OPTIMAL CONTROL OF DIFFERENTIAL SYSTEMS WITH DISCONTINUOUS RIGHT-HAND SIDE

PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN DOCTOR IN DE TECHNISCHE WETENSCHAPPEN AAN DE TECHNISCHE HOGESCHOOL TE EINDHOVEN OP GEZAG VAN DE RECTOR MAGNIFICUS DR. IR. A.A. TH. M. VAN TRIER, HOOGLEERAAR IN DE AFDELING DER ELECTROTECHNIEK, VOOR EEN COMMISSIE UIT DE SENAAT TE VERDEDIGEN OP MAANDAG 29 JUNI 1970 TE 16.00 UUR.

DOOR

MATHEUS LODEWJK JOHANNES HAUTUS

GEBOREN TE HELDEN
DIT PROEFSCHRIFT IS GOEDGEKEURD DOOR
DE PROMOTOR PROF. DR. N.G. DE BRUIJN
Aan Hilde

Aan mijn ouders
# CONTENTS

List of symbols and subject index ................................................................. I-IV

CHAPTER I  INTRODUCTION

I.1. General remarks on optimal control theory .............................................. 1
I.2. Summary of the thesis .......................................................................... 3
I.3. Notations ................................................................................................. 4
I.5. A theorem for linear control systems ...................................................... 14

CHAPTER II  THE YO-YO EQUATION

II.1. Derivation of the equation ....................................................................... 16
II.2. Behaviour of the trajectories, summary of Chapter II ......................... 21
II.3. A basic optimal control problem for the yo-yo equation ................. 27
II.4. Time optimal p-start, N-step control .................................................... 35
II.5. Time optimal null-control ...................................................................... 41
II.6. Switching curves ..................................................................................... 45
II.7. Maximal average acceleration ................................................................. 53
II.8. Time optimal p-start, N-step control with fixed endpoint ................. 56
II.9. An auxiliary optimization problem ....................................................... 59
II.10. The case α = 1 ...................................................................................... 71

CHAPTER III  APPLICATION OF THE DISCRETE MAXIMUM PRINCIPLE

III.1. The discrete maximum principle ............................................................. 74
III.2. Application of the maximum principle to problems of Chapter II .... 79
III.3. A controlled system with Coulomb friction ......................................... 81
III.4. The case of more alternative points ...................................................... 86

References ..................................................................................................... 91

Samenvatting .................................................................................................. 93

Curriculum vitae ............................................................................................ 95
# List of Symbols

\begin{align*}
\mathbb{R}^n, \mathbb{R}^n_+ & \quad 4, 5 \\
S, \hat{S} & \quad 5 \\
S \rightarrow T, C(S \rightarrow T), C^P(S \rightarrow T), L(I \rightarrow S) & \quad 5 \\
x \mapsto f(x) & \quad 5 \\
\exists f(x), \exists \exists f(x, u) & \quad 6 \\
\Lambda & \quad 14, 75 \\
\kappa & \quad 14, 22, 75 \\
\alpha & \quad 19, 21 \\
P_k(\lambda_1, \ldots, \lambda_c) & \quad 24 \\
\beta & \quad 24, 30 \\
f(x) & \quad 24, 32 \\
P_k, \tau_k, \tau_k & \quad 25 \\
P_1, P'_1(\pi, \eta) & \quad 27 \\
P_2, P'_2(\pi, \eta) & \quad 33 \\
P_3, P'_3(\pi) & \quad 41, 42 \\
P_4, P'_4(\pi, \eta) & \quad 53 \\
P_5, P'_5(\pi, \eta, \lambda) & \quad 56, 57 \\
P_6, P'_6(\pi, \sigma, \eta, \lambda) & \quad 59 \\
\pi(\alpha, \eta) & \quad 28 \\
\rho & \quad 31 \\
\pi(\pi, \eta) & \quad 31 \\
\Lambda, \Lambda' & \quad 34 \\
P_k(\pi, \lambda), \tau_k(\pi, \lambda) & \quad 34, 42 \\
\emptyset, (0\text{-step strategy}) & \quad 34 \\
I_N(p) & \quad 35 \\
S_k & \quad 36 \\
\nu(s), h(s) & \quad 37
\end{align*}
\[ \begin{align*}
\delta^* & \quad 39 \\
J^* & \quad 42 \\
T, T(p, x) & \quad 42 \\
T(p) & \quad 43 \\
\tau^* & \quad 43 \\
Y_N(p, x) & \quad 53 \\
\n_X & \quad 54 \\
J_N & \quad 57 \\
I_N(p, q) & \quad 57 \\
S_N(q) & \quad 58 \\
\theta_N(p, q, x), \Theta_N(p, q, x), \Theta_N(q) & \quad 59 \\
\omega_k(p) & \quad 60 \\
v_N(x), \pi_N & \quad 61 \\
\hat{\eta}, \hat{\eta} & \quad 63 \\
\hat{\eta}, \hat{\eta} & \quad 74, 75 \\
H(\mathbb{R}^\infty \times \mathbb{R} \times \hat{\mathbb{S}}) & \quad 75 \\
\hat{q}(a, b, x), \hat{q}(a, b) & \quad 82 \\
\alpha_k, \tau_k & \quad 82 \\
S(x, y), f(x, z), g(x, z) & \quad 87
\end{align*} \]
III

