t,T,x_0} - 2C^T Du_{t,T,x_0} & \phi(T) &= 0 \end{aligned} \right\} \tag{18}$$

Moreover, the optimal control is continuous in time.

**Proof:** This lemma is a straightforward application of the maximum principle. By (16),  $u_{t,T,x_0}(s)$  is the unique pointwise minimizer of the costs

$$\mu^T D^T D\mu + \mu^T g(s) = \|\mu + \frac{1}{2}(D^T D)^{-1} g(s)\|_{D^T D}^2 - \frac{1}{4} g^T(s) (D^T D)^{-1} g(s)$$

taken over all  $\mu \in \Omega$ , where  $g(s) := B^T \phi_{t,T,x_0}(s) + 2D^T Cx_{t,T,x_0}(s)$ .

According to Lemma 1 the projection  $P_\Omega$  is continuous. Since  $x_{t,T,x_0}$  and  $\phi_{t,T,x_0}$  are (absolutely) continuous in time, the continuity of  $u_{t,T,x_0}$  in time follows.  $\square$

Substituting (17) in (18) gives a two-point boundary value problem. Its solutions

provide us with a set of candidates containing the optimal control, because the maximum principle is a necessary condition for optimality.

In the case  $D^T D = \frac{1}{2} I$  with  $I$  the identity matrix, the expression  $P_\Omega(-\frac{1}{2}(D^T D)^{-1}\{B^T \phi(s) + 2D^T Cx(s)\})$  simplifies to

$$\max \{0, -B^T \phi(s) - 2D^T Cx(s)\}$$

where ‘max’ for vectors means taking the maximum componentwise.

### 5.3. Dynamic programming

In dynamic programming the value function satisfies the Hamilton–Jacobi–Bellman equation. The difficulty in using this partial differential equation is often that the value function is not smooth and classical results do not apply. Recently, an extended solution concept has been used, called the ‘viscosity solution’ of the HJB-equation (Fleming and Soner 1993). We show in this subsection that this complicated solution concept is not required for the LQ-problem with positive controls: the value function is continuously differentiable and a classical solution to the HJB equation. This simplifies both analytical and numerical approaches of using dynamic programming for solving the problem at hand.

We start with a short overview of the technique of dynamic programming. To do so, we introduce the function  $L$  (called ‘Lagrangian’) for  $(x, \mu) \in \mathbb{R}^n \times \Omega$  by

$$L(x, \mu) = z^T z = x^T C^T Cx + 2x^T C^T D\mu + \mu^T D^T D\mu$$

The HJB equation is given for  $x \in \mathbb{R}^n$  and  $t \in [0, T]$  by

$$W_t(t, x) + \tilde{H}(x, W_x(t, x)) = 0 \quad (19)$$

where  $\tilde{H}$  is given by

$$\tilde{H}(x, p) = \inf_{\mu \in \Omega} \{p^T Ax + p^T B\mu + L(x, \mu)\} \quad (20)$$

for  $(x, p) \in \mathbb{R}^n \times \mathbb{R}^n$ .  $W$  is a function with domain  $[0, T] \times \mathbb{R}^n$  taking values in  $\mathbb{R}$ , where  $W_t$  and  $W_x$  (considered to be a column vector valued function) denote its partial derivatives with respect to time and state  $x$ , respectively (provided they exist). Before continuing, we introduce the function space  $L_\infty[t, T]^n$  as the normed space of all essentially bounded Lebesgue measurable functions on  $[t, T]$ . The so-called ‘verification theorem’ (Fleming and Soner 1993) states: if  $W$  is a continuously differentiable solution of the HJB equation, satisfying the boundary conditions  $W(T, x) = 0$ ,  $x \in \mathbb{R}^n$ , then  $W(t, x) \leq V(t, x)$ , where  $V$  denotes the value function as introduced in section 2. Moreover, if there exists  $u^* \in L_\infty[t, T]^n$  such that

$$u^*(s) = \arg \min_{\mu \in \Omega} \{W_x(s, x^*(s))^T Ax^*(s) + W_x(s, x^*(s))^T B\mu + L(x^*(s), \mu)\} \quad (21)$$

for almost all  $s \in [t, T]$  then  $u^*$  is optimal for initial data  $(t, x_0)$ . In the above expression,  $x^*$  is the state trajectory corresponding to initial condition  $(t, x_0)$  and input  $u^*$ . If such a function  $W$  is found, (21) is a *sufficient* condition for optimality.

By manipulating (20), we get (analogously as in the previous section)

$$\begin{aligned} \tilde{H}(x, p) &= x^T C^T Cx + p^T Ax - \frac{1}{4} g^T(x, p) (D^T D)^{-1} g(x, p) \\ &\quad + \inf_{\mu \in \Omega} \left\| \mu + \frac{1}{2} (D^T D)^{-1} g(x, p) \right\|_{D^T D}^2 \end{aligned} \quad (22)$$

with  $g(x, p) = B^T p + 2D^T Cx$ . The minimizing  $\mu$  equals  $P_\Omega(-\frac{1}{2}(D^T D)^{-1} g(x, p))$ .

We shall now prove the continuous differentiability of the value function and show that it satisfies (19). Note that for the usage of the verification theorem as stated above, with  $W = V$ , the continuous differentiability of  $V$  is required.

**Theorem 3—Differentiability of  $V$  with respect to  $x$ :** *The value function  $V$  is continuously differentiable with respect to  $x$  and its directional derivative at the point  $(t, x_0) \in [t, T) \times \mathbb{R}^n$  with increment  $h \in \mathbb{R}^n$  is given by*

$$(V_x(t, x_0) \mid h) = 2(z_{t,T,x_0} \mid \mathcal{M}_{t,T}h)_2 \tag{23}$$

or, explicitly

$$V_x(t, x_0) = 2 \int_t^T e^{A^T(s-t)} C^T z_{t,T,x_0}(s) ds \tag{24}$$

$z_{t,T,x_0}$  is the output corresponding to initial condition  $(t, x_0)$  with optimal control  $u_{t,T,x_0}$ . Notice that in (23) the left inner product is the Euclidean inner product in  $\mathbb{R}^n$  and the right inner product is the  $L_2$ -inner product.

**Proof:** The proof is separated into two parts, giving upper and lower bounds of the expression  $V(t, x_0 + h) - V(t, x_0) - 2(z_{t,T,x_0} \mid \mathcal{M}_{t,T}h)_2$ , respectively.

(i) We have  $V(t, x_0 + h) \leq J(t, x_0 + h, u_{t,T,x_0}) = (z \mid z)_2$  with  $z = \mathcal{M}_{t,T}(x_0 + h) + \mathcal{L}_{t,T}(u_{t,T,x_0}) = \mathcal{M}_{t,T}h + z_{t,T,x_0}$ . Writing out the above inequality yields

$$V(t, x_0 + h) - \underbrace{(z_{t,T,x_0} \mid z_{t,T,x_0})}_{=V(t,x_0)} - 2(z_{t,T,x_0} \mid \mathcal{M}_{t,T}h)_2 \leq \|\mathcal{M}_{t,T}h\|_2^2 \leq \|\mathcal{M}_{t,T}\|^2 \|h\|^2 \tag{25}$$

(ii) Since  $J$  is quadratic in  $u$ , it can be expanded as

$$J(t, x, u) = V(t, x) + 2(\mathcal{L}_{t,T}u_{t,T,x} + \mathcal{M}_{t,T}x \mid \mathcal{L}_{t,T}u - \mathcal{L}_{t,T}u_{t,T,x})_2 + \|\mathcal{L}_{t,T}u - \mathcal{L}_{t,T}u_{t,T,x}\|_2^2$$

For ease of exposition, we omit the subscripts  $t$  and  $T$ . For example, we write  $u_{x_0}$  instead of  $u_{t,T,x_0}$  for the optimal controls and  $z_{x_0}$  instead of  $z_{t,T,x_0}$  for the corresponding optimal output.

It is easy to derive that

$$V(t, x_0 + h) = J(t, x_0 + h, u_{x_0} + u_h) - 2(\mathcal{L}u_{x_0+h} + Mx_0 + Mh \mid \mathcal{L}u_h + \mathcal{L}u_{x_0} - \mathcal{L}u_{x_0+h})_2 - \|\mathcal{L}u_h + \mathcal{L}u_{x_0} - \mathcal{L}u_{x_0+h}\|_2^2 \tag{26}$$

and

$$J(t, x_0 + h, u_{x_0} + u_h) = V(t, x_0) + 2(z_{x_0} \mid Mh + \mathcal{L}u_h)_2 + V(t, h)$$

Substituting this in (26) gives

$$V(t, x_0 + h) = V(t, x_0) + 2(z_{x_0} \mid Mh) - 2(\mathcal{L}u_{x_0} + Mx_0 \mid \mathcal{L}u_{x_0} - \mathcal{L}u_{x_0+h})_2 + R(t, x_0, h) \tag{27}$$

where the function  $R$  is given by

$$R(t, x_0, h) := - \|\mathcal{L}u_h + \mathcal{L}u_{x_0} - \mathcal{L}u_{x_0+h}\|_2^2 + V(t, h) + 2(\mathcal{L}u_{x_0+h} - \mathcal{L}u_{x_0} + Mh \mid \mathcal{L}u_{x_0} + \mathcal{L}u_h - \mathcal{L}u_{x_0+h})_2 \tag{28}$$

Because of (11), the last inner product in (27) is non-positive. So, we get

$$V(t, x_0 + h) - V(t, x_0) - 2(z_{x_0} \mid Mh)_2 \geq R(t, x_0, h) \quad (29)$$

We claim that the function  $R$  satisfies

$$|R(t, x_0, h)| \leq E \|h\|^2 \quad (30)$$

for certain positive constant  $E$ . This fact will be verified in a lemma in the appendix. Combining (25) and (29) then completes the proof. Note that continuity of the partial derivative follows by some simple calculations. For a formal proof, see Heemels (1998).  $\square$

Comparing (24) with the adjoint equation of the Maximum Principle yields a connection between the adjoint variable and the gradient of  $V$

$$\phi_{t, T, x_0}(s) = V_x(s, x_{t, T, x_0}(s)) \quad (31)$$

where  $\phi_{t, T, x_0}$  is the solution to the adjoint equation corresponding to  $u_{t, T, x_0}$ .

