# The totally singular linear quadratic problem with indefinite cost

EINDHOVEN UNIVERSITY OF TECHNOLOGY
Department of Mathematics and Computing Science

Memorandum COSOR 89-05

# THE TOTALLY SINGULAR LINEAR QUADRATIC PROBLEM WITH INDEFINITE COST

H.L. Trentelman

Eindhoven University of Technology
Department of Mathematics and Computing Science
P.O. Box 513
5600 MB Eindhoven
The Netherlands

Eindhoven, March 1989
The Netherlands

# THE TOTALLY SINGULAR LINEAR QUADRATIC PROBLEM WITH INDEFINITE COST

Harry L. Trentelman
Department of Mathematics and Computing Science
Eindhoven University of Technology
P.O. Box 513
5600 MB Eindhoven
The Netherlands

**Abstract**

In this paper we study the most general version of the stationary, infinite horizon linear quadratic optimal control problem. In the literature that has appeared on this problem up to now, mostly one (or both) of the following two assumptions are made: (i) the integrand of the cost criterion is given by a *positive semi-definite* quadratic form, (ii) the latter quadratic form is *positive definite in the control variable* alone. In the present paper we propose a problem formulation for the case that neither (i) nor (ii) are imposed. Subsequently, we treat the case that the problem is completely singular.

## 1. INTRODUCTION

The subject of this paper is the time-invariant, infinite horizon linear quadratic optimal control problem. We consider the time-invariant linear system

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

with $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times m}$. For a given control input $u$, the corresponding state trajectory is denoted by $x(u, x_0)$. In addition to (1.1) we consider a quadratic cost functional $J(x_0, u)$ defined by

(1.2) $$J(x_0, u) := \int_0^\infty \omega(x(u, x_0), u)\, dt \ .$$

Here, $\omega$ is a real quadratic form on $\mathbb{R}^n \times \mathbb{R}^m$ defined by

(1.3) $$\omega(x, u) := x^T Q\, x + 2u^T S\, x + u^T R\, u \ ,$$

where $Q \in \mathbb{R}^{n \times n}$ and $R \in \mathbb{R}^{m \times m}$ are symmetric and where $S \in \mathbb{R}^{m \times n}$. We stress that (1.3) is the general shape of a quadratic form on $\mathbb{R}^n \times \mathbb{R}^m$. The optimization problem that we consider in this paper is the following: find the infimum

$$V^+(x_0) := \inf \{ J(x_0, u) \mid u \text{ is such that } \lim_{t \to \infty} x(u, x_0)(t) = 0 \}$$

and find, if one exists, all optimal inputs, that is, all $u^*$ such that $V^+(x_0) = J(x_0, u^*)$.

In the literature that has appeared on this optimization problem up to now, mostly one (or both) of the following two assumptions are made:

1. The quadratic form $\omega$ is positive semi-definite, i.e., $\omega(x, u) \geq 0$ for all $(x, u) \in \mathbb{R}^n \times \mathbb{R}^m$ or, equivalently,

$$\begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \geq 0 \ .$$

2. The quadratic form $\omega$ is positive definite in the variable $u$, i.e. $\omega(0, u) > 0$ for all $u \in \mathbb{R}^m$, $u \neq 0$, or, equivalently, $R > 0$.

Under the Assumptions 1 and 2, the problem has become quite standard in the literature and is usually referred to as the linear quadratic regulator problem [1], [7], [17], [6]. Deleting either one of the above two assumptions introduces its own particular intrinsic difficulties into the problem. If we do assume 2 but delete Assumption 1 we arrive at the problem formulation as studied extensively in [15] (see also [14] and [13]). The linear quadratic problem is then called *regular*. Intrinsic difficulties that arise in this case are especially the existence of solutions to the underlying algebraic Riccati equation and the boundedness from below of the cost functional $J(x_0, u)$. If, on the other hand, we do assume 1 but delete Assumption 2 we obtain a *singular* linear quadratic problem (see [8], [3], [5], [16], [4]). Intrinsic difficulties in this case are the facts that the algebraic Riccati equation is no longer defined and that in general optimal control inputs do not exist unless we extend the class of admissible control inputs to contain also distributions. To the author's knowledge, a treatment of the problem formulation in which *both* Assumption 1 as well as Assumption 2 have been deleted has, up to now, not been given in the literature. In the present paper we shall study this most general formulation of the linear quadratic problem. It will turn out that it is not clear a *priori* how one should formulate this optimal control problem in a mathematically rigorous way. In Section 4 we shall give a precise mathematical formulation of the problem to be considered. In Sections 6 to 9 we shall give a complete solution to this optimal control problem under the assumption that the problem is *completely singular*, i.e., that the matrix $R$ is equal to zero.

## 2. DISTRIBUTIONAL INPUTS

Since $\omega$ is a general real quadratic form, the matrix $R$ can, in principle, be any real symmetric matrix. Typically, $R$ is a singular matrix. As shown in [5], it is then natural to allow the control inputs to be distributions. Let $D'_+$ denote the space of all distributions with support in $\mathbb{R}^+ := [0, \infty)$ (see [11]). We denote by $L_{2,\mathrm{loc}}(\mathbb{R}^+)$ the space of locally square-integrable functions with support in $\mathbb{R}^+$. The latter space can be identified with a subspace of $D'_+$ by defining for $\phi \in L_{2,\mathrm{loc}}(\mathbb{R}^+)$

$$(\phi, \psi) := \int_{\mathbb{R}} \phi(t)\,\psi(t)\,dt \ , \ \psi \in D \ .$$

Here, $D$ denotes the testfunction space of all smooth functions with compact support. With this identification, distributions in $L_{2,\mathrm{loc}}(\mathbb{R}^+)$ will be called *regular* distributions. A regular distribution is called *smooth* if it corresponds to a function which is infinitely often differentiable on $\mathbb{R}^+$ (i.e. derivatives of all orders exist on $(0, \infty)$ and right-derivatives of all orders exist in $t = 0$). A smooth distribution is called *Bohl* if on $\mathbb{R}^+$ it is equal to a finite linear combination of functions of the form $t^k e^{\lambda t}$, where the $k$'s are nonnegative integers and $\lambda \in \mathbb{C}$. A regular distribution $x$ is called stable if $x(t) \to 0$ $(t \to \infty)$.

Let $\delta$ denote the Dirac distribution and let $\delta^{(i)}$ denote its $i$th distributional derivative. Linear combinations of $\delta$ and its higher order derivatives will be called *impulsive* distributions. If $x$ is the smooth distribution corresponding to the $C^\infty$ function, say $\tilde{x}$, on $\mathbb{R}^+$ then it is easily checked that its derivative $\dot{x}$ is given by

$$(2.1) \qquad \dot{x} = \tilde{x}(0^+)\,\delta + \frac{d\tilde{x}}{dt} \;.$$

Here, $\frac{d\tilde{x}}{dt}$ is the derivative of the function $\tilde{x}$ on $\mathbb{R}^+$ and $\tilde{x}(0^+) := \lim_{t \downarrow 0} \tilde{x}(t)$ is the "jump" that $\tilde{x}$ makes at $t = 0$. In the sequel, the symbol $*$ denotes convolution of distributions. The unit element of the convolution operation is $\delta$, i.e. $\delta * x = x$ for all $x \in D'_+$. A distribution $x$ is called *invertible* if there exists a distribution $x^{-1}$ such that $x * x^{-1} = x^{-1} * x = \delta$. It is well-known that $\delta^{(1)} * x = \dot{x}$, where $\dot{x}$ denotes the derivative of $x$.

If $m, n \in \mathbb{N}$, then any $m$-vector or $m \times n$-matrix of regular (smooth, Bohl, impulsive) distributions is again called a regular (smooth, Bohl, impulsive) distribution. If $A \in \mathbb{R}^{n \times n}$ and if $I$ is the $n \times n$-identity matrix then $I\delta^{(1)} - A\delta$ can be shown to be invertible. Moreover, $(I\delta^{(1)} - A\delta)^{-1}$ is equal to the Bohl-distribution corresponding to $e^{At}$ $(t \geq 0)$.

