SOME CONSIDERATIONS ON DEGENERATE CONTROL SYSTEMS: THE LQR PROBLEM

B. THALLER AND S. THALLER *

Institut für Mathematik, Universität Graz, Heinrichstrasse 36, A-8010 Graz, Austria

Abstract. We investigate linear inhomogeneous Cauchy problems in a Hilbert space which are degenerate in the sense that there is a non-invertible operator M at the time derivative. Under certain conditions the degenerate system is replaced with an "effective" non-degenerate system in the subspace given by the orthogonal complement of the kernel of M. The connection between the solutions of the nondegenerate system and the solutions of the degenerate system is in general given by an unbounded and sometimes not closable linear operator. The previously developed theory is extended and criteria are obtained when the solutions can be described in terms of a C_0-semigroup. We describe some examples and introduce the concept of a mild solution in order to deal with the LQR-problem in degenerate control theory.

* Supported by SFB F003 "Optimierung und Kontrolle"

1. Introduction

In this paper we consider the degenerate abstract Cauchy problem

$$\frac{d}{dt} M\, z(t) = A\, z(t) + f(t), \qquad z(0) = z_0, \tag{1.1}$$

where M and A are linear operators, which are densely defined in a Hilbert space $\mathfrak{H}$ and map into a Hilbert space $\mathfrak{K}$. We are interested in the case where the operator M is not invertible. Such problems occur, e.g., for systems of differential equations if M is a matrix with nontrivial kernel. A very simple and canonical example is provided by the Dirac equation which in one dimension and after a similarity transformation can be written as

$$-\mathrm{i}\frac{d}{dt}\begin{pmatrix} 1 & 0 \\ 0 & 1/c^2 \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} V(x) & -\mathrm{i}d/dx \\ -\mathrm{i}d/dx & -2m - V(x)/c^2 \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}. \tag{1.2}$$

We see that in the nonrelativistic limit $c \to \infty$ the Dirac equation becomes a degenerate Cauchy problem. Other examples are obtained if A is some differential operator and M is a multiplication operator with a function that vanishes in a certain region of the configuration space. Consider, e.g., the one-dimensional heat equation in an inhomogeneous media with heat capacity vanishing in some interval.

Of course one wants to be able to factor out the kernel of the operator M. The main problem is that the operator A will not simply map $(\operatorname{Ker} M)^\perp$ into $\operatorname{Ran} M$. Under certain conditions to be discussed in Section 2 it is nevertheless possible to replace A by some "effective operator" $A_0 : (\operatorname{Ker} M)^\perp \to \operatorname{Ran} M$. This replacement can be done in such a way that the non-degenerate Cauchy problem

$$\frac{d}{dt} M^\perp\, x(t) = A_0\, x(t) + f(t), \qquad x(0) = x_0, \tag{1.3}$$

is in some sense equivalent to Eq. (1.1). Here $M^\perp$ denotes the part of M in $(\operatorname{Ker} M)^\perp$, and it is assumed that $f(t) \in \operatorname{Ran} M$ for all t (in [17] more general functions f are treated).

In case of the Dirac equation in the nonrelativistic limit, the effective operator A_0 is easily seen to be the Schrödinger operator

$$A_0 = -\frac{1}{2m}\frac{d^2}{dx^2} + V(x). \tag{1.4}$$

In case of the heat equation with vanishing heat-capacity in an interval I, A_0 is the Laplace operator in $L^2(\mathbb{R} \setminus I)$ defined on a domain with suitable self-adjoint boundary conditions (see [17], Example 2.1).

Degenerate problems of this type have been considered in [17], where we gave criteria for existence and uniqueness of strict solutions to (1.1). Some of the results in that paper will be extended and generalized here. We give criteria for the existence of the factorization $A_0 = AZ_A$ with a suitable linear operator Z_A and investigate also the case that Z_A is not closable (Sections 2 and 3). We show under which conditions the solution of (1.1) is described by a C_0-semigroup acting on a closed subspace of $\mathfrak{H}$ which consists of suitable initial values z_0 (Theorem 2.2).

Furthermore we define and discuss the concept of mild solutions of (1.1) within our framework (Section 4) and investigate the relation with the corresponding definitions for non-degenerate systems. As an example, where the concept of a mild solution is needed we discuss the linear quadratic regulator problem in the framework of control theory.

Degenerate Cauchy problems frequently occur in the applications of control theory. We refer the reader to the book of L. Dai [4], which, however, only treats the finite dimensional case. Control theory for degenerate systems has also been treated in [5], and in [17]. In control theory one considers degenerate systems of the form

$$\begin{aligned} \frac{d}{dt}Mz(t) &= Az(t) + Bu(t) \\ v(t) &= Cz(t) \\ z(0) &= z_0 \end{aligned} \tag{1.5}$$

where B is a bounded linear operator from a Banach space U called the control space into $\operatorname{Ran} M$, and C is a bounded linear operator mapping $\mathfrak{H}$ into a Banach space V, the output space. In [17] we defined the basic notions of degenerate control systems like controllability, observability, detectability, and stabilizability. Furthermore we investigated the relations with the dual system and the corresponding nondegenerate system and its dual. Now it is our goal to investigate the linear quadratic regulator problem (LQR-problem) with infinite time horizon for the system (1.5). We have to find a control function $\hat{u}$ which minimizes the quadratic cost criterion

$$J_\infty(z_0, u) = \int_0^\infty (\|Cz(s)\|^2 + \|u(s)\|^2)ds \tag{1.6}$$

over all controls u which are square integrable with respect to t. Since the strict solutions are in general only defined if $Bu(t)$ is differentiable with respect to t, the concept of mild solution is needed here. Under some simplifying conditions we derive a Riccati equation, a feedback equation for the optimal control, and a degenerate closed loop equation for the optimal state in analogy to the corresponding results for nondegenerate systems (cf. Section 5).

Let us finally make some remarks concerning the relation of our approach with other investigations in the literature on degenerate Cauchy problems. Our approach differs from other treatments in several respects.

In the book of Carroll and Showalter [2] and in the articles [14], [15], [16] the problem (1.1) is formulated in a quite general vector space. The operator M is assumed to define a nonnegative sesquilinear form which is then used to define the Hilbert space structure of the problem. Once the metric has been defined with the help of M, this operator simply becomes the Riesz automorphism in the Hilbert space of the system. In this setting there is no topological structure to describe the behaviour of the part of the solution z in the subspace $\operatorname{Ker} M$. In contrast to this investigation, we describe the degenerate system in a given state space (which we assume to be a Hilbert space). We focus our interest from the very beginning on the relation between the nondegenerate problem (1.3) and the degenerate problem (1.1). This relation can be very nontrivial. Although there is a linear operator Z_A mapping strict solutions x of (1.3) to solutions z of (1.1), this operator turns out to be unbounded and even not closable in many cases of interest (see, e.g., the examples in Section 3). This of course makes the definition of mild solutions of (1.1) a delicate problem, cf. Section 4.

Recently, degenerate systems have been considered by Favini and Plazzi [6], [7], [8], and Favini and Yagi [9], [10]. In the linear case they investigate, among other things, the maximum regularity property of solutions in a situation where the corresponding nondegenerate problem becomes a parabolic equation. These authors, like Showalter in [14], also consider the nonlinear and the nonautonomous case, or even more general situations [9], but always under some sort of parabolicity assumptions. In our setting the operators M and A are linear, but it is not necessary to assume that they are self-adjoint or semi-definite, or that the nondegenerate system is solved by an analytic semigroup (although this assumption is very convenient for some purposes).

Let us finally fix some notation which will be used below.

Notation . Let T be a linear operator. By $\mathfrak{D}(T)$, $\operatorname{Ker} T$, and $\operatorname{Ran} T$ we denote its domain, kernel, and range, respectively. The adjoint operator will be written as T^*, the restriction of T to a set $\mathfrak{D}$ smaller than the domain of T will be denoted by $T \restriction \mathfrak{D}$, and the closure of T will be denoted by T^c.

2. Basic Concepts

In this section we introduce some basic concepts for the degenerate abstract Cauchy problem Eq. (1.1) and extend the treatment given in [17]. The operators M and A are linear, densely defined in a Hilbert space $\mathfrak{H}$, and map into a Hilbert space $\mathfrak{K}$. The operator M need not be invertible, but for convenience we assume that M is bounded, $\mathfrak{D}(M) = \mathfrak{H}$, and $\operatorname{Ran} M$ closed in $\mathfrak{K}$. Let P denote the orthogonal projection onto $\operatorname{Ker} M$, Q the projection onto $\operatorname{Ker} M^*$ and let $P^\perp = \mathbf{1} - P$, $Q^\perp = \mathbf{1} - Q$. The closedness of $\operatorname{Ran} M = Q^\perp \mathfrak{K}$ implies that the operator $M^\perp = M \restriction P^\perp \mathfrak{H}$ has a bounded inverse.

