TA 8l.

Singular Optimal Control – the State of the Art

Volker Michel *

December 4, 1996

Abstract

The purpose of this paper is to present the state of the art in singular optimal control. If the Hamiltonian in an interval $[t_1, t_2]$ is independent of the control we call the control in this interval singular. Singular optimal controls appear in many applications so that research has been motivated since the 1950s. Often optimal controls consist of nonsingular and singular parts where the junctions between these parts are mostly very difficult to find. One section of this work shows the actual knowledge about the location of the junctions and the behaviour of the control at the junctions. The definition and the properties of the orders (problem order and arc order), which are important in this context, are given, too. Another chapter considers multidimensional controls and how they can be treated. An alternate definition of the orders in the multidimensional case is proposed and a counterexample, which confirms a remark given in the 1960s, is given. A voluminous list of optimality conditions, which can be found in several publications, is added. A strategy for solving optimal control problems numerically is given, and the existing algorithms are compared with each other. Finally conclusions and an outlook on the future research is given.

Key words: singular optimal control, Pontrjagin, Minimum Principle, Hamiltonian, junction

1 Introduction

The most important theorem in the theory of optimal control is Pontrjagin's Minimum Principle (theorem 2.1). But sometimes this theorem is useless. This is especially the case if singular optimal controls appear. Inspired by applications in aerospace the first investigations of such controls were done in the 1950s. Thereupon many works were published in which singular optimal controls and their properties were investigated. As well analytical methods as also numerical methods for solving problems of such kind were given. In this work a digest of the state of the art is given. This paper is especially adressed to people who try to solve problems which might have a partially singular optimal control. Examining all the literature will not be necessary any more.

In section 2 the fundamentals of the theory of optimal control which are needed to understand this work are presented. A definition of singular controls is given. Section 3 considers the junctions between singular and nonsingular parts of optimal controls. While onedimensional controls are considered in this section, in section 4 multidimensional controls are examined. A counterexample which shows that a

*University of Kaiserslautern, Department of Mathematics, 67653 Kaiserslautern, Germany

property of onedimensional controls can't be transfered to multidimensional controls is given. Optimality conditions are treated in section 5. In section 6 the usual numerical methods are listed. A strategy for chosing the method which fits most to the considered problem is proposed. Finally in section 7 some supplements are given and in section 8 conclusions are presented.
We declare that 0 is a natural number. The abbreviation f_x is used for the partial derivative of an arbitrary function f with respect to x. If x is a vector and f a scalar function then f_x is the gradient written as a column vector, i. e. ∇_x f. If f is a vectorial function then f_x is the Jacobian matrix. The (total) derivative of x with respect to t is denoted by $x'(t)$. By $H_u^{(k)}$ we denote the k-th total time derivative of H_u (see section 3.2). Every integral is a Lebesqueian integral. The interior of a set M is denoted by *int* M.

2 Fundamentals of the Theory of Optimal Control

2.1 The Bolza Problem

The following problem is called the Bolza problem:
(P) A system be given by
1) state $x(t) \in \mathbb{R}^n$ at time $t \in [t_0, t_f] =: T$,
2) control $u(t) \in U(t) \subset \mathbb{R}^s$, where u **is piecewise continuous** and $U(t)$ is compact for every $t \in T$,
3) state equation $x'(t) = f(t, x(t), u(t))$ almost everywhere,
4) initial condition $x(t_0) = x_0$ and **final condition** $x(t_f) \in Z \neq \emptyset$, where (t_0, t_f) is fixed,
5) scalar functions **g** and **L** with suitable domain.
A pair (x, u) which satisfies the above conditions 1) to 4) is called **admissible pair**, and u is called admissible control if condition 2) is satisfied.
The problem is:

Find an admissible pair (x,u) on T such that the functional
$J(u) = g(x(t_f)) + \int_{t_0}^{t_f} L(\tau, x(\tau), u(\tau))d\tau$ **becomes minimal.**

Here t may also be a generalized parameter and need not be the time such that systems with variable final time can be represented. If $g \equiv 0$ we call **(P)** a Lagrange problem while **(P)** is called a Mayer problem if $L \equiv 0$.

2.2 Pontrjagin's Minimum Principle

Before formulating the theorem we define the Hamiltonian:

Definition 2.1 $H(u, x, \lambda, t) := L(t, x, u) + \lambda^T f(t, x, u)$ *is called* ***Hamiltonian****, where* $\lambda(t) \in \mathbb{R}^n$ *holds.*

Definition 2.2 *Let* (x^*, u^*) *be a solution of a Bolza problem and* u *be an arbitrary admissible control. Then* $\Delta H(u, x^*, \lambda, t) := H(u, x^*, \lambda, t) - H(u^*, x^*, \lambda, t)$
[1]

Sometimes the Hamiltonian is more generally defined as $H(u, x, \lambda, t) := \lambda_0 L(t, x, u) + \lambda^T f(t, x, u)$, where λ_0 is a constant. Some authors use constants $\lambda_0 < 0$. The resulting Hamiltonian has in the case $\lambda_0 = -1$ the same absolute value but a different sign. This has to be kept in mind as some theorems change in this case. The Minimum Principle e. g. becomes a Maximum Principle.

The following theorem is often very useful when trying to solve Bolza problems:

Theorem 2.1 (Pontrjagin's Minimum Principle)
Condition: *Every f_i $(i = 1, ..., n)$ is continuous in (t, x, u). The derivatives $\frac{\partial}{\partial t} f_i$ and $\nabla_x f_i$ exist and are continuous in (t,x,u) for every i. Further on $g \in C^1$.* ***(P)*** *has a solution (x^*, u^*). $Z = \mathbb{R}^n$.*
Statement: *Then there exists an absolutely continuous function $\lambda : T \to \mathbb{R}^n$ with the following properties:*
a) $x' = H_\lambda$ and $\lambda' = -H_x$ along (x^*, u^*)
b) $H(u^*(t), x^*(t), \lambda(t), t) = min\{H(u, x^*(t), \lambda(t), t) \mid u \in U(t)\}$ for every $t \in T$
c) $\lambda \neq 0$ on T
d) $\lambda(t_f)dx(t_f) - dg = 0$ (transversality condition) [2], [3], [4]

Definition 2.3 *A triple (x^*, u^*, λ) is called* ***extremal,*** *if (x^*, u^*) is admissible and the equations $x' = H_\lambda$ and $\lambda' = -H_x$ hold along (x^*, u^*).*

We consider the following example for illustrating the definitions and theorems given in this work:

Example 2.1 *Let $x'(t) = u(t)$ with the boundary conditions $x(0) = 0$ and $x(2) = -1$, where $t_0 = 0, t_f = 2$ and $U \equiv [-1, 1]$ hold. The functional $\int_0^2 x^2(t)(u(t) + 1)dt$ has to be minimized.*
According to definition 2.1 the Hamiltonian is given by
$H(u, x, \lambda, t) = x^2(u+1) + \lambda\, u$. According to theorem 2.1 a) the differential equation $\lambda'(t) = -2x(t)(u(t) + 1)$ holds. Theorem 2.1 b) states that H is minimal for the optimal control u^*. This implies that the optimal control - if it exists - has the following form:

$$u^* = \begin{cases} +1 & , \quad if\ \lambda + x^2 < 0 \\ -1 & , \quad if\ \lambda + x^2 > 0 \\ ? & , \quad if\ \lambda + x^2 = 0 \end{cases}$$

If $\lambda + x^2$ disappears we get $H_u = 0$, i. e. there might be an extremum in the interior of $U \equiv [-1, 1]$. But Pontrjagin's Minimum Principle doesn't help us to find the optimal control. Such controls are called singular.

Another example is the Goddard problem (see e. g. [5]).

2.3 Singular Controls

In the above examples, which are both linear in u, we find that u^* has values on the boundaries of U if H_u does not disappear. If $H_u = 0$, i. e. if H does not explicitly depend on u, u^* can't be determined so easily. This problem doesn't matter if H_u has only isolated roots. A zero set of undetermined control values is neglectable. But a serious problem occurs if H_u disappears on a whole interval. In this interval we have to clarify how u^* looks like. Conventional conditions don't suffice in this case.
This situation results in the following definitions:

Definition 2.4 *Let* ***(P)*** *be given and $T_1 = [a,b] \subset T$, where $a < b$. A control u for which the property $u(t) \in \{\ min\ U,\ max\ U\ \}$ holds in the interior of the whole interval T_1 is called* ***bang bang control*** *on T_1.*

The definitions of singularities are very different in literature. We propose the following compromise (see e. g. [3], [6], [7] and [1]):

Definition 2.5 *Let (P) be given, (x^*, u^*, λ) be an extremal and $[a,b] \subset T$ $(a<b)$.*
a) A control u^* is called ***singular control in the sense of the Minimum Principle*** *on [a,b], if there exists a set $W(t) \subset U(t)$ with several elements such that*

$$\Delta H(u, x^*(t), \lambda(t), t) = 0 \tag{1}$$

holds for every $u \in W(t)$ and every $t \in [a,b]$.
b) Let H be linear in u. Then u^* is called singular control in the classical sense or ***classically singular control*** *on [a,b] if*

$$H_u(u^*(t), x^*(t), \lambda(t), t) = 0 \tag{2}$$

holds for every $t \in]a,b[$.
c) Let H be linear in the component u_k of u. Then u_k^* is classically singular on [a,b] if

$$H_{u_k}(u^*(t), x^*(t), \lambda(t), t) = 0 \tag{3}$$

holds for every $t \in]a,b[$.
d) Let H be non-linear in every component of u. Then u^* is called classically singular control on [a,b] if

$$det(H_{uu})(u^*(t), x^*(t), \lambda(t), t) = 0 \tag{4}$$

holds for every $t \in]a,b[$.
e) Let H be non-linear in u_k. Then u_k is called classically singular on [a,b] if

$$H_{u_k u_k}(u^*(t), x^*(t), \lambda(t), t) = 0 \tag{5}$$

holds for every $t \in]a,b[$.

The expression **purely singular** is used if $[a,b] = T$. In this work the designation **singular** refers to singularity in the classical sense. If H is linear in u resp. u_k the equations (4) resp. (5) hold automatically. The following sections treat conditions which are supposed to enable us to determine optimal controls which have a singular part.

3 Junctions between bang bang and singular control

3.1 Introduction

Definition 3.1 *A function $g : [a,b] \to \mathbb{R}$ is called* ***analytical*** *on (a,b) if there exists a point $\tau \in (a,b)$ in which g can be developed into a Taylor expansion.*
It is called ***piecewise analytical*** *if the following property holds: For every $t_c \in (a,b)$ there exist $t_1 \in (a, t_c)$ and $t_2 \in (t_c, b)$ such that in every point of (t_1, t_c) and (t_c, t_2) g can be developed into a Taylor expansion which converges in the whole associate interval.*

In this section we restrict ourselves to the following linear problem:

> **(LP)** $x'(t) = f_0(t, x(t)) + f_1(t, x(t))\, u(t)$ almost everywhere,
> **1)** and **4)** like in (P).
> $u(t) \in [-K(t), K(t)] \subset \mathbb{R}$, where $K > 0$ is an analytical function.
> Minimize $J(u) = g(x(t_f)) + \int\limits_{t_0}^{t_f} [L_0(\tau, x(\tau)) + L_1(\tau, x(\tau)) u(\tau)] d\tau$, where
> the solution u^* has to be piecewise analytical.
> The functions f_0, f_1, L_0 and L_1 are analytical in both arguments.

We assume in this section that a solution exists and that in the singular control part the control is situated in the interior of $U(t)$ for almost every t.

Optimal controls often consist of nonsingular and singular parts. The junctions between these parts are a part of the research in singular optimal control theory. But the examination of the junctions is not very advanced, which is one of the main reasons for the difficulties in the investigations of singular optimal controls.

3.2 Problem and Arc Order

Definition 3.2 *Let (LP) be given. Then $\Phi(t) := H_u(u(t), x(t), \lambda(t), t)$ is called* ***switching function****.*

There are two concepts of order which resulted from ambiguous definitions given for the order in the beginning. Unfortunately there are authors who do not discern between these concepts. A first clarification was undertaken by Lewis ([8]) and later on repeated by Powers ([9]), who used different names. We use the denomination suggested by Powers. The names suggested by Lewis are given in brackets.

Definition 3.3 *Let (LP) be given. Let $\Phi^{(2q)}$ be the first derivative with respect to t of Φ (resp. H_u) to contain u explicitly, where after each differentiation x' is replaced by f and λ' by $-H_x$. Then q is called* ***problem order*** *(intrinsic order). If there's no such derivative, we define $q := +\infty$.*

Definition 3.4 *Let (LP) be given. The* ***arc order*** *(local order) of an extremal (x^*, u^*, λ) on a subinterval (t_c, t_d) is the smallest number p to satisfy*

$$\left(\frac{d}{du}\frac{d^{2p}}{dt^{2p}}H_u\right)(u^*, x^*, \lambda, t) = 0 \tag{6}$$

for every $t \in]t_c, t_d[$, where the derivatives are computed like in definition 3.3. If there's no such derivative, we define $p := +\infty$.

The denotation stresses that the problem order is given by the problem itself, whereas the arc order depends on the chosen extremal. Different extremals of the same problem can have different arc orders. Powers gives such an example ([9]).
By $H_u^{(k)}$ we will denote the k-th total derivative of H_u with respect to t. It's easy to see that $H_u^{(2p)}$ is a polynomial in u of order $2(p-q)+1$, i. e. $H_u^{(2q)}$ is linear in u. ([8])

Definition 3.5 *We write $H_u^{(2q)}(u, x, \lambda, t) =: A(x, \lambda, t) + u\ B(x, \lambda, t)$, $\alpha(t) := A(x(t), \lambda(t), t)$ and $\beta(t) := B(x(t), \lambda(t), t)$.*

We will now prove that p and q are natural numbers, i. e. $2p$ and $2q$ are even. Robbins was the first to give a proof for this. It's obvious that $p \geq q > 0$ holds. First we need a lemma, whose proof consists of easy calculations and can be found in [5].

Lemma 3.1 *Let $F(x, \lambda)$ and $G(x, \lambda)$ be scalar C^2-functions, where $x' = H_\lambda$ and $\lambda' = -H_x$ hold. Then the following equalities hold:*

$$\frac{d}{dt}F(x, \lambda) = -(\nabla H)^T S\ \nabla F \tag{7}$$

$$\frac{d}{dt}(\nabla F) = \nabla\left(\frac{d}{dt}F\right) + [\nabla(\nabla H)^T]\ S\ \nabla F \tag{8}$$

$$\frac{d}{dt}[(\nabla F)^T S\ \nabla G] = \left(\nabla\frac{d}{dt}F\right)^T S\ \nabla G + (\nabla F)^T S\ \nabla\frac{d}{dt}G, \tag{9}$$

where

$$\nabla = \left(\frac{\partial}{\partial x_1}, ..., \frac{\partial}{\partial x_n}, \frac{\partial}{\partial \lambda_1}, ..., \frac{\partial}{\partial \lambda_n} \right)^T$$

and

$$S = \begin{pmatrix} 0_n & Id_n \\ -Id_n & 0_n \end{pmatrix} \begin{matrix} \}n & (Id_n: \; n \times n \textit{ identity matrix}) \\ \}n & (0_n: \; n \times n \textit{ zero matrix}) \end{matrix}$$

hold.

With this lemma we can prove

Theorem 3.1 *Let $\gamma_k := (H_u^{(k)})_u$ and $\gamma_k \equiv 0$ for every $k = 1, ..., p-1$. If p is odd then $\gamma_p \equiv 0$. (Here p need not be the arc order.)*

Proof: Generally $\gamma_0 \equiv 0$ holds as H is linear in u.
According to Lemma 3.1 (7) we get:

$\gamma_p = (H_u^{(p)})_u = (\frac{d}{dt}(H_u^{(p-1)}))_u = -[(\nabla H)^T S \; \nabla(H_u^{(p-1)})]_u =_{\text{(product rule)}}$
$= -[(\nabla H)^T]_u S \; \nabla(H_u^{(p-1)}) - (\nabla H)^T S \; [\nabla(H_u^{(p-1)})]_u =$
$= -[\nabla(H_u)]^T S \; \nabla(H_u^{(p-1)}) - (\nabla H)^T S \; \nabla[(H_u^{(p-1)})_u] =$
Using 3.1 (7) once again we get:
$= -[\nabla(H_u)]^T S \; \nabla(H_u^{(p-1)}) + \frac{d}{dt}[(H_u^{(p-1)})_u] =$
$= -[\nabla(H_u)]^T S \; \nabla(H_u^{(p-1)}) + \frac{d}{dt}(\gamma_{p-1})$

$$\Rightarrow \gamma_p = -[\nabla(H_u)]^T S \; \nabla(H_u^{(p-1)}), \tag{10}$$

as $\frac{d}{dt}(\gamma_{p-1}) \equiv 0$ holds by assumption.