In this paper, we restrict ourselves to the following class of inputs:

$$U_{\text{dist}} := \{u \mid u = u_1 + u_2, \text{ with } u_1 \text{ impulsive and } u_2 \text{ regular}\} \;.$$

An element of $U_{\text{dist}}$ is called an *impulsive-regular* distribution. If $u_2$ is smooth (Bohl) then $u$ is called *impulsive-smooth (impulsive-Bohl)*. An impulsive-regular distribution is called stable if its regular part is stable. It follows from (2.1) that if an impulsive-Bohl distribution $u$ is stable then also $\dot{u}$ is stable (and impulsive-Bohl).

We now briefly discuss what is meant by the solution of $\dot{x} = Ax + Bu$, $x(0) = x_0$ if $u \in U_{\text{dist}}$. This is a non-trivial matter, since distributions do not have a well-defined value at a particular time instant $t_0$. Now, for $u \in U_{\text{dist}}$ this solution is *defined* as the unique solution $x(u, x_0)$ of the distributional differential equation $\dot{x} = Ax + Bu + \delta x_0$. It is easy to show that this equation indeed has an unique solution, given by

$$x(u, x_0) = (I\delta^{(1)} - A\delta)^{-1} * (Bu + \delta x_0) \;.$$

Moreover, $x(u, x_0)$ is again impulsive-regular. For $x_0 \in \mathbb{R}^n$ we define

$$U_{\text{dist}}^{\text{stab}}(x_0) := \{u \in U_{\text{dist}} \mid x(u, x_0) \text{ is stable}\} \;.$$

Finally, we introduce the following notation. If $P(s) = P_n s^n + P_{n-1} s^{n-1} + \cdots + P_1 s + P_0$ is a polynomial matrix, then $P(\delta^{(1)})$ denotes the impulsive distribution given by

$$P(\delta^{(1)}) = P_n\delta^{(n)} + P_{n-1}\delta^{(n-1)} + \cdots + P_1\delta^{(1)} + P_0\delta \ .$$

If $x$ is a stable impulsive-Bohl distribution and $P(s)$ is a polynomial matrix of compatible dimension, then obviously $P(\delta^{(1)}) * x$ is again a stable impulsive-Bohl distribution. If $G(s) = C(Is - A)^{-1}B$, then $G(\delta^{(1)})$ will denote the Bohl distribution $C(I\delta^{(1)} - A\delta)^{-1}B$ (corresponding to the function $Ce^{At}B$ on $\mathbb{R}^+$).

## 3. THE STRONGLY CONTROLLABLE SUBSPACE

In this section we recall briefly the concept of strongly controllable subspace [5]. Temporarily consider the system

$$\begin{aligned} \dot{x} &= Ax + Bu \ , \\ y &= Cx \ , \end{aligned} \tag{3.1}$$

where $A$ and $B$ are as before and $C \in \mathbb{R}^{p \times n}$. As noted in Section 2, if $x_0 \in \mathbb{R}^n$ and $u \in U_{\text{dist}}$, then the resulting state trajectory $x(x_0, u)$ is impulsive-regular. Thus, $x(x_0, u) = x_1 + x_2$ with $x_1$ impulsive and $x_2$ regular. Denote

$$x(x_0, u)(0^+) := \lim_{t \downarrow 0} x_2(t) \ . \tag{3.2}$$

Intuitively, (3.2) represents the point in state space to which the state trajectory "jumps" from $x_0$ instantaneously. Given $x_0$ and $u$, let $y(x_0, u) = Cx(x_0, u)$ denote the corresponding output of (3.1).

**Definition 3.1.** A point $x_0 \in \mathbb{R}^n$ is called strongly controllable if there exists $u \in U_{\text{dist}}$ such that $x(x_0, u)(0^+) = 0$ and $y(x_0, u) = 0$. The subspace of all strongly controllable points is called *the strongly controllable subspace* of (3.1).

The strongly controllable subspace of (3.1) is denoted by $T(A, B, C)$. It is well-known (see [12], [5], [10]) that $T(A, B, C)$ is equal to the smallest subspace $V$ of $\mathbb{R}^n$ such that $\operatorname{im} B \subset V$ and such that there exists a matrix $G$ with $(A + GC)V \subset V$. From the latter property it is easily seen that $T(A, B, C)$ is invariant under state feedback, i.e., $T(A, B, C) = T(A + BF, B, C)$ for all $F$. The system (3.1) is called *strongly controllable* if $T(A, B, C) = \mathbb{R}^n$ (see [9]). It is easily seen that if (3.1) is strongly controllable, then $(A, B)$ is controllable. Let $G(s) = C(Is - A)^{-1}B$ be the transfer matrix of (3.1). It is well-known (see [9]) that if (3.1) is strongly controllable and if $C$ is surjective then $G$ has a polynomial right-inverse, i.e., there exists a real polynomial matrix $P$ such that $GP = I_p$, the $p \times p$ identity matrix.

A final result that will be needed in the sequel is that $T(A, B, C)$ is generated by a recursive algorithm. For $i = 0, 1, 2, \cdots$ define:

$$
(3.3) \qquad \begin{aligned} T_0 &:= 0 , \\ T_{i+1} &:= \operatorname{im} B + A(T_i \cap \ker C) . \end{aligned}
$$

It can be checked immediately that the sequence $T_i$ is non-decreasing and that it becomes stationary after at most $n$ iterations. It was shown in [5] (see also [12]) that $T_n = T(A, B, C)$.

## 4. ADMISSIBLE INPUTS

Let $u \in U_{\text{dist}}$. The question that we want to consider in this section is: how should we define $J(x_0, u)$ if $u$ is not a regular distribution? The problem here is, that a satisfactory definition of *multiplication of distributions* does not exist in mathematics. Thus it is not clear how one should interpret the expression $\omega(x(x_0, u), u)$, since the products $x(x_0, u)^T Q\, x(x_0, u)$, $u^T S\, x(x_0, u)$ and $u^T R u$ are, in general, undefined objects. To illustrate this conceptual difficulty, consider the following example:

**Example 4.1** $\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u, \quad \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ ,

$\omega(x_1, x_2, u) = 3x_1^2 - 4x_1 x_2 - 2x_2^2 - 2ux_1 + 4ux_2$ .

Clearly, $\omega$ is an indefinite quadratic form. The $R$-matrix of $\omega$ is equal to 0. Let us take $u = -\delta^{(1)}$. Then $x_2 = -\delta$ and $x_1$ is given by $x_1(0) = 1$, $x_1(t) = 0$ $(t > 0)$. Note that $\omega(x_1, x_2, u)$ is not defined for these distributions: the products $x_1 x_2$, $x_2^2$, $ux_1$ and $ux_2$ are not defined. Thus, the question arises: how should we interpret $J(x_0, u)$ for this choice of input $u$? In this example we propose to do this as follows. First, let us restrict ourselves to *regular* inputs, i.e., assume $u \in L_{2,\text{loc}}(\mathbb{R}^+)$. Using the relations $\dot{x}_1 = x_2$, $\dot{x}_2 = u$ we then have

$$
\omega(x_1, x_2, u) = 3x_1^2 - 4x_1\dot{x}_1 - 2x_2^2 - 2\dot{x}_2 x_1 + 4\dot{x}_2 x_2 \ ,
$$

and hence

$$
\omega(x_1, x_2, u) = 3x_1^2 - \frac{d}{dt}(2x_1^2 + 2x_1 x_2 - 2x_2^2) \ .
$$

Consequently, if $x_1(t) \to 0$ and $x_2(t) \to 0$ $(t \to \infty)$ we find

$$
\int_0^\infty \omega(x_1, x_2, u)\, dt = \int_0^\infty 3x_1^2 dt + 2x_1(0)^2 + 2x_1(0)\, x_2(0) - 2x_2(0)^2 \ .
$$

From this it is easily seen that the infimum of $\int_0^\infty \omega(x_1, x_2, u)\, dt$ over all *regular* inputs such that $x_1(t) \to 0$ and $x_2(t) \to 0$ $(t \to \infty)$, is equal to $2x_1(0)^2 + 2x_1(0)x_2(0) - 2x_2(0)^2$, which in our example is equal to 2. Clearly an input $u$ is *optimal* if and only if $\int_0^\infty x_1^2\,dt = 0$. Since $x_1(0) = 1$, this can, in our example, not be achieved by a regular input $u$. Since however $u = -\delta^{(1)}$ yields $\int_0^\infty x_1^2\,dt = 0$, it is reasonable to call $u = \delta^{(1)}$ optimal! This example clearly suggests to call an input $u \in U_{\text{dist}}^{\text{stab}}(x_0)$ *admissible* if $x_1$ is regular and for those $u$ to *define*

$$J(x_0, u) := \int_0^\infty 3x_1^2 dt + 2x_{10}^2\,dt + 2x_{10}x_{20} - 2x_{20}^2 \;.$$