SUBJECT INDEX

admissible control 1, 14, 75
admissible function 51
adjoint equation 15
alteration modulus 25, 82
alteration point 25, 82
alteration time 25, 82
a-start, N-step 82
attainable set 61
continuous maximum principle 74
control (function) 1, 14, 75
discrete maximum principle 74
dynamic programming 3, 25
feedback control 47
final time 82
locally integrable 5
maximum principle 2, 15
N-step strategy 34
one-sided solution 7
optimal control 1, 14, 75, 76
optimal control theory 2
performance index 1
piecewise optimal 3, 25, 34, 86
p-start, N-step 24
p-start, m-step 41
response 75
smoothly discontinuous 8
solution (of a differential equation) 6
state variable 1
strategy 34
switching curve 46
switching point 31, 45
time-optimal control 14
<table>
<thead>
<tr>
<th>Term</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>trajectory</td>
<td>1, 14</td>
</tr>
<tr>
<td>unisolution</td>
<td>7</td>
</tr>
<tr>
<td>yo-yo equation</td>
<td>19</td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION

1.1. General remarks on optimal control theory

In the theory of optimal control of differential systems one deals with systems of differential equations of the form

\[ \dot{x}(t) = f(x(t), u(t), t), \quad (t_1 < t < t_2), \]

\[ x(t_1) = a, \]

where \( a \) is a given \( n \)-vector, \( x \) is an \( n \)-vector valued function on \([t_1, t_2]\), and \( u \) is an \( m \)-vector valued function on the same interval. The function \( u \) is called the control function and \( x \) is called the state variable. A solution of (1) is called a trajectory corresponding to the control \( u \) and is denoted by \( x_u \).

Usually, only control functions are allowed which satisfy some conditions, like \( u(t) \in U \) for \( t \in [t_1, t_2] \), where \( U \) is some subset of \( \mathbb{R}^m \). Also, there are often restrictions on the state variables, such as \( x(t_2) \in S \) or \( x(t) \in X(t) \) \( (t_1 \leq t \leq t_2) \) where \( S \) is a subset of \( \mathbb{R}^n \) and \( X \) is a mapping of \([t_1, t_2]\) into the class of subsets of \( \mathbb{R}^n \). A control \( u \) is called admissible if it satisfies the conditions mentioned above, and if \( x_u \) satisfies the conditions posed upon the state variable.

Furthermore, a performance index \( J \) is given, which attaches a real number to each admissible control \( u \). For example,

\[ J(u) = \int_{t_1}^{t_2} h(x_u(t), u(t), t) \, dt \quad \text{or} \quad J(u) = F(x_u(t_2)) \]

where \( h, F \) are real valued functions.

Then an admissible control \( \bar{u} \) is called optimal if for all admissible controls \( u \) we have \( J(\bar{u}) \leq J(u) \), that is, if \( J(\bar{u}) \) is
minimal for \( \mathbf{u} = \mathbf{u}^* \). The task of optimal control theory is to determine optimal controls.