In the case $p = 1$ we can immediately continue with equation (15) ($p = 2\nu + 1$, so $\nu = 0$), which then coincides with equation (10). If $p > 1$ further examinations are necessary: Analogously to (10) we get

$$\gamma_{p-1} = -[\nabla(H_u)]^T S \; \nabla(H_u^{(p-2)}). \tag{11}$$

Deriving (11) with respect to t and using Lemma 3.1 (9) one gets:

$$\frac{d}{dt}\gamma_{p-1} = -[\nabla(H_u^{(1)})]^T S \; \nabla(H_u^{(p-2)}) - [\nabla(H_u)]^T S \; \nabla(H_u^{(p-1)}) = 0 \text{ (by ass.)} \tag{12}$$

Analogously:

$$\frac{d}{dt}\gamma_{p-2} = -[\nabla(H_u^{(1)})]^T S \; \nabla(H_u^{(p-3)}) - [\nabla(H_u)]^T S \; \nabla(H_u^{(p-2)}) = 0,$$

where the last summand dissapears according to (11).
So we get according to Lemma 3.1 (9):

$$\left(\frac{d}{dt}\right)^2 \gamma_{p-2} = -[\nabla(H_u^{(2)})]^T S \; \nabla(H_u^{(p-3)}) - [\nabla(H_u^{(1)})]^T S \; \nabla(H_u^{(p-2)}) = 0 \text{ (by ass.)} \tag{13}$$

With induction we get for $p > k \geq 1$:

$$\left(\frac{d}{dt}\right)^k \gamma_{p-k} = -[\nabla(H_u^{(k)})]^T S \; \nabla(H_u^{(p-k-1)}) - [\nabla(H_u^{(k-1)})]^T S \; \nabla(H_u^{(p-k)}) = 0 \tag{14}$$

We now assume that p is odd, i. e. p has the shape $p = 2\nu + 1$, $\nu \in \mathbb{N}$. By (10) and (12) we get:

$$\gamma_p = -[\nabla(H_u)]^T S\ \nabla(H_u^{(p-1)}) = [\nabla(H_u^{(1)})]^T S\ \nabla(H_u^{(p-2)}) =_{by\ (13)}$$

$$= -[\nabla(H_u^{(2)})]^T S\ \nabla(H_u^{(p-3)})$$

Hence for $\nu \in \{1,2\}$ we have shown (15). Else we get after applying (14) $(\nu-2)$ times $(p-3 = 2\nu-2,\ (-1)^{\nu-2} = (-1)^{\nu})$ the following equation:

$$\gamma_p = (-1)^{\nu+1}[\nabla(H_u^{(\nu)})]^T S\ \nabla(H_u^{(\nu)}) \tag{15}$$

(In [5] we find $(-1)^\nu$, which is an error.)
Now chose an arbitrary $z \in \mathbb{R}^{2n}$. Then $z^T S z = 0$ holds, as $Sz = (z_{n+1}, ..., z_{2n}, -z_1, ..., -z_n)$ and $z^T S z = z_1 z_{n+1} + z_2 z_{n+2} + ... + z_n z_{2n} - z_{n+1} z_1 - z_{n+2} z_2 - ... - z_{2n} z_n = 0$ hold. If we chose $z = \nabla(H_u^{(\nu)})$, we get using (15) the equality $\gamma_p = 0$, which holds for every t, so $\gamma_p \equiv 0$. q. e. d. ([5] and [10])

Corollary 3.1 *If an **(LP)** with finite problem order q is given, then q is a positive natural number. If additionally the arc order p of a solution is finite, then p is a positive natural number, too.*

The corollary is an immediate result of theorem 3.1.
We define:

Definition 3.6 *An **(LP)** with solution (x^*, u^*) and junction t_c be given. Then $\boldsymbol{r} \geq 0$ be the smallest order of derivation such that ${u^*}^{(r)}$ is discontinuous in t_c and $\boldsymbol{m} \geq 0$ be the smallest order of derivation such that $\beta^{(m)}(t_c) = 0$ does <u>not</u> hold.*

Example 3.1 *We consider example 2.1:*
We have $H(u, x, \lambda, t) = x^2(u+1) + \lambda u$, $\lambda' = -2x(u+1)$ and $x' = u$. Hence $\Phi(t) = x^2(t) + \lambda(t)$. So we get: $\Phi'(t) = 2x(t)x'(t) + \lambda'(t) = 2x(t)u(t) - 2x(t)[u(t) + 1] = -2x(t) \Rightarrow \Phi''(t) = -2x'(t) = -2u(t)$.
Hence our example has the odd problem order $q = 1$, which here coincides with the arc order p. Furtheron we get $\alpha \equiv 0$ and $\beta \equiv -2$, thus $m = 0$. As we don't know the solution we can't compute r.

3.3 Junction Conditions

An important theorem in singular optimal control theory is

Theorem 3.2 (Generalized Legendre-Clebsch-Condition (GLC))
*If (x^*, u^*, λ) is an optimal extremal of **(LP)** then the following properties hold:*
a) If the problem order q is finite then $(-1)^q \beta(t) \geq 0$ for every $t \in T$.
b) If the arc order p is finite on (t_a, t_b) then $(-1)^p \beta(t) > 0$ for every $t \in (t_a, t_b)$.

The theorem was first proved by Robbins (1966). The original proof can be found in [10], another proof can be read in [5]. The classical Legendre-Clebsch-Condition states that $H_{uu} \geq 0$ holds, which is here trivially satisfied as H is linear in u. The GLC is also true for non-linear Bolza problems with sufficient differentiability, where q is defined analogously and β is defined to be $(H_u^{(2q)})_u$ (see [10] and [8]).

Example 3.2 *In example 2.1 we get $(-1)^q \beta(t) = (-1)^p \beta(t) = 2 > 0$. This is satisfied for every extremal.*

Definition 3.7 *Let **(P)** be given. A junction is called **non-analytical junction** if the control u^* is not piecewise analytical in any neighbourhood of the junction.*

The following junction condition was proved by McDanell and Powers in 1971. A detailed version of the proof is given here once again because we need the proof for further considerations.

Theorem 3.3 (1st Junction Cond. by McDanell and Powers (1971))
*The optimal control u^* of an **(LP)** be piecewise analytical in a neighbourhood of the junction t_c. The problem order q be finite and the strict GLC $(-1)^q \cdot \beta(t_c) > 0$ hold.*
Then $q + r$ is odd.

Proof: If we derive H with respect to every variable, then we get sums and products consisting of derivates of the analytical functions f_0, f_1, L_0 and L_1 as well as derivatives of λ with $\lambda' = -H_x$ and derivatives of u. Thus α and β are at least continuous in a neighbourhood of t_c.
As $u^{(\nu)}$ can only appear in $\alpha^{(j)}$ and $\beta^{(j)}$ for $0 < j \leq r$ if $\nu \leq r - 1$, α and β are r times continuously differentiable in t_c. Analogously $\Phi \equiv H_u$ has exactly $2q + r - 1$ continuous derivatives with respect to t in t_c, as $H_u^{(2q+r)}$ contains the term $\beta \cdot u^{(r)}$, which is discontinuous in t_c according to definition 3.6 and condition $\beta(t_c) \neq 0$. We define

$$k := 2q + r. \tag{16}$$

We chose a real number $\epsilon \neq 0$ arbitrarily close to 0 such that $t_c + \epsilon$ is situated in the non-singular part and $t_c - \epsilon$ is situated in the singular part. The sign of ϵ is given by the sequence of both parts. We use the denotations u_n for the control in the non-singular part and u_s for the control in the singular part. Another denotation is:

$$u_n^{(i)}(t_c) := \lim_{\epsilon \to 0} u^{(i)}(t_c + \epsilon),\ u_s^{(i)}(t_c) := \lim_{\epsilon \to 0} u^{(i)}(t_c - \epsilon),$$

where each limes is taken in the associate domain.
We now develop $\Phi(t_c + \epsilon)$ into a Taylor expansion around t_c: In the singular part $\Phi \equiv 0$ holds. Because of (16) we consequently get $\Phi^{(j)}(t_c) = 0$ for every $j = 0, ..., k-1$. Moreover we get $\lim\limits_{\epsilon \to 0} \Phi^{(k)}(t_c + \epsilon) \neq 0$, where according to definitions 3.5 and 3.6 the equality $\Phi^{(k)} \equiv \left(\frac{d}{dt}\right)^r [\alpha + \beta u]$ holds.
Thus according to the Taylor theorem we get the equation

$$\Phi(t_c + \epsilon) = \frac{\epsilon^k}{k!} \left[\alpha^{(r)}(t_c) + \sum_{i=0}^{r} \binom{r}{i} \beta^{(r-i)}(t_c) \cdot u_n^{(i)}(t_c) \right] + o(\epsilon^k) \tag{17}$$

In the singular part $\Phi^{(2q)} \equiv \alpha + \beta u_s \equiv 0$ holds, hence $\alpha \equiv -\beta u_s$.
Thus we get the following equation for the singular part:

$$\alpha^{(r)} = \left(\frac{d}{dt}\right)^r (-\beta u_s) = -\sum_{i=0}^{r} \binom{r}{i} \beta^{(r-i)} \cdot u_s^{(i)} \tag{18}$$

We substitute (18) in (17) and use the continuity of $\alpha^{(r)}$ in t_c. The result is:

$$\Phi(t_c + \epsilon) = \frac{\epsilon^k}{k!} \sum_{i=0}^{r} \binom{r}{i} \beta^{(r-i)}(t_c) \cdot \left[u_n^{(i)}(t_c) - u_s^{(i)}(t_c) \right] + o(\epsilon^k). \tag{19}$$

If $r > 0$ we get by definition 3.6 for $i = 0, 1, ..., r-1$

$$u_n^{(i)}(t_c) = u_s^{(i)}(t_c) \tag{20}$$

Thus in equation (19) only the summand for $i = r$ is left, and we consequently get for every $r \geq 0$ the equation

$$\Phi(t_c + \epsilon) = \frac{\epsilon^k}{k!} \cdot \beta(t_c) \cdot \left[u_n^{(r)}(t_c) - u_s^{(r)}(t_c) \right] + o(\epsilon^k). \tag{21}$$

We now define $\sigma := -sign\Phi(t_c + \epsilon)$.

$$\Rightarrow u_n(t) = \sigma \cdot K(t) \Rightarrow u_n^{(i)}(t_c) = \sigma \cdot K^{(i)}(t_c) \text{ for every } i = 0, 1, ..., r \tag{22}$$

We now develop a Taylor sum in the singular part and get

$$\sigma K(t_c - \epsilon) - u(t_c - \epsilon) = \sum_{i=0}^{r} \frac{(-\epsilon)^i}{i!} \left[\sigma K^{(i)}(t_c) - u_s^{(i)}(t_c)\right] + o(\epsilon^r) \tag{23}$$

Using (20), (22) and (23) we get:

$$\sigma K(t_c - \epsilon) - u(t_c - \epsilon) = \frac{(-1)^r \epsilon^r}{r!} \left[u_n^{(r)}(t_c) - u_s^{(r)}(t_c)\right] + o(\epsilon^r) \tag{24}$$

This is an important equation, to which we will refer later on.

We substitute equation (24) in (21):

$$\Phi(t_c + \epsilon) = \frac{\epsilon^{2q+r}}{k!} \beta(t_c) \frac{(-1)^r r!}{\epsilon^r} \left[\sigma K(t_c - \epsilon) - u(t_c - \epsilon) - o(\epsilon^r)\right] + o(\epsilon^k)$$

We summarize and get ($\epsilon^{2q} \cdot o(\epsilon^r) = o(\epsilon^k)$):

$$\Phi(t_c + \epsilon) = \frac{\epsilon^{2q} r!}{k!} (-1)^r \beta(t_c) \left[\sigma K(t_c - \epsilon) - u(t_c - \epsilon)\right] + o(\epsilon^k) \tag{25}$$

In the non-singular part two cases appear, which we want to discern:
(i) $\Phi(t_c + \epsilon) > 0 \Rightarrow \sigma = -1$, (ii) $\Phi(t_c + \epsilon) < 0 \Rightarrow \sigma = +1$.
Thus according to equation (25) we get for ϵ sufficiently close to zero in each case:
(i) $(-1)^r \cdot \epsilon^{2q} \cdot \beta(t_c) \left[-K(t_c - \epsilon) - u(t_c - \epsilon)\right] > 0 \; / \cdot (-1)$
(ii) $(-1)^r \cdot \epsilon^{2q} \cdot \beta(t_c) \left[+K(t_c - \epsilon) - u(t_c - \epsilon)\right] < 0$

Altogether the following unequation holds:

$$(-1)^r \cdot \epsilon^{2q} \cdot \beta(t_c) \left[K(t_c - \epsilon) \pm u(t_c - \epsilon)\right] < 0 \tag{26}$$

By assumption the GLC is strictly satisfied: $(-1)^q \cdot \beta(t_c) > 0$. We now multiply unequation (26) with this positive term and get:

$$(-1)^{q+r} \cdot \epsilon^{2q} \cdot \beta^2(t_c) \left[K(t_c - \epsilon) \pm u(t_c - \epsilon)\right] < 0 \tag{27}$$

As $|u(t)| \leq K(t)$ holds for every t and $|u(t)| < K(t)$ holds almost everywhere in the singular part (see **(LP)**), there exists an ϵ sufficiently close to 0 such that $K(t_c - \epsilon) \pm u(t_c - \epsilon) > 0$ holds. Two further terms in (27) are positive: $\epsilon^{(2q)} = (\epsilon^q)^2$ and $\beta^2(t_c)$. The second term is different from zero because of the strict GLC. Thus the unequation (27) gets the shape $(-1)^{(q+r)} \cdot \omega < 0$, where $\omega > 0$ holds. Consequently the sum $q + r$ must be odd. q.e.d. [11]

The use of this theorem is demonstrated by:

Corollary 3.2 *Under the assumptions of theorem 3.3 the following properties hold:*
a) If q is even, then u^* is continuous in the junction.
b) If q is odd, then u^* either has a jump in the junction or is continuously differentiable in the junction.
[11]

The corollary is an immediate result of theorem 3.3.

The assumed piecewise analyticity of the solution is a signifying disadvantage of theorem 3.3 and corollary 3.2. Especially in the non-singular part this is often not the case. That's why McDanell and Powers proved ([11]) another junction condition with essentially weaker requirements on u^*.

Theorem 3.4 (2nd Junction Cond. by McDanell and Powers (1971)) *An **(LP)** with finite problem order q be given, where we only require that the optimal control is piecewise continuous in the singular part. If u^* is such a solution with junction t_c, then the following properties hold:*
(i) If the one-sided limes of $H_u^{(2q)}$ in t_c, taken from the non-singular side, does not disappear, u^* is discontinuous in t_c.
(ii) If $A = 0$ and $B \neq 0$ in t_c, then u^* is discontinuous in t_c.
(iii) If u^* is also piecewise continuous in the non-singular part and $\beta(t_c) \neq 0$ holds and the one-sided limes of $H_u^{(2q)}$ in t_c, taken from the non-singular side, disappears, then u^* is continuous in t_c.

Corollary 3.3 *An **(LP)** with finite problem order q be given. A solution u^* with junction t_c exist. If q is even and $A \equiv 0$ and $\beta(t_c) \neq 0$ then the junction is non-analytical.*

The proof of corollary 3.3 can also be found in [11].
Of course, the question arises if the assumption, that u^* be piecewise analytical, in theorem 3.3 and corollary 3.2 is really necessary. The answer is given by the Fuller problem:

Example 3.3 (Fuller Problem)
minimize $\frac{1}{2}\int\limits_0^{t_f} x_1^2(t)dt$ *(t_f sufficiently large), where*
$x_1' = x_2,\ x_1(0) = \xi_1 \neq 0,\ x_2' = u,\ x_2(0) = \xi_2,\ K \equiv 1$
The Hamiltonian is $H = \frac{1}{2}x_1^2 + \lambda_1 x_2 + \lambda_2 u$.
Hence we get $\lambda' = -H_x = (-x_1, -\lambda_1)^T$.
$H_u = \lambda_2 \Rightarrow H_u^{(1)} = \lambda_2' = -\lambda_1 \Rightarrow H_u^{(2)} = -\lambda_1' = x_1 \Rightarrow H_u^{(3)} = x_1' = x_2 \Rightarrow H_u^{(4)} = x_2' = u$.
Thus the Fuller problem has the even problem order $q = 2$, which coincides with the arc order. Furtheron we get $A \equiv 0$ and $B \equiv 1$, thus $m = 0$. Hence the strict GLC holds: $(-1)^q\beta(t) = (-1)^2 \cdot 1 = 1 > 0$. According to corollary 3.3 every junction is non-analytical. This means that every assumption of theorem 3.3 and corollary 3.2a) is satisfied except the piecewise analyticity of the solution.
The optimal control u^* has exactly one junction t_c, where the control is at first bang bang and then singular. It's easy to prove that u^* is discontinuous in t_c, thus $r = 0$, but this means that $q + r$ is even.
Consequently in theorem 3.3 and corollary 3.2a) piecewise analytical controls are needed.
[11], [12]

We consider again example 2.1:

Example 3.4 *We know already that $q = 1$ holds and the GLC is strictly satisfied. If there exists a piecewise analytical solution, then r is even according to theorem 3.3 respectively corollary 3.2b). Consequently u^* is either discontinuous or continuously differentiable in the junction, if one exists. Assuming that a junction exists we get according to theorem 3.4(i) or (ii) that u^* is discontinuous in the junction. Part*

(iii) and corollary 3.3 are not applicable.
Thus $q = 1$, $r = 0$, $m = 0$ holds. Hence these parameters could be derived without knowledge of the solution. Consequently the question arises if there's a general connection between q, r and m.

3.4 McDanell's Conjecture

McDanell and Powers examined a possible connection between the three parameters in their publication ([11]). But their proof contained an error. As up to now it was neither possible to prove their theorem completely nor to disprove it completely, the following conjecture is left:

Conjecture 3.1 (McDanell's Conjecture)
*An **(LP)** with $m \geq 2$ and optimal control u^* which is piecewise analytical in the neighbourhood of a junction t_c be given. The parameters q, r and m be finite. Then the following properties hold:*
(i) If $m \leq r$ then $q + r + m$ is odd.
[(ii) If $m > r$ then $-sign[\beta^{(m)}(t_c+) \cdot \beta^{(m)}(t_c-)] = (-1)^{q+r+m}$]

For $m = 0$ we have the already proved theorem 3.3.
For $m = 1$ the conjecture was proved by Bell and Boissard (1979) and by Bortrins (1983).
For $m = 2$, $r = 0$ and $q = 1$ Pan and Bell gave a counterexample in 1987. That's why part (ii) has only been put into brackets in this paper. The validity of part (i) is also being doubted.
If the conjecture was true this would not be a big help. The fact that this criterion was investigated during the last 20 years without any big success shows how less advanced the research in this section is.
We will give a short sketch of the actual knowledge about McDanell's conjecture.

Lemma 3.2 *An **(LP)** with finite problem order q be given.*
Then the functions α and β are continuous.
If there exists a junction t_c of an optimal solution, then α and β have at least r continuous derivatives in t_c. If additionally $m = 0$, then Φ has exactly $2q + r - 1$ continuous derivatives in t_c.

A proof for this lemma has already been given at the beginning of the proof of theorem 3.3.

Definition 3.8 *An **(LP)** with $q < +\infty$ be given and $H_u^{(2q)}(u, x, \lambda, t) = A(x, \lambda, t) + uB(x, \lambda, t)$. Then we define:*

$$\frac{d}{dt}A(x, \lambda, t) = a_0(x, \lambda, t) + u\ a_u(x, \lambda, t) \text{ and} \tag{28}$$

$$\frac{d}{dt}B(x, \lambda, t) = b_0(x, \lambda, t) + u\ b_u(x, \lambda, t). \tag{29}$$

Lemma 3.3 *An **(LP)** with optimal extremal (x^*, u^*, λ) and $q = 1$ be given. Then $a_u \equiv b_0$.*

The explicitly written proof is very long. We refer to Bell and Boissard ([13]).

The following lemma completes lemma 3.2.