Definition 2.1. A strict solution of the degenerate Cauchy problem Eq. (1.1) is a continuous function $z : [0, \infty) \longrightarrow \mathfrak{H}$ such that $z(t) \in \mathfrak{D}(A) \cap \mathfrak{D}(M)$ for all $t \geq 0$, Mz is continuously differentiable, and Eq. (1.1) holds.

We define a linear subspace $\mathfrak{D}_A$ of $\mathfrak{H}$ by

$$\mathfrak{D}_A := \{z \in \mathfrak{D}(A) \mid QAz = 0\}. \tag{2.1}$$

If f has values in $\operatorname{Ran} M = Q^\perp \mathfrak{K}$, any strict solution z of Eq. (1.1) must satisfy $z(t) \in \mathfrak{D}_A$ for all $t \geq 0$.

In order to be able to reduce the problem to an ordinary Cauchy problem, we make the following crucial assumption.

Assumption 2.1. *For every $x \in P^\perp \mathfrak{D}_A$ the set $(P^\perp)^{-1}\{x\} \cap \mathfrak{D}_A$ consists of precisely one element $z \in \mathfrak{D}_A$.*

This element z clearly depends linearly on x. We write $z = Z_A x$. The linear operator Z_A is defined on $P^\perp \mathfrak{D}_A$ and will occasionally be written as

$$Z_A = P^\perp + T_A \qquad \text{on } P^\perp \mathfrak{D}_A. \tag{2.2}$$

Notice that $z = Pz + P^\perp z \in \mathfrak{D}_A$ implies $Pz = T_A P^\perp z$, i.e., for $z \in \mathfrak{D}_A$ the part of z in $(\operatorname{Ker} M)^\perp$ uniquely determines the part of z in $\operatorname{Ker} M$. Hence the projection $P^\perp$ is injective on $\mathfrak{D}_A$, because if $P^\perp z = 0$ for some $z \in \mathfrak{D}_A$, then also $z = Z_A P^\perp z = 0$. The operator Z_A on $P^\perp \mathfrak{D}_A$ is the inverse of the projection $P^\perp$ on $\mathfrak{D}_A$, i.e.,

$$Z_A P^\perp = \mathbf{1} \text{ on } \mathfrak{D}_A, \qquad P^\perp Z_A = \mathbf{1} \text{ on } P^\perp \mathfrak{D}_A. \tag{2.3}$$

Next we define

$$A_0 = A Z_A, \quad \text{on } \mathfrak{D}(A_0) = P^\perp \mathfrak{D}_A, \tag{2.4}$$

and find immediately that

$$Az = A_0 x, \quad \text{for all } z \in \mathfrak{D}_A \text{ and } x = P^\perp z. \tag{2.5}$$

We will assume that $\mathfrak{D}(A_0) = P^\perp \mathfrak{D}_A$ is dense in $P^\perp \mathfrak{H}$.

Assumption 2.2. *There is a real constant ω such that for all λ with $Re\,\lambda > \omega$ the operator $(A - \lambda M) \restriction \mathfrak{D}_A$ has a bounded inverse which is defined on all of $Ran\,M$. Moreover, there exists a constant $0 < K \leq 1$ such that*

$$\|P^\perp (A - \lambda M)^{-1} M\| \leq \frac{K}{Re\,\lambda - \omega} \tag{2.6}$$

for all λ with $Re\,\lambda > \omega$.

Remark 2.1. If $(A - \lambda M) \restriction \mathfrak{D}_A$ has an inverse which is bounded on a closed domain, then $A - \lambda M$ is closed on $\mathfrak{D}_A$. As in [17], Lemma 2.3, one can show that $(A - \lambda M) Z_A = A_0 - \lambda M^\perp$ is closed, which implies that A_0 itself is closed on $P^\perp \mathfrak{D}_A$. Since $M^\perp$ is bounded and boundedly invertible, also the operators $A_1 = A_0 (M^\perp)^{-1}$ and $A_2 = (M^\perp)^{-1} A_0$ are closed on their natural domains $\mathfrak{D}(A_1) = M \mathfrak{D}_A$ and $\mathfrak{D}(A_2) = P^\perp \mathfrak{D}_A$. A little calculation shows, e.g., the relation

$$(A_2 - \lambda)^{-1} = P^\perp (A - \lambda M)^{-1} M \quad \text{on } P^\perp \mathfrak{H}. \tag{2.7}$$

Hence Eq. (2.6) implies with the help of the Hille-Yoshida theorem that A_2 is the generator of a C_0-semigroup on $P^\perp\mathfrak{H}$. By our assumptions on M, A_2 generates a semigroup if and only if A_1 does. We have the relations

$$A_2 = (M^\perp)^{-1} A_1 M^\perp$$

and

$$e^{A_2 t} = (M^\perp)^{-1} e^{A_1 t} M^\perp.$$

The following theorem can be proved as in [17]:

Theorem 2.1. *Let Assumptions 2.1 and 2.2 be fulfilled and let f be a continuously differentiable function with values in* $\operatorname{Ran} M$*.*

Then the degenerate inhomogeneous equation (1.1)

$$\frac{d}{dt} M\, z(t) = A\, z(t) + f(t), \qquad z(0) = z_0,$$

has a unique strict solution z for each initial value $z_0 \in \mathfrak{D}_A$. It is given by

$$z(t) = Z_A (M^\perp)^{-1} y(t) = Z_A x(t),$$

where

$$y(t) = e^{A_1 t} M z_0 + \int_0^t e^{A_1 (t-s)} f(s)\, ds, \quad \textit{resp.,}$$

$$x(t) = e^{A_2 t} P^\perp z_0 + \int_0^t e^{A_2 (t-s)} \, (M^\perp)^{-1} f(s)\, ds.$$

The idea of the proof is that the degenerate Cauchy problem (1.1) can be reduced to the equation

$$\frac{d}{dt} M^\perp x(t) = A_0 x(t) + f(t). \tag{2.8}$$

Setting $y(t) = M^\perp x(t)$, equation (2.8) becomes

$$\frac{d}{dt} y(t) = A_1 y(t) + f(t) \tag{2.9}$$

or, applying $(M^\perp)^{-1}$ on (2.8), we get

$$\frac{d}{dt} x(t) = A_2 x(t) + (M^\perp)^{-1} f(t). \tag{2.10}$$

If A_2 generates a C_0-semigroup, Eq. (2.10) has the solution

$$x(t) = e^{A_2 t} P^\perp z_0 + \int_0^t e^{A_2 (t-s)} (M^\perp)^{-1} f(s) ds$$

which for $z_0 \in \mathfrak{D}_A$ is in $\mathfrak{D}(A_2) = P^\perp \mathfrak{D}_A$, because f is continuously differentiable. Hence we can apply the operator Z_A on $x(t)$ to get $z(t)$. Notice that while $x(t)$ can be defined for all z_0, the expression $Z_A x(t)$ might be meaningless if the initial value z_0 is not in $\mathfrak{D}_A$.

Continuity of z follows easily from the continuity of x and the fact that $(A - \lambda M) \restriction \mathfrak{D}_A$ has a bounded inverse for a suitable λ.

By Remark 2.1, an equivalent method of solving Eq. (2.8) would be via (2.9) (cf. [17], Remark 2.11).

We will frequently assume that Z_A is closable (or equivalently, that T_A is closable). In the next section we present an example and a counter-example to this assumption.

Lemma 2.1. *The operator Z_A is closable if and only if $Ker\, M \cap \mathfrak{D}_A^c = \{0\}$. Here $\mathfrak{D}_A^c$ denotes the closure of $\mathfrak{D}_A$ with respect to the norm in $\mathfrak{H}$.*

Proof. Since $P^\perp z = 0$ if and only if $z \in \operatorname{Ker} M$, the statement $\operatorname{Ker} M \cap \mathfrak{D}_A^c = \{0\}$ is equivalent to $P^\perp$ injective as an operator from $\mathfrak{D}_A^c$ into $P^\perp \mathfrak{H}$. The bounded operator $P^\perp$ is of course closed on the closed domain $\mathfrak{D}_A^c$. Hence if $P^\perp$ is also injective, then its inverse is a closed extension of Z_A (which is the inverse of $P^\perp \restriction \mathfrak{D}_A$).