We would like to generalize the above to the situation that we have an arbitrary system $(A, B)$ together with a quadratic form $\omega(x, u) = x^TQx + 2u^TSx + u^TRu$. In the following, denote

$$V_{\text{reg}}^+(x_0) := \inf\{J(x_0, u) \mid u \in U_{\text{dist}}^{\text{stab}}(x_0) \text{ is regular}\} \;.$$

In order to make sure that we are not performing infimization over an empty set, as a standing assumption we assume that $(A, B)$ *is controllable*. If this is the case then we have $V_{\text{reg}}^+(x_0) < +\infty$ for all $x_0 \in \mathbb{R}^n$. On the other hand we want to have $V_{\text{reg}}^+(x_0) > -\infty$ for all $x_0 \in \mathbb{R}^n$. A necessary and sufficient condition for this was established in [15]:

**Theorem 4.2.** Assume that $(A, B)$ is controllable. Then $V_{\text{reg}}^+(x_0) > -\infty$ for all $x_0 \in \mathbb{R}^n$ if and only if there exists a real symmetric matrix $K \in \mathbb{R}^{n\times n}$ such that

$$F(K) := \begin{bmatrix} A^TK+KA+Q & KB+S^T \\ B^TK+S & R \end{bmatrix} \geq 0 \;. \tag{4.1}$$

□

If $K$ satisfies (4.1), it is said to satisfy the linear matrix inequality. We denote $\Gamma := \{K \in \mathbb{R}^{n\times n} \mid K = K^T \text{ and } F(K) \geq 0\}$. Motivated by this theorem, a second standing assumption in this paper will be that $\Gamma \neq \varnothing$.

Let $K \in \Gamma$. Then we can factorize

$$F(K) = \begin{bmatrix} C_K^T \\ D_K^T \end{bmatrix} (C_K\; D_K) \;. \tag{4.2}$$

A useful fact is the following (see [15]):

**Lemma 4.3.** Let $x_0 \in \mathbb{R}^n$. For every regular $u \in U_{\text{dist}}^{\text{stab}}(x_0)$ we have

$$J(x_0, u) = \int_0^\infty \|C_Kx(x_0, u)(t) + D_Ku(t)\|^2 dt + x_0^TKx_0 \;. \tag{4.3}$$

□

The latter equality will be used to define the value of the cost-functional for inputs $u \in U_{\text{dist}}^{\text{stab}}(x_0)$ that are *not* regular. The idea is to call an input $u \in U_{\text{dist}}^{\text{stab}}(x_0)$ *admissible* if $C_K x(x_0, u) + D_K u$ is *regular*. The class of all those inputs is denoted by $U_{\text{adm}}(x_0)$. Next, for $u \in U_{\text{adm}}(x_0)$ we *define* $J(x_0, u)$ by (4.3). Of course, there is one slight complication that we have to deal with: if $\Gamma \neq \varnothing$ then, in general, it has more than one element $K$. Thus, our class $U_{\text{adm}}(x_0)$ and the value $J(x_0, u)$ in principle depend on the solution $K$ of $F(K) \geq 0$ for which we perform the factorization (4.2). Temporarily, denote by $U_{\text{adm}}^{K}(x_0)$ the class of all $u \in U_{\text{dist}}^{\text{stab}}(x_0)$ such that $C_K x(x_0, u) + D_K u$ is regular and by $J_K(x_0, u)$ the value defined by (4.3). It can be shown that, in fact, $U_{\text{adm}}^{K}(x_0)$ and $J_K(x_0, u)$ are independent of $K$ for $K \in \Gamma$:

**Theorem 4.4.** Let $K_1, K_2 \in \Gamma$. Let $x_0 \in \mathbb{R}^n$. Then we have $U_{\text{adm}}^{K_1}(x_0) = U_{\text{adm}}^{K_2}(x_0)$. Denote this class by $U_{\text{adm}}(x_0)$. Then for all $u \in U_{\text{adm}}(x_0)$ we have $J_{K_1}(x_0, u) = J_{K_2}(x_0, u)$.

**Proof.** This is a consequence of [5, Lemma 6.21]. □

We are now in a position to give a precise mathematical statement of the optimization problem we consider in this paper. Consider the system (1.1) together with the quadratic form (1.3). Assume $(A, B)$ controllable and $\Gamma \neq \varnothing$. Then our problem is: given $x_0 \in \mathbb{R}^n$, find the infimum

$$V_{\text{dist}}^{+}(x_0) := \inf\{J(x_0, u) \mid u \in U_{\text{adm}}(x_0)\}$$

and find, if one exists, all optimal inputs $u^*$, i.e., find all $u^* \in U_{\text{adm}}(x_0)$ such that $V_{\text{dist}}^{+}(x_0) = J(x_0, u^*)$.

It was shown in [15] that if $\Gamma \neq \varnothing$ and if $(A, B)$ is controllable, then it has a largest element, that is, there is $K^+ \in \Gamma$ such that for all $K \in \Gamma$ we have $K \leq K^+$. Also, if was proven in [15] that $V_{\text{reg}}^{+}(x_0) = x_0^T K^+ x_0$ for all $x_0 \in \mathbb{R}^n$. Now, *by definition*, for $u \in U_{\text{adm}}(x_0)$ we have

$$J(x_0, u) = \int_0^\infty \|C_{K^+} x(u, x_0) + D_{K^+} u\|^2 \, dt + x_0^T K^+ x_0 \ . \tag{4.4}$$

From this it follows immediately that also $V_{\text{dist}}^{+}(x_0) = x_0^T K^+ x_0$. Thus we have dealt with the first part of the problem posed above: we have characterized the infimum $V_{\text{dist}}^{+}(x_0)$.

In order to find an optimal input $u^*$ one could now proceed as follows. Calculate any real symmetric solution $K$ to $F(K) \geq 0$. Factorize $F(K)$ according to (4.2). Assuming that it exists, calculate an optimal control $u^*$ for the singular linear quadratic problem with positive semi-definite cost-functional $\int_0^\infty \|C_K x(u, x_0) + D_K u\|^2 dt$ (this can be done using the theory developed in [16]). Clearly, $u^*$ is then also optimal for our original problem. However, this construction would give us $u^*$ in terms of the transformed data $(A, B, C_K, D_K)$ rather than in terms of the original system parameters $(A, B, Q, S, R)$. In the sequel, we want to develop a theory in which the optimal controls (and conditions for their existence) are expressed in terms of the original data

$A$, $B$, $Q$, $S$ and $R$.

## 5. THE REGULAR LINEAR QUADRATIC PROBLEM

In this section we shall show that if $R > 0$ then the linear quadratic problem as defined in Section 4 reduces to the "classical" regular linear quadratic problem as treated in [15]. Again assume that $(A, B)$ is controllable and that $\Gamma \neq \emptyset$. It is well-known that if $R > 0$ then the largest element $K^+$ of $\Gamma$ is equal to the largest real symmetric solution of the algebraic Riccati equation

$$A^T K + KA + Q - (KB + S^T) R^{-1} (B^T K + S) = 0 \ . \tag{5.1}$$

Thus, $F(K^+)$ can in that case be factorized as

$$F(K^+) = \begin{bmatrix} (K^+ B + S^T) R^{-\frac{1}{2}} \\ R^{\frac{1}{2}} \end{bmatrix} (R^{-\frac{1}{2}} (B^T K^+ + S) \;\; R^{\frac{1}{2}}) \ .$$

