A note on "Families of linear-quadratic problems: continuity properties"
by
A.H.W. Geerts

Eindhoven, Netherlands
February 1988
A NOTE ON: "Families of Linear-Quadratic Problems: Continuity properties"

ABSTRACT

In the paper "Families of Linear-Quadratic Problems: Continuity properties" one-parameter (ε) families of these problems with stability are investigated. In this technical note we will show that the set of invariant zeros of the ε-problem is contained in the set of invariant zeros of the "boundary" problem (the problem with ε = 0) if the "boundary" system is left invertible. This result is obtained only by applying the assumptions on the continuity and the monotonicy of the weighting matrices that are made in the above-mentioned paper. Consequently, we can replace the third assumption in that paper by a much weaker one.

February 1988

Research supported by the Netherlands organization for scientific research (NWO).
1 Introduction

In a recent paper ([1]) one-parameter (ε) families of linear time-invariant finite dimensional systems with corresponding linear-quadratic control problems with stability ([1, Sec. II]) are introduced. Then convergence of the optimal cost for the ε-problem (ε > 0) to the optimal cost for the "boundary" problem (i.e. the problem with ε = 0) is established only by assuming continuity and (in a certain sense) monotonicity of the weighting matrices w.r.t. ε. In addition, convergence (in distributional sense, see [1, Sec. II]) of the optimal inputs for positive ε to the optimal controls for ε = 0 is proven, under the assumption that for each initial condition the latter (and thus the former) are unique. In this case also convergence (in distributional sense) of the corresponding state trajectories is obtained and, moreover, the optimal outputs are proven to converge strongly ([1, Sec. IV, p. 325]).

Now, it is well known that for all initial conditions the optimal inputs for the "boundary" problem exist and are unique (e.g. [1, Prop. 4.1 (ii)]) if and only if the "boundary" system (i.e. the system with ε = 0) is left invertible and there are no invariant zeros (see e.g. [1, Sec. III, p. 325]) on the imaginary axis. Indeed these assumptions are expressed in assumptions A.3 and A.4 on page 325 of [1]. Also, it is remarked there that if the intersection of the set of invariant zeros for the "boundary" system, $\sigma^* (\Xi_0)$, and the imaginary axis $\mathbb{C}^0$ is empty, then in general the intersection of the set of invariant zeros of the ε-system, $\sigma^* (\Xi_\varepsilon)$, (ε > 0) and $\mathbb{C}^0$ need not be empty as well. The latter condition, $\sigma^* (\Xi_\varepsilon) \cap \mathbb{C}^0 = \emptyset$ for small $\varepsilon > 0$, is necessary to ensure that also in the ε-problem optimal inputs exist for all initial conditions.
However, in this technical note we will show that if the
"boundary" system $\Sigma_0$ is left invertible then the continuity
assumption and the monotonicy assumption on the weighting
matrices imply that $\sigma^*(\Sigma_\epsilon) \subseteq \sigma^*(\Sigma_0)$ for all $\epsilon$ in a sufficiently
small closed interval $[0, \delta]$ with $\delta > 0$. Thus we establish that
the assumption " $\exists \eta > 0 : \sigma^*(\Sigma_\epsilon) \cap \mathbb{C}^0 = \emptyset$ for all $\epsilon \in [0, \eta]$ "
of [1] can in fact be replaced by the much weaker assumption
" $\sigma^*(\Sigma_0) \cap \mathbb{C}^0 = \emptyset$ ". 
2 Basic problem formulation and our additional results

Consider the finite-dimensional linear time-invariant systems $\Sigma_\varepsilon$ determined by the system equation

$$\dot{x}(t) = Ax(t) + Bu(t), \quad x(0) = x_0,$$

(2.1a)

and with the $\varepsilon$-depending output equations

$$y_\varepsilon(t) = C(\varepsilon)x(t) + D(\varepsilon)u(t),$$

(2.1b)

and with associated cost-functionals

$$J_\varepsilon(x_0, u) = \int_0^\infty w_\varepsilon(t) \|u\|^2 dt.$$  (2.1c)

Here, $u(t), x(t)$ and $y(t)$ are all real vectors of $m, n$ and $p$ components and $C(\varepsilon), D(\varepsilon)$ are assumed to be linear mappings from $\mathbb{R}^n$ to $\mathbb{R}^m$ and $\mathbb{R}^m$ to $\mathbb{R}^p$, respectively, for $\varepsilon$ lying in an interval $[0, \delta]$ where $\delta > 0$ will be specified below. The pair $(A, B)$ is stabilizable.