Lemma 3.4 *Let (x^*, u^*, λ) be an optimal extremal of an **(LP)** with $q = 1$, where t_c is a junction. Further on $m > 1$ hold and $\beta^{(r+1)}$ be continuous in t_c. Then $\alpha^{(r+1)}$ is continuous in t_c.*

Again we refer to Bell and Boissard ([13]).

Corollary 3.4 *Let (x^*, u^*, λ) be an optimal extremal of an **(LP)** with $q = 1$, where t_c is a junction. If $m > r + 1$ then $\alpha^{(r+1)}$ is continuous in t_c.*

It's easy to see that lemma 3.4 can be applied ([13]).

Remark 3.1 *Now Bell and Boissard ([13]) assert that in the case $q = 1$, $m > r+1$ McDanell's proof ([11]) can be used. But this is wrong, because the existence of $\alpha^{(r+m)}$ and $\beta^{(r+m)}$ in t_c is needed for this argumentation. This shows the following consideration:*
We start like in the proof of theorem 3.3 and define ϵ, k, σ, $u_n^{(i)}$ and $u_s^{(i)}$ in the same way. Analogously we define $\beta_n^{(i)}$, $\beta_s^{(i)}$, $\alpha_n^{(i)}$ and $\alpha_s^{(i)}$.
McDanell and Powers propose in their proof: "The proof is similar to that for Theorem 1 [= theorem 3.3 in this paper]; however, in order to obtain a nontrivial term in the Taylor series expansion for $\Phi(t_c + \epsilon)$, one must consider higher order terms ..." ([11]). The unequation that follows afterwards will be considered later in this remark.
In order to achieve a suitable analogon for equation (21) of the shape

$$\Phi(t_c + \epsilon) = \frac{\epsilon^{k+m}}{(k+m)!} \cdot \beta^{(m)}(t_c) \cdot \left[u_n^{(r)}(t_c) - u_s^{(r)}(t_c)\right] \left(\begin{array}{c} r+m \\ r \end{array} \right) + o(\epsilon^{k+m}) \quad (30)$$

one has to start in (17) with

$$\Phi(t_c+\epsilon) = \sum_{j=k}^{k+m} \frac{\epsilon^j}{j!} \left[\alpha^{(j-2q)}(t_c) + \sum_{i=0}^{j-2q} \left(\begin{array}{c} j-2q \\ i \end{array} \right) \beta_n^{(j-2q-i)}(t_c) \cdot u_n^{(i)}(t_c)\right] + o(\epsilon^{k+m}) \quad (31)$$

in the non-singular part. Here (31) is a Taylor series expansion for Φ around t_c in the nonsingular part. The series is considered up to the $(k+m)$-th summand. Mind that $\Phi^{(2q)} = \alpha + u\,\beta$ holds. Analogously to (18) we get in the singular part

$$\alpha_s^{(j)} = \frac{d^j}{dt^j}[-\beta_s u_s] = -\sum_{i=0}^{j} \left(\begin{array}{c} j \\ i \end{array} \right) \beta_s^{(j-i)} \cdot u_s^{(i)} \quad (32)$$

for every $j = r, r+1, ..., r+m$.
In the proof of theorem 3.3 we now substitute (18) in (17). In order to be able to do this also here, we need the $(r+m)$-times continuous differentiability of α in t_c. Then we get:

$$\begin{aligned} \Phi(t_c + \epsilon) &= \sum_{j=k}^{k+m} \frac{\epsilon^j}{j!} \sum_{i=0}^{j-2q} \left(\begin{array}{c} j-2q \\ i \end{array} \right) \left[\beta_n^{(j-2q-i)}(t_c) u_n^{(i)}(t_c) - \beta_s^{(j-2q-i)}(t_c) u_s^{(i)}(t_c)\right] \\ &+ o(\epsilon^{k+m}) \end{aligned}$$

An index transformation $j \to j + k$ results in:

$$\begin{aligned} \Phi(t_c + \epsilon) &= \sum_{j=0}^{m} \frac{\epsilon^{j+k}}{(k+j)!} \sum_{i=0}^{r+j} \left(\begin{array}{c} r+j \\ i \end{array} \right) \left[\beta_n^{(r+j-i)}(t_c) u_n^{(i)}(t_c) - \beta_s^{(r+j-i)}(t_c) u_s^{(i)}(t_c)\right] \\ &+ o(\epsilon^{k+m}) \end{aligned} \quad (33)$$

We have $\beta_n^{(\nu)}(t_c) = \beta_s^{(\nu)}(t_c) = 0$ for every $\nu = 0, 1, ..., m-1$.
In order to be able to argue like in the proof of theorem 3.3 with $u_n^{(i)}(t_c) = u_s^{(i)}(t_c)$

$\forall i = 0, 1 ..., r-1$ we need the $(r+m)$-times continuous differentiability of β in t_c. In equation (33) derivation orders of u which are greater than or equal r appear plurally. For $j < m$ in (33) the following properties hold:

$$\begin{array}{ll} \text{If} & i < r, \text{ then } u_n^{(i)}(t_c) = u_s^{(i)}(t_c) \\ \text{If} & i \geq r, \text{ then } r + j - i \leq j < m \Rightarrow \beta^{(r+j-i)}(t_c) = 0 \end{array}$$

Thus only for $j = m$ a non-disappearing summand in (33) appears. For $j = m$ we get:

$$\begin{array}{ll} \text{If} & i < r, \text{ then } u_n^{(i)}(t_c) = u_s^{(i)}(t_c) \\ \text{If} & i > r, \text{ then } r + j - i < j = m \Rightarrow \beta^{(r+j-i)}(t_c) = 0 \end{array}$$

Hence only the summand for $i = r$ is left:

$$\Phi(t_c + \epsilon) = \frac{\epsilon^{k+m}}{(k+m)!} \cdot \beta^{(m)}(t_c) \cdot \left[u_n^{(r)}(t_c) - u_s^{(r)}(t_c)\right] \begin{pmatrix} r+m \\ r \end{pmatrix} + o(\epsilon^{k+m}) \quad (34)$$

This is the already mentioned equation (30).
Equation (24) can be taken over directly. We insert it in (34):

$$\begin{aligned} \Phi(t_c + \epsilon) &= \frac{\epsilon^{2q+r+m}}{(k+m)!r!m!}(r+m)! \cdot \beta^{(m)}(t_c) \cdot \\ &\cdot \frac{(-1)^r r!}{\epsilon^r}[\sigma K(t_c - \epsilon) - u(t_c - \epsilon) - o(\epsilon^r)] + o(\epsilon^{k+m}) \end{aligned}$$

We summarize and get:

$$\Phi(t_c+\epsilon) = \frac{\epsilon^{2q+m}}{(k+m)!m!}(r+m)!\beta^{(m)}(t_c)(-1)^r[\sigma K(t_c-\epsilon)-u(t_c-\epsilon)]+o(\epsilon^{k+m}) \quad (35)$$

Again we continue analogously with the distinction of the cases $\Phi(t_c + \epsilon) > 0$ ($\Rightarrow \sigma = -1$) and $\Phi(t_c + \epsilon) < 0$ ($\Rightarrow \sigma = +1$). The result is:

$$(-1)^r \epsilon^{2q+m} \beta^{(m)}(t_c)[K(t_c - \epsilon) \pm u(t_c - \epsilon)] < 0 \quad (36)$$

for ϵ sufficiently close to 0.
This coincides with the unequation (4.19) in the publication by Mc Danell and Pwowers ([11]) with the only exception that McDanell and Powers use $\beta_n^{(m)}(t_c)$ instead of $\beta^{(m)}(t_c)$. Under the additional assumption that β is $(r+m)$ times continuously differentiable in t_c this is irrelevant.
Hence we see that α and β have to be $(r+m)$ times continuously differentiable in t_c such that the proof by McDanell and Powers can be applied. The assertion by Bell and Boissard that the continuity of $\alpha^{(r+1)}$ in t_c suffices for the McDanell-Powers-proof is consequently wrong. It's only true, if $m \leq 1$ holds, which obviously contradicts $m > r + 1 \geq 1$ whereby corollary 3.4 can't be applied (the continuity of $\beta^{(r+1)}$ would result from $m > r + 1$). Apparently this error has not been mentioned yet in literature.

Another not yet considered problem is the following fact: If $m > r$, then $\beta^{(m)}$ could be discontinuous in t_c, if the $(r+m)$ times continuous differentiability of β is not assumed. It's imaginable that then $\beta^{(m)}(t_c-) = 0$ and $\beta^{(m)}(t_c+) \neq 0$ (or vice versa) holds. This also aggravates the treatment of singular control problems.

We summarize our cognitions and get

Theorem 3.5 *An (LP) with optimal extremal (x^*, u^*, λ), which have the junction t_c, be given, where the parameters q, r and m be finite. Further on α and β be $(r+m)$ times continuously differentiable in t_c.*
Then $q+r+m$ is odd.

Proof: We start with equation (36). Now we can continue similarly as described by McDanell and Powers ([11]): A Taylor series expansion for β around t_c in the singular part results in:

$$\beta(t_c - \epsilon) = (-\epsilon)^m (m!)^{-1} \beta^{(m)}(t_c) + o(\epsilon^m) \tag{37}$$

In the singular part of a neighbourhood of t_c the function β must not disappear identically, as $\beta_s^{(m)}(t_c) = 0$ would hold then. Because of the assumed continuity of $\beta^{(m)}$ in t_c the equality $\beta_n^{(m)}(t_c) = 0$ would hold then, which obviously contradicts the definition of m. Thus there exists a sufficiently small, i. e. sufficiently close to 0 situated, ϵ such that $\beta(t_c - \epsilon) \neq 0$ holds. Because of the continuity of β there exists a set of such values for ϵ which has measure greater than zero. We denote this set by $\mathcal{E}$. The intersections of $\mathcal{E}$ with arbitrarily small neighbourhoods of 0 mustn't have measure zero either, as in $\mathcal{E}$ there exist values sufficiently close to zero and every element of $\mathcal{E}$ is interior point of $\mathcal{E}$ because of the continuity of β.
As the GLC holds here, we get the unequation $(-1)^q \beta(t_c - \epsilon) > 0$. Consequently for a set of sufficiently small values for ϵ which has measure greater than zero the unequation

$$\epsilon^m (-1)^{q+m} \beta^{(m)}(t_c) > 0$$

holds because of (37). Multiplying the left side with (36) we get:

$$(-1)^{q+r+m} (\epsilon^{q+m})^2 [\beta^{(m)}(t_c)]^2 [K(t_c - \epsilon) \pm u(t_c - \epsilon)] < 0.$$

This implies:

$$(-1)^{q+r+m} [K(t_c - \epsilon) \pm u(t_c - \epsilon)] < 0. \tag{38}$$

As almost everywhere in the singular part

$$|u(\tau)| < K(\tau) \tag{39}$$

holds and $\mathcal{E}$ has measure greater than zero, there exists an $\epsilon \in \mathcal{E}$, such that $K(t_c - \epsilon) \pm u(t_c - \epsilon) > 0$ holds. Consequently from (38) we get: $(-1)^{q+r+m} < 0$. Thus $q + r + m$ is odd. **q. e. d.**

Thus under the strict differentiability assumptions in this theorem both cases of Mc Danell's Conjecture reduce to $q + r + m$ being odd.
The assumption of $(r+m)$ times differentiability is very restrictive, as especially for large values of m the appearance of $u^{(i)} (i \geq r)$ in $\beta^{(j)}$ becomes more and more probable with increasing $j \geq r$. Hence a quest for weaker assumptions is necessary.

In 1979 Bell and Boissard managed to prove the validity of McDanell's Conjecture for $q = 1$, $m = 1$:

Theorem 3.6 *An (LP) with piecewise analytical solution u^* and junction t_c be given, where $q = 1$ and $m = 1$ hold. Then the following properties hold:*
(i) If $r > 0$ then r is odd.
(ii) If $r = 0$ then $sign[\beta'(t_c+)\beta'(t_c-)] = -1$

The proof can be found in [13].
Hence Bell and Boissard made a step towards the clarification of the conjecture.

The next step is to concern the other cases for $m = 1$, i. e. extremals and problems which satisfy $m = 1$ and $q > 1$. In fact Richard Bortrins managed to prove the validity of McDanell's Conjecture for these cases in 1983. For explaining this we need more properties:

Definition 3.9 *Let $f, g : G \to \mathbb{R}^n, x \mapsto f(x)$ resp. $g(x)$ be differentiable functions, where $G \subset \mathbb{R}^n$ is open and non-empty.*
*a) We define the **Lie bracket** in the following way:*
$[f,g](x) := g_x(x)f(x) - f_x(x)g(x)$ *(f_x, g_x: Jacobi matrices)*
b) In this paper we introduce the abbreviation: $< fg > (x) := [f,g](x) (\in \mathbb{R}^n)$.
c) If f is i times ($i \in \mathbb{N} \setminus \{0\}$) differentiable, then we define recursively:
$< f^i g > (x) := [f, < f^{i-1} g >](x)$, $< f^0 g > (x) := g(x)$
d) For $(m-1)$ times ($m \geq 2$) differentiable functions $f_i : G \to \mathbb{R}^n$ ($i = 1, ..., m$) we write:
$< f_1 f_2 ... f_m > := < f_1 < f_2 ... f_m >>$, *where* $< f_m > := f_m$

Theorem 3.7 (Properties of the Lie bracket)
a) $[f,g] = -[g,f]$ ***antisymmetry***
b) $[f,h] + [g,h] = [f+g,h]$, $\lambda[f,g] = [\lambda f, g] = [f, \lambda g]$ *for every* $\lambda \in \mathbb{R}$ ***bilinearity***

Corollary 3.5 $[f,f] = 0$ *and* $[f,0] = 0$ *for every* f

The properties are easy to prove.
The following example shows that in general no associative law holds for the Lie bracket.

Example 3.5 *Let* $f(x) = x$, $g(x) = x^2$, $h(x) = \log x$.
$[[f,g],h] = x(1 - 2\log x)$, $[f,[g,h]] = -2x$

The following theorem in not new. Nevertheless a proof is given here as none was found in the literature during the research for this paper.

Theorem 3.8 (Jacobi identity)
Let $f, g, h : G \to \mathbb{R}^n$ be twice continuously differentiable functions and $G \subset \mathbb{R}^n$ be open and non-empty. Then the following identity holds for every $x \in G$:

$$< fgh > (x) = < gfh > (x) + < hgf > (x) \tag{40}$$

Proof: The function f_{xx} is triple subscripted and has the shape:

$$f_{xx} = \begin{pmatrix} grad\, f_1 \\ grad\, f_2 \\ \cdot \\ \cdot \\ grad\, f_n \end{pmatrix}_x = \begin{pmatrix} grad\,[(f_1)_{x_1}](x) & \cdots & grad\,[(f_1)_{x_n}](x) \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ grad\,[(f_n)_{x_1}](x) & \cdots & grad\,[(f_n)_{x_n}](x) \end{pmatrix}$$

Mind that $grad\, f_k$ is a line vector, whereas $(a_i)_{i=1,\dots,n}$ denotes a column vector. We have:
$(f_x(x)g(x))_x = f_{xx}(x)g(x) + f_x(x)g_x(x)$, as

$$(f_x(x)g(x))_x = \left[\left(\sum_{j=1}^{n} (f_i)_{x_j}(x) g_j(x) \right)_{i=1,\dots,n} \right]_x = \left(grad \sum_{j=1}^{n} (f_i)_{x_j}(x) g_j(x) \right)_{i=1,\dots,n}$$

and $f_{xx}(x)g(x) + f_x(x)g_x(x) = \left(\sum_{j=1}^{n} grad\,[(f_i)_{x_j}](x) g_j(x) \right)_{i=1,\dots,n} +$

$$+ \left(\sum_{j=1}^{n} (f_i)_{x_j}(x)\, grad\, g_j(x) \right)_{i=1,\dots,n}$$

Further on $f_{xx}gh = f_{xx}hg$ can be proved by using the theorem by H. A. Schwartz.
If we now apply definition 3.9 to $< fgh >$ and $< gfh > + < hgf >$ and use the above properties we get the Jacobi identity. **q. e. d.**

Our further considerations will be made for the following Mayer problem. This is no restriction at all as we will soon see.

Problem 3.1 (MLP) *We define:*

$$\begin{aligned} x'(t) &= f(x(t)) + g(x(t))u(t) \text{ and} && (41) \\ J(u) &= F(x(t_f)) && (42) \end{aligned}$$

*Except this the **(MLP)** coincides with the **(LP)** (with the changed denominations).*

Theorem 3.9 *Every **(LP)** can be transformed into an **(MLP)** and vice versa.*

The proof is easy and can e. g. be found in [7].
The following lemma is helpful for our further work.