Consequently, if an input $u$ is admissible then necessarily

$$z := R^{-\frac{1}{2}} (B^T K^+ + S) x(x_0, u) + R^{\frac{1}{2}} u$$

is a regular distribution. This implies that if $u$ is admissible then it must satisfy

$$u = R^{-\frac{1}{2}} z - R^{-1} (B^T K^+ + S) x(x_0, u)$$

for some regular $z$. Now, if the latter holds then of course $x(x_0, u)$ satisfies

$$\dot{x} = (A - BR^{-1}(B^T K^+ + S)) x + BR^{-\frac{1}{2}} z + \delta x_0$$

and therefore, since $z$ is regular, $x(x_0, u)$ is regular. In turn, this implies that $u$ is regular. We conclude that if $R > 0$ then every admissible input is regular and, in particular, for all $x_0 \in \mathbb{R}^n$ we have

$$U_{\mathrm{adm}}(x_0) = \{u \in L^m_{2,\mathrm{loc}}(\mathbb{R}^+) \mid \lim_{t \to \infty} x(x_0, u)(t) = 0\} \ .$$

(Here, $L^m_{2,\mathrm{loc}}(\mathbb{R}^+)$ denotes the space of $m$-vectors with components in $L_{2,\mathrm{loc}}(\mathbb{R}^+)$). Furthermore, in that case for all $u \in U_{\mathrm{adm}}(x_0)$ we have *by definition*

$$J(x_0, u) = \int_0^\infty \| R^{-\frac{1}{2}} (B^T K^+ + S) x(x_0, u) + R^{\frac{1}{2}} u \|^2 \, dt + x_0^T K^+ x_0 \ ,$$

which is easily seen to be equal to $\int_0^\infty \omega(x(x_0, u)(t), u(t))\, dt$. The conclusion we draw from all this is that if $R > 0$ then the problem as we formulated it in Section 4 coincides with the "classical" problem formulation as studied in [15]. We shall briefly recall its solution here. Let $K^-$ be the smallest solution of the algebraic Riccati equation (5.1) and let $\Delta := K^+ - K^-$ (the "gap" of (5.1)).

**Theorem 5.1.** [15] Let $R > 0$, $(A, B)$ controllable and $\Gamma \neq \varnothing$. Then we have the following:

(i) For all $x_0 \in \mathbb{R}^n$ $V^+_{\text{dist}}(x_0) = x_0^T K^+ x_0$.

(ii) For all $x_0 \in \mathbb{R}^n$ there exists an optimal input $u^* \in U_{\text{adm}}(x_0)$ if and only if $\Delta > 0$.

(iii) Let $\Delta > 0$. Then for each $x_0 \in \mathbb{R}^n$ there is exactly one optimal input $u^* \in U_{\text{adm}}(x_0)$. This input is given by the feedback law

$$u = -R^{-1}(B^T K^+ + S)x \ .$$

(iv) $A - BR^{-1}(B^T K^+ + S)$ is asymptotically stable if and only if $\Delta > 0$. □

## 6. A STRONGLY CONTROLLABLE SUBSPACE FOR (A, B, ω)

In the remainder of this paper we assume that the matrix $R$ appearing in the cost functional (1.2), (1.3) is equal to zero. It is easily verified that a real symmetric matrix $K$ is an element of $\Gamma$ if and only if $A^T K + KA + Q \geq 0$ and $KB + S^T = 0$. In the sequel, for $K \in \Gamma$ we shall denote

$$(6.1) \qquad L(K) := A^T K + KA + Q \ .$$

For $x_0 \in \mathbb{R}^n$ the class of admissible inputs $U_{\text{adm}}(x_0)$ is in this case given by

$$U_{\text{adm}}(x_0) = \{u \in U_{\text{dist}} \mid L(K)x(x_0, u) \text{ is regular and } x(x_0, u) \text{ is stable}\} \ .$$

(where $K \in \Gamma$ is arbitrary). Furthermore, the cost $J_{\text{dist}}(x_0, u)$ is given by

$$J_{\text{dist}}(x_0, u) = \int_0^\infty \|C_K x(x_0, u)\|^2 dt + x_0^T K x_0 \ ,$$

where, again, $K \in \Gamma$ is arbitrary and where $C_K$ is any matrix such that $C_K^T C_K = L(K)$ (of course, $L(K)x$ is regular if and only if $C_K x$ is regular).

For $K \in \Gamma$, let $T(K)$ denote the strongly controllable subspace associated with the system $\dot{x} = Ax + Bu$, $y = L(K)x$, i.e.,

$$(6.2) \qquad T(K) := T(A, B, L(K)) \ .$$

Furthermore, for $i = 0, 1, 2, \cdots$, let $T_i(K)$ be defined recursively by

$$
(6.3) \qquad \begin{aligned} &T_0(K)=0\,, \\ &T_{i+1}(K)=\operatorname{im}B+A\,(T_i(K)\cap\ker L(K))\,. \end{aligned}
$$

As already noted in Section 3 we have $T_n(K)=T(K)$. It turns out that $T_i(K)$ is, in fact, independent of $K$ and that we have:

**Lemma 6.1.** Let $K_1, K_2 \in \Gamma$. Then for all $i$ we have

(i) $T_i(K_1)=T_i(K_2)=: T_i$ ,

(ii) $T_i \subset \ker(K_1-K_2)$ ,

(iii) $T_i \cap \ker L(K_1)=T_i \cap \ker L(K_2)$ .

**Proof.** We first prove (i) and (ii). This is done by induction. For $i=0$ the claims are obvious. Now assume they are true up to $i$, that is, assume $T_i(K_1)=T_i(K_2)=: T_i$ and $T_i \subset \ker(K_1-K_2)$. Let $x \in T_{i+1}(K)$. There exists $u \in I\!R^m$ and $w \in T_i$ such that $x=Bu+Aw$ and $L(K_1)w=0$. Obviously we have

$$
(6.4) \qquad L(K_1)=L(K_2)+A^T(K_1-K_2)+(K_1-K_2)A
$$

and hence

$$
L(K_2)\,w+A^T(K_1-K_2)w+(K_1-K_2)Aw=0\ .
$$

Since $w \in T_i$ we have $(K_1-K_2)\,w=0$. Thus we find $w^TL(K_2)w=0$ and therefore $w \in \ker L(K_2)$. We conclude that $x \in T_{i+1}(K_2)$. By again applying (6.4) and using the fact that $K_1B=K_2B$, we obtain

$$
\begin{aligned} (K_1-K_2)x &= (K_1-K_2)\,(Aw+Bu) \\ &= (K_1-K_2)\,Aw=0\ . \end{aligned}
$$

By interchanging $K_1$ and $K_2$ we also find the converse inclusion. Put $T_{i+1} := T_{i+1}(K_1)=T_{i+1}(K_2)$. Then $T_{i+1} \subset \ker(K_1-K_2)$. This proves (i) and (ii) for all $i$. Finally, (iii) follows immediately from the facts that for all $x \in T_i$ we have $K_1x=K_2x$ and that $L(K_j)\geq 0$ $(j=1,2)$. ▯

As a consequence of the above result we also find that the limiting subspaces $T(K_1)$ and $T(K_2)$ (see (6.2)) coincide for $K_1, K_2 \in \Gamma$. Thus we have assigned to the quadruple $(A, B, Q, S)$ a *unique* subspace $T(K)$, independent of $K \in \Gamma$. This subspace will be denoted by $T^*$. Also by the previous lemma we know that for $K_1, K_2 \in \Gamma$ we have $K_1 \mid T^* = K_2 \mid T^*$. Finally, we know that the subspace $T^* \cap \ker L(K)$ is independent of $K$ for $K \in \Gamma$. This subspace is therefore also uniquely determined by $(A, B, Q, S)$. It will be denoted by $W^*$. It follows from (6.3) that

(6.5) $\qquad AW^* \subset T^*,\ \mathrm{im}\, B \subset T^*$ .

In the sequel, the subspaces $T^*$ and $W^*$ and the mapping $K \mid T^*$ for $K \in \Gamma$ will play an important role in characterizing the optimal cost and optimal controls for the optimization problem under consideration. One possible way to calculate $T^*, W^*$ and $K \mid T^*$ is of course first to calculate some $K \in \Gamma$ and then to apply the algorithm (6.3). However, it is not clear in general how to calculate an element of $\Gamma$ (see e.g. [2]). Instead, we shall therefore develop a *recursive algorithm* to calculate $T^*, W^*$ and $K \mid T^*$ without having to calculate an element of $\Gamma$ first. At each iteration, the algorithm calculates a subspace $R_i$ of the state space, a matrix $V_i$ such that $\mathrm{im}\, V_i = R_i$ and a matrix $S_i$. We will show that, for all $K \in \Gamma, T_i(K) = R_i$ and that the map $K \mid T_i(K)$ is given by $-S_i^T$. Consequently, $T^* = R_n$ and $K \mid T^*$ is given by $-S_n^T$.