For proving the differentiability of  $V$  with respect to  $t$  we need an auxiliary result, which we formulate in the next lemma. It is an extension of the mean value theorem for functions of a real variable (see Thomas and Finney 1996).

**Lemma 4:** *Let  $f$  be a function from  $\mathbb{R}^n$  to  $\mathbb{R}$ , which is differentiable on  $\mathbb{R}^n$  with gradient  $f_x$ . Let  $x_0, x_1 \in \mathbb{R}^n$ . Then, there exists a  $\beta \in (0, 1)$  such that for*

$$\zeta_\beta = \beta x_1 + (1 - \beta)x_0$$

it holds that

$$f(x_1) - f(x_0) = (f_x(\zeta_\beta) \mid x_1 - x_0) \quad (32)$$

**Proof:** Consider the function  $h$  with domain  $[0, 1]$  defined by

$$h: \alpha \mapsto f(x_0 + \alpha[x_1 - x_0])$$

This function is continuous on  $[0, 1]$  and differentiable on  $(0, 1)$  with derivative

$$\frac{dh(\alpha)}{d\alpha} = (f_x(x_0 + \alpha[x_1 - x_0]) \mid x_1 - x_0)$$

Using the mean value theorem for functions of one real variable (Thomas and Finney 1996) we arrive at the existence of a  $\beta \in (0, 1)$  such that  $h(1) - h(0) = (dh(\alpha)) / (d\alpha)(\beta)$ . Translating this result to  $f$  proves the lemma.  $\square$

**Theorem 4—Differentiability of  $V$  with respect to  $t$ :** *The value function  $V$  is continuously differentiable with respect to  $t$  and the partial derivative in the point  $(t, x_0) \in [t, T) \times \mathbb{R}^n$  equals*

$$V_t(t, x_0) = - (V_x(t, x_0) \mid Ax + Bu_{t, T, x_0}(t)) - z_{t, T, x_0}^T(t) z_{t, T, x_0}(t) \quad (33)$$

**Proof:** Let  $\delta$  be such that  $0 \leq t < t + \delta \leq T$ , and write  $x^* = x_{t, T, x_0}$  for short. In the second step of the derivation below, we use Lemma 4, where

$\zeta_\beta(\delta) = \beta x_0 + (1 - \beta)x^*(t + \delta)$  for certain  $\beta \in (0, 1)$

$$\begin{aligned} \frac{V(t + \delta, x_0) - V(t, x_0)}{\delta} &= \frac{V(t + \delta, x_0) - V(t + \delta, x^*(t + \delta))}{\delta} \\ &\quad + \frac{V(t + \delta, x^*(t + \delta)) - V(t, x_0)}{\delta} \\ &= - \left( V_x(t + \delta, \zeta_\beta(\delta)) \left| \frac{x^*(t + \delta) - x_0}{\delta} \right. \right) \\ &\quad - \frac{1}{\delta} \int_t^{t+\delta} z_{t,T,x_0}^T(s) z_{t,T,x_0}(s) \, ds \end{aligned}$$

Because  $x^*(t + \delta) \rightarrow x_0(\delta \downarrow 0)$  it follows that  $\zeta_\beta(\delta) \rightarrow x_0(\delta \downarrow 0)$ . The continuity of  $V_x$  implies now that

$$\frac{V(t + \delta, x_0) - V(t, x_0)}{\delta} \rightarrow - (V_x(t, x_0) \mid \dot{x}^*(t)) - z_{t,T,x_0}^T(t) z_{t,T,x_0}(t)$$

if  $\delta \downarrow 0$ . Since  $u_{t,T,x_0}$  is continuous,  $x^*$  is continuously differentiable in time and  $\dot{x}^*(t)$  equals  $Ax + Bu_{t,T,x_0}(t)$ .

We proved the right differentiability of  $V$  with respect to  $t$ . The left differentiability can be deduced almost analogously and gives the same derivative. With this last remark the proof is completed.

By using (31) and the maximum principle, which states that the  $u_{t,T,x_0}$  minimizes the Hamiltonian, one recognizes the Hamilton–Jacobi–Bellman equation in (33). In other words, the value function satisfies the HJB equation (19) in a classical sense. Since  $V_x$  and  $\tilde{H}$  are continuous in its arguments, the right-hand side of (19) is continuous and thus so is the left-hand side. Hence,  $V_t$  is continuous in its arguments. □

We extract the following statement from the proof.

**Remark 1:** The value function satisfies the HJB equation (19) in the classical sense. □

Since it is clear that  $V(T, x) = 0, x \in \mathbb{R}^n$ , the verification theorem states that

$$u_{t,T,x_0}(s) = P_\Omega(-\frac{1}{2}(D^T D)^{-1} [B^T V_x(s, x_{t,T,x_0}(s)) + 2D^T Cx_{t,T,x_0}(s)]) \tag{34}$$

This time-varying optimal feedback can also be obtained by combining (31) and (17). Actually, this yields (34) without using the verification theorem.

Moreover, we can even conclude that both dynamic programming and the maximum principle are necessary and sufficient conditions for optimality. This follows from the link (31) between the two methods. The necessity of the maximum principle in terms of Lemma 3 translates via (31) into the necessity of dynamic programming in terms of (34), and vice versa, the sufficiency of dynamic programming leads to sufficiency of the maximum principle along the same lines.

**Remark 2:** The results obtained so far hold also for time-varying systems of the form

$$\begin{aligned} \dot{x}(s) &= A(s)x(s) + B(s)u(s) \\ z(s) &= C(s)x(s) + D(s)u(s) \end{aligned}$$

with control restraint set  $\Omega$  being an arbitrary closed convex set containing the origin. However, we have to assume smoothness conditions like  $A(\cdot), B(\cdot), C(\cdot), D(\cdot)$  are continuous and bounded,  $D(s)$  has full column rank for all  $s$  and  $(D^T(\cdot)D(\cdot))^{-1}$  bounded and continuous.  $\square$

## 6. Approximative aspects

Both the maximum principle and dynamic programming lead to a 2-point boundary value problem and a partial differential equation, respectively, which provide methods to extract (candidates for) the optimal control when solved. A drawback of solving a 2-point boundary value problem arising from the maximum principle is that it specifies only an open-loop optimal control function for one particular initial condition. For practical implementation of a controller, a feedback law specifying the optimal control value for all states and times is much more effective, in particular when disturbances are active. Moreover, solving a 2-point boundary value problem is not trivial at all. From this point of view a feedback like (34) is preferred. Since solving a partial differential equation is, in general, very complicated and time consuming, an alternative approach resulting in a feedback law will be investigated. We would like to stress that our aim is not to give a complete treatise of the numerical solution of the problem at hand, but to specify a possible framework and justify its effectiveness by convergence results.

Closer inspection of the expression  $u_{l,T,x_0} = \mathcal{L}_{l,T} \mathcal{P}(-\mathcal{M}_{l,T}x_0)$  reveals that only  $\mathcal{P}$ —the projection on  $\mathcal{L}_{l,T}(\mathcal{P}[l, T])$ —cannot explicitly be computed. The other two operators are given by a state-space description. Hence, to get an approximation using this description of the optimal control function, we must focus on approximating projections  $P_K$  with a complex convex set  $K$  (in our case:  $K = \mathcal{L}_{l,T}(\mathcal{P}[l, T])$ ). One way of doing this is considering a sequence of closed convex subsets  $\{K_n\}_{n \in \mathbb{N}}$  of  $K$ . This sequence has to fill up the set  $K$  from the inside and  $P_{K_n}$  should—of course—be easier to compute. To make this more specific, we write  $K_n \uparrow K$  ( $n \rightarrow \infty$ ), if  $K_n \subseteq K$ ,  $n \in \mathbb{N}$  and  $\text{dist}(k, K_n) := \inf \{\|k - l\| \mid l \in K_n\} \rightarrow 0$  ( $n \rightarrow \infty$ ) for all  $k \in K$ .

Since  $K_n$  is a subset of  $K$ , it is obvious that  $\|x - P_K x\| \leq \|x - P_{K_n} x\|$  and in general it is even strictly smaller. However, for increasing  $n$ ,  $K_n$  becomes closer and closer to  $K$ , so one may guess that the difference between  $\|x - P_K x\|$  and  $\|x - P_{K_n} x\|$  vanishes. By using this result and the parallelogram law, it can be shown that  $P_{K_n} x$  converges to  $P_K x$  for increasing  $n$ .