Lemma 3.5 *(Gabasov and Kirillova 1972) An **(MLP)** with piecewise analytical solution u and finite problem order q be given. Then the following identities hold:*

$$\begin{aligned} H_u^{(2q)} &= \lambda^T(< f^{2q}g > (x) + u < gf^{2q-1}g > (x)) && (43) \\ H_u^{(2q+1)} &= \lambda^T(< f^{2q+1}g > (x) + u < gf^{2q}g > (x) + \\ &\quad +u < fgf^{2q-1}g > (x) + u^2 < g^2f^{2q-1}g > (x) + \\ &\quad +u' < gf^{2q-1}g > (x)) && (44) \\ \frac{d}{dt}((H_u^{(2q)})_u) &= \lambda^T(< fgf^{2q-1}g > (x) + u < g^2f^{2q-1}g > (x)) && (45) \\ < gf^kg > &= 0 \text{ for every } k \in \{0,1,...,2q-2\} && (46) \\ H_u^{(k)} &= \lambda^T < f^kg > \text{ for every } k \in \{0,1,...,2q-1\} && (47) \end{aligned}$$

Proof: We have: $H = \lambda^T f + u\lambda^T g$, $(\lambda^T)' = -\lambda^T f_x - u\lambda^T g_x$, $x' = f + ug$.
First we show (47). For this we use an in [6] given sketch of a proof.
Property (47) can be shown by induction:
$k = 0 : H_u^{(0)} = H_u = \lambda^T g = \lambda^T < f^0 g >$
$k \to k+1$: (In [6] this part is only represented by "...").
Let $H_u^{(k)} = \lambda^T < f^kg >$ be proved for a $k \in \{0,1,...,2q-2\}$. Then we get by the assumption of the induction:

$$\begin{aligned} H_u^{(k+1)} &= \frac{d}{dt}(\lambda^T < f^kg >) = (\lambda^T)' < f^kg > + \lambda^T < f^kg >_x x' = \\ &= (-\lambda^T f_x - u\lambda^T g_x) < f^kg > + \lambda^T < f^kg >_x (f + ug) = \\ &= \lambda^T(< f^kg >_x f - f_x < f^kg >) + u\lambda^T(< f^kg >_x g - g_x < f^kg >) = \\ &= \lambda^T[f, < f^kg >] + u\lambda^T[g, < f^kg >] = \lambda^T < f^{k+1}g > + u\lambda^T < gf^kg > \end{aligned}$$

As $k + 1 < 2q$ holds, u mustn't explicitly appear in $H_u^{(k+1)}$. We get:

$$\lambda^T < gf^kg > = 0 \text{ for } k \in \{0,1,...,2q-2\}$$

As the respresentation of $H_u^{(k)}$ as a function with the variables x, λ and u is used for the determination of q, the factor of λ^T must disappear. Thus we get:

$$< gf^kg > = 0 \text{ for } k \in \{0,1,...,2q-2\}.$$

Thus if the induction is completed, (46) will also be shown. We further get:

$$H_u^{(k+1)} = \lambda^T < f^{k+1}g > .$$

Hence the induction is complete and (47) is proved.
Thus we get: $H_u^{(2q-1)} = \lambda^T < f^{2q-1}g >$. Consequently:

$$\begin{aligned}
H_u^{(2q)} &= (\lambda^T)' < f^{2q-1}g > + \lambda^T < f^{2q-1}g >_x x' = \\
&= (-\lambda^T f_x - u\lambda^T g_x) < f^{2q-1}g > + \lambda^T < f^{2q-1}g >_x (f + ug) = \\
&= \lambda^T(< f^{2q-1}g >_x f - f_x < f^{2q-1}g > + \\
&\quad + u(< f^{2q-1}g >_x g - g_x < f^{2q-1}g >)) = \\
&= \lambda^T(< f^{2q}g > + u < gf^{2q-1}g >). \text{ This is (43).}
\end{aligned}$$

We derive the result:

$$\begin{aligned}
H_u^{(2q+1)} &= (\lambda^T)' < f^{2q}g > + \lambda^T < f^{2q}g >_x x' + u'\lambda^T < gf^{2q-1}g > \\
&\quad + u\left(\frac{d}{dt}\left(\left(H_u^{(2q)}\right)_u\right)\right)
\end{aligned} \tag{48}$$

We consider:

$$\begin{aligned}
\frac{d}{dt}((H_u^{(2q)})_u) &= \frac{d}{dt}(\lambda^T < gf^{2q-1}g >) \\
&= (\lambda^T)' < gf^{2q-1}g > + \lambda^T < gf^{2q-1}g >_x x' = \\
&= (-\lambda^T f_x - u\lambda^T g_x) < gf^{2q-1}g > + \lambda^T < gf^{2q-1}g >_x (f + ug) = \\
&= \lambda^T(< gf^{2q-1}g >_x f - f_x < gf^{2q-1}g > + \\
&\quad + u(< gf^{2q-1}g >_x g - g_x < gf^{2q-1}g >)) = \\
&= \lambda^T(< fgf^{2q-1}g > + u < g^2 f^{2q-1}g >). \text{ This is (45).}
\end{aligned}$$

We substitute the result in (48):

$$\begin{aligned}
H_u^{(2q+1)} &= (\lambda^T)' < f^{2q}g > + \lambda^T < f^{2q}g >_x x' + u'\lambda^T < gf^{2q-1}g > + \\
&\quad + u\lambda^T(< fgf^{2q-1}g > + u < g^2 f^{2q-1}g >) = \\
&= (-\lambda^T f_x - u\lambda^T g_x) < f^{2q}g > + \lambda^T < f^{2q}g >_x (f + ug) + \\
&\quad + u'\lambda^T < gf^{2q-1}g > + u\lambda^T(< fgf^{2q-1}g > + u < g^2 f^{2q-1}g >) = \\
&= \lambda^T(< f^{2q}g >_x f - f_x < f^{2q}g > + u < f^{2q}g >_x g - ug_x < f^{2q}g >) + \\
&\quad + u'\lambda^T < gf^{2q-1}g > + u\lambda^T(< fgf^{2q-1}g > + u < g^2 f^{2q-1}g >) = \\
&= \lambda^T(< f^{2q+1}g > + u < gf^{2q}g > + u < fgf^{2q-1}g > + \\
&\quad + u^2 < g^2 f^{2q-1}g > + u' < gf^{2q-1}g >)
\end{aligned}$$

This is (44). **q. e. d.**

Mind that $(H_u^{(k+1)})_u = \lambda^T < gf^k g >$ holds for $k \in \{0, 1, ..., 2q-1\}$. Thus the problem order q is given by:
(i) $< gf^i g >= 0$ for every $i \in \{0, 1, ..., 2q-2\}$ and (ii) $< gf^{2q-1}g > \neq 0$.
If q coincides with the arc order p, then $\lambda^T < gf^{2q-1}g >$ must not disappear on the considered interval. In other words:
$p \neq q \Leftrightarrow \lambda \perp < gf^{2q-1}g >$ in a part of the analysed extremal.
Pontrjagin's Minimum principle (theorem 2.1) excludes $\lambda = 0$.

The following lemma and its corollary were proved by Richard Bortrins in 1983, see [14].

Lemma 3.6 *Let $f, g : \mathbb{R}^n \to \mathbb{R}^n$ be functions of the class C^{m+1}. Then the following identity holds for $j \in \{0, 1, ..., m\}$ and $k \in \{j, j+1, ..., m\}$:*

$$0 = \sum_{i=0}^{j} (-1)^i \binom{j}{i} < f^i g f^{k-i} g > -(-1)^j [< f^j g >, < f^{k-j} g >] \tag{49}$$

Corollary 3.6 *Let $f, g : \mathbb{R}^n \to \mathbb{R}^n$ be functions of the class C^{m+1}.*
a) Let $q \in \mathbb{N} \setminus \{0\}$ such that $2q \le m$ and $< g f^i g >= 0$ for $i \in \{0, 1, ..., 2q-2\}$ holds. Then:

$$0 = \; < g f^{2q} g > -q < f g f^{2q-1} g > \tag{50}$$

b) Let $q \in \mathbb{N} \setminus \{0; 1\}$ such that $2q - 1 \le m$ and $< g f^i g >= 0$ for $i \in \{0, 1, ..., 2q-2\}$ holds. Then:

$$0 = \; < g^2 f^{2q-1} g > \tag{51}$$

Lemma 3.6 is proved by induction. The proof of corollary 3.6 uses Lemma 3.6, the Jacobi Identity (theorem 3.8) and the properties of the Lie Bracket (theorem 3.7). [14]

Remark 3.2 *Bortrins ([14]) remarks without proof that part a) of corollary 3.6, i. e. (50), results for $q = 1$ in lemma 3.3. This has the following reason:*
For $q = 1$ corollary 3.6 says that

$$0 =< g f^2 g > - < f g f g > \tag{52}$$

holds. According to lemma 3.5 we get for $q = 1$:

$$\begin{aligned} H_u^{(2q+1)} &= \lambda^T (< f^3 g > (x) + u < g f^2 g > (x) + u < f g f g > (x) + \\ &\quad + u^2 < g^2 f g > (x) + u' < g f g > (x)) && (53) \\ \frac{d}{dt}((H_u^{(2q)})_u) &= \lambda^T (< f g f g > (x) + u < g^2 f g > (x)) && (54) \\ &\left(= \frac{d}{dt} B = b_0 + u b_u \right) \end{aligned}$$

As $H_u^{(2q+1)} = a_0 + u a_u + u' B + u(b_0 + u b_u) = a_0 + u(a_u + b_0) + u^2 b_u + u' B$ must hold, we can read from (53) and (54):

$$a_0 = \lambda^T < f^3 g >, b_u = \lambda^T < g^2 f g >, B = \lambda^T < g f g >,$$
$$a_u = \lambda^T < g f^2 g >, b_0 = \lambda^T < f g f g >$$

Thus (52) means that $a_u = b_0$ holds. ***q. e. d.***

Now we can show that McDanell's conjecture also holds for $q > 1$, $m = 1$. But we need an additional assumption.

Theorem 3.10 *(Richard Bortrins 1983) An **(MLP)** with piecewise analytical solution u^* and junction t_c be given. The problem order q and the arc order p be finite and $q = p > 1$ hold. Further on $m = 1$ hold. Then $q + r + 1$ is odd.*

A sketch of a proof is given in [14].
Hence McDanell's conjecture has been proved for $m = 1$, $p = q > 1$. Concerning part (i) this is obvious. Part (ii) ($r = 0$) has been proved because $\beta^{(m)} = \beta'$ is continuous in t_c under the assumptions of theorem 3.10. ([14])
No counterexample is known for $m = 1$, $1 < q < p$ yet.
Now we've finished the discussion of the case $m = 1$. A transfer to larger values of m is - if it's possible at all - not possible without additional assumptions. This shows the following counterexample:

Example 3.6 (Counterexample by Pan and Bell (1987))
Pan Guo-Lin and D. J. Bell constructed the following problem:

$$\text{minimize } J(u) = \int_{-1}^{1} \left(x_2 + \frac{t}{2}\right)^2 (x_2 + 2t)^2 (1 + u)dt \geq 0, \text{ where}$$

$$K \equiv 1,\ x_1' = x_2 u,\ x_1(-1) = \frac{3}{4}(e + e^{-1}) - 1,\ x_1(1) = \frac{1}{2}$$

$$x_2' = u - x_1,\ x_2(-1) = \frac{3}{4}(e - e^{-1}),\ x_2(1) = -\frac{1}{2}$$

An optimal solution starts nonsingular with $u^* \equiv -1$ on $[-1, 0[$ and ends singular with $u^* \equiv 0$ on $]0, 1]$. The control u^* is optimal as $J(u^*) = 0$ holds. As u^* is discontinuous in the junction $t_c = 0$ we get $r = 0$. Deriving the switching function we get $q = 1$ and $m = 2\,(> r)$. Hence case (ii) of McDanell's conjecture is relevant. Further on we get:

$$\beta^{(2)}(0+) = -9,\ \beta^{(2)}(0-) = 3 \Rightarrow -sign[\beta^{(m)}(t_c+)\beta^{(m)}(t_c-)] = 1$$

But this contradicts McDanell's conjecture, as $(-1)^{q+r+m} = (-1)^{1+0+2} = -1$ holds. [15]

Let's summarize the results: The validity of McDanell's conjecture has been proved in the following four cases:

- $m = 0$ (theorem 3.3)
- α and β are $(r+m)$ times continuously differentiable in the junction (theorem 3.5)
- $q = 1$ and $m = 1$ (theorem 3.6)
- $q = p > 1$ and $m = 1$ (theorem 3.10)

The general validity of case (ii) was disproved by a counterexample by Pan and Bell. In 1993 Bell ([12]) published evidences according to which the construction of a problem with an optimal solution that satisfies $(q, m, r) = (1, 2, 3)$ might be possible. This would contradict part (i) of the conjecture, as $q + r + m$ is even in this case. The construction of such a counterexample has not been possible yet. Apparently McDanell's conjecture has to be rejected for problems with optimal solutions that satisfy $m > 1$. A proof of the conjecture with more restrictive assumptions might be possible. But one mustn't forget that the conjecture is not very helpful for the solution of singular optimal control problems. The question arises, whether the examination of the properties of the parameters q, r and m won't prove to be an impasse.

3.5 The Location of the Junctions

For a long time it has been supposed that there's a connection between the location of the junctions and the boundary conditions of the problem. But no proof could be made. In 1995 such a connection was proved by Ruxton and Bell ([16]). They considered the following problem:

Problem 3.2 ***(LQP)*** *minimize* $\frac{1}{2}\int_{t_0}^{t_f} x^T C x\, dt$ *subject to*
$x' = Ax + bu,\ x(t_0) = x_0,\ x(t_f) = x_f$
where $U(t) = [-1, 1] \subset \mathbb{R}\,\forall t,\ t_f$ *fixed,* $b \in \mathbb{R}^n$ *and* $A, C \in \mathbb{R}^{n \times n}$.

Two problems of kind **(LQP)**, who have both partially singular optimal controls and only differ in the final state $x(t_f)$ are considered. The following assumptions are made:

- The optimal control consists of exactly two parts, where one is nonsingular (and constant) and the other one singular, independent of the sequence. The junction is t_c.
- For one choice of $x(t_f)$ the optimal control u^* is known on whole T. The corresponding extremal be (x^*, u^*, λ^*). The control u^* is called reference control.
- For an arbitrary choice of the final state $\bar{x}(t_f)$ the corresponding optimal control be $\bar{u}(t)$ (variational control) with junction $\bar{t}_c$ and extremal $(\bar{x}, \bar{u}, \bar{\lambda})$. Then $\bar{u}$ switches to the same bound $|u^*(t)| = \pm 1$ as the reference control but at $\bar{t}_c$ where $t_c = \bar{t}_c$ only if $\bar{x}(t_f) = x^*(t_f)$

The question is: If we know (x^*, u^*, λ^*), $\bar{x}(t_f)$, $x^*(t_f)$ and t_f, are we able to compute $\bar{t}_f$?
Using the abbreviations $\delta u = \bar{u} - u^*$, $\delta x = \bar{x} - x^*$ and $\delta\lambda = \bar{\lambda} - \lambda^*$ the following theorem holds ([16]):

Theorem 3.11

$$\textit{Let } R = \frac{b\,b^T}{b^T b}, X = \begin{pmatrix} (I-R)A + RA^T & RA^{2^T} \\ -I & -A^T \end{pmatrix}, \tau = t_f - \bar{t}_c,$$

$$Q(.) = \exp X(.),\ \hat{Q} = Q^{-1} = \begin{pmatrix} \hat{Q}_{11} & \hat{Q}_{12} \\ \hat{Q}_{21} & \hat{Q}_{22} \end{pmatrix} \in \mathbb{R}^{2n \times 2n}, \textit{ where } Q_{ij} \in \mathbb{R}^{n \times n}$$

a) bang/singular case with $\bar{t}_c > t_c$. *The following equation holds:*

$$\hat{Q}_{11}(\tau)\delta x(t_f) + \hat{Q}_{12}(\tau)\delta\lambda(t_f) = \int_{t_c}^{\bar{t}_c} exp\,(A(\bar{t}_c - \tau))\, b\, \delta u(\tau)\, d\tau \tag{55}$$

b) singular/bang case with $t_c > \bar{t}_c$. *The following equation holds:*

$$Q_{12}(\bar{t}_c)\, \delta\lambda(t_0) = exp\,(A(\bar{t}_c - t_f))\, \delta x(t_f) + \int_{t_c}^{\bar{t}_c} exp\,(A(\bar{t}_c - \tau))\, b\, \delta u(\tau)\, d\tau \tag{56}$$

If the final costate variation $\delta\lambda(t_f)$ resp. the initial costate variation $\delta\lambda(t_0)$ is calculable then $\bar{t}_c$ may be calculated as a function of $x(t_f)$ ([16]).
Although the theorem is only applicable to a very special type of problems it gives hope that more theorems of that kind can be proved in the future.

4 Multidimensional Controls

4.1 Introduction

In section 3 we restricted ourselves to one-dimensional controls. In pratice however more complicated systems can have several independent control variables. This results in additional problems:
The anyway very difficult treatment of the junctions e. g. becomes even more complicated, as e. g. the problem order need not be an integer any more (we will soon see that).

Further on it's posssible that a control is only singular in certain components, where junctions need not refer to all control components either.
This section shows some properties of multidimensional controls, that might be helpful for the treatment of such problems.

4.2 Mixed-linear-nonlinear Problems

Here we consider problems with the Hamiltonian

$$\begin{aligned} H(u,x,\lambda,t) &= \Theta(x,\lambda,t)+\omega(x,\lambda,t)v+\sigma(w,x,\lambda,t), \\ \text{where } u &= (v,w) \end{aligned}$$

In such cases we treat the linear and the nonlinear part seperately as far as possible. Thus singularity means that $H_v = \omega = 0$ and $H_{ww} = \sigma_{ww} = 0$ hold. Hence we get equations in the linear control variables and equations in the nonlinear control variables. [7]

4.3 Controls with Singular and Nonsingular Components

Definition 4.1
*Let **(P)** be given.*
*a) The Hamiltonian H be linear in u and at least one component of H_u be zero. Then u is called **singular of rank** $\sharp\{i|H_{u_i}=0\}$.*
b) The Hamiltonian H be nonlinear in u and rkH_{uu} be not maximal. Then u is called singular of rank $(s-rkH_{uu})$.
*If the rank of singularity is maximal, i. e. equal to s, then u is called **totally singular**. (cf. [10])*

In [10] a possibility to eliminate nonsingular control components in order to create Hamiltonians with totally singular controls is described. In this work this will be substantiated and further developed:
a) nonlinear case:
For simplification we use Einstein's convention to sum over every index that appears more than once in a summand.
In the nonsingular case we assume that $H_u = 0$ holds along the considered interval of time. If u is in the interior of U this is satisfied because of theorem 2.1.
Thus we get s equations. Because of the nonlinearity we get in the singular case $detH_{uu} = 0$, thus $\rho := rkH_{uu} < s$.
If we make a linear transformation of the control variables by $u = Zv$, where Z is a regular $s\times s$-matrix, and define $H^0(v,x,\lambda,t) := H(Zv,x,\lambda,t)$, then we get $H^0_{v_k} = H_{u_i}z_{ik}$, as $u_i = z_{ij}v_j$. Hence $H^0_v = Z^T H_u$. Further on we get:

$$(H^0_{vv})_{k,j} = (H_{uu})_{i\mu}z_{ik}z_{\mu j} = z^T_{ki}(H_{uu})_{i\mu}z_{\mu j} \Rightarrow H^0_{vv} = Z^T H_{uu} Z$$

As H_{uu} is symmetrical according to the theorem by H. A. Schwartz, we can make a principal axis transformation such that

$$Z^T H_{uu} Z = \begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix} \tag{57}$$

holds, where C is a regular $\rho\times\rho$-(diagonal)-matrix. Now we use the matrix Z, built by orthonormal eigenvectors, for the described transformation $u = Zv$. Then $H^0_v = Z^T H_u = 0$ holds and

$$H^0_{vv} = \begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix}.$$

We write $v = (v_{non}, v_{sing})$, where $v_{non}(t) \in \mathbb{R}^\rho$, and interpret v_{non} as the nonsingular part and v_{sing} as the singular part. Then we get:

$$H^0_{v_{non}} = 0 \quad (58)$$

$$H^0_{v_{non} v_{non}} = C \quad (59)$$

As we assume the existence of a solution and C is regular, the implicit function theorem allows us to resolve equation (58) such that we get a function $v_{non} = \psi(v_{sing}, x, \lambda, t)$. This way we can reduce the dimension of the control space from s to $s - \rho$, and we get a new Hamiltonian

$$\mathcal{H}(v_{sing}, x, \lambda, t) := H^0(\psi(v_{sing}, x, \lambda, t), v_{sing}, x, \lambda, t).$$