For fixed $\varepsilon \in [0, \delta]$ the linear-quadratic problem with stability associated with (2.1) is the problem of finding the infimal value of $J_\varepsilon(x_0, u)$ with respect to an appropriate class of inputs (in [1, Sec. II] $u_{\text{dist}}^\text{stab}(x_0)$) and to compute, if it exists, an optimal control. The optimal cost in [1] is denoted $J_\varepsilon^*(x_0)$. It is proven in [1, Th. 3.2] that $\lim_{\varepsilon \to 0} J_\varepsilon^*(x_0) = J_0^*(x_0)$ for all $x_0$ if the following two assumptions hold:

A.1: $\varepsilon \to C(\varepsilon)$ and $\varepsilon \to D(\varepsilon)$ are continuous at 0.

A.2: For all $0 \leq \varepsilon_1, \varepsilon_2 \leq \delta$ we have $Q(\varepsilon_1) \leq Q(\varepsilon_2)$.

Here

$$Q(\varepsilon) = (C(\varepsilon), D(\varepsilon)) (C(\varepsilon), D(\varepsilon))$$

(2.2)

and thus the cost-functionals in (2.1c) can be rewritten as

$$J_\varepsilon(x_0, u) = \int_0^\infty [x'(t), u'(t)]Q(\varepsilon)[x(t), u(t)]dt.$$  (2.3)

The Assumption A.2 may just as well be replaced by
(A.2)' : For all $0 \leq \varepsilon_1 \leq \varepsilon_2 \leq \delta$ we have $Q(\varepsilon_1) \subseteq Q(\varepsilon_2)$ and there exists a subspace $M$ such that $\ker(Q(\varepsilon)) = M$ for all $\varepsilon \in (0, \delta]$. This is shown in Lemma 2.1.

Lemma 2.1.

Assume that A.2 holds. Then there is a $\delta' \leq \delta$ and there exists a subspace $M$ such that $\ker(Q(\varepsilon)) = M$ for all $\varepsilon \in (0, \delta']$.

Proof. From A.2 we observe that, if $0 \leq \varepsilon_1 \leq \varepsilon_2 \leq \delta$, then $\ker(Q(0)) \supseteq \ker(Q(\varepsilon_1)) \supseteq \ker(Q(\varepsilon_2))$. Hence $\ker(Q(\varepsilon))$ is constant for $\varepsilon$ sufficiently small and positive. Note that not necessarily $\ker(Q(0)) = M$.

Remark.

Of course, the assumption $\ker(Q(\varepsilon)) = M$ ($\varepsilon \in (0, \delta]$) is equivalent to: For all $\varepsilon \in (0, \delta]$, $\ker([C(\varepsilon), D(\varepsilon)]) = M$.

Next, let $\varphi(\Sigma_\varepsilon)$ and $\psi(\Sigma_\varepsilon)$ be the weakly unobservable and the strongly reachable subspace corresponding to $\Sigma_\varepsilon$ ($\varepsilon \in [0, \delta]$), respectively, see e.g. [2, Def. 3.8 and Def. 3.13]. Then from the algorithms [2, (3.20)] for computing $\varphi(\Sigma_\varepsilon)$ and [2, (3.22)] for computing $\psi(\Sigma_\varepsilon)$ with the Remark we establish that

Proposition 2.2.

Under the assumption (A.2)' we have: There are subspaces $\varphi$ and $\psi$ such that for all $\varepsilon \in (0, \delta]$, $\varphi(\Sigma_\varepsilon) = \varphi$ and $\psi(\Sigma_\varepsilon) = \psi$. 
It turns out that these subspaces are contained in their respective versions for the "boundary" system:

**Proposition 2.3.**

\[ \nabla \subseteq \nabla(\Sigma_0) =: \nabla, \quad \mathcal{W} \subseteq \mathcal{W}(\Sigma_0) =: \mathcal{W}, \quad \mathcal{K} := \nabla \cap \mathcal{W} \subseteq \mathcal{K} := \nabla \cap \mathcal{W}. \]