**Theorem 5—Approximation of projections:** *Let  $K$  be a closed convex set in a Hilbert space  $H$ . Let  $\{K_n\}_{n \in \mathbb{N}}$  be a sequence of closed convex subsets of  $K$  with  $K_n \uparrow K$  ( $n \rightarrow \infty$ ). Then for all  $x \in H$  we have  $P_{K_n} x \rightarrow P_K x$  ( $n \rightarrow \infty$ ).*

**Proof:** This proof resembles the proof of Theorem 1 in Luenberger (1969).

Let  $\{k_n\}_{n \in \mathbb{N}}$  be a sequence that converges to  $P_K x \in K$  with the property that  $k_n \in K_n$  for all  $n \in \mathbb{N}$ . Since  $K_n \subseteq K$ , we get for all  $n \in \mathbb{N}$  the inequality

$$\delta := \|x - P_K x\| \leq \|x - P_{K_n} x\| \leq \|x - k_n\| \rightarrow \|x - P_K x\| =: \delta \quad (n \rightarrow \infty)$$

Hence,  $\lim_{n \rightarrow \infty} \|x - P_{K_n} x\| = \|x - P_K x\|$ .

Next, we show that  $\{P_{K_n} x\}_{n \in \mathbb{N}}$  is a Cauchy sequence. By the parallelogram law

$$\|P_{K_n} x - P_{K_m} x\|^2 = 2\|P_{K_n} x - x\|^2 + 2\|P_{K_m} x - x\|^2 - 4 \left\| x - \frac{P_{K_n} x + P_{K_m} x}{2} \right\|^2$$

The convexity of  $K$  yields  $\|x - (P_{K_n}x + P_{K_m}x)/2\| \geq \delta$ . Hence

$$\|P_{K_n}x - P_{K_m}x\|^2 \leq 2\|P_{K_n}x - x\|^2 + 2\|P_{K_m}x - x\|^2 - 4\delta^2 \rightarrow 0 (n, m \rightarrow \infty)$$

The sequence  $\{P_{K_n}x\}_n \in \mathbb{N}$  is Cauchy and let  $k_0 \in K$  be its limit. By continuity,  $\|x - k_0\| = \|x - P_Kx\|$ . The uniqueness of the minimizing vector implies that  $k_0 = P_Kx$ . □

To apply the above procedure a sequence of closed convex subsets has to be constructed, which fill up the admissible control set as described. One method that is often used in practice, is discretization by requiring additionally that the input  $u$  is piecewise constant. Formally, let  $h := (T - t)/N$  for some  $N \in \mathbb{N}$  and  $t_i = t + ih$ ,  $i = 0, \dots, N$ . Define  $P_h[t, T]$  as

$$\{u \in P[t, T] \mid u|_{[t_i, t_{i+1})} \text{ is constant, } i = 0, \dots, N - 1\}$$

Analogous to subsection 5.1, it can be shown that there exists a unique control function in  $P_h[t, T]$  which minimizes the cost function for initial condition  $(t, x_0)$  over the set  $P_h[t, T]$ . This optimal discrete control with time-step  $h$  will be denoted by  $u_{t,T,x_0}^h$ .

Since the collection of all step-functions with different time-steps forms a dense subset of  $L_2[t, T]^n$  and  $\mathcal{L}_{t,T}$  is a continuous, linear transformation, we see that

$$\mathcal{L}_{t,T}(P_h[t, T]) \uparrow \mathcal{L}_{t,T}(P[t, T]) (h \downarrow 0)$$

Theorem 5 shows that  $P_{\mathcal{L}_{t,T}(P_h[t, T])}$  converges pointwise to  $\mathcal{P} := P_{\mathcal{L}_{t,T}(P[t, T])}$ . Using (12) and realizing that

$$u_{t,T,x_0}^h = \tilde{\mathcal{L}}_{t,T} P_{\mathcal{L}_{t,T}(P_h[t, T])} (-\mathcal{M}_{t,T}x_0)$$

yields Theorem 6.

**Theorem 6:** *The optimal discrete controls converge in the norm of  $L_2$  to the exact optimal control when the time-step converges to zero. Put in a formula*

$$u_{t,T,x_0}^h \xrightarrow{L_2} u_{t,T,x_0} \quad (h \downarrow 0)$$

The last question of this section to be answered is how to compute the optimal discrete control for a fixed time-step  $h$ . In fact, by the discretizations, the problem is transformed to an optimal control problem of a discrete time system in a way that the original techniques of Bellman's dynamics programming can be used, see Bellman (1967) and Kirk (1970).

Every control  $u^h \in P_h[t, T]$  can be parametrized as

$$u^h = \sum_{i=1}^N u_i^h 1_{[t_i, t_{i+1})} \tag{35}$$

with  $u_i^h \in \Omega$ . For such a discrete control  $u^h$ , we denote the corresponding solution of (1) with initial condition  $(t, x_0)$  by  $x$  and introduce  $x^h(i) := x(t_i)$ ,  $i = 0, \dots, N$ . For  $x^h$  we can write

$$\begin{aligned} x^h(i+1) &= e^{A_h} x^h(i) + \int_{t_i}^{t_{i+1}} e^{A(t_{i+1}-\tau)} B u^h(\tau) d\tau = e^{A_h} x^h(i) + \int_0^h e^{A(h-\theta)} B d\theta u_i^h \\ &=: A_h x^h(i) + B_h u_i^h \end{aligned} \tag{36}$$

with initial condition  $x^h(0) = x_0$ .



Furthermore, we introduce a discrete version of the value function.

**Definition 3:** Fix times  $t, T$  such that  $t < T < \infty$ . Fix time-step  $h = (T - t)/N$  for some positive integer  $N$ . The function  $V^h$  from  $\{0, \dots, N\} \times \mathbb{R}^n$  to  $\mathbb{R}$  is for  $(i, x_0) \in \{0, \dots, N\} \times \mathbb{R}^n$  defined by

$$V^h(i, x_0) = \min_{u^h \in \mathcal{P}_h[t_i, T]} J(t_i, x_0, u^h), \quad (37)$$

where  $J$  is defined as in (4). □

We set up a recursive scheme that is based on the *optimality principle* which states that

$$u_{t_i, T, x_0}^h \big|_{[t_i, T]} = u_{t_i, T, x^h(i)}^h \quad (38)$$

where  $x^h$  is the solution to (36) with initial condition  $x^h(0) = x_0$  and control  $u_{t_i, T, x_0}^h$ . In words, this says that the optimal control stays optimal along its trajectory. So, we can optimize backwards in time, because the tail part of the optimal control is optimal as well. More specifically for  $i = 1, \dots, N$

$$V^h(i-1, x) = \min_{v \geq 0} \left\{ V^h(i, A_h x + B_h v) + \int_{t_{i-1}}^{t_i} z_{i,v}^T(s) z_{i,v}(s) ds \right\} \quad (39)$$

where  $z_{i,v}$  is the output of (1) and (2) with initial conditions  $(t_{i-1}, x)$  and control identically equal to  $v \in \mathbb{R}^m$  on the interval  $[t_{i-1}, t_i]$ .

It is clear that  $V^h(N, x) = 0$  for all  $x \in \mathbb{R}^n$ . So, we can now recursively determine the value function  $V^h$  and store the optimal control values for every point  $(i, x)$ . The integral in (39) can be expressed explicitly in terms of  $x$  and the chosen value of  $v$  (Åström and Wittenmark 1984), thereby facilitating the calculations.

The above approximation avoids the problem of solving the Hamilton–Jacobi–Bellman partial differential equation. However, we obviously cannot store the value function into a computer without discretization (also called ‘gridding’) of the state space. The explicit expressions for the derivatives of the value function can be helpful in choosing how dense the gridding of the state space should be in order to get accurate estimations for the states that are not stored. Using (13) and (14), it is possible to store one dimension less, because knowledge of the optimal control values on the unit circle of the state space is sufficient to characterize the optimal control for all states. Note that finishing the complete recursive scheme gives the optimal controls for all stored initial states. Since the optimal control values depend continuously on the initial conditions (as is proven by Heemels 1995), interpolation between stored values gives good approximations of the optimal control values for the non-stored states. For similar techniques and more details, we refer to monographs like Kushner and Dupuis (1992).

A problem is of course how small to choose the time-step  $h$ . The time constants of the system  $(A, B)$  provide us with some information about the size to make efficient control possible. Explicit formulas specifying upper bounds on the time-step corresponding to a certain performance are not available. Rules of thumb based on engineer’s insight indicate suitable choices of  $h$ , where the explicit expressions of the time derivative of the value functions are quite valuable.

Finally, we would like to remark that physical implementation of continuously updating of control values as is required by a feedback law like (34) is often impossible in practice due to complex computations. So it could be possible that a

lower bound is put on the size of the time-step in order to provide enough time for the controller to update its control values. In this case the optimal control problem optimizing over  $P[t, T]$  is not relevant. From the beginning we should focus on optimizing over  $P_h[t, T]$  with  $h$  the lower bound on the time-step.

A drawback of the above method is the storage of the optimal control values in some kind of look-up table. The time-varying behaviour of the feedback law is the main cause of the large amount of stored data. To get rid of the time-variance, the infinite horizon case will be investigated in the next section.

## 7. Infinite horizon

### 7.1. Problem formulation

In this section, the object of study is the infinite horizon case ( $T = \infty$ ). The time-invariance of the problem guarantees that we can take  $t = 0$  without loss of generality. On an infinite horizon it seems natural to require some boundedness properties of the state. In this paper the boundedness is expressed by a finite  $L_2$ -norm of the state trajectory. The infinite horizon problem is then defined as follows.