Instead of (1) one can have a discrete system of the form
\[
\mathbf{x}(k+1) = f(\mathbf{x}(k), \mathbf{u}(k), k) \quad (k = 0, 1, 2, \ldots),
\]
\[\mathbf{x}(0) = \mathbf{a}. \tag{2}\]

Here \( \mathbf{x} \) is a sequence of \( n \)-vectors and \( \mathbf{u} \) is a sequence of \( n \)-vectors.

For (2) one can define an optimal control problem analogous to the one of (1). Discrete systems can occur as approximations for continuous systems and also they appear directly in applications. In this thesis we will consider another type of problem which gives rise to discrete systems (see section 2).

REMARK. Instead of (1) or (2) one can have more general systems, governed for example by difference-differential equations, integral equations or partial differential equations, but we will not discuss them in this thesis.

The method for finding an optimal control usually consists of two parts: First, one shows that there exists an optimal control (mostly by a non-constructive method). Then one derives necessary conditions for a control to be optimal, that is, properties that are satisfied by an optimal control. Sometimes there exists just one admissible control which satisfies the necessary conditions and then this control must be optimal. Sometimes there is a finite number of admissible controls satisfying the necessary conditions. In that case, an optimal control can be found by comparison. Anyhow, by means of the necessary conditions, one restricts the set of possibly optimal controls.

For a general class of differential systems of the form (1) necessary conditions for optimality are given by the maximum principle of Pontryagin. (See [1], Chapter IV, V; [2], Chapter I.) (A special case of this theorem will be given in Chapter I, section 4.) For discrete systems there exists a necessary condition which is analogous to the maximum principle and which was first given in its
general form by Halkin [3]. (We will state this theorem in Chapter III, section 1.) For discrete systems there exists a second method for finding optimal controls, namely, the dynamic programming method (see [4]). Except for simple situations, the result of this method is a complicated functional-recurrence equation which cannot be solved explicitly. On the other hand, the discrete maximum principle results in a discrete boundary value problem.

1.2. Summary of the thesis

In this thesis we will study systems of the form (1.1) where the right-hand side is discontinuous for some values of \((x,t)\). For these systems the maximum principle ceases to hold. Problems of this kind are also considered in [5] where the maximum principle is shown to be valid in a modified form. We will follow a different method, however. Loosely speaking, in order to solve optimal control problems for these systems, we will divide the trajectory into pieces on which \(f\) is continuously differentiable. For these pieces the maximum principle holds. A trajectory which is optimal on each of these pieces will be called piecewise optimal. In order to find a control which is optimal for the whole trajectory, we have to fit the pieces together. This results in a discrete optimal control problem which can be solved by one of the methods discussed in section 1. Instead of creating a general theory, we will restrict ourselves to a few simple examples which are carried out in detail.

In Chapter II we will discuss a simple equation, and we will solve a number of optimal control problems for this equation. We will use the dynamic programming method for the "discrete parts" of these problems. A more complete summary of Chapter II can be found in section 2 of Chapter II.

In Chapter III we will use the discrete maximum principle for the discrete part of some of the problems of Chapter II. Also, a different system of the form (1.1) will be solved by the discrete maximum principle.

We give a summary of the remaining part of this chapter. In
section 3 we will develop some notations to be used throughout the thesis. In section 4 we will discuss some existence theorems for differential equations. When we require the controls to be continuous or even piecewise continuous, we cannot give existence theorems for optimal controls; hence we have to admit measurable controls. Therefore, we need a more general definition of the solution of a differential equation than the classical one. Such a definition (section 4, definition 1) was first given by Caratheodory [6] (p. 665-688). Existence and uniqueness theorems for the solutions of these generalized differential equations can be given similar to the ones for classical differential equations (for example Theorem 1 of section 4). For our purposes however, even these general existence and uniqueness theorems are not sufficient, since they require the right-hand side to be a continuous function of the state variable $x$. If the right-hand side does not depend continuously on $x$, a solution need not exist, and if we have existence, we do not necessarily have uniqueness. We will give conditions on the right-hand side which assure existence and uniqueness of the solution (section 4, Theorem 2).