**Proof.** Let \( x_0 \in \nabla \) and \( u(t) \) be a smooth input (e.g. [2, Prop. 3.9]) with resulting state trajectory \( x(t) \) such that \( y_0(t) = 0 \). Then for all \( t \) also (see (2.3) and recall that \( Q(\epsilon) \supseteq Q(0) \))

\[ \begin{bmatrix} x'(t) \\ u'(t) \end{bmatrix} Q(0) \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} = 0, \]

i.e. \( y_0(t) = 0 \) and \( x_0 \in \nabla \). Next, if \( \mathcal{W}_0 = \{0\} \) and \( \mathcal{W}_{i+1} = [A, B]\{ (\mathcal{W}_i \otimes \mathbb{R}^m) \cap \mathcal{K} \} \), then \( \mathcal{W} = \mathcal{W}_n \) ([2, (3.22)]). If, in addition, \( \mathcal{W}_0 = \{0\} \) and \( \mathcal{W}_{i+1} = [A, B]\{ (\mathcal{W}_i \otimes \mathbb{R}^m) \cap \mathcal{K}_0 \} \) with \( \mathcal{K}_0 = \ker(Q(0)) \), then \( \mathcal{W} = \mathcal{W}_n \) and it is easy to show that for all \( i = 0, \ldots, n \), \( \mathcal{W}_i \subseteq \mathcal{W}_i \). Thus \( \mathcal{W} \subseteq \mathcal{K} \). The last assertion now is obvious.

Observe from the previous Proposition that if \( \mathcal{K} = \{0\} \) then \( \mathcal{K} = \{0\} \) (in [1] \( \mathcal{K} \) is called the **controllable output-nulling** subspace and denoted by \( \mathcal{K}^* \) and \( \nabla \) is denoted by \( \nabla^* \)). Hence if \( \Sigma_0 \) is left invertible ([2, Th. 3.26]) then \( \Sigma_0 \) is left invertible. This is also remarked in [1] on page 325. Note that not necessarily \( \mathcal{W} = \mathcal{W} \) and \( \nabla = \nabla \).

We are now ready to state our main result. Assume, as in [1, Sec. IV], that \( \Sigma_0 \) is left invertible. Then \( \Sigma_0 \) is left invertible. If we denote the set of invariant zeros of \( \Sigma_0 \) by \( \sigma^*(\Sigma_0) \), then we have
Theorem 2.4.

\[ \sigma^*(\Sigma_\varepsilon) = \sigma^*(\Sigma_0) \text{ for all } \varepsilon \in [0, \varepsilon]. \]

Proof. To start, define for a fixed \( \varepsilon \in (0, \varepsilon] \) the set of mappings \( \Phi(\varepsilon) := \{ F : \mathbb{R}^n \to \mathbb{R}^m \mid (A + BF)v < v, \ (C(\varepsilon) + D(\varepsilon)F)v = 0 \} \). It is well known (e.g. [11]) that the spectrum \( \sigma(A + BF)v \) is independent of \( F \in \Phi(\varepsilon) (\varepsilon = 0) \). Now it is easy to see that \( \sigma(\varepsilon_2) < \sigma(\varepsilon_1) (\varepsilon_2 > \varepsilon_1 > 0) \) since \( Q(\varepsilon_2) \supset Q(\varepsilon_1) \). Therefore we observe that \( \sigma^*(\Sigma_\varepsilon) \) does not depend on \( \varepsilon \) for \( \varepsilon \in (0, \varepsilon] \). Let \( F \in \Phi(\varepsilon) \). It follows from \( Q(\varepsilon) \supset Q(0) \) that \( (C(0) + D(0)F)v = 0 \). Next, let the regular transformation \( S_0 \) be such that \( D(0)S_0 = [D_0, 0] \) with \( D_0 \) of full column rank and write \( F = S_0[F_0|0] \) and \( BS_0 = [B_0, B_0] \). Then we find that \( F_0|v = (-D_0'D_0)^{-1}D_0'C(0)|v \) and thus with

\[
\begin{align*}
C(0)_0 &:= (I - D_0(D_0'D_0)^{-1}D_0'C(0))v, \quad (2.4a) \\
A(0)_0 &:= A - B_0(D_0'D_0)^{-1}D_0'C(0), \quad (2.4b)
\end{align*}
\]

it holds that

\[
(A(0)_0 + B_0F_0_0)v < 0 \text{ and } C(0)_0v = 0. \quad (2.5)
\]