Conversly, let Z_A be closable. Let $z = \lim z_n \in \mathfrak{D}_A^c$, $z_n \in \mathfrak{D}_A$, and let $P^\perp z = 0$. Clearly, $P^\perp z = \lim x_n$, where $Z_A x_n = z_n$. But if $\lim x_n = 0$ and $Z_A x_n$ is convergent, then the closability of Z_A implies that $z = \lim Z_A x_n = 0$. Hence $P^\perp \restriction \mathfrak{D}_A^c$ is injective. □

If Z_A is closable, the domain of the closure Z_A^c will be denoted by $\mathfrak{H}_0$. Since $P^\perp \mathfrak{D}_A$ is assumed to be dense, also $\mathfrak{H}_0$ is a dense subspace of $P^\perp \mathfrak{H}$.

Lemma 2.2. *Let Z_A be closable. Then $\mathfrak{H}_0$ equipped with the graph norm*

$$\|x\|_G^2 = \|x\|^2 + \|T_A^c x\|^2 \tag{2.11}$$

is a Hilbert space which is isometrically isomorphic to the closure of $\mathfrak{D}_A$ with respect to the norm in $\mathfrak{H}$. The isomorphism from $\mathfrak{H}_0$ onto $\mathfrak{D}_A^c$ is given by Z_A, its inverse by $P^\perp$.

Proof. Each z can be uniquely decomposed into the orthogonal parts $x = P^\perp z$ and $T_A x$. Obviously, (z_n) is a Cauchy sequence in $\mathfrak{D}_A$, if and only if the elements $x_n = P^\perp z_n$ form a Cauchy sequence with respect to the graph norm in $P^\perp \mathfrak{D}_A$. Moreover, $\|Z_A^c x\|^2 = \|x\|^2 + \|T_A x\|^2 = \|x\|_G^2$ holds for all $x \in \mathfrak{H}_0 = P^\perp \mathfrak{D}_A^c$ which shows that Z_A is unitary as an operator from $(\mathfrak{H}_0, \|\cdot\|_G^2)$ onto $(\mathfrak{D}_A^c, \|\cdot\|)$. □

Theorem 2.2. *If $\mathfrak{H}_0$ is an invariant set for the semigroup* $\exp(A_2 t)$*, i.e.,*

$$e^{A_2 t} \mathfrak{H}_0 \subset \mathfrak{H}_0 \tag{2.12}$$

then the operators

$$T(t) = Z_A^c e^{A_2 t} P^\perp \tag{2.13}$$

form a C_0-semigroup on the Hilbert space $\mathfrak{D}_A^c$.

Proof. The first step is to show that $\exp(A_2 t)$ is bounded with respect to the graph norm in $\mathfrak{H}_0$. Let $x_n \in \mathfrak{H}_0$ be a Cauchy sequence which converges in the graph norm to x_0, i.e., $\|x_n - x_0\|_G \to 0$ as $n \to \infty$. Assume that also $\exp(A_2 t)x_n$ converges with respect to $\|\cdot\|_G$. If we call this limit y_0, then in particular $\lim \|\exp(A_2 t)x_n - y_0\| = 0$ and hence $y_0 = \exp(A_2 t)x_0$. Since $\mathfrak{H}_0$ is a Hilbert space with respect to the graph norm, x_0 and y_0 belong to $\mathfrak{H}_0$. Hence $\exp(A_2 t)$ is closed as a mapping from $\mathfrak{H}_0$ to $\mathfrak{H}_0$ and therefore bounded with respect to $\|\cdot\|_G$. By Lemma 2.2 the operators $T(t) = Z_A^c e^{A_2 t} P^\perp$ form a semigroup of bounded operators. For initial values z_0 in $\mathfrak{D}_A$ the solution $z(t) = T(t)z_0$ is continuous. Since $\mathfrak{D}_A$ is dense in $\mathfrak{D}_A^c$, the strong continuity of $T(t)$ follows by an approximation argument. $\square$

Remark 2.2. Under the conditions of the previous theorem the homogeneous degenerate equation (1.1) is completely equivalent to an ordinary Cauchy problem on the Hilbert space $(\mathfrak{H}_0, \|\cdot\|_G)$. The situation is more complicated for the inhomogeneous problem, because f takes values in a set larger than $\mathfrak{H}_0$.

Remark 2.3. The operator $Z_A = P^\perp + T_A$ can be defined more explicitly under the following conditions. Assume that $P\mathfrak{D}_A \subset \mathfrak{D}(A)$ and that the operator $QAP \restriction P\mathfrak{D}_A$ is invertible. Then (cf. [17]) $z \in \mathfrak{D}_A$ if and only if $z \in \mathfrak{D}(A)$ and $QAPz = QAP^\perp z$. Hence we define

$$T_A = -(QAP)^{-1}QA \quad \text{on } P^\perp\mathfrak{D}_A. \tag{2.14}$$

3. Two examples

3.1. **Example.** Let $\mathfrak{H} = \mathfrak{K} = L^2(\mathbb{R})$. Let M be the operator of multiplication by a characteristic function,

$$(M\psi)(x) = \begin{cases} 0 & \text{if } x \in [-1,1], \\ \psi(x) & \text{if } x \notin [-1,1], \end{cases} \qquad \text{for all } \psi \in L^2(\mathbb{R}). \tag{3.1}$$

Hence $P^\perp = Q^\perp = M$, and $P = Q$ is multiplication by the characteristic function of the interval $[-1,1]$.

Given some real-valued function $\phi \in L^2(\mathbb{R})$, we define

$$\mathfrak{D}(A) = \{\psi \in L^2(\mathbb{R}) \mid \psi \text{ absolutely continuous}, \ \psi' \in L^2(\mathbb{R})\}, \tag{3.2}$$

$$(A\psi)(x) = -\psi'(x) - \mathrm{i}\,(\phi,\psi)\,\phi(x) \quad \text{for all } \psi \in \mathfrak{D}(A). \tag{3.3}$$

Here (ϕ,ψ) denotes the L^2-scalar product. Note that $\mathrm{i}A$ is self-adjoint.

For later use we define

$$\Phi(x) = \int_{-1}^{x} \phi(y)\,dy, \qquad \alpha = 1 + \mathrm{i}\Phi(1)^2/2, \tag{3.4}$$

and note the relation

$$(P\phi, \Phi) = \tfrac{1}{2}\Phi(1)^2 = \mathrm{i}(1-\alpha). \tag{3.5}$$

The states in the subspace $\mathfrak{D}_A$ are characterized by $(A\psi)(x) = 0$ for $x \in [-1, 1]$, hence

$$\psi(x) = \psi(-1) - \mathrm{i}(\phi, \psi)\Phi(x), \quad \text{for all } x \in [-1, 1]. \tag{3.6}$$

Hence we can evaluate

$$(P\phi, \psi) = \int_{-1}^{+1} \phi(x)\psi(x)\,dx = \Phi(1)\psi(-1) + (1-\alpha)(\phi, \psi) \tag{3.7}$$

$$\alpha(P\phi, \psi) = \Phi(1)\psi(-1) + (1-\alpha)(P^{\perp}\phi, \psi). \tag{3.8}$$

Inserting into Eq. (3.6) we find for all $\psi \in \mathfrak{D}_A$

$$\alpha\psi(x) = \alpha\psi(-1) - \mathrm{i}[\Phi(1)\psi(-1) + (P^{\perp}\phi, \psi)]\Phi(x), \quad x \in [-1, 1]. \tag{3.9}$$

Since $\psi \in \mathfrak{D}_A$ is absolutely continuous, the boundary value $\psi(-1)$ is given by $\lim_{x\to-1-0} P^{\perp}\psi(x)$, and $(P^{\perp}\phi, \psi) = (P^{\perp}\phi, P^{\perp}\psi)$. Hence we see that the part $P^{\perp}\psi$ of $\psi \in \mathfrak{D}_A$ completely and uniquely determines $P\psi$, i.e., the part of ψ within the interval $[-1, 1]$. In particular, any $\psi \in \mathfrak{D}_A$ satisfies

$$\alpha\psi(+1) - \overline{\alpha}\psi(-1) = -\mathrm{i}\Phi(1)(P^{\perp}\phi, P^{\perp}\psi) \tag{3.10}$$

Hence $P^{\perp}\mathfrak{D}_A \subset L^2(\mathbb{R} \setminus [-1, 1])$ consists of functions which are absolutely continuous on $(-\infty, -1] \cup [1, \infty)$, satisfy the "nonlocal" boundary condition (3.10), and have a square integrable derivative. Any function $\psi^{\perp} \in P^{\perp}\mathfrak{D}_A$ defines a unique function $\psi^0 = T_A\psi^{\perp}$ (given by (3.9)) such that $\psi = \psi^0 + \psi^{\perp}$ is in $\mathfrak{D}_A$.

Following the theory of Section 1 we define the operator $A_0 = A_1 = A_2$ on $P^{\perp}\mathfrak{D}_A$ by $AZ_A = A(1 + T_A)$, which gives

$$A_0\psi^{\perp} = -(\psi^{\perp})' - (\mathrm{i}/\alpha)\,[\Phi(1)\psi^{\perp}(-1) + (P^{\perp}\phi, \psi^{\perp})]\,P^{\perp}\phi. \tag{3.11}$$

for all $\psi^{\perp} \in P^{\perp}\mathfrak{D}_A$.