Now we can continue with this function without problems, as:

$$\mathcal{H}_x = (H^0_{v_{non}})^T \psi_x + H^0_x =_{(58)} H^0_x = H_x = -\lambda'$$

$$\mathcal{H}_\lambda = (H^0_{v_{non}})^T \psi_\lambda + H^0_\lambda =_{(58)} H^0_\lambda = H_\lambda = x'$$

Further on we get $\mathcal{H}_{v_{sing}} = 0$ and $\mathcal{H}_{v_{sing} v_{sing}} = 0$. Hence the transformed problem is totally singular.
The method once again in a short version:
(i) Make a principal axis transformation:

$$Z^T H_{uu} Z = \begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix}$$

and define $H^0(v, x, \lambda, t) := H(Zv, x, \lambda, t)$.
(ii) If C is a $\rho \times \rho$-matrix then resolve the equation system $(H^0_{v_i})_{i=1,\ldots,\rho} = (0)_{i=1,\ldots,\rho}$ w. r. t. $(v_1, \ldots, v_\rho)^T$. Result: $(v_1, \ldots, v_\rho)^T = \psi(v_{\rho+1}, \ldots, v_s, x, \lambda, t)$
(iii) The new Hamiltonian is

$$\mathcal{H}(v_{\rho+1}, \ldots, v_s, x, \lambda, t) := H^0(\psi(v_{\rho+1}, \ldots, v_s, x, \lambda, t), v_{\rho+1}, \ldots, v_s, x, \lambda, t)$$

with the totally singular control $(v_{\rho+1}, \ldots, v_s)$.
The set $U(t)$ of admissible control values is transformed correspondingly.
b) linear case:
We use the restriction $|u_k(t)| \leq K_k(t) > 0$ (cf. **(LP)**).
The control be singular of rank $s - \rho (> 0)$. W. l. o. g. exactly the last $s - \rho$ components of H_u dissappear: $H_u = (*, \ldots, *, 0, \ldots, 0)^T$. Thus $u_1, \ldots, u_\rho$ are nonsingular and $u_{\rho+1}, \ldots, u_s$ are singular. We have:

$$u_k(t) = -K_k(t) \cdot sign\, H_{u_k}(u(t), x(t), \lambda(t), t) \text{ for } k = 1, \ldots, \rho$$

As H_u doesn't depend on u any more, we have the representation

$$(u_1, \ldots, u_\rho) = \psi(x, \lambda, t),$$

with which we proceed like in the nonsingular case. Hence we define:

$$\mathcal{H}(u_{\rho+1}, \ldots, u_s, x, \lambda, t) := H(\psi(x, \lambda, t), u_{\rho+1}, \ldots, u_s, x, \lambda, t)$$

4.4 Problem Order and Other Concepts

In case of multidimensional controls we have the situation that $H_u^{(k)}$ is now a vector in $\mathbb{R}^s$ and $(H_u^{(k)})_u$ is an $s \times s$-matrix. The different controls can appear primarily in each component of $H_u^{(k)}$ for different values of k. Apparently only Robbins ([10], 1967) as well as Bell and Jacobson ([3],1975) have examined such problems up to now. Robbins proposed a definition for the order, but at this time there was no distinction between problem order and arc order. Today the definition would probably be the following:

Definition 4.2 *Let **(P)** be given. Then*

$$\mathbf{q} := \min\{k \geq 0 \mid H_u^{(2k)} \text{ has at least one component in which a component of } u \text{ appears explicitly.}\}$$

is called **problem order** *of **(P)**.*
*Further on (x, λ, t) be an extremal of **(P)** and $[a, b] \subset T$ $(a < b)$. Then*

$$\mathbf{p} := \min\{k \geq 0 \mid (H_u^{(2k)})_u(u(t), x(t), \lambda(t), t) \text{ has at least one component which does not dissappear for any } t \in]a, b[.\}$$

is called **arc order** *of (x, λ, t) w. r. t. **(P)**.*
The derivatives are computed as described in definition 3.3.

The generalized definition 4.2 is consistent with the definitions 3.3 and 3.4. Now a new concept for the treatment of the order in the case of multidimensional controls will be introduced.

Definition 4.3 *Let **(P)** be given. The matrix $\mathbf{Q} = (q_{ij})_{i,j=1,\dots,s}$, where the control u_i appears explicitly in H_{u_j} for the first time in the $2q_{ij}$-th total derivative with respect to t, i. e. in $H_{u_j}^{2q_{ij}}$, is called* **problem order matrix**.
Further on (x, u, λ) be an extremal and $[a, b] \subset T$ $(a < b)$ be a given interval. Then we call $\mathbf{P} = (p_{ij})_{i,j=1,\dots,s}$ **arc order matrix**, *where*

$$p_{ij} = \min\{k \geq 0 | (H_{u_j}^{(2k)})_{u_i}(u(t), x(t), \lambda(t), t) \neq 0 \text{ for every } t \in]a, b[\}.$$

It is $q = \min\{q_{ij} | i, j = 1, ..., s\}$ and $p = \min\{p_{ij} | i, j = 1, ..., s\}$. In the onedimensional case the matrices become the usual orders.
It's easy to see the validity of the following theorem:

Theorem 4.1 *The diagonal elements of Q and P are natural numbers.*

Proof: As the diagonal elements refer to only one control component we can proceed analogously like in the proof of theorem 3.1 with $(H_{u_i}^{(k)})_{u_i}$ for q_{ii} resp. p_{ii} and receive the desired result. **q. e. d.**

The remaining elements of Q and P need not be natural. If one starts in the proof of theorem 3.1 with $\gamma_p = (H_{u_i}^{(p)})_{u_j}$, one would get instead of (15) the relation

$$\gamma_p = (-1)^{\nu+1}[\nabla(H_{u_j}^{(\nu)})]^T S \ \nabla(H_{u_i}^{(\nu)}) \tag{60}$$

Thus here we don't have the shape $z^T S z$ any more. Hence the proof can be applied to the diagonal elements of the matrices and only to these.

Robbins ([10]) as well as Bell and Jacobson ([3]) point out that in the multidimensional case non-natural orders can appear, and at least Robbins means probably the todays arc order. But examples are not given. Robbins points out that such extremals (with odd $2p$) are not optimal however.
Non-natural problem orders on the other hand are independent of extremals and can indeed appear, as the following hitherto unknown counterexample shows, which is also important in another respect.

Example 4.1

$$\textit{minimize } J(u) = \int_0^2 (x_1^2 + x_2^2 - 1)^2 dt, \textit{ where}$$
$$x_1' = u_1 x_2, x_1(0) = 0, x_2' = u_2 x_1, x_2(0) = 1$$
$$u_1(t), u_2(t) \in [-1, 1] \textit{ piecewise continuous}$$

We get $H = \lambda_1 u_1 x_2 + \lambda_2 u_2 x_1 + (x_1^2 + x_2^2 - 1)^2$.
$\Rightarrow \lambda_1' = -\lambda_2 u_2 - 2(x_2^2 + x_2^2 - 1) \cdot 2x_1$ *and* $\lambda_2' = -\lambda_1 u_1 - 2(x_1^2 + x_2^2 - 1) \cdot 2x_2$
The transversality condition results in:

$$\lambda(2) = (0, 0)^T \tag{61}$$

We get: $H_u = (\lambda_1 x_2, \lambda_2 x_1)^T$.

$$\begin{aligned} H_u^{(1)} &= \begin{pmatrix} -\lambda_2 u_2 x_2 - 4(x_1^2 + x_2^2 - 1)x_1 x_2 + \lambda_1 u_2 x_1 \\ -\lambda_1 u_1 x_1 - 4(x_1^2 + x_2^2 - 1)x_1 x_2 + \lambda_2 u_1 x_2 \end{pmatrix} \\ &= \begin{pmatrix} -4(x_1^2 + x_2^2 - 1)x_1 x_2 + (\lambda_1 x_1 - \lambda_2 x_2)u_2 \\ -4(x_1^2 + x_2^2 - 1)x_1 x_2 + (\lambda_2 x_2 - \lambda_1 x_1)u_1 \end{pmatrix} \end{aligned} \tag{62}$$

Thus $q_{12} = q_{21} = \frac{1}{2}$ *holds. Hence all elements of* Q *outside the diagonal are non-natural!*
Further on we get:

$$H_u^{(2)} = \begin{pmatrix} -4(2x_1 u_1 x_2 + 2x_2 u_2 x_1)x_1 x_2 + \\ +[-4(x_1^2 + x_2^2 - 1)u_1 x_2^2] - 4(x_1^2 + x_2^2 - 1)u_2 x_1^2 + \\ +[-\lambda_2 u_2 x_1 - 4(x_1^2 + x_2^2 - 1)x_1^2 + \lambda_1 u_1 x_2 + \lambda_1 u_1 x_2 + \\ +4(x_1^2 + x_2^2 - 1)x_2^2 - \lambda_2 u_2 x_1]u_2 + (\lambda_1 x_1 - \lambda_2 x_2)u_2' \\ \\ -4(2x_1 u_1 x_2 + 2x_2 u_2 x_1)x_1 x_2 + \\ +[-4(x_1^2 + x_2^2 - 1)u_1 x_2^2] - 4(x_1^2 + x_2^2 - 1)u_2 x_1^2 + \\ +[-\lambda_1 u_1 x_2 - 4(x_1^2 + x_2^2 - 1)x_2^2 + \lambda_2 u_2 x_1 + \lambda_2 u_2 x_1 + \\ +4(x_1^2 + x_2^2 - 1)x_1^2 - \lambda_1 u_1 x_2]u_1 + (\lambda_2 x_2 - \lambda_1 x_1)u_1' \end{pmatrix}$$

Thus we get the following matrix:

$$(H_u^{(2)})_u = \begin{pmatrix} \begin{matrix} -8x_1^2 x_2^2 + \\ +\left[-4(x_1^2 + x_2^2 - 1)x_2^2\right] + \\ +2\lambda_1 u_2 x_2 \end{matrix} & \begin{matrix} -8x_1^2 x_2^2 - 4(x_1^2 + x_2^2 - 1)x_1^2 + \\ +(-4\lambda_2 u_2 x_1) + 2\lambda_1 u_1 x_2 + \\ +\left[-4(x_1^2 + x_2^2 - 1)(x_1^2 + x_2^2)\right] \end{matrix} \\ \\ \begin{matrix} -8x_1^2 x_2^2 - 4(x_1^2 + x_2^2 - 1)x_2^2 + \\ +(-4\lambda_1 u_1 x_2) + 2\lambda_2 u_2 x_1 + \\ +\left[-4(x_1^2 + x_2^2 - 1)(x_1^2 + x_2^2)\right] \end{matrix} & \begin{matrix} -8x_1^2 x_2^2 + \\ \left[-4(x_1^2 + x_2^2 - 1)x_1^2\right] + \\ +2\lambda_2 u_1 x_1 \end{matrix} \end{pmatrix} \tag{63}$$

Consequently

$$Q = \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix} \tag{64}$$

holds. Thus theorem 4.1 has been validated.
Now we construct an optimal solution. A very simple solution would be $u_1 \equiv u_2 \equiv 0$.
But it's not the only one:

$$\begin{aligned} \textit{(i) } 0 \le t < \frac{\pi}{2} \quad : \quad & u_1(t) = +1,\ u_2(t) = -1 \\ & x_1(t) = \sin t,\ x_2(t) = \cos t \\ & \lambda_1' = \lambda_2,\ \lambda_2' = -\lambda_1 \end{aligned}$$

$$(ii)\ \frac{\pi}{2} < t \le 2 \quad : \quad \begin{aligned} & u_1(t) = u_2(t) = 0 \\ & x_1(t) = 1, x_2(t) = 0 \\ & \lambda_1' = \lambda_2' = 0 \end{aligned}$$

Using the transversality condition (61) we get: $\lambda_1 \equiv \lambda_2 \equiv 0$.
As generally $J(u) \ge 0$ *holds and along the constructed extremal* $J = 0$ *holds, we have an optimal solution. If we insert this extremal in* $(H_u^{(1)})_u$ *(see (62)) then we get the zero matrix. Thus* $p_{ij} > \frac{1}{2}$ *holds for every* $i, j \in \{1, 2\}$. *If we insert the extremal in* $(H_u^{(2)})_u$ *(see (63)) then we get at* $]0, \frac{\pi}{2}[$*:*

$$(H_u^{(2)})_u(u(t), x(t), \lambda(t), t) = -8 \sin^2 t \cos^2 t \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

Hence at $[0, \frac{\pi}{2}]$ *the problem order matrix of the considered optimal solution is*

$$P = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

and consequently $p = 1$.
The considered optimal control function u *is totally singular on* T. *Here a part of the singular control is situated on the boundary of* U. *But changing the set* U *to* $[1 - \epsilon, 1 + \epsilon]$ $(\epsilon > 0)$ *is possible without problems.*
Further on $(-1)^p (H_u^{(2)})_u(u(t), x(t), \lambda(t), t)$ *is positive semidefinite on* $[0, \frac{\pi}{2}]$, *as the eigenvalues are* 0 *and* $8 \sin^2 t \cos^2 t$.

This looks like a multidimensional GLC. Indeed such a theorem can be proved.

Theorem 4.2 (Generalized Legendre Clebsch Condition (GLC))
*A **(P)** with analytical functions* f *and* L *be given and* (x^*, u^*, λ) *be an optimal extremal with arc order* $p < +\infty$ *on* $[a, b] \subset T$ $(a < b)$. *Then* $p \in \mathbb{N}$ *and along* $(x^*(t), u^*(t), \lambda(t))$ *the matrix* $(-1)^p \cdot (H_u^{(2p)})_u$ *is positive semidefinite for every* $t \in]a, b[$.

(see [3] and [10])

This shows that indeed odd values for $2p$ allow to conclude that nonoptimality appears, as Robbins discovered. Mind that the onedimensional GLC for arc orders $((-1)^p (H_u^{(2p)})_u > 0)$ doesn't result in the positive definiteness of $(-1)^p (H_u^{(2p)})_u$ but merely in its positive semidefiniteness, as counterexample 4.1 has shown.
Basically the onedimensional GLC also only consists of a statement with the arc order, as in the case $p = q$ both statements are identical and in the other case the unequation $(-1)^q (H_u^{(2q)})_u \ge 0$ is trivially satisfied.

5 Optimality Conditions for Singular Controls

5.1 Introduction

The purpose of this section is to present conditions for the optimality of a singular control. We restrict ourselves to purely singular controls, i. e. controls which are singular on the whole interval. Although optimal controls often consist of singular and nonsingular parts, this is no essential restriction, as optimal controls are also optimal in any subinterval if the associate boundary conditions are given. The determination of optimal purely singular control segments can furthermore bear informations on the location of junctions, if e. g. a criterion excludes the optimality of singular controls in certain time segments.
Because of scarcity of room we forgo the proofs of the quoted theorems.

5.2 Determination of the Singular Control

First of all we describe ways to determine the singular control without taking care of optimality. We restrict our considerations to linear Mayer problems of the shape **(MLP)**.
The control we consider is supposed to be singular on the whole interval T. Thus in the interior of T the identity $H_u^{(k)} = 0$ must hold per def. for every $k \in \mathbb{N}$. Thereby we get basically infinitely many equations, which shouldn't be left unused. We could use them e. g. for the attempt to determine the singular control. But the question is how many independent equations we have. In lemma 3.5 we've already seen that in case of finite problem order q the equation $H_u^{(k)} = \lambda^T < f^k g >$ holds for every $k \in \{0, 1, ..., 2q-1\}$.

Theorem 5.1 *An* ***(MLP)*** *with associate extremal (x^*, u^*, λ) be given, where u^* be purely singular. The problem order q be finite and be equal to the arc order p on $[a, b] \subset T$ $(a < b)$. Then $\{< f^k g > | k = 0, 1, ..., 2q-1\}$ are linear independent along the extremal (on whole T).*

The proof uses lemma 3.5 and can be found in [7].
Using this theorem we get the following interesting result:

Corollary 5.1 *An* ***(MLP)*** *with associate extremal (x^*, u^*, λ) be given, where u^* be purely singular and $q = p < +\infty$ hold on whole T. Then $2q \leq n$ holds.*

The proof can be found in [7].

The statement of corollary 5.1 appears only at first sight to contradict example 2.1, in which $2q = 2 > 1 = n$ holds. But example 2.1 is not a Mayer problem. If we transform it into a Mayer problem, then the dimension n increases by 1 and the corollary is confirmed.

By lemma 3.5 we know that

$$H_u^{(2q)} = \lambda^T(< f^{2q} g > (x) + u < g f^{2q-1} g > (x))$$

holds if $q < +\infty$. Thus in segments with $q = p$ we can determine the singular control from the equation $H_u^{(2q)} \equiv 0$ as

$$u = -\frac{< f^{2q} g > (x)}{< g f^{2q-1} g > (x)} \tag{65}$$

But for high orders the usage of (65) is very expendable and numerically unusable as the differentiation is an ill-posed problem.
If $q < p < +\infty$ then we have besides $H_u^{(k)} = 0$ $(k = 0, 1, ..., 2q-1)$ additionally the following equations:

$$H_u^{(k)} = 0 \text{ for } k \in \{2q, 2q+1, ..., 2p-1\} \text{ (because of the singularity)}$$
$$(H_u^{(k)})_u = 0 \text{ for } k \in \{2q, 2q+1, ..., 2p-1\} \text{ (because of the definition of } p\text{)}$$

Of course, $H_u^{(k)} = 0$ also holds for $k \geq 2p$. [7]

Therewith we have possibilities to find candidates for singular controls. On these candidates we then apply optimality conditions like the GLC in order to e. g. exclude certain candidates.

Example 5.1 *In example 2.1 the singular control can be determined in the described way. If we use the (optimality) condition $\lambda \neq 0$ (see theorem 2.1) then we get $x_1 \equiv 0$ and $u \equiv 0$ in case of singularity.*

5.3 First and Second Variation of J

The computation of the variations is at length described in [3]. We simply quote the results. At first we construct another in comparision to **(P)** slightly changed problem:

Problem 5.1 *By* $(\mathbf{P_2})$ *we mean a problem of kind (P) mit free final time* t_f, *where* $x^*, f, g, L \in C^2$ *holds and additionally the final condition* $\psi(x(t_f), t_f) = 0$ $(\in \mathbb{R}^\sigma)$ *has to be satisfied. Here* x^* *is the state function which belongs to the optimal piecewise continuous control* u^*.