**Problem 2:** *The objective is to find for all initial states  $x_0$  a control function  $u \in P[0, \infty)$  minimizing*

$$J_\infty(x_0, u) := \int_0^\infty z_{x_0, u}^T(s) z_{x_0, u}(s) ds \tag{40}$$

*subject to the relations (1), (2),  $x(0) = x_0$  and the additional constraint that the corresponding state trajectory  $x$  is contained in  $L_2[0, \infty)^n$ .*

Notice that  $x \in L_2[0, \infty)^n$  and  $u \in L_2[0, \infty)^m$  imply  $\dot{x} \in L_2[0, \infty)$ . If both  $x$  and  $\dot{x}$  are  $L_2$ -integrable, then  $x(s) \rightarrow 0$  ( $s \rightarrow \infty$ ).

Naturally, this leads to the following concept concerned with the existence of admissible controls. We introduce the function space  $L_1^{loc}[0, \infty)^m$  as the collection of all locally integrable functions taking values in  $\mathbb{R}^m$  with domain  $[0, \infty)$ . More specifically,  $u \in L_1^{loc}[0, \infty)^m$ , if for all  $a, b \in [0, \infty)$ ,  $a < b$  the integral  $\int_a^b u(s) ds$  exists in Lebesgue sense. For locally integrable inputs, the system given by (1) and (2) results in well-defined state and output trajectories.

**Definition 4—Stabilizability and positive stabilizability:** A control  $u \in L_1^{loc}[0, \infty)$  is said to be stabilizing for initial state  $x_0$ , if the corresponding state trajectory  $x = x_{0, x_0, u}$  satisfies  $x \in L_2[0, \infty)^n$ .  $(A, B)$  is said to be stabilizable, if for every  $x_0$  there exists a stabilizing control  $u \in L_2[0, \infty)^m$ .  $(A, B)$  is said to be positive stabilizable, if for every  $x_0$  there exists a stabilizing control  $u \in P[0, \infty)$ .

Notice that  $u \in L_2[0, \infty)^m$  and  $x \in L_2[0, \infty)^n$  imply that  $z \in L_2[0, \infty)^p$  and hence, the finiteness of  $J_\infty$ . The positive stabilizability of  $(A, B)$  guarantees the existence of positive controls, that keep the cost function finite.

In Heemels (1998) the following necessary and sufficient conditions are proven for positive stabilizability of a system  $(A, B)$ .

**Theorem 7:** *The system  $(A, B)$  is positively stabilizable if and only if  $(A, B)$  is stabilizable and all real eigenvectors  $v$  of  $A^T$  corresponding to a non-negative eigenvalue of  $A^T$  have the property that  $B^T v$  has both positive and negative components.*

**Corollary 1:** Consider the system  $(A, B)$  with scalar input ( $m = 1$ ). Then  $(A, B)$  is positively stabilizable if and only if  $(A, B)$  is stabilizable and  $A$  has no real non-negative eigenvalues.

### 7.2. Existence and uniqueness

In the remainder of this paper it will be assumed that:

- (1)  $(A, B)$  is positively stabilizable;
- (2)  $D$  has full column rank;
- (3)  $(A, B, C, D)$  is minimum phase.

To introduce the concept of minimum phase, we first define what we mean by the transmission zeros of a system  $(A, B, C, D)$ . A transmission zero is a complex number  $\lambda$  such that the system matrix

$$\begin{pmatrix} \lambda I - A & -B \\ C & D \end{pmatrix} \quad (41)$$

loses rank. If all zeros have real parts smaller than zero, the system  $(A, B, C, D)$  is called minimum phase. A system  $(C, A)$  is detectable, if the matrix  $(A^T - \lambda I; C^T)$  does not lose rank for any complex number  $\lambda$  with real part larger than or equal to zero. A well-known result in systems theory is that an equivalent characterization of detectability of the pair  $(C, A)$  is the existence of a matrix  $G$  such that  $A + GC$  is stable (Hautus 1969).

Since  $D$  has full column rank,  $(A, B, C, D)$  is minimum phase iff  $(C + DF_2, A + BF_2)$  is detectable, where  $F_2$  is uniquely determined by  $(C + DF_2)^T D = 0$ . This fact is observed, if you postmultiply the matrix in (41) by the invertible matrix

$$\begin{pmatrix} I & 0 \\ F_2 & I \end{pmatrix}$$

resulting in

$$\begin{pmatrix} \lambda I - A - BF_2 & -B \\ C + DF_2 & D \end{pmatrix} \quad (42)$$

The left and right blocks in the resulting matrix are orthogonal due to  $(C + DF_2)^T D = 0$ . The equivalence to the detectability of  $(C + DF_2, A + BF_2)$  is obtained by noting that  $D$  has full column rank.

The assumption of minimum phase is needed to get convergence between the finite and infinite horizon problem. Under the assumptions stated above the optimal controls with infinite horizon exist and are unique. These results will be proven automatically in establishing the convergence results. The optimal control on the interval  $[t, \infty)$  will be denoted by  $u_{t, \infty, x_0}$ . In some subsections abbreviations will be used.

### 7.3. Connection between finite and infinite horizons

In the following two subsections, connections between the finite and infinite horizon problem will be investigated. We will prove the convergence of the optimal costs for the finite horizon to the optimal costs for the infinite horizon when the horizon approaches infinity. Also, the optimal control, state trajectory, output for the finite horizon converges to the optimal control, state trajectory, output for the

infinite horizon, respectively in the norm of  $L_2$ . The convergence of the optimal controls in the sense of pointwise convergence can be shown by using the maximum principle for the finite horizon and extending it to the infinite horizon.

We start with an auxiliary result, which states the existence of a bounded causal operator from the  $L_2$ -output to corresponding control. An operator  $\mathcal{L}$  from  $L_2[0, \infty)^p$  to  $L_2[0, \infty)^m$  is called causal, if for all  $z_1, z_2 \in L_2[0, \infty)^p$  with  $z_1(s) = z_2(s)$  for almost all  $s \leq s_0$  we have  $(\mathcal{L} z_1)(s) = (\mathcal{L} z_2)(s)$  for almost all  $s \leq s_0$ .

**Lemma 5:** Consider the system given by (1) and (2). If for input  $u \in L_1^{\text{loc}}[0, \infty)^m$  and initial conditions  $(0, x_0)$  we have that the output  $z_{1, x_0, u} \in L_2[0, \infty)^p$ , then it must hold that  $u \in L_2[0, \infty)^m$  and moreover, the corresponding state trajectory  $x$  is contained in  $L_2[0, \infty)^n$ .

In fact, there exists a causal, bounded operator  $\mathcal{G}_{x_0}$  from  $L_2[0, \infty)^p$  to  $L_2[0, \infty)^m$ , which maps an output of system  $(A, B, C, D)$  to the corresponding control input  $u$  for given initial state  $x_0$ . This operator can be written as

$$u = \mathcal{G}_{x_0} z = \tilde{\mathcal{G}} x_0 + \mathcal{G}_0 z$$

for  $z \in L_2[0, \infty)^p$ .  $\tilde{\mathcal{G}}$  is a linear bounded operator from  $\mathbb{R}^n$  to  $L_2[0, \infty)^m$  and  $\mathcal{G}_0$  is a linear, bounded operator from  $L_2[0, \infty)^p$  to  $L_2[0, \infty)^m$ .

**Proof:** Let  $F_2$  be such that  $(C + DF_2)^T D = 0$ . Apply the preliminary feedback  $u(s) = F_2 x(s) + w(s)$  to get  $z(s) = (C + DF_2)x(s) + Dw(s)$  and premultiply by  $D^T$  to get

$$D^T Dw(s) = D^T [z(s) - (C + DF_2)x(s)] = D^T z(s)$$

So, there holds

$$w(s) = (D^T D)^{-1} D^T z(s) \tag{43}$$

from which we conclude  $w \in L_2[0, \infty)^m$ .

Since  $(C + DF_2, A + BF_2)$  is detectable, there exists  $G$  such that  $A + BF_2 + G(C + DF_2)$  is stable. But then

$$\dot{x} = [A + BF_2 + G(C + DF_2)]x - Gz + (B + GD)w$$

Since  $z, w$  are  $L_2$ -functions, we conclude  $x \in L_2[0, \infty)^n$ .

Combining (43) with the above equation results in

$$\left. \begin{aligned} \dot{x} &= \underbrace{[A + BF_2 + G(C + DF_2)]}_{\tilde{A}} x - \underbrace{[G + (B + GD)(D^T D)^{-1} D^T]}_{\tilde{B}} z \\ u &= F_2 x + w = F_2 x + (D^T D)^{-1} D^T z \end{aligned} \right\} \tag{44}$$

which is a description of the transformation  $\mathcal{G}_{x_0}$ . More explicitly

$$u(s) = (\mathcal{G}_{x_0} z)(s) = \underbrace{e^{\tilde{A}s}}_{\tilde{\mathcal{G}}_{x_0}} x_0 + \underbrace{\int_0^s e^{\tilde{A}(s-\tau)} \tilde{B} z(\tau) d\tau}_{(\mathcal{G}_0 z)(s)} \tag{45}$$

□

We prove now the uniqueness of optimal controls. Suppose  $u_0$  and  $u_1$  are both optimal for initial state  $x_0$  with corresponding outputs  $z_0$  and  $z_1$ , respectively. Consider the admissible controls  $u_\alpha$  given by  $u_\alpha = (1 - \alpha)u_0 + \alpha u_1$  with output  $z_\alpha$

for  $0 \leq \alpha \leq 1$ . Hence

$$\|z_\alpha\|^2 = \|z_0 + \alpha(z_1 - z_0)\|^2 = \|z_1 - z_0\|^2 \alpha^2 + 2(z_0 | z_1 - z_0)\alpha + \|z_0\|^2$$

$\|z_1 - z_0\| = 0$ , because otherwise  $\|z_\alpha\| < \|z_0\|$  for all  $\alpha \in (0, 1)$ , contradicting optimality of  $z_0$ . Hence,  $z_1 = z_0$ . The previous lemma implies now that  $u_0 = u_1$ .