Proposition 3.1. *The operator iA_0 is self-adjoint on $\mathfrak{D}(A_0) = P^{\perp}\mathfrak{D}_A$.*

Proof. Define $N = \mathbb{R} \setminus [-1, 1]$. Consider the maximal operator $\tilde{A}_0$ on the domain $\mathfrak{D}(\tilde{A}_0)$ of absolutely continuous functions f in $L^2(N)$ with finite boundary values $f(\pm 1)$, whose derivative f' is again in $L^2(N)$. Clearly, $\tilde{A}_0$ is an extension of A_0. A partial integration shows for all $g \in \mathfrak{D}(A_0)$ and all $f \in \mathfrak{D}(\tilde{A}_0)$

$$(f, A_0 g) + (\tilde{A}_0 f, g) = \frac{1}{\overline{\alpha}}\left\{\overline{\alpha f(+1)} - \overline{\overline{\alpha} f(-1)} - \mathrm{i}\Phi(1)\,(f, P^{\perp}\phi)\right\}\,g(+1), \tag{3.12}$$

where we have used the boundary condition for g. We see that the right side of this equation vanishes for all $g \in \mathfrak{D}(A_0)$ if and only if f satisfies the boundary condition (3.10), i.e., $f \in \mathfrak{D}(A_0)$. This proves $(A_0)^* = -A_0$. $\square$

Proposition 3.2. *A_0 can be factorized as $A_0 = AZ_A$, where for all $f \in P^\perp\mathfrak{D}_A$*

$$(Z_A f)(x) = \begin{cases} f(-1) - (i/\alpha)[\Phi(1)f(-1) + (P^\perp\phi, f)]\Phi(x) & \text{if } x \in [-1,1], \\ f(x) & \text{if } x \notin [-1,1]. \end{cases} \tag{3.13}$$

The operator Z_A is unbounded and not closable.

Proof. The explicit form of Z_A follows from Eq. (3.9). In order to see that Z_A is not closable, take a sequence of functions $f_n \in P^\perp\mathfrak{D}_A$ which vanish outside $[-1-1/n, 1+1/n]$ and interpolate linearly between 0 and the boundary value $f_n(-1) = 1$ in $(-1-1/n, -1)$ (resp. between $f_n(+1)$ and 0 in $(1, 1+1/n)$, where $f_n(+1)$ is determined according to (3.10)). Then $f_n \to 0$ in $L^2(\mathbb{R} \setminus [-1,1])$, but $Z_A f_n$ converges in $L^2(\mathbb{R})$ to

$$g(x) = \begin{cases} 1 - (\mathrm{i}/\alpha)\Phi(1)\Phi(x) & \text{if } x \in [-1,1], \\ 0 & \text{if } x \notin [-1,1]. \end{cases} \tag{3.14}$$

□

3.2. **Example.** Our second example concerns a system of two first order equations. We note that the degenerate Cauchy problem with the operators M and A below coincides with the nonrelativistic limit of the one-dimensional Dirac equation [17]. This system has been investigated in [13].

Let $\mathfrak{H} = L^2(\mathbb{R}) \oplus L^2(\mathbb{R})$ and define the matrix-operators

$$M = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad A = \begin{pmatrix} -\mathrm{i}V & -\frac{d}{dx} \\ -\frac{d}{dx} & \mathrm{i} \end{pmatrix}. \tag{3.15}$$

Again we have $M = P^\perp = Q^\perp$. Here we assume that V is a symmetric operator, which is defined on $\mathfrak{D}(V) \supset W^{2,2}(\mathbb{R})$ and satisfies

$$\|Vf\|^2 \le a\|f''\|^2 + b\|f\|^2, \tag{3.16}$$

for some constants $a < 1$ and $b > 0$ and all $f \in W^{2,2}(\mathbb{R})$. The operator $\mathrm{i}A$ is symmetric on the domain

$$\mathfrak{D}(A) = \big(\mathfrak{D}(V) \cap W^{1,2}(\mathbb{R})\big) \oplus W^{1,2}(\mathbb{R}). \tag{3.17}$$

Writing $z = (f, g)^\top$, the condition $QAz = 0$ means $g = -\mathrm{i}f'$, hence

$$\mathfrak{D}_A = \left\{ z = \begin{pmatrix} f \\ g \end{pmatrix} \in \mathfrak{D}(A) \;\middle|\; g = -\mathrm{i}f' \right\} \tag{3.18}$$

This implies that in fact $f \in \mathfrak{D}(V) \cap W^{2,2}(\mathbb{R}) = W^{2,2}(\mathbb{R})$, i.e.,

$$P^\perp\mathfrak{D}_A = W^{2,2}(\mathbb{R}), \quad Z_A f = \begin{pmatrix} f \\ -\mathrm{i}f' \end{pmatrix} \quad \text{on } P^\perp\mathfrak{D}_A. \tag{3.19}$$

Proposition 3.3. *The operator*

$$iA_0 = iAZ_A = -\frac{d^2}{dx^2} + V, \quad \text{on } \mathfrak{D}(A_0) = W^{2,2}(\mathbb{R}), \tag{3.20}$$

is self-adjoint, hence $\exp(A_0 t)$ *is a unitary group on* $P^\perp\mathfrak{H}$*. The operator* Z_A *defined in Eq. (3.19) is closable, the domain of the closure* Z_A^c *is* $W^{1,2}(\mathbb{R})$*. The range of* Z_A^c *is* $\mathfrak{D}_A^c$*, the closure of* $\mathfrak{D}_A$ *with respect to the norm in* $\mathfrak{H}$*. Moreover,*

$$P^\perp Z_A^c = \mathbf{1} \restriction W^{1,2}(\mathbb{R}), \qquad Z_A^c P^\perp = \mathbf{1} \restriction \mathfrak{D}_A^c. \tag{3.21}$$

Proof. The self-adjointness of A_0 follows from the relative boundedness condition Eq. (3.16). The closedness of the mapping $f \in W^{1,2} \to -\mathrm{i}f'$ is clear because the set $W^{1,2}(\mathbb{R})$ is the domain where $-\mathrm{i}d/dx$ is self-adjoint. Therefore T_A and likewise Z_A are closable with closure defined on $W^{1,2}$. Since

$$\|z\|^2 = \|f\|^2 + \|f'\|^2 \quad \text{for all } z \in \mathfrak{D}_A, \text{ with } f = P^\perp z, \tag{3.22}$$

the relations (3.21) follow by approximation from the corresponding relations on $\mathfrak{D}_A$ and $P^\perp\mathfrak{D}_A$. □

We also want to stress that the usual Sobolev-norm on $W^{1,2}(\mathbb{R})$ is equivalent to the graph norm of the operator T_A^c, i.e.,

$$\|f\|_G^2 \equiv \|f\|^2 + \|f'\|^2 = \|Z_A^c f\|^2, \quad \text{for all } f \in W^{1,2}(\mathbb{R}). \tag{3.23}$$

Hence Z_A^c is an isometry between the Sobolev space $(W^{1,2}(\mathbb{R}), \|\cdot\|_G)$ and the Hilbert space $(\mathfrak{D}_A^c, \|\cdot\|)$. In particular, for all $z \in \mathfrak{D}_A^c$ with $f = P^\perp z$ we have $\|z\| = \|f\|_G$.

Proposition 3.4. $T : t \to T(t) = Z_A^c \exp(A_0 t) P^\perp$, $t \in \mathbb{R}$, *defines a strongly continuous group on the closed subspace* $\mathfrak{D}_A^c \subset \mathfrak{H}$*. The operators* $T(t)$ *are bounded uniformly in* t *(with respect to the norm inherited from* $\mathfrak{H}$*). If* $V = 0$*, then* T *is even a unitary group.*

Proof. For any self-adjoint operator H we can write $H = (\operatorname{sgn} H)\,|H|$, where $\operatorname{sgn} H$ is a unitary involution and $|H|$ is positive on $\mathfrak{D}(|H|) = \mathfrak{D}(H)$. Moreover, $\exp(-\mathrm{i}Ht)$ leaves $\mathfrak{D}(|H|^{1/2})$ invariant, and

$$\exp\Big(-\mathrm{i}Ht\Big)\,|H|^{1/2} = |H|^{1/2}\,\exp\Big(-\mathrm{i}Ht\Big) \quad \text{on } \mathfrak{D}(|H|^{1/2}). \tag{3.24}$$

All these assertions follow easily from the spectral theorem.