In section 5 we will give a theorem which assures the existence of an optimal control for a control system of the form (1.1) where $f$ depends linearly on $x$. Furthermore, we give a maximum principle for such a system. These results will be used in Chapter II, section 3 and in Chapter III, section 3.

1.3. Notations

A. Formula indication

We will number the formulas and theorems in each section independently. When referring for example to formula 1 of section I of Chapter I we will write (1) in section I of Chapter I, (1.1) in Chapter I and not in section I, and (1.1.1) in Chapters II, III.

B. Vectors

The $n$-dimensional real euclidean space is denoted by $\mathbb{R}^n$. We will
distinguish between row and column vectors. Elements of $\mathbb{R}^n$ are considered column vectors. The set of row vectors will be denoted by $\mathbb{R}^n_\times$. The norm of $x$, denoted by $|x|$, is defined by

$$|x| := \left(\sum_{i=1}^{n} x_i^2\right)^{1/2},$$

where $x_1, \ldots, x_n$ are the entries of $x$. In the two-dimensional case we define $x = (x_1, x_2)$.

If $S \subset \mathbb{R}^n$, then $\partial S$ and $\overline{S}$ are the closure and the boundary of $S$. The transpose of a vector or matrix will be denoted by an accent. For example, if $x$ is a column vector then $x = (x_1, \ldots, x_n)^T$. In the text however we will often omit the accent when there is no danger of confusion. The set of real numbers is denoted by $\mathbb{R}^1$ and the set of positive real numbers by $\mathbb{R}^+_1$.

### C. Notations

If $S$ and $T$ are sets, the set of mappings from $S$ into $T$ is denoted by $S \rightarrow T$. If $S \subset \mathbb{R}^n$ and $T \subset \mathbb{R}^m$ then $C(S \rightarrow T)$ is the set of continuous mappings from $S$ into $T$. If in addition $S$ is open then $C^p(S \rightarrow T)$ denotes the set of $p$ times continuously differentiable functions for $p = 1, 2, \ldots$. If $I$ is a set of real numbers and $S \subset \mathbb{R}^n$, $L(I \rightarrow S)$ denotes the set of all functions in $I \rightarrow S$, which are integrable on every compact subset of $I$. We will call these functions locally integrable. In particular we will encounter $L([0,\infty) \rightarrow S)$ and $L([\xi_1,\xi_2) \rightarrow S)$.

The expression "almost everywhere" (abbreviated a.e.) always refers to a one-dimensional independent variable, usually denoted by $t$ or $s$.

The function which attaches the value $f(x)$ to the variable $x$ will be denoted by $f$ or by $x \mapsto f(x)$. For example, if $g \in (\mathbb{R}^n \times \mathbb{R}^m)$ and $f \in (\mathbb{R}^n \times (\mathbb{R}^m \times \mathbb{R}^p))$, then $x \mapsto f(x, g(x))$ is a function in $\mathbb{R}^n \times \mathbb{R}^p$.

**Remark.** If $T$ is a discrete set (for example the natural numbers) and $S$ is an open set in $\mathbb{R}^n$, then we say $\xi \in C^p(S \times T - U)$ if $(x \mapsto f(x, \xi)) \in C^p(S + U)$ for all $t \in T$. 
D. Partial derivatives

If \( f \in C^1(S \times T) \) where \( S \) is an open set in \( \mathbb{R}^n \) and \( T \subset \mathbb{R}^m \), then for a given \( x \in S \) the functional matrix of \( f \) at the point \( x \) will be denoted by \( \partial f(x) \), hence

\[
(\partial f(x))_{ij} := \frac{\partial f_i}{\partial x_j} \quad (i = 1, \ldots, m; j = 1, \ldots, n).
\]

In particular, if \( m = 1 \), then \( \partial f(x) \) is the gradient of \( f \) at \( x \).

(Note that this is a row vector.)