We consider $(\mathbf{P_2})$ with optimal admissible pair (x^*, u^*) and construct a one-parameter family of control vector functions $u(.,\epsilon)$ such that $u(.,0) = u^*(.)$. The associate family of state vector functions is denoted by $x(.,\epsilon)$, where analogously $x(.,0) = x^*(.)$ holds. In case of variable final time we write $t_f(\epsilon)$. Further on we use the denotations $\delta x := x_\epsilon\, d\epsilon$, $\delta x' := \delta(x') = (x')_\epsilon\, d\epsilon$, $\delta^2 x := \delta(\delta x)$, $\xi_f := (t_f)_\epsilon|_{\epsilon=0}$, $\eta := x_\epsilon|_{\epsilon=0}$ and $\beta := u_\epsilon|_{\epsilon=0}$. By $\nu \in \mathbb{R}^\sigma$ we denote an arbitrary constant vector. Mind that η and β are functions of t and λ is independent of ϵ.
Thus the optimal trajectory (i. e. the case $\epsilon = 0$) can be described by:

$$dt_f = \xi_f\, d\epsilon,\ \delta x = \eta\, d\epsilon \tag{66}$$

Definition 5.1 *As* J *is in our considerations a function of* ϵ *we can (under the assumption of sufficient differentiability) develop* J *into a Taylor expansion:*

$$J(\epsilon) = J(0) + J_1\,\epsilon + J_2\,\epsilon^2 + o(\epsilon^2) \tag{67}$$

Here J_1 $(= \frac{d}{d\epsilon}J(0))$ *is called the* ***first variation of J*** *and* J_2 *is called the* ***second variation of J***.

By means of the calculus of variations we get:

$$\begin{aligned} J_1(\xi,\eta,\beta) &= (g_t + \nu^T\psi_t + H)|_{t=t_f}\,\xi_f + \left[(g_x + \nu^T\psi_x - \lambda^T)(x'\xi_f + \eta)\right]\Big|_{t=t_f} + \\ &\quad + \int_{t_0}^{t_f} [(H_x + (\lambda')^T)\eta + H_u\beta]dt \end{aligned} \tag{68}$$

$$\begin{aligned} J_2 &= \frac{1}{2}\Phi_{t_f t_f}\,\xi_f^2 + \Phi_{t_f x_f}\,\xi_f(x'\xi_f + \eta(t_f)) + \\ &\quad + \frac{1}{2}(x'_f\xi_f + \eta(t_f))^T\Phi_{x_f x_f}(x'_f\xi_f + \eta(t_f)) + \\ &\quad + \left[\frac{1}{2}(H_t - H_x x')\xi^2 + H_x\xi(x'\xi + \eta)\right]\Big|_{t=t_f} + \\ &\quad + \int_{t_0}^{t_f}\left(\frac{1}{2}\eta^T H_{xx}\eta + \beta^T H_{ux}\eta + \frac{1}{2}\beta^T H_{uu}\beta\right)dt \end{aligned} \tag{69}$$

where $\Phi(x(t_f), t_f) = (g + \nu^T\psi)|_{t=t_f}$.

The formula for the first variation is also valid for non-optimal extremals. With (68) and (69) we can prove Pontrjagin's Minimum Principle (theorem 2.1). Necessary for the minimality of J is the non-negativity of J_2 for every variation, i. e. for every family $(x(.,\epsilon), u(.,\epsilon))$ of admissible pairs satisfying $u(.,0) = u^*(.)$ and $x(.,0) = x^*(.)$, whereas the so called strong positivity (for every variation), which will be defined later on, is sufficient - always under the assumption $J_1 = 0$.

Theorem 5.2 *A* $(\mathbf{P_2})$ *be given. Let* (x^*, u^*, λ) *be an extremal which satisfies* $J_1 = 0$ *and* $J_2 \geq 0$ *for every variation.*
Then the set of variations which satisfy $\xi_f = 0$, $\eta \equiv 0$ *and* $\beta \equiv 0$ *minimizes* J_2.

Proof: If $\xi_f = 0$, $\eta \equiv 0$ and $\beta \equiv 0$ holds, then $J_2 = 0$ holds, whereas generally $J_2 \geq 0$ holds in case of the given extremal. **q. e. d.**

Hence we get the following problem:

Problem 5.2 ((AMP) = acessory minimum problem)
minimize J_2 *among all variations* ξ_f, η, β, *which satisfy*

$$\begin{aligned} \eta' &= (f_x \eta + f_u \beta)|_{\epsilon=0} && (70)\\ \eta(t_0) &= 0 && (71)\\ 0 &= [\psi_{t_f}\xi_f + \psi_{x_f} \cdot (x'\xi_f + \eta)]\,|_{\epsilon=0,t=t_f} && (72) \end{aligned}$$

As the optimum (x^*, u^*) is reached with $\xi_f = 0$, $\eta \equiv 0$ and $\beta \equiv 0$, the seeked solution of $(\mathbf{P_2})$ is consequently also a solution of **(AMP)**.
The conditions (70), (71) and (72) are necessary for the admissibility of each pair $(x(., \epsilon)), u(., \epsilon))$. Now we need the following abbreviations:

$$Q := H_{xx},\ C := H_{ux},\ R := H_{uu},\ Q_f := \Phi_{x_f x_f},\ A := f_x,\ B := f_u,\ D := \psi_{x_f} \quad (73)$$

Instead of η we will now write x and instead of β we choose u. Hence from (70) we get

$$x' = Ax + Bu\ , \quad (74)$$

from (71)

$$x(t_0) = 0 \quad (75)$$

and from (72)

$$Dx(t_f) = 0 \quad (76)$$

in case of fixed final time. Considering the **(AMP)** we minimize the second variation and denote it as new cost function $J_2[u(.)]$. Further on we now assume that $\xi_f = 0$ holds, i. e. we fix the final time. According to (69) we get:

$$J_2[u(.)] = \frac{1}{2}x^T(t_f)Q_f x(t_f) + \int_{t_0}^{t_f} \left(\frac{1}{2}x^T Qx + u^T Cx + \frac{1}{2}u^T Ru\right) dt \quad (77)$$

Hence the **(AMP)** consists of the minimization of (77) under the conditions (74), (75) and (76).

5.4 Kelley, Jacobson and Generalized Legendre-Clebsch-Condition

Sometimes one finds in literature the following theorem:

Theorem 5.3 (Kelley Condition)
Let (x^*, u^*, λ) *be an optimal extremal of a* $(\mathbf{P_2})$ *with fixed final time, where* u *is singular on whole* T. *Then*

$(H_u^{(2)})_u(u^*(t), x^*(t), \lambda(t), t)$ *is negative semidefinite for every* $t \in [t_0, t_f]$. *[3]*

The Kelley Condition is now only a special case of the GLC (see section 3).

Definition 5.2
a) $J_2[u(.)]$ is called ***non-negative*** *if its value is non-negative for every admissible $u(.)$.*
b) $J_2[u(.)]$ is called ***positive definite*** *if $J_2[u(.)] > 0$ ($\in \mathbb{R}$) holds for every admissible $u(.)$ with $u(.) \neq 0$ (0: zero function).*
c) $J_2[u(.)]$ is called ***strongly positive*** *if there exists a $k > 0$ such that $J_2[u(.)] \geq k\|u(.)\|^2$ holds for every admissible $u(.)$, where $\|.\|$ is a suitable norm in the set of the bounded piecewise continuous functions.*

We've already considered the GLC. Let's add that in the proof of the GLC the assumption, that we have a so called normal problem, is used, which is e. g. not required in case of Pontrjagin's Minimum Principle. (see [3] and [17])
Another condition is still missing:

Theorem 5.4 (Jacobson Condition)
Let $(\mathbf{P_2})$ be given with free final point and fixed final time t_f. Further on (x^*, u^*, λ) be an extremal such that $J_2[u(.)]$ is non-negative for every admissible variation $u(.)$ and $J_1 = 0$ holds. The control u^* be singular on whole T. Then the following property holds along (x^*, u^*, λ):

$$H_{ux} f_u + f_u^T W f_u \text{ is positive semidefinite for every } t \in]t_0, t_f[, \tag{78}$$
$$\text{where } -W' = H_{xx} + f_x^T W + W f_x,\ W(t_f) = (g(x_f))_{x_f x_f} \tag{79}$$

holds and W is a symmetrical and continuously differentiable $n \times n$ matrix function.

Regarding the proof we refer to [3].

Now we will use the denotations u and x for the control and the state again and not for the variations any more.

Example 5.2 *Jacobson ([18]) constructed the following problem:*

$$\begin{aligned} \text{minimize } J &= \int_0^2 x^2 dt - \frac{1}{2} S x(2)^2 \text{ for a given } S > 0, \text{ where} \\ x' &= u,\ x(0) = 1,\ U \equiv [-1, 1] \end{aligned}$$

We get $H_u = \lambda$. Thus the nonsingular control appears if $\lambda \neq 0$ and is given by $u = -sign\ \lambda$. The singular control satisfies:

$$u \equiv \lambda \equiv x \equiv 0$$

The GLC is always satisfied. Jacobson considers the following admissible control: $u(t) = -1$ for $t \in [0, 1]$ and $u(t) = 0$ for $t \in]1, 2]$, hence $J(u) = \frac{1}{3}$. The matrix function W in the Jacobson Condition is then given by $W(t) = -S + 2(2 - t)$. Property (78) is equivalent to $t \leq 2 - \frac{1}{2}S$. This condition is independent of the extremal, i. e. the optimal control may only be singular up to $t = 2 - \frac{1}{2}S$. Hence Jacobson's candidate is not optimal.
Gerald M. Anderson ([19]) instances another admissible control for $0 < S < 1$: $u(t) = -1$ for $t \in [0, 1]$, $u(t) = 0$ for $t \in]1, 2 - S]$ and $u(t) = 1$ for $t \in]2 - S, 2]$. The Jacobson Condition is satisfied for this candidate and we get $J(u) = \frac{1}{3} - \frac{1}{6}S^3 < \frac{1}{3}$.

Anderson adds an explanation why the Jacobson Condition is more meaningful than the GLC in example 5.2. In the proof of the Jacobson Condition (see [3]) the following control variation is used:

$$\begin{aligned} \beta(t) &= 0 \text{ if } t \notin [t_1, t_1 + \Delta] \subsetneq T \\ \beta(t) &= \Gamma \text{ if } t \in [t_1, t_1 + \Delta] \end{aligned}$$

This is called a Single Pulse Control Variation (SPCV) whereas in the proof of the GLC a Double Pulse Control Variation (DPCV) is used:

$$\begin{aligned}\beta(t) &= +\Gamma \text{ if } t \in [t_1, t_1+\Delta] \\ \beta(t) &= -\Gamma \text{ if } t \in [t_2, t_2+\Delta] \\ \beta(t) &= 0 \text{ else; } t_1+\Delta < t_2\end{aligned}$$

We will denote these variations here as δu_S and δu_D.
If $x'_S = u^* + \delta u_S$ holds, where (x^*, u^*) is an arbitrary admissible pair, then we get here: $x_S(t) = x^*(t) + \delta x_S(t)$ with

$$\delta x_S(t) = 0 \text{ for } t \leq t_1,\ \delta x_S(t) = \Gamma(t-t_1) \text{ for } t_1 < t \leq t_1+\Delta,$$

$$\delta x_S(t) = \Gamma\Delta \text{ for } t > t_1+\Delta \Rightarrow x_S(2) = x^*(2) + \Gamma\Delta$$

On the other hand $x'_D = u^* + \delta u_D \Rightarrow x_D(t) = x^*(t) + \delta x_D(t)$ holds with

$$\begin{aligned}\delta x_D(t) &= 0 \text{ for } t \leq t_1,\ \delta x_D(t) = \Gamma(t-t_1) \text{ for } t_1 < t \leq t_1+\Delta, \\ \delta x_D(t) &= \Gamma\Delta \text{ for } t_1+\Delta < t \leq t_2, \\ \delta x_D(t) &= \Gamma(\Delta - t + t_2) \text{ for } t_2 < t \leq t_2+\Delta, \\ \delta x_D(t) &= 0 \text{ for } t > t_2+\Delta \\ \Rightarrow x_D(2) &= x^*(2).\end{aligned}$$

Hence the Mayer term is invariant with respect to a DPCV but not with respect to a SPCV. Thus the Jacobson Condition is so to speak more sensitive than the GLC in example 5.2.
However Jacobson and Anderson apparently overlooked the fact that u is singular iff $\lambda = 0$ holds, which is an obvious contradiction to Pontrjagin's Minimum Principle (theorem 2.1). Hence an optimal control can't have a singular part.

5.5 More Necessary Conditions

At first we define a general Mayer Problem:

Problem 5.3 *(**MP**) be a special case of (**P**), where $\operatorname{int} U(t) \neq \emptyset$ for every t, $Z = \mathbb{R}^n$ and $U(t) \subset \mathbb{R}$ hold. The functional, that has to be minimized, be $J(u) = g(x(t_f))$ and the functions f and g be piecewise analytical. Again we're looking for a piecewise continuous control.*

We first consider the following theorem for Mayer Problems, which are of course not only in the linear case equivalent to Bolza Problems.

Theorem 5.5 *(Skorodinskii 1979) An (**MP**) with optimal control u^* and respective extremal (x^*, u^*, λ) be given such that the following properties hold:*

$$u(t) \in \operatorname{int} U(t) \text{ for every } t \in [a, b[\subset T \ (a < b) \tag{80}$$
$$\text{and } H_{uu}(u^*, x^*, \lambda, t) = 0 \text{ for every } t \in [a, b[. \tag{81}$$

Then the following property holds along the extremal for every $p \in \mathbb{R}$ and $t \in [a, b[$:

$$S(t,p) := (H_u^{(2)})_u - 3p\left(\frac{d}{dt}(H_{uu})\right)_u - p^2\frac{\partial^4}{\partial u^4}H \leq 0 \tag{82}$$

The proof uses DPCVs and can be read in [20]. Skorodinskii also proved the following corollary:

Corollary 5.2 *Under the assumptions of theorem 5.5 the following properties hold:*
a) $(H_u^{(2)})_u \leq 0$ *and* $\frac{\partial^4}{\partial u^4} H \geq 0$ *along the extremal for every* $t \in [a, b[$.
b) If $t = \tau \in [a, b[$ *and* $\frac{\partial^4}{\partial u^4} H > 0$ *then*

$$\left(\left[\frac{\frac{3}{2} \left(\frac{d}{dt}(H_{uu}) \right)_u}{\frac{\partial^4}{\partial u^4} H} \right]^2 + \frac{(H_u^{(2)})_u}{\frac{\partial^4}{\partial u^4} H} \right) \Bigg|_{t=\tau} \leq 0$$

c) Let $t = \tau \in [a, b[$ *and the properties* $(H_u^{(2)})_u = 0$ *or* $\frac{\partial^4}{\partial u^4} H = 0$ *hold along the extremal for every* $t \in [a, b[$. *Then the following identity holds:*

$$\left(\frac{d}{dt}(H_{uu}) \right)_u (u^*(\tau), x^*(\tau), \lambda(\tau), \tau) = 0$$

Regarding the proof we again refer to [20]. The statement $(H_u^{(2)})_u \leq 0$ in part a) of corollary 5.2 is the Kelley Condition (theorem 5.3).
From the theory of differential equations we know the following lemma ([21]):

Lemma 5.1 (Gronwall Lemma)
If $\Phi(t) = h(t) + \int_0^t k(\tau)\Phi(\tau)d\tau$ *holds on* $[a, b]$ $(a < b)$; $h, k \in C^0([a, b])$, $k \geq 0$, *then:*

$$\Phi(t) \leq h(t) + \int_0^t k(\tau)h(\tau)exp \left(\int_\tau^t k(s)ds \right) d\tau$$

For the next theorem we need another definition and some lemmas.

Definition 5.3 *Let* u *and* $\hat{u}$ *be two admissible controls of a* ***(P)***. *Then we define:*
$\mathbf{E} := \{t \in T \mid u(t) \neq \hat{u}(t)\}$, $\mathbf{d_1}(u, \hat{u}) := \lambda(E)$ *(*λ*: Lebesque measure)*

Sometimes we will need the following assumption:

Assumption 5.1 *(with respect to* ***(P)****)*
(i) $f, g, L \in C^0$; f_x, g_x *and* L_x *exist and be continuous.*
(ii) There exist an $M \in \mathbb{R}$ *such that for every* $u \in U(t)$, $x \in \mathbb{R}^n$ *and every* $t \in T$ *the following inequality holds:*

$$\|f(x, u, t)\| \leq M(\|x\| + 1) \tag{83}$$

By $\|.\|$ *we mean the euklidean norm in* $\mathbb{R}^n$. *[1]*

Lemma 5.2 *(Mayne 1973)*
Let (x, u) *and* $(\hat{x}, \hat{u})$ *be two arbitrary admissible pairs for* ***(P)*** *under the assumption 5.1. Let* $\Delta u := u - \hat{u}$ *and* $\Delta x := x - \hat{x}$. *Then there exists a* $c \in \mathbb{R}^+$ *such that* $\|\Delta x(t)\| \leq c \cdot d_1(u, \hat{u})$ *for every* $t \in T$.

Regarding the proof we refer to [22].

Lemma 5.3 *Let* (x^*, u^*, λ) *be an optimal extremal for* ***(P)*** *under the assumption 5.1. Let* $\Delta u := u - u^*$ *and* $\Delta x := x - x^*$ *and*

$$\begin{aligned} \Delta \mathcal{J} \quad := \quad & (g_x^T(x^*) - \lambda^T)\big|_{t=t_f} \Delta x(t_f) + \\ & + \int_{t_0}^{t_f} \left(\Delta H + \left[H_x^T(u^* + \Delta u, x^*, \lambda, t) + (\lambda')^T(t) \right] \Delta x(t) \right) dt, \end{aligned}$$

*where $\Delta H = H(u, x^*, \lambda, t) - H(u^*, x^*, \lambda, t)$. Then $\Delta \mathcal{J} \geq 0$ for every admissible pair (x, u) of **(P)** with sufficiently small $d_1(u, u^*)$. In other words: There exists a $d > 0$ such that for every **(P)**-admissible pair (x, u) that satisfies $d_1(u, u^*) < d$ the inequality $\Delta \mathcal{J} \geq 0$ holds.*

Regarding the proof we refer to [1]. As Pontrjagin's Minimum Principle, which is not helpful in the singular case, is a first order condition, it appears to be suggesting to look for a second order property.