**Definition 5:** A sequence  $\{x_n\}_{n \in \mathbb{N}}$  in a Hilbert space  $H$  is said to converge weakly to  $x \in H$ , if for every  $h \in H$  we have  $(x_n|h) \rightarrow (x|h)$  ( $n \rightarrow \infty$ ). Notation:  $x_n \xrightarrow{w} x$  ( $n \rightarrow \infty$ ). □

The following properties of weak limits are classical (see for example Yosida 1980).

**Lemma 6:**

- (1) *A weakly convergent sequence is bounded.*
- (2) *Each bounded sequence in a Hilbert space, has a weakly convergent subsequence.*
- (3) *In a Hilbert space  $x_n \rightarrow x$  if and only if  $x_n \xrightarrow{w} x$  and  $\|x_n\| \rightarrow \|x\|$ .*
- (4) *In a Hilbert space  $x_n \xrightarrow{w} x$  implies  $\|x\| \leq \liminf_{n \rightarrow \infty} \|x_n\|$ .*
- (5) *If a bounded sequence in a Hilbert space has the property that all weakly convergent subsequences have the same limit, then the sequence itself converges weakly to this limit.*

We introduce the operators  $\Pi_T : L_2[0, \infty)^m \rightarrow L_2[0, T]^n$  for  $u \in L_2[0, \infty)^m$  by

$$(\Pi_T u)(s) = u(s), \quad s \in [0, T]$$

and  $\Pi_T^* : L_2[0, T]^n \rightarrow L_2[0, \infty)^m$  for  $u \in L_2[0, T]^n$  by

$$(\Pi_T^* u)(s) = \begin{cases} u(s) & s \leq T \\ 0 & s > T \end{cases}$$

In what follows, we denote by  $u_T, x_T$  and  $z_T$  the optimal control, the optimal state trajectory and the optimal output on  $[0, T]$  with fixed initial condition  $(0, x_0)$  for  $T > 0$  or  $T = \infty$ . Furthermore, to distinguish between the value function for different horizons,  $V^T(t, x_0)$  denotes the minimal costs over the time interval  $[t, T]$  with initial state  $x_0$ .

**Theorem 8:** *For all  $x_0 \in \mathbb{R}^n$ , there holds that  $V^T(x_0) \rightarrow V^\infty(x_0)$  ( $T \rightarrow \infty$ ), where  $V^T(x_0) = V^T(0, x_0)$ . For  $T \rightarrow \infty$  we have  $\Pi_T^* u_T \rightarrow u_\infty$ ,  $\Pi_T^* z_T \rightarrow z_\infty$  and  $\Pi_T^* x_T \rightarrow x_\infty$  in the  $L_2$ -norm.*

The proof of the next auxiliary lemma can be found in the appendix.

**Lemma 7:** *Let  $\{T_j\}_{j \in \mathbb{N}}$  be a sequence of positive numbers with  $T_j \rightarrow \infty$  ( $j \rightarrow \infty$ ). If the sequence  $\{\Pi_{T_j}^* z_{T_j}\}_{j \in \mathbb{N}}$  is weakly convergent with limit  $\bar{z}$ , then there exists a control  $\bar{u} \in P[0, \infty)$  such that  $z_{x_0, \bar{u}} = \bar{z}$  and*

$$\Pi_{T_j}^* u_{T_j} \xrightarrow{w} \bar{u} \quad (j \rightarrow \infty)$$

**Proof of Theorem 8:** We know that  $V^T(x_0)$  is increasing in  $T$  and bounded by  $V^\infty(x_0)$  for fixed initial state  $x_0$ . Hence, it has a limit, say  $\alpha$ . Obviously,  $\{\Pi_T^* z_T\}_{T \in \mathbb{R}_+}$  is bounded in the  $L_2$ -sense. Hence, there exists a weakly convergent sequence  $\{\Pi_{T_j}^* z_{T_j}\}_{j \in \mathbb{N}}$  with limit  $\bar{z}$ . According to the previous lemma, there exists an admissible

control  $\bar{u}$  with  $z_{x_0, \bar{u}} = \bar{z}$ . Finally

$$V^\infty(x_0) \leq \|\bar{z}\|^2 \leq \liminf_{j \rightarrow \infty} \|\Pi_{T_j}^* z_{T_j}\|^2 = \liminf_{j \rightarrow \infty} V^{T_j}(x_0) = \alpha \leq V^\infty(x_0)$$

Note that  $\bar{u}$  is optimal because  $\|z_{x_0, \bar{u}}\|^2 = V^\infty(x_0)$ . Hence, the existence of an optimal control is established. The second part of the theorem is proven as follows. Let  $\{T_j\}_{j \in \mathbb{N}}$  be any sequence of positive numbers with  $T_j \rightarrow \infty$  ( $j \rightarrow \infty$ ). Since  $V^{T_j}(x_0)$  is bounded by  $V^\infty(x_0)$ , we have  $\{\Pi_{T_j}^* z_{T_j}\}_j$  is a bounded sequence. Take any two weakly convergent subsequences  $\{\Pi_{n_j}^* z_{n_j}\}_j$  and  $\{\Pi_{m_j}^* z_{m_j}\}_j$  of  $\{\Pi_{T_j}^* z_{T_j}\}_j$  with limits  $z^n$  and  $z^m$ , respectively. We will prove  $z^n = z^m$ .

On account of Lemma 7, we conclude the existence of admissible controls  $u^n$  and  $u^m$  with  $z_{x_0, u^n} = z^n$  and  $z_{x_0, u^m} = z^m$ .

Then  $\|z^n\|_{2, [0, \infty)}^2 \leq \liminf_{j \rightarrow \infty} \|\Pi_{n_j}^* z_{n_j}\|_{2, [0, \infty)}^2 = \liminf_{j \rightarrow \infty} \|z_{n_j}\|_{2, [0, n_j)}^2 = V^\infty(x_0)$  and similar for  $z^m$ . We find  $u^n$  and  $u^m$  are both optimal inputs. By uniqueness of optimal controls, we see that  $u^n = u^m$  and hence  $z^n = z^m = z_\infty$ . Using Lemma 6 we conclude that  $\{\Pi_{T_j}^* z_{T_j}\}_j$  is a weakly convergent sequence with limit  $z_\infty$ . Since  $\|\Pi_{T_j}^* z_{T_j}\|^2 = V^{T_j}(x_0) \rightarrow V^\infty(x_0) = \|z_\infty\|^2$ , we get now from Lemma 25 that  $\{\Pi_{T_j}^* z_{T_j}\}_j$  actually converges to  $z_\infty$  in the sense of  $L_2$ .

Since  $\Pi_{T_j}^* z_{T_j} \xrightarrow{L_2} z_\infty$ , it follows from the continuity of  $\mathcal{G}_0$  that  $\mathcal{G}_0 \Pi_{T_j}^* z_{T_j} \xrightarrow{L_2} \mathcal{G}_0 z_\infty$ . Hence,  $\mathcal{G}_{x_0} \Pi_{T_j}^* z_{T_j} \xrightarrow{L_2} \mathcal{G}_{x_0} z_\infty = u_\infty$ , which implies by causality of  $\mathcal{G}_{x_0}$

$$\|u_{T_j} - \Pi_{T_j} u_\infty\|_{2, [0, T_j]} = \|\Pi_{T_j} \mathcal{G}_{x_0} \Pi_{T_j}^* z_{T_j} - \Pi_{T_j} u_\infty\|_{2, [0, T_j]} \rightarrow 0 \quad (j \rightarrow \infty)$$

Since  $u_\infty \in L_2[0, \infty)^m$ ,  $\|u_\infty\|_{2, [T, \infty)} \rightarrow 0$  ( $j \rightarrow \infty$ ). Since

$$\|\Pi_{T_j}^* u_{T_j} - u_\infty\|_{2, [0, \infty)} \leq \|u_{T_j} - \Pi_{T_j} u_\infty\|_{2, [0, T_j]} + \|u_\infty\|_{2, (T_j, \infty)}$$

these last two remarks combined lead to  $\Pi_{T_j}^* u_{T_j} \xrightarrow{L_2} u_\infty$  if  $j \rightarrow \infty$ .

The result for the state trajectories is proved analogously by replacing  $\mathcal{G}_{x_0}$  by the operator from the output  $z$  to the trajectory  $x$ , which can be derived from (44).  $\square$

Notice that in the above proof, the existence of optimal controls has been shown.

#### 7.4. Maximum principle with the infinite horizon

In this subsection the maximum principle for the finite horizon together with the convergence results are exploited to arrive at a maximum principle for the infinite horizon. This maximum principle states the smoothness properties of the infinite optimal control function and will result in an additional convergence relation between finite and infinite horizon optimal control functions: the optimal controls also converge pointwise. The precise mathematical formulation will be stated below. This extended convergence lemma justifies the numerical approximation of the stationary infinite horizon optimal feedback in subsection 7.5.