By the Kato-Rellich theorem and Eq. (3.16), the operators $\mathrm{i}A_0$ and $\mathrm{i}B_0 = -d^2/dx^2$ are both defined and self-adjoint on the domain $W^{2,2}(\mathbb{R})$. Hence (see [18], Theorem 9.4)

$$\mathfrak{D}(|\mathrm{i}A_0|^{1/2}) = \mathfrak{D}((\mathrm{i}B_0)^{1/2}) = W^{1,2}(\mathbb{R}). \tag{3.25}$$

This shows that $P^\perp\mathfrak{D}_A^c = W^{1,2}(\mathbb{R})$ is left invariant under $\exp(A_0 t)$. Hence $T(t)$ is a C_0-semigroup by Theorem 2.2. Next we show that $T(t)$ is even bounded uniformly in t.

The set $W^{1,2}(\mathbb{R})$ is a Hilbert space with respect to the norm $\|\cdot\|_G$, which is also the graph norm of the positive operator $R = (\mathrm{i}B_0)^{1/2}$. The equality of domains, Eq. (3.25), implies that the operators $S = |\mathrm{i}A_0|^{1/2}$ and R are bounded with respect to each other (cf. [18], Theorem 5.9) and therefore the corresponding graph norms are equivalent. Hence, with a suitable constant $K > 0$,

$$\begin{aligned}\|e^{A_0 t} f\|_G^2 &\leq K\Big(\|f\|^2 + \|\,|\mathrm{i}A_0|^{1/2}\, e^{A_0 t} f\|^2\Big)\\ &= K\Big(\|f\|^2 + \|\,|\mathrm{i}A_0|^{1/2}\, f\|^2\Big)\\ &\leq K^2\|f\|_G^2.\end{aligned}$$

Hence $\|\exp(A_0 t)\|_G \leq K$, and for all $z \in \mathfrak{D}_A^c$, with $f = P^\perp z$,

$$\|T(t)z\| = \|Z_A^c e^{A_0 t} P^\perp z\| = \|e^{A_0 t} f\|_G \leq K\|f\|_G = K\|z\|. \tag{3.26}$$

This proves the uniform boundedness of $T(t)$, $t \in \mathbb{R}$. With Eq. (3.21) it is easy to see that the operators $T(t)$ form a group, $T(t+s) = T(t) + T(s)$, $T(0) = \mathbf{1}$ on $\mathfrak{D}_A^c$.

For $z_0 \in \mathfrak{D}_A$, the map $z : t \to z(t) = T(t)z_0$ is continuous, because z is the strict solution of the homogeneous equation. Since $T(t)$ is bounded on $\mathfrak{D}_A$, a simple approximation argument proves the continuity of $z(t)$ for all initial values in $\mathfrak{D}_A^c$. □

4. Mild Solutions of Degenerate Control Problems

Let M and A fulfill the conditions for factorization formulated in Section 2 and consider the degenerate abstract Cauchy problem

$$\frac{d}{dt}Mz(t) = Az(t) + f(t), \quad \text{for } f \in L^p([0,\infty), Q^\perp\mathfrak{K}). \tag{4.1}$$

Here we assume that f is a p-integrable function, $1 \leq p \leq \infty$, with values in $\operatorname{Ran} M = Q^\perp\mathfrak{K}$. As we have seen in Section 2, (4.1) can be factorized into a nondegenerate abstract Cauchy problem, e.g.,

$$\frac{d}{dt}x(t) = A_2 x(t) + (M^\perp)^{-1} f(t) \tag{4.2}$$

such that the strict solutions z of the degenerate system Eq. (4.1) can be obtained from the strict solutions x of Eq. (4.2). The strict solutions are defined if f is continuously differentiable. For $f \in L^p$ we can define mild solutions for Eq. (4.2) in the standard way. The problem now is to find an analogon to this mild solution for a degenerate system.

For the nondegenerate problem (4.2) having a mild solution $x(t)$ is equivalent to the existence of a sequence of strict solutions $x_n(t)$ of Cauchy problems

$$\frac{d}{dt}x_n(t) = A_2 x_n(t) + (M^\perp)^{-1} f_n(t) \tag{4.3}$$

with $x_n(t) \longrightarrow x(t)$ uniformly on bounded t-intervals, $f_n \in C^1$ and $f_n \longrightarrow f$ in L^p. In analogy to that we consider a sequence of degenerate Cauchy problems

$$\frac{d}{dt}Mz_n(t) = Az_n(t) + f_n(t), \tag{4.4}$$

and define a mild solution $z(t)$ of (4.1) as the uniform limit of a sequence of strict solutions $z_n(t)$ of (4.4).

On the other hand, if the operator Z_A defined in Section 2 is closable, we can define a mild solution of (4.1) by writing $z(t) = Z_A^c x(t)$, whenever $x(t)$ is the mild solution of (4.2) which satisfies $x(t) \in \mathfrak{D}(Z_A^c) = \mathfrak{H}_0$ for all t. Let us put this together to the following

Definition 4.1. $z(t)$ is called a *mild solution of type A* of (4.1) if and only if there exists a sequence (f_n) with $f_n \in C^1$ and a sequence (z_n) of strict solutions of (4.4) such that $z_n(t) \longrightarrow z(t)$ uniformly on bounded t-intervals and $f_n \longrightarrow f$ in L^p.

Let Z_A be closable. $z(t)$ is called a *mild solution of type B* of (4.1) if and only if there exists a mild solution $x(t)$ of (4.2), $Z_A^c x(t)$ is well defined and $z(t) = Z_A^c x(t)$.

Remark 4.1. Equivalently, we can define $z(t)$ as a mild solution of type B of (4.1) via a mild solution $y(t)$ of (2.9) which is in $M^\perp \mathfrak{H}_0$ for all t, i.e., $z(t) = Z_A^c (M^\perp)^{-1} y(t)$.

Remark 4.2. A mild solution z of type A can be defined even if Z_A is not closable. In any case, $x(t) = P^\perp z(t)$ is a mild solution of the corresponding nondegenerate problem, and there are approximating sequences of strict solutions x_n and differentiable functions f_n. It is well known that the mild solution x does not depend on the choice of f_n, as long as $\int_0^t \|f_n(s) - f(s)\| ds \to 0$. However, if Z_A is not closable, there might exist two sequences of strict solutions $x_n^{(i)}(t)$, $i = 1, 2$, both converging to $x(t)$, such that both limits $z^{(i)}(t) = \lim Z_A x_n^{(i)}(t)$ exist but are not equal. Hence it could well be that the mild solution $z(t)$ of the degenerate system depends on the choice of the approximating functions f_n.

Corollary 4.1. *Let Z_A be closable. $z(t)$ is a mild solution of (4.1) of type A if and only if $x(t)$ is a mild solution of (4.2) with respect to the graph norm $\|\cdot\|_G$ defined in Eq. (2.11), i.e., $x(t) \in \mathfrak{H}_0$ for all t, and there exists a sequence (f_n) with $f_n \longrightarrow f$ in L^p and a sequence of strict solutions $x_n(t)$ of (4.3) such that for all $0 < T < \infty$*

$$\sup_{t \in [0,T]} \|x(t) - x_n(t)\|_G \to 0, \quad \text{as } n \to \infty. \tag{4.5}$$

Proof. By Assumption 2.1 every strict solution $z_n(t)$ of (4.4) can be written uniquely as the sum $x_n(t) + k_n(t)$, where $x_n(t) \in P^{\perp}\mathfrak{H}$ and $k_n(t) = T_A x_n(t) \in P\mathfrak{H}$. Therefore, $z_n(t)$ converges if and only if both $x_n(t)$ and $k_n(t)$ converge, i.e., $x_n(t)$ converges in the graph norm of T_A. □

The next corollary states that type A is a special case of type B whenever Z_A is closable.

Corollary 4.2. *Let Z_A be closable. If $z(t)$ is a mild solution of type A then it is also of type B.*

Proof. $z(t)$ is a mild solution of type A of (4.1) implies that $z_n(t) \longrightarrow z(t)$ and $P^{\perp}z_n(t) = x_n(t)$ converges to the mild solution $x(t)$ of (4.2). All $x_n(t)$ are strict solutions of (4.3), hence $z_n(t) = Z_A x_n(t) = Z_A^c x_n(t)$. The operator Z_A^c is closed, therefore we get $x(t) \in \mathfrak{D}(Z_A^c)$ and $z(t) = Z_A^c x(t)$, i.e., $z(t)$ is the mild solution of type B of (4.1). □

Remark 4.3. In general the converse of Corollary 4.2 need not be true. If $x(t) = \lim x_n(t)$ is a mild solution in $\mathfrak{H}_0 = \mathfrak{D}(Z_A^c)$ such that $Z_A^c x(t)$ is of type B, the limit of $z_n(t) = Z_A x_n(t)$ need not exist due to the unboundedness of Z_A. The problem would be to construct the approximating sequence x_n in such a way that (4.5) holds.