Theorem 5.6 (Gift's Second Order Minimum Principle (1993))
*Let (x^*, u^*) be an optimal pair of **(P)** under the assumption 5.1, where u^* be singular in the sense of the Minimum Principle (see definition 2.5) on whole T. Further on $\frac{d}{dt}H_x$ exist. Then the following properties hold:*
a) $\Delta H_x^T \cdot \Delta H_\lambda \geq 0$ for every admissible pair (x, u) with $u(t) \in W$ and sufficiently small difference $\sup E - \inf E > 0$ and for every $t \in T$.
*b) If **(P)** is linear in scalar u and $H \in C^2$ holds, then $H_{ux}^T H_{u\lambda} \geq 0$ holds along the optimal extremal for every $t \in T$.*
Here we used the abbreviations
$\Delta H_x := H_x(u, x^*, \lambda, t) - H_x(u^*, x^*, \lambda, t)$ and
$\Delta H_\lambda := H_\lambda(u, x^*, \lambda, t) - H_\lambda(u^*, x^*, \lambda, t)$

The theorem was proved in [1]. In [23] Zhou claims to have a counterexample for theorem 5.6. But this counterexample is incorrect. Zhou uses $u^* \equiv 0$ as optimal control and $u \equiv 1$ as variation. But this means that $\sup E - \inf E$ is not sufficiently small but has maximum size.
Another necessary condition is:

Theorem 5.7 *(Kalinin 1985)*
*Condition: Let **(P)** with final condition $\varphi_i(x(t_f)) \leq 0$ $(1 \leq i \leq s)$, $\varphi_i(x(t_f)) = 0$ $(s+1 \leq i \leq m)$, fixed final time t_f and cost function $J(u) = \varphi_0(x(t_f))$ be given. Here $f, \varphi_i \in C^2$ $(0 \leq i \leq m)$ hold and $U(t)$ be open and non-empty for every $t \in T$. For this problem there exist an optimal pair (x^*, u^*), where u^* be singular on whole T and additionally to the piecewise continuity also leftsided or rightsided continuous in the discontinuity points. Further on $\varphi_i(x^*(t_f)) = 0$ also hold for $1 \leq i \leq s$.*
Statement: Then for every $t \in T$ and every with respect to the above problem admissible control u with

$$H_u^T(u^*, x^*, \psi_i, t)u \leq 0 \text{ for } 0 \leq i \leq s,\ H_u^T(u^*, x^*, \psi_i, t)u = 0 \text{ for } s+1 \leq i \leq m \quad (84)$$

the following inequality holds:

$$\min\{u^T[H_{xu}(u^*, x^*, \psi_\alpha, t)f_u(t, x^*, u^*) + \\ + f_u^T(t, x^*, u^*)\Psi_\alpha(t)f_u(t, x^*, u^*)]u \mid \alpha \in A(u^*)\} \geq 0,$$

where $\psi_i(t) \in \mathbb{R}^n$ $(0 \leq i \leq m)$ and $\psi_\alpha(t) \in \mathbb{R}^n$ $(\alpha \in A(u^*))$, $t \in T$, are solutions of

$$\psi_i' = -H_x(u^*, x^*, \psi_i, t),\ \psi_i(t_f) = (\varphi_i)_x(x^*(t_f))$$

$$\psi_\alpha' = -H_x(u^*, x^*, \psi_\alpha, t),\ \psi_\alpha(t_f) = \sum_{i=0}^{m} \alpha_i(\varphi_i)_x(x^*(t_f))$$

and $\Psi_\alpha(t)$, $t \in T, \alpha \in A(u^*)$, are $n \times n$-matrices and solutions of

$$\Psi_\alpha' = -f_x^T(t, x^*, u^*)\Psi_\alpha - \Psi_\alpha f_x(t, x^*, u^*) - H_{xx}(u^*, x^*, \psi_\alpha, t),$$

$$\Psi_\alpha(t_f) = \sum_{i=0}^{m} \alpha_i(\varphi_i)_{xx}(x^*(t_f)).$$

Further on
$A(u^*) :=$
$$\left\{ \alpha \in \mathbb{R}^{m+1} \;\middle|\; \sum_{i=0}^{m} |\alpha_i| = 1;\ \alpha_i \geq 0,\ 0 \leq i \leq s;\ H_u(u^*, x^*, \psi_\alpha, t) = 0\,\forall t \in T \right\}$$

The proof can be found in [24]. An example is given there, which shows that condition (84) is really needed.
Another necessary condition is given in [25]. But this one has the disadvantage that a partial differential equation has to be solved, which is generally different for each extremal. Thus this condition is only useful if certain candidates are to be excluded.

5.6 Sufficient Conditions

We consider again ($\mathbf{P_2}$). With the denotations of (73) we get from (74), (75) and (77) as the second variation of an unbounded ($\mathbf{P_2}$), i. e. with $D \equiv 0$:

$$J_2[u(.)] = \frac{1}{2}x^T(t_f)Q_f x(t_f) + \int_{t_0}^{t_f} \left(\frac{1}{2}x^T Q x + u^T C x + \frac{1}{2}u^T R u\right) dt \quad (85)$$
$$\text{where } x' = Ax + Bu \text{ and } x(t_0) = 0 \quad (86)$$

According to the Legendre-Clebsch-Condition $R(t) \geq 0$ holds for optimal controls. The matrix functions Q, C, R, A and B are continuous on T, and Q_f is constant. We may assume that $H, g \in C^2$ holds (see ($\mathbf{P_2}$)) and consequently Q, R and Q_f are symmetrical.

Definition 5.4 *The dynamical system (86) with respect to* ($\mathbf{P_2}$) *is* ***completely controllable*** *on* $[t_0, t']$, *where* $t_0 < t' \leq t_f$, *if*

$$\int_{t_0}^{t'} \psi(t', \sigma)\, B(\sigma)\, B^T(\sigma)\, \psi^T(t', \sigma)\, d\sigma > 0$$

with $\frac{\partial}{\partial t}\psi(t, \sigma) = A(t)\psi(t, \sigma)$; $\psi(\sigma, \sigma) = I$ *(identity matrix).*
The system is called completely controllable if it is completely controllable on $[t_0, t']$ *for every* $t' \in]t_0, t_f]$. *[3]*

Definition 5.5 *Let* M *be a symmetrical* $n \times n$*-matrix. We write*
a) $M > 0$, *if* M *is positive definite.*
b) $M \geq 0$, *if* M *is positive semidefinite.*

Definition 5.6 *The second variation* $J_2[u(.)]$ *according to (85) is called*
a) nonsingular, if $R(t) > 0$ *holds for every* $t \in T$.
b) totally singular, if $R(t) = 0$ *holds for every* $t \in T$.
c) partially singular, if $R(t) \geq 0$ *holds for every* $t \in T$. *[3]*

Mind that we're talking about singularity of the second variation and not of the control. The case $R(t) < 0$ need not be considered because of the Legendre-Clebsch-Condition.
The strong positivity of $J_2[u(.)]$ is sufficient for optimality. If $J_2[u(.)]$ is nonsingular there exists a very helpful theorem ([3]):

Theorem 5.8
Condition: The system (86) be completely controllable and $J_2[u(.)]$ according to (85) be nonsingular.
Statement: $J_2[u(.)]$ is strongly positive $\Longleftrightarrow$ There exists an $n \times n$-matrix function $S(.)$ on T which satisfies the ***Riccati Equation:***

$$\begin{aligned} -S' &= Q + SA + A^T S - (C + B^T S)^T R^{-1}(C + B^T S) & (87) \\ S(t_f) &= Q_f & (88) \end{aligned}$$

A nonsingular second variation can only exist in the nonlinear case as otherwise $R \equiv 0$ would hold.
A similar condition like theorem 5.8 can't be constructed in the totally singular case for the following reason ([26]):

Theorem 5.9
Condition: $(\mathbf{P_2})$ have a solution u^*.
Statement: A totally singular second variation $J_2[u(.)]$ is never strongly positive.

In the case of a totally singular second variation the Riccati Equation is not applicable, as R^{-1} doesn't exist.
As the space of all piecewise continuous control functions is infinite dimensional, we can't assume the equivalence of the positive definiteness and the strong positivity any more, as the first one is necessary but not sufficient for the second one. An example is given in [3].

Some conditions which are sufficient for non-negativity or positive definiteness of the second variation $J_2[u(.)]$ are given in [3]. A sufficient condition is given in [27].

5.7 The Solution of Example 2.1

The functional has the property $J(u) = \int_0^2 x^2(t)(u(t)+1)dt \geq 0$, as $u(t) \geq -1$ holds.
A purely singular control would not satisfy the final condition $x(2) = -1$. A purely nonsingular control appears not to be optimal. Hence we've got a junction. We already know that the control must be discontinuous in the junction. The sequence bang-bang $\rightarrow$ singular is not optimal as we would get $x \neq 0$ in the singular control part and we've seen before that in optimal singular control parts $x \equiv 0$ must hold. But with singular $\rightarrow$ bang-bang we get $J(u^*) = 0$ if:

$$\begin{aligned} u^*(t) = 0, \text{ if } 0 \leq t \leq 1 \quad &; \quad u^*(t) = -1, \text{ if } 1 < t \leq 2 \\ \Rightarrow x^*(t) = 0, \text{ if } 0 \leq t \leq 1 \quad &; \quad x^*(t) = -t + 1, \text{ if } 1 < t \leq 2 \end{aligned}$$

The junction $t_c = 1$ is given by the boundary conditions and the continuity of the state.
The constructed control is obviously optimal. The way we solved the problem is exemplary for the general treatment of problems with possibly singular parts. First we try to gain as many cognitions as possible from known conditions, then we hope to find the solution by considering specialties of the problem.

6 Numerical Methods

6.1 Proposal of a strategy

We restrict ourselves to problems which are linear in u. Different methods for determining optimal controls with possibly singular part have been developed so

far. The BFGS and the DFP methods from nonlinear optimization e. g. were taken over. These decay methods are, like other decay methods, also applicable to purely nonsingular and nonlinear problems.
Trying to attain a solution numerically we first try to determine the structure of the solution, i. e. the sequence of the bang bang and the singular parts, where we also determine for the bang bang segments on which boundary the solution is situated. In example 2.1 the structure of the considered solution is consequently $(s, -1)$, where s represents the singular part. In subsection 5.2 we've seen that often the singular controls can be determined as functions $u = \psi(t, x, \lambda)$ or even $u = \psi(t, x)$.
Often the junctions are unknown and we only get conditions $S_j(x(t_j)) = 0$ which must be satisfied. We get them e. g. for singular parts from the equations $H_u^{(\nu)} = 0$. They can also be a result of the boundary condtitions and the continuity of the state.
We start at a known junction t_k, e. g. t_0 or t_f, which are also considered here as junctions, and integrate forward and backward until the conditions $S_j(x(t_j)) = 0$ are satisfied. That's why these conditions are also called stopping conditions. A stopping condition of the kind $S_j(x(t_j)) = x(t_j) - x_j$ is called completely specified. Else it's called incompletely specified. In the case $S_j \equiv 0$ the stopping condition is missing.
Hence we start at t_k $(t_0 < t_1 < ... < t_{\mu-1} < t_\mu = t_f)$ and solve the differential equations. As initial value $x(t_k)$ or $\lambda(t_k)$ can serve. Here four types of problems are distinguished according to Fraser-Andrews ([28]):

a) Some initial values of the differential equations are needed. (appears if $u = \psi(t, x, \lambda)$ with incompletely specified $\lambda(t_k)$ or $u = \psi(t, x)$ with incompletely specified $x(t_k)$ occurs)
b) No initial values are needed but some junctions haven't got a stopping condition. (appears if $u = \psi(t, x)$ with completely specified $x(t_k)$ occurs but $S_j \equiv 0$ for at least one j)
c) No initial values are needed and all junctions have got a stopping condition. (appears if $u = \psi(t, x)$ with completely specified $x(t_k)$ occurs and $S_j \not\equiv 0$ for every j)
d) A representation $u = \psi(t, x)$ or $u = \psi(t, x, \lambda)$ is not possible, wherefore x can't be determined by integration.

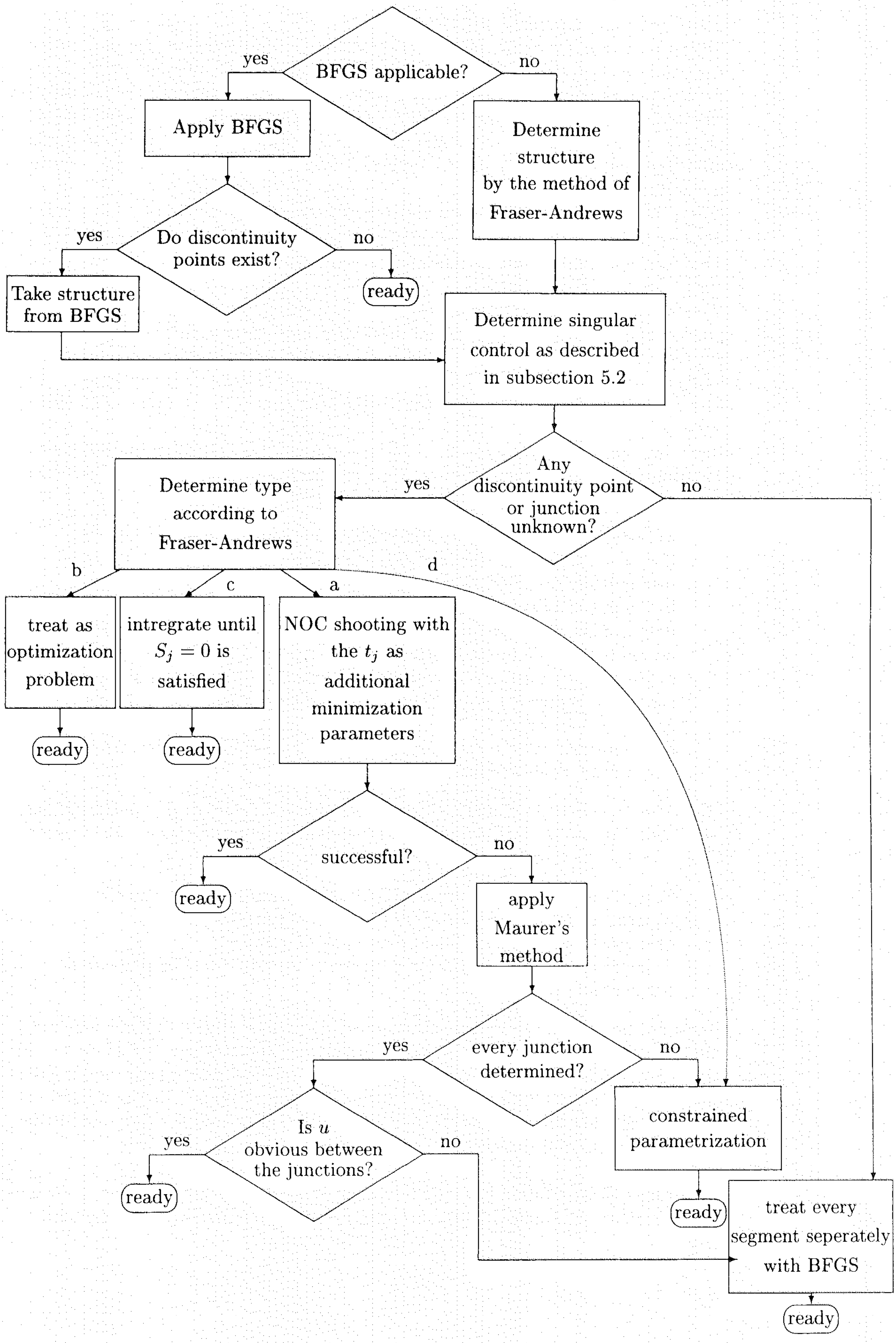
Fraser-Andrews examines in [29] different methods and comes to the conclusion that NOC shooting and constrained parametrization are the best methods. He proposes ([28]) to treat the four types in the following way:

a) NOC shooting with the t_j as additional minimization parameters
b) treat as optimization problem (degenerate case of a))
c) integration until $S_j = 0$ and determine t_j from it
d) constrained parametrization with the t_j als additional minimization parameters

Based on this and on experiences of the authors a general strategy for the numerical solution of optimal control problems is proposed. The strategy is represented in figure 6.1.1. The diagram is also applicable to problems with purely nonsingular optimal control. Of course, if one reaches the point "ready" in the diagram it can't be excluded that no suitable solution has been obtained. For this case further methods are listed in this section.

If the optimal solution has infinitely many discontinuity points, like e. g. in the Fuller Problem (example 3.3), the arcs are called "chattering arcs" and finitely

Figure 6.1.1: Strategy for the numerical solution of optimal control problems



many discontinuity points are chosen for numerical treatment (see [28]).
Constrained Parametrization means that the constraint $u(t) \in [a, b]$ is treated with the ansatz $u(t) = \frac{1}{2}(a+b)+\frac{1}{2}(b-a)\sin v(t)$, where $v(t)$ is an expansion in finitely many orthogonal functions, like e. g. Tschebyscheff polynomials: $v(t) = \sum_{j=0}^{m} a_j T_j(t)$ (m: order of the expansion) (see [29])
The named methods and some others are treated in section 6.2. Two types of methods are distinguished: the direct methods, where first the differential equations and the boundary conditions are satisfied and then the costs are lessened, and the indirect methods, which start with optimality conditions and then try to satisfy the boundary conditions. ([30])

6.2 The Methods

6.2.1 BFGS and Other Decay Methods

The decay methods were taken over from nonlinear optimization. A detailed description of these methods and remarks how to implement them on computers are given by Edge and Powers in [31] and [32]. A part of the algorithm of every decay method is a linear search. The following theorem holds ([33]):

Theorem 6.1 *(Dixon 1971) If an exact linear search is made then DFP and BFGS method generate the same directions for searching.*

Especially in case of singular controls DFP and BFGS method have proved to be the best decay methods, where the experience shows that the BFGS method often has a better convergence than the DFP method if non-exact linear search is made ([32]). The gradient method is problematical in the singular case ([31]).
If junctions appear the gradient method is unuseful, whereas the BFGS method is at least useful for the determination of the structure. That's enough for the task the method has in the proposed strategy.

6.2.2 Maurer's Method and the NOC Shooting

Both methods are shooting methods and belong to the indirect methods.