**Algorithm 6.2**

**data:** $A, B, Q, S$

**step 1:** Put $R_1 := \mathrm{im}\, B,\ V_1 := B,\ S_1 := S$.

**from** $i$ **to** $i + 1$: Put

$$R_{i+1} = \mathrm{im}\, B + AV_i \ker(V_i^T QV_i - S_i AV_i - (S_i AV_i)^T) \ .$$

Let dim $R_{i+1} = m_{i+1}$. Choose $V_{i+1} \in \mathbb{R}^{n \times m_{i+1}}$ such that $R_i = \mathrm{im}\ V_{i+1}$. Let $U_i$ and $X_i$ be matrices such that $R_{i+1} = \mathrm{im}\ V_{i+1}$.

Let $U_i$ and $X_i$ be matrices such that

$$V_{i+1} = BU_i + AV_i X_i$$

and

$$\mathrm{im}\, X_i \subset \ker (V_i^T QV_i - S_i AV_i - (S_i AV_i)^T) \ .$$

Then define a matrix $S_{i+1} \in \mathbb{R}^{m_{i+1} \times n}$ by

$$S_{i+1} = U_i^T S + X_i^T (V_i^T Q - S_i A) \ .$$

**Theorem 6.3** For all $i = 1, 2, 3, \cdots$ and $K \in \Gamma$ we have $R_i = T_i(K)$ and $KV_i = -S_i^T$. Furthermore, $T^* = R_n$,

(6.6) $\qquad W^* = V_n \ker(V_n^T QV_n - S_n AV_n - (S_n AV_n)^T)$

and $KV_n = -S_n^T$ for all $K \in \Gamma$.

**Proof.** We first prove the first two assertions. Clearly, these are true for $i = 1$. Assume now they are true up to $i$. Then we clearly have

$$V_i^T L(K) V_i = V_i^T Q V_i - S_i A V_i - (S_i A V_i)^T \geq 0 \ .$$

Therefore

(6.7) $$T_i(K) \cap \ker L(K)$$
$$= \{V_i \zeta \mid L(K)\, V_i \zeta = 0\}$$
$$= \{V_i \zeta \mid \zeta^T V_i^T L(K)\, V_i \zeta = 0\}$$
$$= V_i \ker(V_i^T Q V_i - S_i A V_i - (S_i A V_i)^T) \ ,$$

which immediately implies that $T_{i+1}(K) = R_{i+1}$. Next we show $KV_{i+1} = -S_{i+1}^T$. Since

$$\operatorname{im} X_i \subset \ker V_i^T L(K)\, V_i$$

and since $L(K) \geq 0$, we find $L(K)\, V_i X_i = 0$. This yields

$$KAV_i X_i = (A^T S_i^T - QV_i)\, X_i$$

and hence

$$KV_{i+1} = KBU_i + KAV_i X_i$$
$$= -S^T U_i + (A^T S_i^T - QV_i)\, X_i$$
$$= -S_{i+1}^T \ .$$

Thus we have proven the first two assertions. Of course, the facts $T^* = R_n$ and $KV_n = -S_n^T$ follow immediately. Finally, (6.6) follows from (6.7). □

## 7. A SUITABLE DECOMPOSITION OF THE STATE SPACE

In this section we shall introduce a decomposition of the state space that will enable us to display some important structural properties of the system $(A, B)$ and the quadratic form $\omega(x, u) = x^T Q x + 2u^T S x$. Let $x_1, \cdots, x_s, x_{s+1}, \cdots, x_t, x_{t+1}, \cdots, x_n$ be an orthonormal basis of $\mathbb{R}^n$ such that $x_{s+1}, \cdots, x_n$ is a basis of $T^*$ and $x_{s+1}, \cdots, x_t$ is a basis of $W^*$. In other words, decompose $\mathbb{R}^n$ as an orthogonal direct sum $\mathbb{R}^n = X_1 \oplus X_2 \oplus X_3$ with $X_2 = W^*$ and $X_2 \oplus X_3 = T^*$. Using (6.5) we find that the matrices of $A$ and $B$ with respect to this basis have the form

$$A = \begin{bmatrix} A_{11} & 0 & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}, \ B = \begin{bmatrix} 0 \\ B_2 \\ B_3 \end{bmatrix} \ .$$

Moreover, the matrix of $Q$ with respect to such basis is symmetric. Let it be given by

$$Q = \begin{bmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{12}^T & Q_{22} & Q_{23} \\ Q_{13}^T & Q_{23}^T & Q_{33} \end{bmatrix} .$$

Also, if $K \in \Gamma$ then its matrix with respect to the above basis is symmetric. Let it be given by

$$K = \begin{bmatrix} K_{11} & K_{12} & K_{13} \\ K_{12}^T & K_{22} & K_{23} \\ K_{13}^T & K_{23}^T & K_{33} \end{bmatrix} .$$

Let $(x_1^T, x_2^T, x_3^T)^T$ be the coordinate vector of an arbitrary $x \in \mathbb{R}^n$. The equations of the system $\dot{x} = Ax + Bu$ can be arranged in such a way that they have the form

(7.1) $$\dot{x}_1 = A_{11}x_1 + A_{13}x_3 ,$$

(7.2) $$\begin{bmatrix} \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} A_{21} \\ A_{31} \end{bmatrix} x_1 + \begin{bmatrix} B_2 \\ B_3 \end{bmatrix} u .$$

Finally, for $K \in \Gamma$, let $L(K)$ be as defined by (6.1). With respect to the basis introduced above

$$L(K) = \begin{bmatrix} L_{11}(K) & 0 & L_{13}(K) \\ 0 & 0 & 0 \\ L_{13}(K)^T & 0 & L_{33}(K) \end{bmatrix} .$$

(Here, the zero blocks appear due to the fact that $X_2 = W^* \subset \ker L(K)$). We shall prove the following lemma:

**Lemma 7.1.**

(i) Let $K \in \Gamma$. Then $L_{33}(K) > 0$.

(ii) The system

$$\Sigma_1 := \left[ \begin{bmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} B_2 \\ B_3 \end{bmatrix} , (0 \; I) \right] ,$$

with state space $X_2 \oplus X_3 (=T^*)$, input space $\mathbb{R}^m$ and output space $X_3$, is strongly controllable.

**Proof.** (i) Let $(x_1^T, x_2^T, x_3^T)^T$ be the coordinate vector of $x \in \mathbb{R}^n$. First note that $L_{33}(K) \geq 0$. Assume $L_{33}(K)x_3 = 0$. Let $\tilde{x}$ be the vector with coordinates $(0^T, 0^T, x_3^T)^T$. Then $\tilde{x}^T L(K) \tilde{x} = 0$ whence $L(K)\tilde{x} = 0$. Consequently, $\tilde{x} \in T^* \cap \ker L(K) = X_2$. Also, $\tilde{x} \in X_3$. Thus $\tilde{x} = 0$ so $x_3 = 0$.