Remark 4.4. In [11] mild solutions of type B were considered. In the setting of Showalter [2] the norm in $\mathfrak{H}$ is defined by M such that there is a continuous relation between $x(t)$ and $z(t)$ and therefore type B is equivalent to type A.

Theorem 4.1. *Let Z_A be closable. Let at least one of the following assumptions be true:*

1. *Z_A is bounded*
2. *A_2 generates an analytic semigroup*
3. *$Z_A e^{A_2 t}$ is bounded*

Then the mild solution $z(t)$ of type B of (4.1) exists and is also a mild solution of type A.

Proof. Clearly, if Z_A is bounded then also $Z_A e^{A_2 t}$ is bounded. If A_2 generates an analytic semigroup then $A_2 e^{A_2 t}$ is bounded for all $t > 0$. Since Z_A is bounded relative to A_2, this implies that also $Z_A e^{A_2 t}$ is bounded. So it is sufficient to prove the theorem in the case that $Z_A e^{A_2 t}$ is bounded. Define $z(t) = Z_A e^{A_2 t} x_0 + \int_0^t Z_A^c e^{A_2(t-s)}(M^{\perp})^{-1} f(s) ds$. Since $f \in L^p$ is in particular integrable on $[0, t]$, $z(t)$ is well defined. Now, consider the mild solution x of (4.2) which can be written as $x(t) = e^{A_2 t} x_0 + \int_0^t e^{A_2(t-s)}(M^{\perp})^{-1} f(s) ds$. There exists a sequence $x_n(t)$ of strict solutions of (4.3), given by $x_n(t) = e^{A_2 t} x_{n,0} + \int_0^t e^{A_2(t-s)}(M^{\perp})^{-1} f_n(s) ds$

with $x_n(t) \longrightarrow x(t)$, $x_{n,0} \in \mathfrak{D}(A_2)$, $x_{n,0} \longrightarrow x_0$, $f_n \in C^1$ and $f_n \longrightarrow f$ in L^p. Now we can apply Z_A on $x_n(t)$ and we get

$$\begin{aligned} Z_A^c x_n(t) &= Z_A x_n(t) = z_n(t) \\ &= Z_A e^{A_2 t} x_{n,0} + Z_A \int_0^t e^{A_2(t-s)} (M^\perp)^{-1} f_n(s) ds \\ &= Z_A e^{A_2 t} x_{n,0} + \int_0^t Z_A^c e^{A_2(t-s)} (M^\perp)^{-1} f_n(s) ds \end{aligned}$$

which converges to $z(t)$ for $n \longrightarrow \infty$. As Z_A^c is closed this implies that $Z_A^c x(t)$ exists and $Z_A^c x(t) = z(t)$, i.e., $z(t)$ is a mild solution of type B. Since $z_n(t) \longrightarrow z(t)$, it is also a mild solution of type A. □

5. Degenerate Control Theory and the LQR Problem

With M and A as in Section 2 we now consider the degenerate control problem

$$\frac{d}{dt} M z(t) = A z(t) + B u(t), \tag{5.1}$$

$$v(t) = C z(t), \tag{5.2}$$

$$z(0) = z_0,$$

where B is a bounded linear operator from a Banach space U called the control space into $\mathfrak{K}$, and C is a bounded linear operator mapping $\mathfrak{H}$ into a Banach space V, the output space. We assume furthermore that B and C satisfy

$$\operatorname{Ran} B \subset (\operatorname{Ran} M)^c, \ \operatorname{Ker} C \supset \operatorname{Ker} M. \tag{5.3}$$

To make notation easier we write the ordered quadrupel (M, A, B, C) instead of equations (5.1) and (5.2).

We are going to make the following simplifying assumption:

Assumption 5.1. *Let A be a closed linear operator with bounded inverse, and assume that $P\mathfrak{D}(A) \subset \mathfrak{D}(A)$. Let the operator $QAP \restriction P\mathfrak{H}$ have a bounded inverse, and let*

$$Z_A = P^\perp - (QAP)^{-1} QAP^\perp \tag{5.4}$$

be closable.

Since we also want to consider the dual system $(-M^*,A^*,C^*,B^*)$ we assume that A^* satisfies an analogous assumption with P and Q exchanged, so that we can define the factorization operator

$$Z_{A^*} = Q^\perp - (PA^*Q)^{-1} PA^* Q^\perp. \tag{5.5}$$

Using Assumption 5.1 the operator A_0 of the corresponding nondegenerate system can be written as

$$A_0 = P^\perp A P^\perp - P^\perp A Q (QAP)^{-1} QAP^\perp = A Z_A = Y_A A, \tag{5.6}$$

where we have defined

$$Y_A = Q^\perp - Q^\perp AP\,(QAP)^{-1}\,Q. \tag{5.7}$$

Remark 5.1. Formally we have

$$Y_A = (Z_{A^*})^*, \qquad Y_{A^*} = (Z_A)^*, \tag{5.8}$$

where the adjoint is taken with respect to the scalar product in $\mathfrak{H}$. The operator $Y_{A^*} = P^\perp - P^\perp A^* Q(PA^*Q)^{-1}P$ has values in $P^\perp\mathfrak{H}$. Let P_A be the orthogonal projection onto the closed subspace $\mathfrak{D}_A^c$. Let $z = x + k \in \mathfrak{D}_A$, where $x = P^\perp z$ and consider $y = Y_{A^*}z$. Then a formal calculation shows that $y - z$ is orthogonal to z, i.e., z is the orthogonal projection of y onto $\mathfrak{D}_A$. Hence $Y_{A^*}^c$ is the inverse of P_A, restricted to a suitable subspace of $P^\perp\mathfrak{H}$.

By Theorem 2.1 Eq. (5.1) has a unique strict solution $z(t,u,z_0)$ whenever Bu is continuously differentiable and $z_0 \in \mathfrak{D}_A$. Here again we have two equivalent methods to get nondegenerate control systems which lead to the same solution $z(t,u,z_0)$ of (5.1):

First we restrict system (5.1), (5.2) to

$$\frac{d}{dt}M^\perp x(t) = A_0 x(t) + Bu(t), \tag{5.9}$$

$$v(t) = Cx(t), \tag{5.10}$$

and then by the first method of factorization we get ($\mathbf{1}$,A_1,B_1,C_1), i.e.

$$\frac{d}{dt}x(t) = A_1 x(t) + B_1 u(t), \tag{5.11}$$

$$v(t) = C_1 x(t), \tag{5.12}$$

with $A_1 = A_0(M^\perp)^{-1}$, $B_1 = B$, $C_1 = C(M^\perp)^{-1}$, or, by using the second method, we get an equivalent system ($\mathbf{1}$,A_2,B_2,C_2), i.e.

$$\frac{d}{dt}x(t) = A_2 x(t) + B_2 u(t), \tag{5.13}$$

$$v(t) = C_2 x(t), \tag{5.14}$$

where $A_2 = (M^\perp)^{-1}A_0$, $B_2 = (M^\perp)^{-1}B$, $C_2 = C$.

Define now the dual degenerate control problem to (5.1), (5.2) by the quadrupel ($-M^*$,A^*,C^*,B^*). The connection between the factorized dual equation and the dual factorized equation is the following ([17]): Factorizing (M,A,B,C) by the first method leads to a nondegenerate system ($\mathbf{1}$,A_1,B_1,C_1), where $A_1 = A_0(M^\perp)^{-1}$ as before, $B_1 = B$, $C_1 = C(M^\perp){-}1$. The dual system to ($\mathbf{1}$,A_1,B_1,C_1) is ($-\mathbf{1}$,$(A_1)^*$,$(C_1)^*$,$(B_1)^*$). A short calculation shows that this is equal to the system ($-\mathbf{1}$,$(A^*)_2$,$(C^*)_2$,$(B^*)_2$) which is the factorized equation of ($-M^*$,A^*,C^*,B^*), the dual system of (M,A,B,C), where we have used $A_2 = (M^\perp)^{-1}A_0$, $B_2 =$

$(M^\perp)^{-1}B$, $C_2 = C$. This shows that we can interchange factorization and dualization if we interchange the method of factorization. In this sense the two methods of factorization are dual.