Maurer's Method In the middle of the 70s Maurer ([34]) developed a method to determine junctions numerically. It's applicable to problems of the kind **(MLP)** with final condition $r(x(t_f)) = 0$ ($\in \mathbb{R}^b$) and $q = 1$. Further on we assume the following basic structure:
nonsingular on $[t_0, t_1[$, singular on $[t_1, t_2[$ and nonsingular on $[t_2, t_f]$.
Moreover we need a representation $u = \psi(t, x, \lambda)$ (or $u = \psi(t, x)$) for the singular control.
The method converts the problem into a two point boundary value problem (tpbvp), which can be solved by the multi shooting method ([35]), which was developed by Bulirsch, Stoer and Deuflhard, who used ideas of Keller.
Newton's method, which is used in the algorithm, can be replaced by the modified Newton's method ([36]) which is less sensitive to bad initial approximations. In [37] a Fortran program for the multi shooting method is given. Some propositions on the solvability of the tpbvp are given in [34] and [35]. Often there's a stabilizing effect if some differential equations or boundary conditions are omitted, which is unfortunately not always possible (see [34] and [28]). If several singular subarcs appear then the corresponding equations and conditions are added to the tpbvp for each singular subarc. ([34])
Maurer's method is non applicable if the optimal control is not piecewise analytical

in a neighbourhood of the junction and in the rare cases with odd q (here $q = 1$) and a control which is continuous in the junction ([28]).

NOC Shooting Classical indirect methods start with an estimation of $\lambda(t_0)$, which is problematical, as the methods are very sensitive to bad initial values and the costates λ often can't be illustrated ([38]). The NOC shooting starts with the initial controls, which are easier to estimate, and is less sensitive to bad initial values. The method is described in [38] and is applicative to problems of the kind **(P)** with final condition $r(x(t_f)) = 0$ and cost function $J = g(x(t_f))$, where $x(t) \in \mathbb{R}^n$ and $u(t) \in \mathbb{R}^s$ holds. The equations $x' = f$ and $\lambda' = -H_x$ must be stable with respect to forward integration. Further on H should be unimodal and convex.
In [30] the trajectory of a sattelite that leaves an earth orbit and enters a mars orbit was optimized.
The last step of the NOC shooting algorithm is a finite-dimensional minimization. The recursive quadratic programming technique is recommended for this ([39]). For NOC shooting the structure must be known or derivable from a more general structure. The switching times should be initially well approximated. Although the method is less sensitive because of the usage of the control instead of the costates λ, good initial values should be chosen nevertheless. For $s > 1$ the control problem may also contain more complicated minimizations. [28]

6.2.3 The Method by Fraser-Andrews for the Determination of the Structure

The method starts with the ansatz $\sum\limits_{i=0}^{m} a_i P_i(t)$, where the P_i are orthogonal functions. Then the interval $[t_0, t_f]$ is partitioned iteratively and a minimization over $\{a_i | i = 1, ..., m\}$, the interval length and eventually t_f is done. The method is described in [28].

6.2.4 Other Methods

The **ϵ-method** by Bell and Jacobson ([3]) transforms the problem **(P)** into a (normally) nonsingular one by adding the perturbance term $\frac{1}{2}\epsilon_k \int\limits_{t_0}^{t_f} u^T u \, dt$ to the functional $J(u)$ and decreasing ϵ_k stepwise.
As expected the method shows numerical instability for $\epsilon \to 0$ ([40]). The ϵ-method was used among others for computations which were made for chemical reactors. It's analogous to a method in fluid dynamics. ([41])

Iterative Dynamic Programming by Rein Luus ([41]) is a direct method for the problem:

$$\text{minimize } g(x(t_f)), \text{ where } x' = f(x,u),\ x(t_0) = x_0 \in \mathbb{R}^n,$$
$$\alpha \leq u(t) \leq \beta,\ t_f \text{ fixed}$$

The control u is approximated as step function and a grid of possible state values is constructed through which a minimal way is searched.
The method is also able to find several optima (local and global). It's an alternative if it's difficult to apply other methods. In case of nonlinear systems another method should be chosen for validation. The advantage of Iterative Dynamic Programming is the easy programmability and the fact that the costate equations $\lambda' = -H_x$ need not be integrated. [41]

Further methods are the modified quasilinearization technique by Aly and Chan ([42]), which is an extension of the quasilinearization technique by Baird ([43]) and can only be used for problems with purely singular optimal controls, and the control averaging technique by Virendra Kumar ([44]), which reduces optimal control problems to fixed point problems.

For the methods treated in section 6 almost no convergence property is known.

7 Supplements

Singular optimal controls have appeared in many applications like ecology, aerospace, theoretical biology, chemical engineering, epidemiology, robotics and economy (cf. e. g. [5] and [45] to [52]).
The set of all points in $\mathbb{R}^n$ that are states with respect to singular controls is sometimes called singular surface. In [53] a problem of kind **(P)** is considered, where the free final time instead of the functional J is to be minimized.
In [54] Jacobson gives switching strategies for quadratic problems. In [55] it's proved that the singular controls need not be unique in case of linear Lagrange problems. The author J. Grasman considers in [56] also non-unique singular solutions and uses so called nearly singular problems.
A condition for the existence of weak solutions if generalized controls are considered is proved in [57]. Especially the case of singular optimal controls is treated.
In [58] Françoise Lamnabhi-Lagarrigue treats a unification of the definition of the arc order. At this a definition is given, which is also applicable to problems with final constraints. Lie-brackets are used for this.
A connection between optimal controls, especially singular ones, and the rank of a certain matrix is derived in [59].
A geometric treatment of singular optimal controls is given in [60]. A maximum principle of higher order is treated in [61]. Numerous interesting and useful cognitions, which were only partially given in this work, are due to Gabasov and Kirillova ([6]).

8 Conclusions

Analytical and numerical methods for the treatment of optimal controls which have singular parts were treated. The state of the art was presented. But the present theory is not sufficient for having a satisfying general concept for solving the problems. Often considerations which are specific for the problem are necessary. A unification of these considerations has not been possible yet.
The GLC and the Jacobson Condition appear to be the most helpful optimality conditions, with which many candidates can be excluded. But we have seen that these conditions are only necessary conditions. The almost complete lack of sufficient conditions is a big problem. Hence we're now able to determine admissible partially singular controls which satisfy every imaginable necessary condition, but we often can't clarify if the considered control is indeed optimal. Cases, like in example 2.1, where the optimality can be proved by giving a bound for the costs, are rather rare. Often a control is found, which is presumed to be relatively close (w. r. t. costs) to the optimum, if one exists.
In this connection one will often use numerical methods. Here we can be content with the existing methods. Numerous numerical treatments of practial problems show this. The most difficult part appears to be the determination of the junctions. This is also true w. r. t. theoretical aspects. Until 1995 nothing was known about

the location of the junctions. The junction conditions were not very helpful. The fact that in 1995 the first theorem about a connection between the location of the junctions and the boundary conditions could be proved gives hope that more knowledge will be gained about this matter in the future.
The main activities concerning singular optimal controls took place in the 1960s and the 1970s. But research on this topic is still active. The manifold applications will put forward the investigation of singular optimal controls. A property which is for singular controls as helpful as Pontragin's Minimum Principle is in the whole theory of optimal control is desirable.

9 Acknowledgements

This work is an extract of the diploma thesis written by Volker Michel with care of Prof. Dr. Horst Becker, who also gave useful comments for this paper. Volker Michel is now member of the Graduiertenkolleg Technomathematik at the university of Kaiserslautern and PhD student in the research group headed by Prof. Dr. Willi Freeden.

References

[1] Gift, S. J. G.: Second-Order Optimality Principle for Singular Optimal Control Problems, JOTA, volume 76 (1993), pages 477-484

[2] Pinch, Enid R.: Optimal Control and the Calculus of Variations, Oxford University Press 1993

[3] Bell, David J. and Jacobson, D. H.: Singular Optimal Control Problems, Mathematics in Science and Engineering, volume 117, Academic Press, London 1975

[4] Pontrjagin, L. S. et al.: L. S. Pontrjagin Selected Works, volume 4, The Mathematical Theory of Optimal Processes, translated from Russian by K. N. Triroyoff, Gordon and Breach Science Publishers, Montreux/Switzerland 1986

[5] Kelley, H. J.; Kopp, R. E. and Moyer, H. G.: Singular Extremals, in: Topics in Optimization, edt. by G. Leitman, Academic Press, New York 1967, pages 63-101

[6] Gabasov, R. and Kirillova, F. M.: High Order Necessary Conditions for Optimality, SIAM Journal Control, volume 10 (1972), pages 127-168

[7] Fraser-Andrews, G.: Finding Candidate Singular Optimal Controls: A State of the Art Survey, JOTA, volume 60 (1989), pages 173-190

[8] Lewis, R. M.: Defintions of Order and Junction Conditions in Singular Optimal Control Problems, SIAM Journal Control and Optimization, volume 18 (1980), pages 21-32

[9] Powers, W. F.: On the Order of Singular Optimal Control Problems, Journal of Optimization Theory and Applications (**JOTA**), volume 32 (1980), pages 479-489

[10] Robbins, H. M.: A Generalized Legendre-Clebsch-Condition for the Singular Cases of Optimal Control, IBM Journal Res. Develop., volume 11 (1967), pages 361-372

[11] McDanell, J. P. and Powers, W. F.: Necessary Conditions for Joining Optimal Singular and Nonsingular Subarcs, SIAM Journal Control, volume 9 (1971), pages 161-173

[12] Bell, D. J.: Optimality Conditions at Junctions of Singular and Nonsingular Controls, JOTA, volume 78 (1993), pages 1-8

[13] Bell, D. J. and Boissard, M.: Necessary conditions at the junction of singular and non-singular control subarcs, International Journal of Control, volume 29 (1979), pages 981-990

[14] Bortrins, Richard: On joining non-singular and singular optimal control subarcs, International Journal of Control, volume 37 (1983), pages 867-872

[15] Pan Guo-Lin and Bell, D. J.: A Counterexample to a Conjecture of McDanell, Systems & Control Letters, volume 9 (1987), pages 85-87

[16] Ruxton, David J. W. and Bell, David J.: Junction Times in Singular Optimal Control, Applied Mathematics and Computation, volume 70 (1995), pages 143-154

[17] Bliss, G. A.: Lectures on the Calculus of Variations, Univ. Chicago Press, Chicago 1946

[18] Jacobson, D. H.: A new necessary condition of optimality for singular control problem, SIAM Journal Control, volume 7 (1969), pages 578-595

[19] Anderson, Gerald M.: A Note on Jacobson's Necessary Condition for an Optimal Singular Control, IEEE Transactions on automatic Control, volume AC-16 (1971), pages 107-108

[20] Skorodinskii, I. T.: Singular Optimal Control, Autom. Remote Control, volume 41 (1981), pages 1210-1213

[21] Walter, Wolfgang: Gewöhnliche Differentialgleichungen: eine Einführung, 4. Auflage; Springer Verlag, Berlin Heidelberg 1990

[22] Mayne, D. Q.: Differential Dynamic Programming: A Unified Approach to the Optimization of Dynamic Systems, Advances in Control Systems, edt. by C. T. Leondes, Academic Press, New York, volume 10 (1973), pages 179-254

[23] Zhou, Q.: Second Order Optimality Principle for Singular Optimal Control Problems, JOTA, volume 88 (1996), pages 247-249

[24] Kalinin, A. I.: Singular Optimal Control Problems, Differential Equations, volume 21 (1985), pages 240-244

[25] Gabasov, R.: Necessary Conditions for Optimality of Singular Control, Engineering Cybernetics, volume 5 (1968), pages 28-37

[26] Jacobson, David H. and Speyer, Jason L.: Necessary and Sufficient Conditions for Optimality for Singular Control Problems: A Limit Approach, Journal of Mathematical Analysis and Applications, volume 34 (1971), pages 239-266

[27] Dykhta, V. A.: Sufficient conditions of the strong local extremum for the singular optimal control problems, Problemy ustojcivosti dvizenija, analiticeskoj mechaniki i upravlenija dvizeniem. Irkutsk, 20. - 25. 6. 1977. Novosibirsk 1979 (=Cetaevskaja Konferencija po Analiticeskoj Mechanike, Ustojcivosti Dvizenija i Optimal nomu Upravleniji. 3,1) pages 305-309 (in Russian)

[28] Fraser-Andrews, G.: Numerical Methods for Singular Optimal Control, JOTA, volume 61 (1989), pages 377-401

[29] Fraser-Andrews, G.: Numerical Techniques for Singular Optimal Trajectories, Hatfield Polytechnic, Hatfield, Herfordshire, England, PhD Thesis, 1986

[30] Dixon, L. C. W and Biggs, M. C.: The advantages of adjoint-control transformations when determining optimal trajectories by Pontrjagin's Maximum Principle, Aeronautical Journal, March 1972, pages 169-174

[31] Edge, Ernest R. and Powers, William F.: Shuttle Ascent Trajectory Optimization with Function Space Quasi-Newton Techniques, AIAA Journal, volume 14 (1976), pages 1369-1379

[32] Edge, Ernest R. and Powers, William F.: Function-Space Quasi-Newton Algorithms for Optimal Control Problems with Bounded Controls and Singular Arcs, JOTA, volume 20 (1976), pages 455-479

[33] Dixon, L. C. W.: Variable Metric Algorithms: Necessary and Sufficient Conditions for Identical Behavior and Sufficient Conditions for Identical Behavior on Nonquadratic Functions, Nottingham Optimization Centre, Nottingham, England, Report Nr. NOC-TR 26, 1971

[34] Maurer, H.: Numerical Solution of Singular Control Problems Using Multiple Shooting Techniques, JOTA, volume 18 (1976), pages 235-257

[35] Stoer, Josef and Bulirsch, Roland: Numerische Mathematik 2, Eine Einführung unter Berücksichtigung von Vorlesungen von F. L. Bauer, 3rd edition, Springer Verlag, Berlin Heidelberg 1990

[36] Stoer, Josef: Numerische Mathematik 1, Eine Einführung unter Berücksichtigung von Vorlesungen von F. L. Bauer, 6nd edition, Springer Verlag, Berlin Heidelberg 1993

[37] Oberle, H. J.; Grimm, W. and Berger, E.: BNDSCO. Rechenprogramm zur Lösung beschränkter optimaler Steuerungsprobleme Benutzeranleitung, Technische Universität München, Mathematisches Institut, TUM-M 8509, August 1985

[38] Dixon, L. C. W. and Bartholomew-Biggs, M. C.: Adjoint-Control Transformations for Solving Practical Optimal Control Problems, Optimal Control Applications & Methods, volume 2 (1981), pages 365-381

[39] Bartholomew-Biggs, M. C.: The optimization of spacecraft orbital manoeuvres, in: Numerical Optimisation of Dynamic Systems, edt. by L. C. W. Dixon and G. P. Szegö, North-Holland, 29-74, 1980

[40] Flaherty, Joseph E. and O'Malley, Robert E., Jr.: On the Computation of Singular Controls, IEEE Transactions on automatic control, volume AC-22 (1977), pages 640-648

[41] Luus, Rein: On the Application of Iterative Dynamic Programming to Singular Optimal Control Problems, IEEE Transactions on automatic control, volume 37 (1992), pages 1802-1806

[42] Aly, G. M. and Chan, W. C.: Application of a modified quasilinearization technique to totally singular optimal control problems, International Journal Control, volume 17 (1973), pages 809-815

[43] Baird, C. A.: Modified Quasilinearization Technique for the Solution of Boundary-Value Problems for Ordinary Differential Equations, JOTA, volume 3 (1969), pages 227 ff.

[44] Kumar, Virendra: A control averaging technique for solving a class of singular optimal control problems, International Journal of Control, volume 23 (1976), pages 361-380

[45] Bulirsch; Montrone and Pesch: Abort Landing in the Presence of Windshear as a Minimax Optimal Control Problem, Part 1, JOTA, volume 70 (1991), pages 1-23

[46] Bulirsch; Montrone and Pesch: Abort Landing in the Presence of Windshear as a Minimax Optimal Control Problem, Part 2, JOTA, volume 70 (1991), pages 221-252

[47] Siebenthal, C. D. and Aris, R.: Studies in optimization - VI: The application of Pontrjagin's methods to the control of a stirred reactor, Chemical Engineering Science, volume 19 (1964), pages 729-746

[48] Goh, B. S.: Optimal Control of a Fish Resource, Malayan Scientist, volume 5 (1969/70), pages 65-71

[49] Antonios, M. N.: Optimal Strategies for Water Quality Control, Numerical functional analysis and optimization, volume 9 (1987/88), pages 1271-1307

[50] Clark, C. W.: Mathematical Bioeconomics: The optimal management of renewable resources, Wiley, New York 1976

[51] Conway, G. R.: Mathematical models in applied ecology, Nature (London), volume 269 (1977), pages 291-297

[52] King, David and Roughgarden, Jonathan: Graded Allocation between Vegetative and Reproductive Growth for Annual Plants in Growing Seasons of Random Length, Theoretical Population Biology, volume 22 (1982), pages 1-16

[53] Melikyan, A. A.: Necessary Optimality Conditions for a Singular Surface in the Form of Synthesis, JOTA, volume 82 (1994), pages 203-217

[54] Jacobson, D. H.: On Singular Arcs and Surfaces in a Class of Quadratic Minimization Problems, Journal of Mathematical Analysis and Applications, volume 37, pages 185-201

[55] Grasman, J.: Non-Uniqueness in Singular Optimal Control, in: Mathematical theory of networks and systems, int. Symp., Delft 1979, volume 3, pages 415-420

[56] Grasman, J.: On a class of nearly singular optimal control problems, Math. Cent., Amsterdam, Afd. Toegespaste Wiskd. TW196/80, 1980

[57] Orlov, Yu V.: Special Solutions in Singular Optimization Problems, in: Lecture Notes in Control and Information Sciences 5., Berlin-Heidelberg-New York, Springer-Verlag 1978

[58] Lamnabhi-Lagarrigue, Françoise: Singular optimal control problems: On the order of a singular arc, Systems & Control Letters, volume 9 (1987), pages 173-182

[59] Schumacher, J. M.: The role of the dissipation matrix in singular optimal control, Systems & Control Letters, volume 2 (1983), pages 262-266

[60] Willems, J. C.; Kitapçi, A. and Silverman, L. M.: Singular Optimal Control: A Geometric Approach, SIAM Journal Control and Optimization, volume 24 (1986), pages 323-337

[61] Krener, Arthur J.: The high-order maximal principle, in: Geometric Methods in System Theory, Proceedings of the NATO Advanced Study Institute held at London, England, August 27 - Sept 7, 1973, edt. by D. Q. Mayne and R. W. Brockett