Let \( \lambda \in \sigma(A + BF) \) and \( v \in \mathcal{V} \) such that \( (A + BF)v = \lambda v \) (i.e. \( \lambda \) is an invariant zero of \( \Sigma_\varepsilon \)). Then from the above \( (A(0)_0 + B_0F_0_0)v = \lambda v \). Now (Prop. 2.3) \( \mathcal{V} < \mathcal{V} \) and it is always possible to extend \( F_0 \) such that for the extension \( F_0_1 \), it holds that \( (A(0)_0 + B_0F_0_1)v < 0 \), \( C(0)_0v = 0 \), and \( F_0_1|v = F_0|v \) (see e.g. [2, (3.12)]; the extension for \( F_0 \), of course, is \( F_0|v = (-D_0'D_0)^{-1}D_0'C(0)|v \)). But since for any \( F_0_1 \), for which it holds that \( (A(0)_0 + B_0F_0_1)v < 0 \) and \( C(0)_0v = 0 \), \( \sigma(A(0)_0 + B_0F_0_1)v \) is fixed (\( \varepsilon = 0 \)) we thus have shown that there is a \( v \in \mathcal{V} \) such that \( (A(0)_0 + B_0F_0_1)v = \lambda v \), namely \( v = v \). In other words, \( \lambda \in \sigma^*(\Sigma_0) \).
From Theorem 2.4 we conclude that if \( \sigma^*(E_o) \cap \mathcal{E}^0 = \emptyset \) then \( \sigma^*(I_e) \cap \mathcal{E}^0 = \emptyset \), provided that \( E_o \) is left invertible. Hence the assumption A.3 in [1, Sec. IV] can be replaced by:

\[(A.3)': \sigma^*(I_e) \cap \mathcal{E}^0 = \emptyset.\]

In addition, in Theorem 4.3 of [1] it suffices to assume that A.1, (A.2)', (or A.2), (A.3)' and A.4 hold in order to prove the stated convergence results of inputs, state trajectories and outputs.

Remark.

In [1, Sec. V] the special case of "cheap control" is considered and in [1, Lemma 5.1] it is stated that A.3 holds if and only if \( \sigma^*(I_o) \cap \mathcal{E}^0 = \emptyset \) and \( \sigma^*(I_e) \cap \mathcal{E}^0 = \emptyset \) with \( \sigma^*(I_e) = \sigma(\mathcal{A}|\ker(\mathcal{C})|\mathcal{A}) \). Observe that if \( E_o \) is left invertible (i.e. A.4 holds) then indeed \( \sigma(\mathcal{A}|\ker(\mathcal{C})|\mathcal{A}) < \sigma^*(I_o) \) as follows from Theorem 2.4.
Discussion

In [1, Remark 3.4] it is noted (by means of a counterexample) that in general there is no convergence of the optimal cost for the \( \varepsilon \)-problem without stability (i.e. with no end-point conditions) to the optimal cost for the "boundary" problem without stability. In a future paper ([5]) we will show that this is indeed the case. By starting from the assumptions A.1, (A.2)', (A.3)', and A.4 we will prove in [5] that the limit of the optimal cost for the \( \varepsilon \)-problem without stability equals \( x_0'Kx_0 \) where \( K \) turns out to be the largest element in the set of real symmetric solutions \( K \) of the dissipation inequality that satisfy \( (\gamma + \omega) \subset \text{ker}(K) \). This dissipation inequality is in [3, Sec. 6] written in the form

\[
F(K) = \left[ \begin{array}{cc}
A'K + KA + C'(O)C(O) & KB + C'(O)D(O) \\
B'K + D'(O)C(O) & D'(O)D(O)
\end{array} \right] \geq 0,
\]

and it can be shown that \( K \) even is the largest real symmetric \( K \) for which \( F(K) \geq 0, (\gamma + \omega) \subset \text{ker}(K) \) and rank \( (F(K)) \) is minimal. Also convergence results concerning inputs, states and outputs will be obtained in [5], similar to the ones in [1, Th. 4.3]. A first contribution to the general problem of determining the limits of optimal cost, inputs, states and outputs for the \( \varepsilon \)-problem without stability sketched above is given in [4], where the "cheap control" problem without stability is studied.
References