(ii) Let $T(\Sigma_1)$ be the strongly controllable subspace associated with the system $\Sigma_1$. We shall prove that $T(\Sigma_1) = X_2 \oplus X_3$. First note that there exists $G_0 = \begin{bmatrix} G_2 \\ G_3 \end{bmatrix}$ such that

$$\left[ \begin{bmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{bmatrix} + \begin{bmatrix} G_2 \\ G_3 \end{bmatrix} (0 \ I) \right] T(\Sigma_1) \subset T(\Sigma_1)$$

and that

$$\operatorname{im} \begin{bmatrix} B_2 \\ B_3 \end{bmatrix} \subset T(\Sigma_1) \ .$$

Now assume that $T(\Sigma_1) \subset X_2 \oplus X_3$ with *strict* inclusion. Define $V \subset \mathbb{R}^n$ by

$$V := \{ \begin{bmatrix} 0 \\ x_2 \\ x_3 \end{bmatrix} \ \Big| \ \begin{bmatrix} x_2 \\ x_3 \end{bmatrix} \in T(\Sigma_1) \} \ .$$

Then clearly $V \subset T^*$ with *strict* inclusion. Now take an arbitrary $K \in \Gamma$. Recall that $T^*$ is equal to the strongly reachable subspace of $(A, B, L(K))$. Now, we claim that there exists $G$ such that

(7.3) $\qquad (A + G\,L(K))\,V \subset V \ .$

Indeed, (7.3) holds if we take

$$G := \begin{bmatrix} -A_{13} \\ G_2 \\ G_3 \end{bmatrix} (0 \ \ 0 \ \ L_{33}(K)^{-1}) \ .$$

We also have

(7.4) $\qquad \operatorname{im} B \subset V \ .$

This however contradicts the fact that $T^*$ is equal to the smallest subspace $V$ for which a $G$ exists such that (7.3) and (7.4) hold. □

We shall also need the following:

**Lemma 7.2.** Let $K_1, K_2 \in \Gamma$, with

$$K_l = \begin{bmatrix} K_{l11} & K_{l12} & K_{l13} \\ K_{l12}^T & K_{l22} & K_{l23} \\ K_{l13}^T & K_{l23}^T & K_{l33} \end{bmatrix} \quad (l = 1, 2) \ .$$

Then $K_{1ij} = K_{2ij}$ for all $(i, j) \neq (1, 1)$.

**Proof.** It was already noted in Section 6 that $K_1 \mid T^* = K_2 \mid T^*$. Thus, $K_1$ and $K_2$ coincide on $X_2 \oplus X_3$. □

For $K \in \Gamma$, denote the fixed blocks $K_{ij}$, $(i, j) \neq (1, 1)$, by $M_{ij}$. Define

$$M := \begin{bmatrix} 0 & M_{12} & M_{13} \\ M_{12}^T & M_{22} & M_{23} \\ M_{13}^T & M_{23}^T & M_{33} \end{bmatrix} . \tag{7.5}$$

We note that the matrix $M$ can be calculated from the system matrices $A$, $B$, $Q$ and $S$ directly, by means of the Algorithm 6.2.

## 8. A REDUCED ORDER REGULAR LINEAR QUADRATIC PROBLEM

In this section we show that in order to tackle the linear quadratic problem associated with our original system $\dot{x} = Ax + Bu$ and quadratic form $\omega(x, u) = x^T Qx + 2u^T Sx$ we can consider the linear quadratic problem associated with the subsystem $\dot{x}_1 = A_{11}x_1 + A_{13}x_3$ (see (7.1)) and quadratic form

$$\omega_r(x_1, x_3) = x_1^T Q_r x_1 + 2x_3^T S_r x_1 + x_3^T R_r x_3 \; . \tag{8.1}$$

Here, $Q_r$, $S_r$ and $R_r$ are defined as follows:

$$Q_r := A_{21}^T M_{12}^T + M_{12}A_{21} + A_{31}^T M_{13}^T + M_{13}A_{31} + Q_{11} \; , \tag{8.2}$$

$$S_r := A_{23}^T M_{12}^T + A_{33}^T M_{13}^T + M_{13}^T A_{11} + M_{23}^T A_{21} + M_{33}A_{31} + Q_{13}^T \; , \tag{8.3}$$

$$R_r := M_{13}^T A_{13} + A_{13}^T M_{13} + M_{23}^T A_{23} + A_{23}^T M_{23} + M_{33}A_{33} + A_{33}^T M_{33} + Q_{33} \; . \tag{8.4}$$

In this reduced order linear quadratic problem $x_3$ is treated as an input. The intuition for this is provided by the following observation:

**Lemma 8.1.** Assume $\Gamma \neq \varnothing$. Then the following statements are equivalent:

(i) $K \in \Gamma$,

(ii) $K = \begin{bmatrix} K_{11} & M_{12} & M_{13} \\ M_{12}^T & M_{22} & M_{23} \\ M_{13}^T & M_{23}^T & M_{33} \end{bmatrix}$, with $K_{11}$ a real symmetric matrix satisfying the linear matrix inequality

$$
(8.5) \qquad \begin{bmatrix} A_{11}^T K_{11} + K_{11} A_{11} + Q_r & K_{11} A_{13} + S_r^T \\ A_{13}^T K_{11} + S_r & R_r \end{bmatrix} \geq 0 \ .
$$

**Proof.** As already noted before, $K \in \Gamma$ if and only if $L(K) \geq 0$ and $KB + S^T = 0$. This implies

$$
\begin{bmatrix} L_{11}(K) & L_{13}(K) \\ L_{13}(K)^T & L_{33}(K) \end{bmatrix} \geq 0 \ .
$$

By writing out the expression $A^T K + KA + Q$ in terms of our decomposition and using the fact that $K$ can be written as

$$
K = \begin{bmatrix} K_{11} & M_{12} & M_{13} \\ M_{12}^T & M_{22} & M_{23} \\ M_{13}^T & M_{23}^T & M_{33} \end{bmatrix}
$$

(see Lemma 7.2), we obtain

$$
(8.6) \qquad L_{11}(K) = A_{11}^T K_{11} + K_{11} A_{11} + Q_r \ ,
$$

$$
(8.7) \qquad L_{13}(K) = K_{11} A_{13} + S_r^T \ ,
$$

$$
(8.8) \qquad L_{33}(K) = R_r \ .
$$

Thus we find that $K_{11}$ satisfies the linear matrix inequality (8.5). The converse implication can be proven analogously, using the assumption that $\Gamma \neq \varnothing$. $\square$

As noted in Section 4, the optimal cost $V_{\text{dist}}^+$ for our linear quadratic problem is represented by the matrix $K^+$, the largest element of $\Gamma$. Using the previous lemma it can be proven easily that, in fact,

$$
K^+ = \begin{bmatrix} K_{11}^+ & M_{12} & M_{13} \\ M_{12}^T & M_{22} & M_{23} \\ M_{13}^T & M_{23}^T & M_{33} \end{bmatrix} \ ,
$$

with $K_{11}^+$ the largest real symmetric solution to the reduced order linear matrix inequality (8.5). (Note that $(A_{11}, A_{13})$ is controllable if $(A, B)$ is controllable, so a *largest* solution to (8.5) indeed exists). Now, recall from (8.8) that $R_r = L_{33}(K)$ and hence by Lemma 7.1 that $R_r > 0$. Thus, the linear quadratic problem associated with the system $(A_{11}, A_{13})$ and quadratic form $\omega_r$ (with $x_3$ treated as an input) is *regular*. It is a basic result from [15] that in that case the largest solution of the linear matrix inequality coincides with the largest solution to the corresponding algebraic Riccati equation. Thus we obtain the following nice characterization of the optimal cost $V_{\text{dist}}^+(x_0)$ of our singular indefinite linear quadratic problem:

**Theorem 8.2.** Assume $(A, B)$ is controllable and $\Gamma \neq \emptyset$. Then

$$K^+ = \begin{bmatrix} K_{11}^+ & M_{12} & M_{13} \\ M_{12}^T & M_{22} & M_{23} \\ M_{13}^T & M_{23}^T & M_{33} \end{bmatrix} ,$$

where $K_{11}^+$ is the largest real symmetric solution of the algebraic Riccati equation

$$A_{11}^T K_{11} + K_{11} A_{11} + Q_r - (K_{11} A_{13} + S_r^T) R_r^{-1} (A_{13}^T K_{11} + S_r) = 0 \ . \tag{8.9}$$

□

We recall that the matrices $M_{ij}$ can be calculated *recursively* using the Algorithm 6.2. In order to calculate $K_{11}^+$ one first has to calculate $Q_r$, $S_r$ and $R_r$ and, subsequently, the largest real symmetric solution of (8.9).