If \( f \in C^1(S \times U \to T) \) with \( S \subset \mathbb{R}^n \), \( U \subset \mathbb{R}^s \) open, \( T \subset \mathbb{R}^m \), and if generic elements of \( S \times U \) are denoted by \( (s,u) \), then \( \partial_s f(s,u) \) is the partial functional matrix with respect to \( s \), at the point \( (s,u) \); that is,

\[
(\partial_s f(s,u))_{ij} := \frac{\partial f_i}{\partial s_j} \quad (i = 1, \ldots, m; j = 1, \ldots, n).
\]

and \( \partial_u f(s,u) \) is defined similarly.

I.4. Existence and uniqueness of solutions of differential equations

In this section \( S \) is an open set in \( \mathbb{R}^n \) and \( T_c, T_1, T_2 \) are real numbers, satisfying \( T_1 < T_0 < T_2 \).

DEFINITION 1. If \( a \in S \), \( f \in (S \times (T_1, T_2) \to \mathbb{R}^n) \), and if \( t_1 \) and \( t_2 \) are real numbers, then \( x \in ((t_1, t_2) \to S) \) is called a solution of the differential equation (with initial condition)

\[ x = f(x, t), \quad x(T_0) = a \tag{1} \]

on \((t_1, t_2)\) if \( T_1 \leq t_1 < T_0 < t_2 < T_2 \),

\[ (t \mapsto f(x(t), t)) \in L((t_1, t_2) \to \mathbb{R}^n) \]

and

\[
x(t) = a + \int_{T_0}^{t} f(x(s), s)ds \quad (t_1 < t < t_2).
\]

REMARK 1. If \( f \) is continuous, it is easily seen that every solution of (1) is a solution in the classical sense. In general, a solution
of \( (1) \) is an absolutely continuous function which satisfies \( (1) \) almost everywhere. Note that the phrase "solution of the differential equation \( (1) \) on \( (t_1, t_2) \)" implies conditions on the numbers \( t_1 \) and \( t_2 \). Accordingly, these conditions will not be repeated in the sequel.

**Remark 2.** Sometimes solutions, defined on \([T_0, t_2)\) or \((t_1, T_0]\) (with \( T_1 \leq t_1 < T_0 \) and \( T_0 < t_2 < T_2 \)) are considered. The definitions of these solutions are similar to definition 1 and they are called one-sided solutions.

**Definition 2.** A solution of \( (1) \) on \((t_1, t_2)\) is called a unisolution of \( (1) \) on \((t_3, t_4)\) if for every \( t_3 \) and \( t_4 \) and every solution \( x^* \) of \( (1) \) on \((t_3, t_4)\) we have \( x^*(t) = x(t) \) \((t \in (t_1, t_2) \cap (t_3, t_4))\).

**Remark 3.** If all solutions of \( (1) \) are unisolations and if we identify functions with their graphs, then the solutions of \( (1) \) are totally ordered by inclusion. It is easily seen in that case that the union of all solutions of \( (1) \) is the maximal solution of \( (1) \) (with respect to inclusion). (Furthermore, this solution is also a unisolution.)

**Remark 4.** If \( x \) is a unisolution and \( x \) is a one-sided solution of \( (1) \), then \( x \) and \( \bar{x} \) coincide on their common domain. In order to see this, assume that \( \bar{x} \) is defined on \((T_0, c_3)\) and \( x \) on \((t_1, t_2)\) and observe that the function \( x^* \), defined by \( x^*(t) = x(t) \) \((t_1 < t < t_0)\), \( x^*(t) = \bar{x}(t) \) \((T_0 < t < t_3)\), is a solution of \( (1) \).

**Definition 3.** If \( f \in \langle S \times (T_1, T_2) \rightarrow \mathbb{R}^N \rangle \), we say that \( f \in \text{Li}(S \times (T_1, T_2) \rightarrow \mathbb{R}^N) \) if

i) \( (t \mapsto f(x, t)) \in L((T_1, T_2) \rightarrow \mathbb{R}^N) \) for all \( x \in S \).

ii) For every compact set \( K \subseteq S \) and every compact subinterval \([t_1, t_2]\) of \((T_1, T_2)\) there exists \( M > 0 \) such that for almost all \( t \in [t_1, t_2] \) we have \( |f(x, t)| \leq M \) \((x \in K)\) and \( |f(x, t) - f(y, t)| \leq M|x - y| \) \((x, y \in K)\).