For a fixed initial state  $x_0$  the corresponding optimal controls  $u_T$  for finite horizon  $T$  satisfy

$$u_T(s) = P_\infty (-\frac{1}{2}(D^T D)^{-1} \{B^T \phi_T(s) + 2D^T C x_T(s)\}) \quad (46)$$

for all  $s \in [0, T]$  where the functions  $x_T$  and  $\phi_T$  are given by

$$\left. \begin{aligned} \dot{x}_T &= Ax_T + Bu_T, & x_T(0) &= x_0 \\ \dot{\phi}_T &= -A^T \phi_T - 2C^T z_T, & \phi_T(T) &= 0 \end{aligned} \right\} \quad (47)$$

**Lemma 8:** For all  $T > 0$  it holds that  $\|\phi_T(0)\| \leq M_0\|x_0\|$  for certain constant  $M_0$ .

**Proof:** Define  $\bar{x}_T$  by  $\bar{x}_T = A\bar{x}_T + B\bar{u}_T$ ,  $\bar{x}_T(0) = \phi_T(0)$ , where  $\bar{u}_T$  is the optimal control achieving  $V^T(\phi_T(0)) = V^T(0, \phi_T(0))$ . That is  $\bar{u}_T = u_{0,T,\phi_T(0)}$ . We get

$$\begin{aligned} \frac{d}{dt}(\phi_T | \bar{x}) &= (B^T \phi_T | \bar{u}_T) - 2(C\bar{x}_T | Cx_T + Du_T) \\ &= (B^T \phi_T + 2D^T Cx_T + 2D^T Du_T | \bar{u}_T) \\ &\quad + -2(C\bar{x}_T | Cx_T + Du_T) - (2D^T Cx_T + 2D^T Du_T | \bar{u}_T) \\ &\geq -2(C\bar{x}_T + D\bar{u}_T | Cx_T + Du_T) \end{aligned}$$

We used, according to (9) and (46)

$$(B^T \phi_T + 2D^T Cx_T + 2D^T Du_T | \bar{u}_T) = -2(-\frac{1}{2}(D^T D)^{-1} [B^T \phi_T + 2D^T Cx_T] - u_T | \bar{u}_T)_{D^T D} \geq 0$$

Since  $\|\phi_T(0)\|^2 = (\phi_T(0) | \bar{x}(0)) = \int_0^T -d/dt(\phi_T(s) | \bar{x}(s)) ds$ , we have

$$\begin{aligned} \|\phi_T(0)\|^2 &\leq 2(C\bar{x}_T + D\bar{u}_T | Cx_T + Du_T)_{2, [\bar{0}, T]} \\ &\leq 2\|Cx_T + Du_T\|_{2, [\bar{0}, T]} \|C\bar{x}_T + D\bar{u}_T\|_{2, [\bar{0}, T]} \\ &= 2(V^T(\phi_T(0))V^T(x_0))^{1/2} \\ &\leq 2(V^\infty(\phi_T(0))V^\infty(x_0))^{1/2} \end{aligned}$$

$(\cdot | \cdot)_{2, [\bar{0}, T]}$  denotes the standard inner product in  $L_2[\bar{0}, T]$ . Clearly, there exists a constant  $M$  such that  $V^\infty(x) \leq M\|x\|^2 \forall x \in \mathbb{R}^n$ . But then

$$\|\phi_T(0)\|^2 \leq 2M\|\phi_T(0)\|\|x_0\|$$

□

**Theorem 9:** The optimal control  $u_\infty$  corresponding to initial condition  $(t, x_0)$  satisfies

$$u_\infty = P_\Omega(-\frac{1}{2}(D^T D)^{-1} \{B^T \phi_\infty(s) + 2D^T Cx_\infty(s)\}) \quad (48)$$

where the continuous function  $\phi_\infty \in L_2[\bar{0}, \infty)^n$  is given by

$$\dot{\phi}_\infty = -A^T \phi_\infty - 2C^T z_\infty \quad (49)$$

for some initial condition  $(0, \phi_\infty(0))$ . Moreover,  $u_\infty$  is a continuous function.

**Proof:** Take an arbitrary  $\tau$ . Lemma 8 ensures the existence of a convergent sequence  $\{\phi_{T_i}(0)\}_i$ , where  $T_i \rightarrow \infty$  ( $i \rightarrow \infty$ ). Since  $\{\Pi_\tau z_{T_i}\}$  converges to  $\Pi_\tau z_\infty$  in  $L_2[\bar{0}, \tau]$  when  $(T \rightarrow \infty)$ , we see by comparing (47) and (49), that  $\Pi_\tau \phi_{T_i} \rightarrow \Pi_\tau \phi_\infty$  in  $L_\infty[\bar{0}, \tau]$  when  $i \rightarrow \infty$ . Analogously, the  $L_\infty[\bar{0}, \tau]$ -convergence of  $\{\Pi_\tau x_{T_i}\}_i$  to  $\Pi_\tau x_\infty$  can be deduced by using the  $L_2$ -convergence of  $\{\Pi_\tau u_{T_i}\}$ . Equation (46) implies now that  $\{\Pi_\tau u_{T_i}\}_i$  converges uniformly on  $[\bar{0}, \tau]$  to

$$P_\Omega(-\frac{1}{2}(D^T D)^{-1} \{B^T \phi_\infty(s) + 2D^T Cx_\infty(s)\})$$

because the projection  $P_\Omega$  is continuous.

The  $L_2$ -convergence of  $\{\Pi_\tau u_{T_i}\}_i$  to  $\Pi_\tau u_\infty$  implies the existence of a subsequence, which converges a.e. on  $[\bar{0}, \tau]$  to  $\Pi_\tau u_\infty$ . This implies that  $\Pi_\tau u_\infty$  equals the

expression above almost everywhere, because of the uniqueness of limits. Letting  $\tau$  go to infinity, completes the proof.  $\square$

We use the above maximum principle for strengthening the convergence properties between finite and infinite horizon optimal controls.

**Theorem 10:** For all  $\tau > 0$   $\Pi_\tau u_T \rightarrow \Pi_\tau u_\infty$  ( $T \rightarrow \infty$ ) in  $L_\infty[0, \tau]^n$ .

In the proof of this theorem, the following lemma is used.

**Lemma 9:** We consider a fixed initial condition  $(0, x_0)$ . For every  $\tau > 0$ , there exists a constant  $C_\tau$ —depending on  $\tau$ —such that the optimal controls for finite and infinite horizons satisfy, for all  $s, t \in [0, \tau]$

$$\|u_T(s) - u_T(t)\| \leq C_\tau |s - t|^{1/2}$$

**Proof:** See the appendix for the proof.  $\square$

**Proof of Theorem 10:** Define the continuous difference function  $w_T \in L_2[0, \tau]^n$  for  $T \geq \tau$  by

$$w_T := \Pi_\tau (u_T - u_\infty)$$

Suppose

$$\exists \varepsilon > 0 \quad \forall T_0 \geq \tau \quad \exists T > T_0 \quad \exists s \in [0, \tau] \quad \|w_T(s)\| \geq \varepsilon \tag{50}$$

Then, there exists a subsequence  $\{w_{T_i}\}_{i \in \mathbb{N}}$  with  $T_i \rightarrow \infty$  ( $i \rightarrow \infty$ ) and a sequence of points  $\{s_i\}$  in  $[0, \tau]$  such that  $\|w_{T_i}(s_i)\| \geq \varepsilon$  for a certain  $\varepsilon$ .

Since for  $s, t \in [0, \tau]$

$$\begin{aligned} \|w_T(s) - w_T(t)\| &\leq \|u_T(s) - u_T(t)\| + \|u_\infty(s) - u_\infty(t)\| \\ &\leq 2C_\tau |s - t|^{1/2} \end{aligned}$$

we can construct for each  $i$  an interval around  $s_i$  with a measure larger than a constant  $\delta > 0$ , where  $\|w_{T_i}\| \geq \frac{1}{2}\varepsilon$ . This contradicts the  $L_2$ -convergence of  $w_{T_i}$  to the zero function, because  $\|w_{T_i}\|_{2, [0, \tau]} \geq \frac{1}{2}\varepsilon\sqrt{\delta}$ .  $\square$

### 7.5. Approximation

In the finite horizon problem, we saw that the optimal control could be given by a time-varying feedback of the form (cf. (34))

$$u(s, x) = P_\Omega(-\frac{1}{2}(D^T D)^{-1} [B^T V_x(s, x) + 2D^T Cx]) \tag{51}$$

As mentioned in section 6, the time-varying behaviour of the feedback is, for implementation purposes, not convenient due to the large storage requirements of the control system. In particular, when the time-step  $h$  is chosen small, giant storage capacities are needed to store the optimal control values for all discrete time instants. In contrast with the finite horizon, the time-dependence vanishes in the infinite horizon case. It is obvious that for all  $s \geq 0$   $u_{t, \infty, x_0}(s+t) = u_{0, \infty, x_0}(s)$ . From this, it follows that there exists a time-invariant optimal feedback  $u_{f, db}$  defined by

$$u_{f, db}(x_0) := u_{t, \infty, x_0}(t) = u_{0, \infty, x_0}(0) = \lim_{T \rightarrow \infty} u_{0, T, x_0}(0) \tag{52}$$

The limit in (52) is the result of the technical and analytical work we performed in the previous (sub)section. It provides us with a method to approximate the optimal feedback. We proposed a method to compute the finite horizon optimal controller in



section 6. If  $T$  is large enough,  $u_{0,T,x_0}$  can be used as an approximation for  $u_{f\text{ db}}(x_0)$ . An unsolved problem at the moment is how to choose  $T$  such that it is large enough. Of course, the time constants of the system under study give an indication. It is also clear that the choice of weightings  $C$  and  $D$  influence the choice of  $T$ . If  $D^T D$  increases, the control intensity reduces, resulting in a slower closed-loop system. Consequently, a larger value of  $T$  is required. Since the numerical algorithm recedes recursively in time, one could also compare two successive time steps and base a stopping criterion on the difference between the computed feedbacks on  $t_i$  and  $t_{i-1}$ . If this difference is small compared to a specified tolerance the algorithm is stopped (no further increase of  $T$ ), otherwise one continues. Another method is using the explicit expression for the time derivative of the value function:  $T$  has to be chosen so large that  $V_t(0, \cdot)$  becomes close to zero, indicating that the control does not change too much any more. An alternative could be simulation of the closed-loop system to verify whether the chosen  $T$  is large enough. In the example below,  $T = 15$  would be sufficiently large to approximate the feedback, because the system is stabilized within this time span.