Let us now consider a feedback-control system

$$\frac{d}{dt}Mz(t) = (A+BK)z(t), \tag{5.15}$$

where K is an operator defined on $\mathfrak{D}(K) \supset \mathfrak{D}(A)$ and maps into U. If we factorize (5.15) we get a nondegenerate control system which again is of feedback type:

$$\frac{d}{dt}y(t) = (A_1 + B_1K_1)y(t), \tag{5.16}$$

where $K_1 = KZ_A(M^\perp)^{-1}$, A_1, B_1 as before, or

$$\frac{d}{dt}x(t) = (A_2 + B_2K_2)x(t), \tag{5.17}$$

with $K_2 = K$. The solution of the degenerate system (5.15) is therefore given by

$$z(t) = S_{A+BK}(t)z_0$$

with the (possibly unbounded) evolution operator

$$S_{A+BK}(t) = Z_A e^{(A_2+B_2K_2)t}P^\perp = Z_A(M^\perp)^{-1}e^{(A_1+B_1K_1)t}M^\perp P^\perp. \tag{5.18}$$

Remark 5.2. If A_1 (resp. A_2) generates an analytic semigroup then $A_1 = B_1K_1$ (resp. $A_2 + B_2K_2$) also generates an analytic semigroup.

Definition 5.1. The degenerate control system (M, A, B, C) is called *approximately controllable*, if for all $\epsilon > 0$, all $T > 0$ and all $z_0, z_1 \in (\mathfrak{D}_A)^c$ there exists a control function u such that $\|z(T, u, z_0) - z_1\| < \epsilon$.

It is called *observable*, if the dual system $(-M^*, A^*, C^*, B^*)$ is approximately controllable.

The degenerate control system is called *stabilizable*, if there exists a bounded operator K such that $S_{A+BK}(t)$ is bounded for $t \geq 0$ with $\|S_{A+BK}(t)\| \leq \mu \exp(-\omega t)$, $\mu \geq 1$, $\omega > 0$.

The degenerate control system is called *detectable*, if the dual system is stabilizable.

The following theorem shows the connection between these properties of a degenerate control system and the corresponding nondegenerate control system ([17]).

Theorem 5.1. *Let (M, A, B, C) satisfy the assumptions formulated in this section and assume that the factorization operator Z_A is bounded. Then the degenerate system (M, A, B, C) is controllable (observable, stabilizable, detectable) if and only if the ordinary control system $(1, A_j, B_j, C_j)$ ($j = 1$ or 2) is controllable (observable, stabilizable, detectable).*

Next we turn to a discussion of the linear quadratic regulator (LQR) problem. Let $z(t)$ be a mild solution of (1.5). Then $z(t)$ is continuous for all inputs $u \in L^2$. The LQR problem with infinite time horizon consists of minimizing the functional

$$J_\infty(z_0, u) = \int_0^\infty (\|Cz(s)\|^2 + \|u(s)\|^2)ds \tag{5.19}$$

over all controls $u \in L^2(0, \infty; U)$. A control $u \in L^2(0, \infty; U)$ is called *admissible* if $J_\infty(z_0, u) < \infty$. An admissible control u^* is called *optimal* if $J_\infty(z_0, u^*) \leq J_\infty(z_0, u)$ for all $u \in L^2(0, \infty; U)$. If we consider the corresponding nondegenerate control problem

$$\begin{aligned} \frac{d}{dt}x(t) &= A_2 x(t) + B_2 u(t), \\ v(t) &= Cx(t), \\ x(0) &= x_0, \end{aligned} \tag{5.20}$$

we have the cost functional

$$J_\infty(x_0, u) = \int_0^\infty (\|Cx(s)\|^2 + \|u(s)\|^2)ds. \tag{5.21}$$

As we have $Cz(t) = Cx(t)$, i.e., the output is the same in the degenerate control problem (1.5) and the corresponding nondegenerate control problem (5.20), and the control $u(t)$ also is the same for both systems, the cost functionals (5.19) and (5.21) are the same. Therefore, if we apply the known results to the LQR problem for the nondegenerate control problem (5.20) for finding a solution which minimizes (5.21) we get a corresponding solution of (1.5) which also minimizes the cost functional (5.19). So we start with some definitions:

Definition 5.2. The nondegenerate system (5.20) is *C-stabilizable* if and only if for all x_0 there exists a control u such that the cost $J_\infty(x_0, u)$ given by (5.21) is finite.

System (5.20) is *I-stabilizable* if and only if for all x_0 there exists a control u such that $\int_0^\infty (\|x(t)\|^2 + \|u(t)\|^2)dt < \infty$.

The degenerate system (1.5) is *C-stabilizable* if and only if for all z_0 there exists a control u such that the cost $J_\infty(z_0, u)$ given by (5.19) is finite.

It is called *I-stabilizable* if and only if for all z_0 there exists a control u such that $\int_0^\infty (\|z(t)\|^2 + \|u(t)\|^2)dt < \infty$.

Proposition 5.3.

a) System (5.20) is I-stabilizable $\Rightarrow$ System (5.20) is C-stabilizable.

b) System (5.20) is I-stabilizable $\Leftrightarrow$ System (5.20) is stabilizable.

c) System (1.5) is I-stabilizable $\Rightarrow$ System(1.5) is C-stabilizable.

Proof. a) and c) follow immediately from the fact that C is bounded. For b) see e.g.[1]. $\square$

Proposition 5.4. *Consider the degenerate System (M, A, B, C) and the corresponding nondegenerate System $(\mathbf{1}, A_2, B_2, C_2)$ where A_2, B_2, C_2 is defined as in Section 2. Then the following relations hold:*

a) (M, A, B, C) C-stabilizable $\Leftrightarrow$ $(\mathbf{1}, A_2, B_2, C_2)$ C_2-stabilizable,

b) $(\mathbf{1}, A_2, B_2, C_2)$ I-stabilizable $\Leftrightarrow$ (M, A, B, C) $P^{\perp}$-stabilizable,

c) If Z_A is bounded we get

$$(\mathbf{1}, A_2, B_2, C_2) \text{ } I\text{-stabilizable} \Leftrightarrow (M, A, B, C) \text{ } I\text{-stabilizable.}$$

Proof. a) $C_2 = C$ on $P^{\perp}\mathfrak{H}$ and by (5.3) $C = 0$ on $P\mathfrak{H}$. Therefore, $Cz(t) = Cx(t)$ and the cost functionals (5.19) and (5.21) are equal.

b) If we take the identity I instead of C in the degenerate system (1.5) then I in general does not fulfill (5.3), i.e. $\operatorname{Ker} I \supset \operatorname{Ker} M$. $P^{\perp}$ fulfills (5.3) and therefore can be taken as output operator.

c) follows immediately from the definition of I-stabilizable. □

Proposition 5.5. *Let Z_A be bounded. Then we have*

$$(M, A, B, C) \text{ } I\text{-stabilizable} \Leftrightarrow (M, A, B, C) \text{ stabilizable}$$

Proof. Equality follows immediately from Prop. 5.4 c, Prop. 5.3 b, and Theorem 5.1. □

In the following theorem we sum up the results for the infinite time horizon problem for nondegenerate control systems (see e.g. [1],[3],[12]):

Theorem 5.2. *Consider the optimal control problem $(\mathbf{1}, A, B, C)$ with initial value $x(0) = x_0$, let A generate a C_0-semigroup and let $(\mathbf{1}, A, B, C)$ be C-stabilizable. Then there exists a unique optimal pair $(\hat{u}, \hat{x})$ for the optimal problem and*

a) $\hat{x}$ is the mild solution of the closed loop equation

$$\frac{d}{dt}x(t) = (A - BB^*P^{\infty}_{min})x(t), \qquad (5.22)$$
$$x(0) = x_0.$$

b) $\hat{u}$ is given by the feedback formula

$$\hat{u}(t) = -B^*P^{\infty}_{min}\hat{x}(t). \qquad (5.23)$$

where P^{∞}_{min} is the minimal solution of the algebraic Riccati equation

$$A^*X + XA - XBB^*X + C^*C = O. \qquad (5.24)$$

c) The optimal cost $J_{\infty}(x_0, \hat{u})$ is given by

$$J_{\infty}(x_0, \hat{u}) = (P^{\infty}_{min}x_0, x_0). \qquad (5.25)$$

If $(\mathbf{1}, A, B, C)$ is detectable then $F = A - BB^*P^{\infty}_{min}$ is exponentially stable and P^{∞}_{min} is the unique positive solution of the algebraic Riccati equation (5.24).

If we now put the facts together we get

Theorem 5.3. *Consider the degenerate control system (M, A, B, C) with initial value $z(0) = z_0$ and let the assumptions for the factorization to a nondegenerate system $(\mathbf{1}, A_1, B_1, C_1)$ resp. $(\mathbf{1}, A_2, B_2, C_2)$ be fulfilled. Assume furthermore that Z_A is bounded and let (M, A, B, C) be C-stabilizable and detectable. Then there exists a unique optimal pair $(\hat{u}, \hat{z})$ which minimizes the cost functional (5.19) and*

a) $\hat{u}$ is given by

$$\hat{u}(t) = -(B_2)^* P_2 \hat{x}(t), \tag{5.26}$$

where $\hat{x}(t)$ is the solution to the closed loop equation

$$\begin{aligned} \frac{d}{dt}x(t) &= (A_2 - B_2(B_2)^* P_2)x(t) \\ x(0) &= x_0, \end{aligned} \tag{5.27}$$

P_2 is the unique nonnegative solution of the algebraic Riccati equation

$$(A_2)^* X + X A_2 - X B_2 (B_2)^* X + (C_2)^* C_2 = 0, \tag{5.28}$$

and $(A_2 - B_2(B_2)^* P_2)$ is exponentially stable.

b) $\hat{z}$ is given by $\hat{z} = Z_A \hat{x}$.

c) The optimal cost is given by $J_\infty(z_0, \hat{u}) = (P_2 x_0, x_0)$.