## 9. CHARACTERIZATION OF OPTIMAL CONTROLS

Next, we turn attention to a characterization of optimal controls and optimal state trajectories. It will turn out that these can be characterized in terms of the optimal system $(A_{11}, A_{13})$ with quadratic form $\omega_r$. In the sequel, denote the associated cost functional by

$$J_r(x_{01}, x_3) := \int_0^\infty \omega_r(x_1(t), x_3(t)) \, dt \ , \tag{9.1}$$

where it is understood that $x_1$ and $x_3$ are related by

$$\dot{x}_1 = A_{11} x_1 + A_{13} x_3 , \; x_1(0) = x_{10} \ . \tag{9.2}$$

Recall that since $R_r > 0$ the class of admissible inputs for the latter problem consists only of regular distributions. The optimal cost is equal to $V_r^+(x_{10}) := x_{10}^T K_{11}^+ x_{10}$, where $K_{11}^+$ is the largest real symmetric solution of (8.9).

Now the idea that we want to elaborate is the following. We shall first determine the optimal "control" $x_3^*$ for the linear quadratic problem associated with (9.1), (9.2). Since this is a regular problem we know how to do this. Let $x_1^*$ be the corresponding state trajectory. Next, we determine an input $u$ for the *original* system (1.1) that "generates" these $x_1^*$ and $x_3^*$ (together with some $x_2$). Equivalently, we look for an admissible input $u$ such that the equation

$$\begin{bmatrix} \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} A_{21} \\ A_{31} \end{bmatrix} x_1 + \begin{bmatrix} B_2 \\ B_3 \end{bmatrix} u + \delta \begin{bmatrix} x_{02} \\ x_{03} \end{bmatrix} \tag{9.3}$$

is satisfied with $x_1 = x_1^*$ and $x_3 = x_3^*$ (for some $x_2$). Such input $u$ turns out to exists due to the fact that $\Sigma_1$ (see Lemma 7.1) has a right-inverse. Finally, we shall show that such $u$ is optimal for our original linear quadratic problem.

**Lemma 9.1.** Let $x_1$ and $x_3$ be arbitrary stable Bohl distributions. Then there exists an impulsive-regular distribution $u$ such that $u$, $x_1$ and $x_3$ satisfy (9.3), for some stable $x_2$. If, in addition, $x_1$ and $x_3$ satisfy (9.2) then any such distribution $u$ satisfies $u \in U_{\mathrm{adm}}(x_0)$.

**Proof.** Denote $C_0 := (0 \;\; I)$,

$$A_0 := \begin{bmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{bmatrix}, B_0 := \begin{bmatrix} B_2 \\ B_3 \end{bmatrix}, E_0 := \begin{bmatrix} A_{21} \\ A_{31} \end{bmatrix} .$$

According to Lemma 7.1 the system $(A_0, B_0, C_0)$ is strongly controllable. Hence $(A_0, B_0)$ is controllable. Consequently, there exists $F_1$ such that $\sigma(A_0 + B_0F_1) \subset \mathbb{C}^-$. Also, the system $(A_0 + B_0F_1, B_0, C_0)$ is strongly controllable. Now, put

$$u = F_1 \begin{bmatrix} x_2 \\ x_3 \end{bmatrix} + v .$$

Then (9.3) is equivalent to

$$\begin{bmatrix} \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = (A_0 + B_0F_1) \begin{bmatrix} x_2 \\ x_3 \end{bmatrix} + B_0 v + E_0 x_1 + \delta \begin{bmatrix} x_{02} \\ x_{03} \end{bmatrix} . \tag{9.4}$$

Define

$$G_1(s) := C_0(Is - A_0 - B_0F_1)^{-1}B_0 ,$$

$$H_1(s) := C_0(Is - A_0 - B_0F_1)^{-1} .$$

The transfer matrix $G_1$ has a polynomial right-inverse, say $P$. Consider the equation

$$x_3 = G_1(\delta^{(1)}) * v + H_1(\delta^{(1)}) * (E_0 x_1 + \delta \begin{bmatrix} x_{02} \\ x_{03} \end{bmatrix} ) . \tag{9.5}$$

Clearly, if $v$ satisfies (9.5) then $v$, $x_1$ and $x_3$ satisfy (9.4). Define

$$v := P(\delta^{(1)}) * [x_3 - H_1(\delta^{(1)}) * (E_0 x_1 + \delta \begin{bmatrix} x_{02} \\ x_{03} \end{bmatrix} )] .$$

Then $v$ indeed satisfies (9.5). Since $x_1$ and $x_3$ are stable Bohl distributions and $\sigma(A_0 + B_0F_1) \subset \mathbb{C}^-$ and since convolution with $P(\delta^{(1)})$ represent differenzations, we find that $v$ is impulsive-Bohl and stable. In turn, by (9.4), this implies that $x_2$ is impulsive Bohl and stable. Now define

$$u := v + F_1 \begin{bmatrix} x_2 \\ x_3 \end{bmatrix} \ .$$

Then $u, x_1, x_2$ and $x_3$ satisfy (9.3), with $x_2$ stable. Next we show that if $x_1$ and $x_3$ are stable Bohl distributions and if $u, x_1$ and $x_3$ satisfy the equations (9.2) and (9.3) with $x_2$ stable, then $u$ is admissible. This is however obvious since then $x(x_0, u) = (x_1^T, x_2^T, x_3^T)$ so $L(K)\, x(x_0, u)$ is regular (see Section 7). □

Now, the existence of optimal controls for our singular indefinite linear quadratic problem turns out to depend on the gap of the reduced order Riccati equation (8.9). Let $K_{11}^-$ be the smallest real symmetric solution of (8.9). Define $\Delta_r := K_{11}^+ - K_{11}^-$. Furthermore, define

(9.6) $$A_r := A_{11} - A_{13} R_r^{-1} (A_{13}^T K_{11}^+ + S_r) \ .$$

Recall from Theorem 5.1 that $\sigma(A_r) \subset \mathbb{C}^-$ if and only if $\Delta_r > 0$. We have the following result:

**Theorem 9.2.** Assume that $(A, B)$ is controllable and $\Gamma \neq \varnothing$. Then for all $x_0 \in \mathbb{R}^n$ there exists an optimal input $u$ if and only if $\Delta_r > 0$. Assume that the latter holds. Let $x_0 \in \mathbb{R}^n$ and let $u \in U_{\mathrm{adm}}(x_0)$ with corresponding state trajectory $x = (x_1^T, x_2^T, x_3^T)^T$. Then $u$ is optimal if and only if $x_1$ and $x_3$ are equal to the Bohl distributions given by

(9.7) $$x_3(t) = -R_r^{-1} (A_{13}^T K_{11}^+ + S_r)\, e^{A_r t} x_{10} \quad (t \geq 0) \ ,$$

(9.8) $$x_1(t) = e^{A_r t} x_{10} \quad (t \geq 0) \ ,$$

and $u, x_1, x_2$ and $x_3$ satisfy (9.3) with $x_2$ stable.

**Proof.** If $\Delta_r > 0$ then $x_3$ defined by (9.7) is the unique optimal input for the linear quadratic problem (9.1), (9.2). Furthermore, $x_1$ and $x_3$ are stable Bohl distributions that satisfy (9.2). Let $u \in U_{\mathrm{adm}}(x_0)$ be such that $x_1, x_3$ and $u$ satisfy (9.3) with $x_2$ stable. We claim that any such $u$ is optimal. To prove this, it suffices to show that $J_{\mathrm{dist}}(x_0, u) = x_0^T K^+ x_0$. Now, by definition