In our investigation of connections between finite and infinite horizon problems, the approximation of the stationary feedback as proposed is justified by the convergence results. In Kushner and Dupuis (1992) many more approximation techniques are considered for value functions.

Numerical robustness of the method is guaranteed by choosing sufficiently large  $T$  and sufficiently small  $h$ . The explicit expressions of the derivatives of the value function as given in subsection 5.3 can be helpful in choosing good values for  $T$  and  $h$  as indicated before. These expressions can also be used to determine the error one makes by gridding of the state space.

A point to be mentioned is that the indicated approximation suffers from the 'curse of dimensionality.' Increasing state dimensions of the system result in exponentially more computation time and storage capacities. However, the ongoing evolution of very fast processors in computers makes the method more and more feasible for very large systems.

We illustrate our approximation by the example of the pendulum. The dotted line in figure 2 is the value of the cost function if we apply the control as in figure 3. The optimal control is computed by the algorithm sketched above with  $T = 25$  and time-step  $h = 0.125$ . The resulting state trajectory is also depicted.

We like to stress the advantage of a feedback optimal formulation over open-loop control once more. The open-loop control is determined *a priori* on the initial condition only. Hence, the open-loop control is computed for the complete considered time-interval and adaptation to disturbances active on the system is not possible any more. Feedback, however, determines the control-value on the actual measured state. In this way, it can adapt to all kinds of disturbances. To illustrate the advantage of feedback over open-loop controllers, we implement both control strategies on the pendulum (initial state  $(1, 0)^T$ ) in case a constant wind is blowing. We model this by adding 0.5 to the input  $u$ . In figure 4, we see how the feedback control adapts to this disturbance. We observe that the oscillations of the state die out more quickly in the feedback case, resulting in a better performance.

## 8. Conclusions

In this paper, we addressed three approaches to the Linear Quadratic Regulator

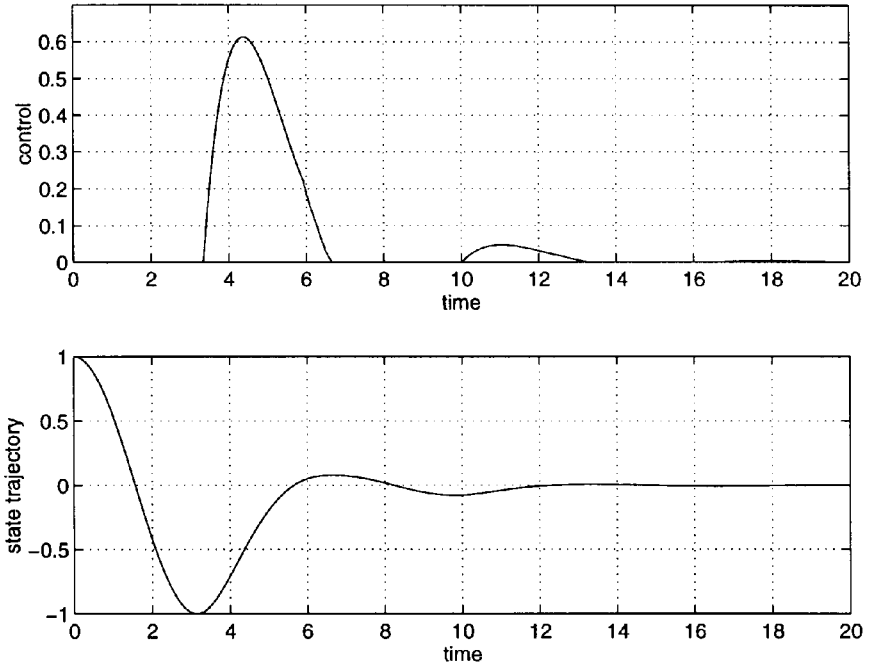


Figure 3. Optimal control and optimal state-trajectory.

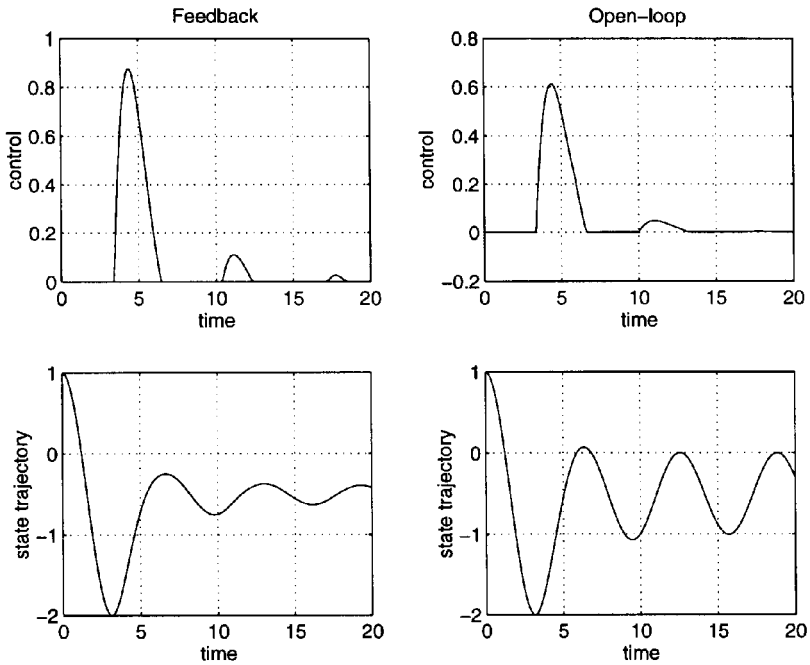


Figure 4. Comparing feedback and open-loop formulations.

problem with a positivity constraint on the admissible control set. The optimal controls were characterized by necessary and sufficient conditions in terms of inner products, the maximum principle and dynamic programming. The connection between dynamic programming and the maximum principle was exploited to show that both the maximum principle and dynamic programming are necessary and sufficient for optimality. The maximum principle indicated the smoothness properties of the optimal control, especially the continuity of the optimal control. The maximum principle is stated by a two-point boundary value problem, dynamic programming by a partial differential equation. If one of these equations is solved, then the solution leads to the optimal controller. Although the value function turned out to be continuously differentiable and a classical solution to the Hamilton–Jacobi–Bellman equation, the partial differential equation is not easily solved. The same holds for the two-point boundary problem with the additional drawback that it gives an open-loop control function for only one particular initial condition. For that reason another approach was taken resulting in a recursive scheme for approximating the optimal control by piecewise constant control functions. Convergence results between the optimal discrete controls, and the exact optimal control justified the effectiveness of the method. However, it was argued that the storage capacities of the controller could become quite large, if a small time step is used. To overcome this problem the infinite horizon problem was studied to arrive at a stationary feedback that requires less data storage.

The investigation of the connections between the finite and infinite horizon problem resulted in convergence results between the optimal controls. To strengthen the  $L_2$ -convergence to pointwise convergence, the maximum principle for the finite horizon was extended to an infinite horizon under the condition that the system is minimum phase. This also guaranteed the continuity of the optimal control on the infinite horizon. These analytical results justified a method to approximate the optimal positive feedback in the considered problem. Currently, the main problem is that explicit bounds for the time-step  $h$  and the horizon  $T$  guaranteeing good approximation of the stationary infinite horizon optimal feedback are not available. However, some rules of thumb were specified to get a rough estimate on how to choose these quantities. We briefly touched upon how to use the explicit expressions for the derivatives of the value function to facilitate these choices. In particular, the specification of these bounds is a topic of current development.

Briefly, we discussed the advantage of feedback: it performs better than the open-loop controllers, because it can adapt to disturbances like measurement noise and unmodelled dynamics.

To conclude, we would like to stress that the questions raised in section 2 are natural in any optimal control problem and that we answered them all in the settings of the optimal control problem at hand. In particular, the proposed approximation method based on discretization is often used in practice, but justification of such a method by means of convergence results is hardly encountered in the literature.

### Acknowledgment

The authors would like to thank the anonymous reviewers for their useful comments and constructive criticism that improved the overall quality of this paper.