Proof. By Theorem (5.2) $\hat{u}$ given by (5.26) is the optimal control of the control system $(\mathbf{1}, A_2, B_2, C_2)$, the factorized system of (M, A, B, C). As $C = C_2$ on $P^\perp \mathfrak{H}$ the cost functionals (5.21) and (5.19) are equal. □

Remark 5.6. Of course we get the same optimal control if we take $(\mathbf{1}, A_1, B_1, C_1)$ instead of $(\mathbf{1}, A_2, B_2, C_2)$: If P_2 is a solution of the Riccati equation (5.28) then $P_1 = (M^\perp)^{-1*} P_2 (M^\perp)^{-1}$ is a solution of the Riccati equation

$$(A_1)^* X + X A_1 - X B_1 (B_1)^* X + C_1^* C_1 = 0. \tag{5.29}$$

$\hat{x}_1 = (M^\perp)^{-1} \hat{x}_2$ is the solution to the closed loop equation

$$\frac{d}{dt}x(t) = (A_1 - B_1(B_1)^* P_1)x(t), \tag{5.30}$$

and $\hat{u}_1(t) = -(B_1)^* P_1 \hat{x}_1(t) = -(B_2)^* P_2 \hat{x}_2(t) = \hat{u}_2(t)$.

In the following theorem we calculate the optimal control and the optimal state via degenerate equations and we get the feedback operator via a Riccati equation acting in the space $\mathfrak{D}_A$.

Theorem 5.4. *Consider the optimal control problem (M, A, B, C) and let the assumptions of Theorem 5.3 be fulfilled. Then the unique optimal control $\hat{u}$ is given by*

$$\hat{u}(t) = -B^* \overline{P} \hat{z}(t), \tag{5.31}$$

where $\hat{z}(t)$ is the solution of the degenerate closed loop equation

$$\frac{d}{dt}Mz(t) = (A - BB^*\overline{P})z(t), \qquad (5.32)$$
$$z(0) = z_0.$$

$\overline{P}$ is the solution of the equation

$$A^*X + X^*A - X^*BB^*X + C^*C = 0. \qquad (5.33)$$

Proof. By Theorem 5.3 there exists a unique nonnegative solution P_2 of the algebraic Riccati equation (5.28). Defining

$$\overline{P} = {(M^\perp)^*}^{-1} P_2 P^\perp$$

and using $C_2 = CZ_A$, $B_2 = (M^\perp)^{-1}B$, $A_2 = (M^\perp)^{-1}AZ_A$ (c.f. Section 2) we get

$$Y_{A^*}(A^*\overline{P} + \overline{P}^*A - \overline{P}^*BB^*\overline{P} + C^*C)Z_A = 0,$$

and as Y_{A^*} and Z_A are bijections we have $\overline{P}$ is a solution of (5.33) if and only if P_2 is a solution of (5.28). Consider now the closed loop equation (5.27). By (5.18) $\hat{x}(t)$ is a solution of (5.27) if and only if $\hat{z}(t) = Z_A\hat{x}(t)$ is a solution of

$$\frac{d}{dt}Mz(t) = (A + BK)z(t),$$

with $K = -(B_2)^*P_2P^\perp$. If we insert $B_2 = (M^\perp)^{-1}B$ and $P_2 = (M^\perp)^*\overline{P}Z_A$ we get $K = -B^*\overline{P}$ and therefore $\hat{z}(t)$ is a solution of (5.32). By Theorem 5.3 the optimal control $\hat{u}(t)$ is given by (5.26). Using again $B_2 = (M^\perp)^{-1}B$ and $P_2 = (M^\perp)^*\overline{P}Z_A$ we get $\hat{u}(t) = -B^*\overline{P}\hat{z}(t)$. □

Remark 5.7. $A - BB^*\overline{P}$ is exponentially stable, i.e., the evolution operator

$$S_{A-BB^*\overline{P}}(t) = Z_A e^{(A_2 - B_2(B_2)^*P_2)} P^\perp \qquad (5.34)$$

is bounded for $t \geq 0$ with $\|S_{A-BB^*\overline{P}}\| \leq \mu \exp(-\omega t)$, $\mu \geq 1$, $\omega > 0$. (c.f. [17], Theorem 5.1).

Remark 5.8. Theorem 5.4 is the exact degenerate analogue to Theorem 5.2 while Theorem 5.3 seems to be somewhere between the degenerate and the nondegenerate case. On the other hand, for solving the closed loop equation (5.32) it is also necessary to calculate the operators A_2, B_2. So it is not possible to avoid the factorization but it depends on the problem whether it is easier to factorize first and then solve the Riccati equation (5.28) or to solve the Riccati equation (5.33) first and then factorize equation (5.32).

References

[1] Bensoussan, A., Da Prato, G., Delfour, M.C., and Mitter, S.K.: *Representation and Control of Infinite Dimensional Systems*, Birkhäuser, Boston, Basel, Berlin 1993

[2] Carroll, R. W. and Showalter, R. E.: *Singular and degenerate Cauchy problems*, Mathematics in Science and Engineering **127**, Academic Press, New York, San Francisco, London 1976.

[3] Curtain, R.F., and Pritchard, A.J., *Infinite dimensional linear systems theory*, Lecture Notes in Control and Information Sciences, No.8, Springer, Berlin, Heidelberg, New York 1978.

[4] Dai, L.: *Singular Control Systems*, Lecture Notes in Control and Information Sciences **118**, Springer-Verlag, Berlin, Heidelberg, New York 1989.

[5] Favini, A., *Controllability conditions of linear degenerate evolution systems*, Appl. Math. Optim. **6**, 153–167 (1980).

[6] Favini, A. and Plazzi, P., *On some abstract degenerate problems of parabolic type–1. The linear case*, Nonlinear Analysis **12**, 1017–1027 (1988).

[7] Favini, A. and Plazzi, P., *On some abstract degenerate problems of parabolic type–2. The nonlinear case*, Nonlinear Analysis **13**, 23–31 (1989).

[8] Favini, A. and Plazzi, P., *On some abstract degenerate problems of parabolic type–3: Applications to linear and nonlinear problems*, Osaka J. Math. **27**, 323–359 (1990).

[9] Favini, A. and Yagi, A., *Multivalued linear operators and degenerate evolution problems*, Annali di Mat. Pura Appl. (1992).

[10] Favini, A. and Yagi, A., *Space and time regularity for degenerate evolution equations*, J. Math. Soc. Japan **44**, 331–350 (1992).

[11] Lamm, P.K., and Rosen, I.G.: *Approximation theory for the estimation of parameters in degenerate Cauchy problems*, J. Math. Anal. Appl. **162**, 13–48 (1991)

[12] Kappel, F., *Approximation of LQR-problems for delay systems: A survey*, in: Computation and Control Vol.2, K. Bowers and J. Lund, Eds., pp. 187–224, Birkhäuser, Boston 1991

[13] Orgler, P.: *Streutheorie für Diracoperatoren im nichtrelativistischen Limes*, Diplomarbeit, Universität Graz 1995

[14] Showalter, R. E., *Nonlinear degenerate evolution equations and partial differential equations of mixed type*, SIAM J. Math. Anal. **6**, 25–42 (1975).

[15] Showalter, R. E., *Degenerate parabolic initial-boundary value problems*, J. Diff. Eq. **31**, 296–312 (1979).

[16] Showalter, R. E., *Initial and final-value problems for degenerate parabolic evolution systems*, Indiana University Mathematics Journal **28**, 883–893 (1979).

[17] Thaller, B. and Thaller, S.: *Factorization of degenerate Cauchy problems: The linear case*, Bericht No.18 des SFB "Optimierung und Kontrolle", University of Graz, 1995

[18] Weidmann, J.: *Linear Operators in Hilbert Spaces*, Springer-Verlag, Berlin, Heidelberg, New York 1978

Inst. f. Mathematik, Universität Graz, Heinrichstrasse 36,, A-8010 Graz, Austria

E-mail address: `bernd.thaller@kfunigraz.ac.at`

E-mail address: `sigrid.thaller@kfunigraz.ac.at`