(9.9) $$J_{\mathrm{dist}}(x_0, u) = \int_0^\infty \| C_{K^+} x \|^2 \, dt + x_0^T K^+ x_0 \ ,$$

where $C_{K^+}$ is such that $C_{K^+}^T C_{K^+} = L(K^+)$. Obviously,

(9.10) $$\begin{aligned} &\| C_{K^+} x \|^2 \\ &= x_1^T L_{11}(K^+) x_1 + 2 x_3^T L_{13}(K^+) x_1 + x_3^T L_{33}(K^+) x_3 \\ &= x_1^T (A_{11}^T K_{11}^+ + K_{11}^+ A_{11} + Q_r) x_1 + 2 x_3^T (K_{11}^+ A_{13} + S_r^T) x_1 \end{aligned}$$

$$+ x_3^T R_r x_3$$

$$= \frac{d}{dt}\,(x_1^T K_{11}^+ x_1) + \omega_r(x_1, x_3) \ .$$

By integrating this, using the facts that $\lim_{t \to \infty} x_1(t) = 0$ and that

$$\int_0^\infty \omega_r(x_1, x_3)\, dt = V_r^+(x_{10}) = x_{10}^T K_{11}^+ x_{10} \ ,$$

we find that

$$\int_0^\infty \|C_{K^+} x\|^2 \, dt = 0 \ .$$

Conversely, assume that $u \in U_{\mathrm{adm}}(x_0)$ is optimal. We first claim that $x_1$ and $x_3$ are regular. Indeed, since $L(K^+)\,x$ is regular, $L_{13}(K^+)^T x_1 + L_{33}(K^+) x_3$ is regular. Hence $x_3 = -L_{33}(K^+)^{-1} L_{13}(K^+)^T x_1$. Since also $\dot{x}_1 = A_{11} x_1 + A_{13} x_3 + \delta x_{10}$, this proves our claim. Now, since $u$ is optimal, we have $J_{\mathrm{dist}}(x_0, u) = x_0^T K^+ x_0$. Using (9.9) and (9.10) this yields

$$\int_0^\infty \|C_{K^+}\ x\|^2 \, dt = 0$$

and

$$\int_0^\infty \frac{d}{dt}\,(x_1^T K_{11}^+ x_1)\, dt + J_r(x_{01}, x_3) = 0 \ ,$$

respectively. Now, since $u$ is admissible we have $\lim_{t \to \infty} x_1(t) = 0$ and hence we find

$$J_r(x_{01}, x_3) = x_{01}^T K_{11}^+ x_{01} = V_r^+(x_{01})$$

so $x_3$ is optimal for the linear quadratic problem associated with (9.1) and (9.2). This however implies that $x_3$ is given by (9.7). Obviously, $u$, $x_1, x_2$ and $x_3$ also satisfy (9.3) with $x_2$ stable. Finally, if an optimal $u$ exists for all $x_0$ then by the above also an optimal $x_3$ exists for all $x_{01}$ (for the linear quadratic problem associated with (9.1) and (9.2)). Hence $\Delta_r > 0$. □

If follows from the above theorem that if $x_0 \in \mathbb{R}^n$ and if $u$ is an optimal input, then the corresponding optimal state trajectory $x = (x_1^T, x_2^T, x_3^T)^T$ looks as follows: $x_3$ is regular and equals the function that takes the value $x_{03}$ in $t = 0$ and the values $x_3(t)$ given by (9.7) for $t > 0$. Note that $x_3$ makes a jump from $x_{30}$ to $-R_r^{-1}(A_{13} K_{11}^+ + S_r) x_{10}$. the component $x_1$ is smooth. It corresponds to the function given by (9.8) for $t \geq 0$. Finally, $x_2$ is an impulsive-regular distribution that in general has a non-zero impulsive part. We note that every optimal input $u$ yields the same $x_1$ and $x_3$. However, the component $x_2$ in general depends on the particular choice of $u$. The regular part of *any* optimal trajectory $x$ however for $t > 0$ corresponds to a movement on the linear subspace given by the equation

$$x_3 = -R_r^{-1}(A_{13}K_{11}^{+} + S_r)x_1 \ .$$

An obvious question that now arises is: when do we have uniqueness of optimal controls? It turns out that a condition for this can be formulated in terms of the real rational matrix

$$H(-s, s) := S(Is - A)^{-1}B + B^T(-Is - A^T)^{-1}S^T$$
$$+ B^T(-Is - A^T)^{-1}Q(Is - A)^{-1}B \ .$$

This matrix also appears in [15]. In fact, it was shown there that the *frequency domain inequality*: $H(-i\omega, i\omega) \geq 0$ for all $\omega \in I\!R$ such that $i\omega$ is not a pole of $H(-s, s)$, is equivalent to: $\Gamma \neq \varnothing$. Without proof we state the following:

**Theorem 9.3.** Assume $(A, B)$ is controllable, $\Gamma \neq \varnothing$ and $\Delta_r > 0$. Then for all $x_0 \in I\!R^n$ there exists exactly one optimal input $u^*$ if and only if $H(-s, s)$ is an invertible real rational matrix. $\square$

## 10. CONCLUSIONS

In this paper we have studied a general version of the infinite horizon linear quadratic optimal control problem. We have proposed a mathematical formulation of the problem in case that the cost criterion is given by the integral of an *indefinite* quadratic form, while *at the same time* the weighting matrix $R$ of the control input is *not necessarily invertible*. In our subsequent treatment of this optimization problem we have restricted ourselves to the totally singular case, i.e., the case that $R = 0$. Under this assumption we have found a characterization of the optimal cost or, equivalently, of the largest real symmetric solution of the corresponding linear matrix inequality. It was shown that this solution can be found by means of the recursive algorithm (6.2) followed by the calculation of the largest real symmetric solution of the reduced order algebraic Riccati equation (8.9). In Section 9 we have given a characterization of all optimal controls. We have also given necessary and sufficient conditions on $A$, $B$ and $\omega$ such that for every initial condition there is exactly one optimal control input.

Of course, the results of this paper are incomplete in the sense that we did not treat the most general case "$R \geq 0$" but only the case "$R = 0$" (the classical case "$R > 0$" is treated in [15]). In the case that we only have $R \geq 0$ instead of $R = 0$ it is indeed also possible to develop an analysis based on a strongly controllable subspace $T^*$ associated with $(A, B, \omega)$. At this moment however it seems difficult to develop an algorithm like 6.2 to calculate this subspace $T^*$, the associated $W^*$ and the fixed mapping $K \mid T^*$ in this most general case. The latter is left for future research.

**References**


[1] B.D.O. Anderson & J.B. Moore, *Linear Optimal Control*, Prentice Hall, Englewood Cliffs, 1971.

[2] B.D.O. Anderson & S. Vongpanitlerd, *Network Analysis and Synthesis*, Prentice Hall, Englewood Cliffs, 1973.

[3] D.J. Clements & B.D.O. Anderson, *Singular Optimal Control: The Linear Quadratic Problem*, Springer Lecture Notes in Contr. & Inf. Sciences, No. 5, Springer, Berlin, 1978.

[4] A.H.W. Geerts, "All optimal controls for the singular linear quadratic problem without stability; a new interpretation of the optimal cost", to appear in *Journal of Lin. Alg. & Appl.*, 116, 1989.

[5] M.L.J. Hautus & L.M. Silverman, "System structure and singular control", *Journal of Lin. Alg. & Appl.*, 50, pp. 369-402, 1983.

[6] O. Hijab, *Stabilizaton of Control Systems*, Springer, New York, 1987.

[7] H. Kwakernaak & R. Sivan, *Linear Optimal Control Systems*, Wiley-Interscience, New York, 1972.

[8] P.J. Moylan & J.B. Moore, "Generalizations of singular optimal control theory", *Automatica*, 7, pp. 591-598, 1971.

[9] J.M. Schumacher, "On the structure of strongly controllable systems", *Int. Journal of Contr.*, Vol. 38, pp. 525-545, 1983.

[10] J.M. Schumacher, "A geometric approach to the singular filtering problem", *IEEE Trans. Aut. Contr.*, Vol. AC-30, No. 11, pp. 1075-1082, 1985.

[11] L. Schwartz, *Theorie des Distributions*, Hermann, Paris, 1978.

[12] L.M. Silverman, "Discrete Riccati equations, alternative algorithms, asymptotic properties, and system theory interpretations, in *Control and Dynamic Systems*, Vol. 12, pp. 381-385, 1970.

[13] J.M. Soethoudt & H.L. Trentelman, "The regular indefinite linear-quadratic problem with linear endpoint constraints", *Systems & Control Letters*, 12, pp. 23-31, 1989.

[14] H.L. Trentelman, "The regular free-endpoint linear quadratic problem with indefinite cost", *SIAM Journal Contr. & Opt.*, Vol. 27, No. 1, pp. 27-42, 1989.

[15] J.C. Willems, "Least squares stationary optimal control and the algebraic Riccati equation", *IEEE Trans. Aut. Contr.*, Vol. AC-16, No. 6, 1971.

[16] J.C. Willems, A. Kitapçi & L.M. Silverman, "Singular optimal control, a geometric approach", *SIAM Journal Contr. & Opt.*, Vol. 24, pp. 323-337, 1986.

[17] W.M. Wonham, *Linear Multivariable Control: a Geometric Approach*, Springer Verlag, New York, 1979.