Appendix

**Lemma 10:** For the function  $R$  as defined in (28), it holds that there exists a constant  $E$  such that for all  $h \in \mathbb{R}^n$

$$|R(t, x_0, h)| \leq E \|h\|^2 \tag{53}$$

**Proof:** By using the following inequalities in combination with the Cauchy–Schwarz inequality, the lemma is proved.

$$V(t, h) \leq J(t, h, 0) = \|\mathcal{M}_{t,T} h\|_2^2 \leq \|\mathcal{M}_{t,T}\|^2 \|h\|^2$$

Lemma 1 and (12) imply that for all  $t \in [0, T]$  and  $x_0, x_1 \in \mathbb{R}^n$

$$\|u_{t,T,x_0} - u_{t,T,x_1}\|_2 \leq \|\mathcal{M}_{t,T}\| \|\tilde{\mathcal{L}}_{t,T}\| \|x_0 - x_1\| \quad \square$$

**Proof of Lemma 7:** There holds,

$$\mathcal{G}_{x_0}(\Pi_{T_j}^* z_{T_j}) \xrightarrow{w} \mathcal{G}_{x_0}(\bar{z}) \quad (j \rightarrow \infty)$$

We define  $\hat{u}_{T_j} := \mathcal{G}_{x_0}(\Pi_{T_j}^* z_{T_j})$  and  $\bar{u} := \mathcal{G}_{x_0}(\bar{z})$ . Thus,  $\{\hat{u}_{T_j}\}_{j \in \mathbb{N}}$  is weakly convergent with limit  $\bar{u}$ . Since the operator  $\mathcal{G}_{x_0}$  is causal,  $\Pi_{T_j} \hat{u}_{T_j} = u_{T_j}$ . We see that for every  $h \in L_2[0, \infty)^m$

$$(h|\hat{u}_{T_j})_{[0, \infty)} = (h|\Pi_{T_j}^* u_{T_j})_{[0, \infty)} + (h|\hat{u}_{T_j})_{[T_j, \infty)} \tag{54}$$

The Cauchy–Schwarz inequality implies

$$|(h|\hat{u}_{T_j})_{[T_j, \infty)}| \leq \|h\|_{[T_j, \infty)} \|\hat{u}_{T_j}\|_{[0, \infty)}$$

Since weak convergence of the sequence  $\{\hat{u}_{T_j}\}_j$  implies  $L_2$ -boundedness, we see that the left-hand side of the above equation tends to zero, if  $j \rightarrow \infty$ . Hence from (54) we get,  $\{\Pi_{T_j}^* u_{T_j}\}_j$  is a weakly convergent sequence with limit  $\bar{u}$ . Notice that  $\bar{u} \in P[0, \infty)$ , because  $P[0, \infty)$  is weakly closed in  $L_2[0, \infty)^m$ .

Left to prove is  $z_{x_0, \bar{u}} = \bar{z}$ . Consider an arbitrary scalar  $\tau > 0$ . From  $\Pi_{T_j}^* u_{T_j} \xrightarrow{w} \bar{u}$  ( $j \rightarrow \infty$ ) and the boundedness and linearity of  $\Pi_\tau$  we get for  $T_j > \tau$ , that  $\Pi_\tau u_{T_j} \xrightarrow{w} \Pi_\tau \bar{u}$ . Hence,  $\mathcal{M}_{0,\tau} x_0 + \mathcal{L}_{0,\tau} \Pi_\tau u_{T_j} \xrightarrow{w} \mathcal{M}_{0,\tau} x_0 + \mathcal{L}_{0,\tau} \Pi_\tau \bar{u}$ . On the other hand,  $\mathcal{M}_{0,\tau} x_0 + \mathcal{L}_{0,\tau} \Pi_\tau u_{T_j} = \Pi_\tau z_{T_j} \xrightarrow{w} \Pi_\tau \bar{z}$ . The uniqueness of weak limits gives  $\mathcal{M}_{0,\tau} x_0 + \mathcal{L}_{0,\tau} \Pi_\tau \bar{u} = \Pi_\tau \bar{z}$ . Since  $\tau$  was arbitrary we conclude  $z_{x_0, \bar{u}} = \bar{z}$ .  $\square$

**Proof of Lemma 9:**  $\|P_\Omega(a) - P_\Omega(b)\|_{D^T D} \leq \|a - b\|_{D^T D}$  can be translated into  $\|P_\Omega(a) - P_\Omega(b)\| \leq \Lambda \|a - b\|$ , where  $\|\cdot\|$  is the usual Euclidean norm and  $\Lambda$  is a certain positive constant. Since the optimal controls  $u_T$  satisfy (46) and (48), we have for  $\tau \geq s > t \geq 0$

$$\begin{aligned} & \|u_T(s) - u_T(t)\| \\ & \leq \frac{1}{2} \Lambda \|(D^T D)^{-1} B\| \|\phi_T(s) - \phi_T(t)\| + \Lambda \|(D^T D)^{-1} D^T C\| \|x_T(s) - x_T(t)\| \end{aligned}$$

First, we note that the collection  $\{\Pi_\tau z_T \mid T > \tau\}$  is bounded in  $L_2[0, \tau]$ , because  $\{V^T(x_0)\}_T$  is bounded. Since we can write  $u_T = \mathcal{G}_{x_0} + \mathcal{G}_0 z_T$ , we see that the collection  $\{\Pi_\tau u_T \mid T > \tau\}$  is also bounded in  $L_2[0, \tau]^n$ . In turn, this implies

that  $\{\Pi_\tau x_T \mid T > \tau\}$  is bounded in  $L_\infty[0, \tau]$ . Let  $s > t$

$$\begin{aligned} \|x_T(s) - x_T(t)\| &\leq \left\| \left[ e^{A(s-t)} - I \right] x_T(t) \right\| + \left\| \int_t^s e^{A(s-\tau)} B u_T(\tau) d\tau \right\| \\ &\leq \left[ e^{\|A\|(s-t)} - 1 \right] \|x_T(t)\| + e^{\|A\|\tau} \|B\| |s - t|^{1/2} \|u_T\|_{2, [0, \tau]} \\ &\leq \frac{e^{\|A\|\tau} - 1}{\tau} |s - t| \|x_T\|_{\infty, [0, \tau]} + e^{\|A\|\tau} \|B\| |s - t|^{1/2} \|u_T\|_{2, [0, \tau]} \\ &\leq \frac{e^{\|A\|\tau} - 1}{\sqrt{\tau}} |s - t|^{1/2} \|x_T\|_{\infty, [0, \tau]} + e^{\|A\|\tau} \|B\| |s - t|^{1/2} \|u_T\|_{2, [0, \tau]} \end{aligned}$$

The second step is a consequence of the Cauchy-Schwarz inequality. The last upper bound can be dominated by a constant which depends on  $\tau$  only, because  $\|u_T\|_{2, [0, \tau]}$  and  $\|x_T\|_{\infty, [0, \tau]}$  are uniformly bounded for all  $T \geq \tau$  and  $T = \infty$ .

The term  $\|\phi_T(s) - \phi_T(t)\|$  can be dealt with in an almost identical way.  $\square$

## References

- ANDERSON, B. D. O., and MOORE, J. B., 1990, *Optimal Control—Linear Quadratic Methods* (Prentice-Hall).
- ÅSTRÖM, K. J., and WITTENMARK, B., 1984, *Computer-controlled Systems—Theory and Design* (Prentice-Hall).
- BELLMAN, R., 1967, *Introduction to the Mathematical Theory of Control Processes* (Academic Press).
- FEICHTINGER, G., and HARTL, R. F., 1986, *Ökonomischer Prozesse: Anwendung des Maximumprinzips in den Wirtschaftswissenschaften* (Berlin: De Gruyter).
- FLEMING, W. H., and RISHEL, R. W., 1975, *Deterministic and Stochastic Optimal Control* (Springer-Verlag).
- FLEMING, W. H., and SONER, H. M., 1993, *Controlled Markov Processes and Viscosity Solutions* (Springer-Verlag).
- HAUTUS, M. L. J., 1969, Stabilization, controllability and observability of linear autonomous systems. *Nederl. Akad. Wetensch., Proc., Ser. A73*, pp. 448–455.
- HEEMELS, W. P. M. H., 1995, *Optimal Positive Control & Optimal Control with Positive State Entries*. Master's thesis, Eindhoven University of Technology.
- HEEMELS, W. P. M. H., and STOOBVOGEL, A. A., 1998, Positive stabilizability of a linear continuous-time system. Technical Report 981/01. Eindhoven University of Technology.
- HIRIART-URRUTY, J.-B., and LEMARÉCHAL, C., 1993, *Convex Analysis and Minimization Algorithms I* (Springer-Verlag).
- KAILATH, T., 1980, *Linear Systems* (Prentice-Hall).
- KIRK, D. E., 1970, *Optimal Control Theory, An Introduction* (Prentice-Hall).
- KUSHNER, H. J., and DUPUIS, P. G., 1992, *Numerical Methods for Stochastic Control Problems in Continuous Time* (Springer-Verlag).
- LEE, E. B., and MARKUS, L., 1967, *Foundations of Optimal Control Theory* (John Wiley & Sons).
- LUENBERGER, D. G., 1969, *Optimization by Vector Space Methods* (John Wiley & Sons).
- MACKI, J., and STRAUSS, A., 1982, *Introduction to Optimal Control Theory* (Springer-Verlag).
- PACHTER, M., 1980, The linear-quadratic optimal control problem with positive controllers. *International Journal of Control*, **32**, 589–608.
- PONTRYAGIN, L. S., BOLTYANSKII, V. G., GAMKRELIDZE, R. V., and MISHCHENKO, E. F., 1962, *The Mathematical Theory of Optimal Processes* (Interscience Publishers, John Wiley & Sons).
- THOMAS, G. B., and FINNEY, R. L., 1996, *Calculus and Analytic Geometry* (Addison-Wesley).
- YOSIDA, K., 1980, *Functional Analysis* (Springer-Verlag).