Now we have:

**Theorem 1.** If \( a \in S \) and \( f \in \text{Li}(S \times (T_1, T_2) \rightarrow \mathbb{R}^N) \), then there exist \( t_1, \ldots, t_n \)
\( t_2 \) such that (1) has a solution on \((t_1, t_2)\). Furthermore, every solution of (1) is a unisolution.

For a proof see [7], Chapter IX.

**Definition 4.** A function \( f \in (S \times (T_1, T_2) \to \mathbb{R}^n) \) is called smoothly discontinuous if there exists \( g \in C^1(S \to \mathbb{R}^1) \) such that for almost all \( t \in (T_1, T_2) \) we have

\[
  f(x, t) = f_j(x, t) \quad (x \in G_j) \quad (j = 0, 1, 2),
\]

where

\[
  f_j \in Li(G_j^+ \times (T_1, T_2) + \mathbb{R}^n) \quad (j = 0, 1, 2),
\]

\[
  G_0 := \{ x \in S \mid g(x) = 0 \},
\]

\[
  G_1 := \{ x \in S \mid g(x) > 0 \},
\]

\[
  G_2 := \{ x \in S \mid g(x) < 0 \},
\]

and where \( G_j^+ \) is some neighborhood of \( G_j \cup G_0 \) \((j = 0, 1, 2)\).

For a smoothly discontinuous function \( f \), a solution of (1) does not necessarily exist and if a solution does exist, then it is not necessarily a unisolution. This is illustrated by the following examples:

**Example 1.** Consider the differential equation in \( \mathbb{R}^1 \):

\[
  \dot{x} = 1 - 2 \text{ sgn } x, \ x(0) = 0 \quad (-1 < t < 1).
\]

Suppose that \( x \) is a solution of (2) on \((t_1, t_2)\), where \(-1 \leq t_1 < 0 < t_2 \leq 1\). Then \( x \) cannot be identically zero on \([0, t_2)\), since otherwise we would have \( \dot{x} = 1 \) a.e. on \([0, t_2)\), which is contradictory. Let us assume that \( x(t_0) > 0 \) for some \( t_0 \in (0, t_2) \). Let

\[
  t_3 := \max(t \leq t_0 \mid x(t) = 0).
\]

Then we have

\[
  x(t_0) = x(t_3) + \int_{t_3}^{t_0} (1 - 2 \text{ sgn } x(s))ds = -(t_0 - t_3) < 0
\]
1.4

which is also a contradiction. In a similar way \( x(t_0) < 0 \) can be proved to be impossible. Therefore (2) does not have a solution.

EXAMPLE 2. Consider the differential equation in \( \mathbb{R}^1 \):

\[ \dot{x} = \text{sgn} x \quad (-1 < t < 1) , \quad x(0) = 0 . \]  

(3)

Then the functions \( \phi(t - a) \) (restricted to \((-1,1)\)), where \( \phi(t) = \max(0, t) \quad (t \in \mathbb{R}^1) \) and \( a \in \mathbb{R}^1 \) are solutions of (3) according to definition 1.

The following theorem gives a sufficient condition for the existence of solutions and unisolusions of (1) if \( f \) is smoothly discontinuous:

THEOREM 2. If \( f \in \mathcal{G}(S \times (T_1, T_2) \to \mathbb{R}^n) \) is smoothly discontinuous, with \( g, f_j, G_j, G_j^s \) as defined in definition 4 and if \( h_j \in \mathcal{G}(G_0 \times (T_1, T_2) \to \mathbb{R}^1) \) is defined by

\[ h_j(c,t) := (3g(c))f_j(c,t) \quad (j = 0,1,2) , \]  

(4)

then we have the following results:

\[ \text{i) If for every } c \in G_0 \text{ there exist } \delta > 0 \text{ such that at least one of the statements:} \]

\[ "h_1(c,t) \geq \delta \land h_2(c,t) \geq \delta \quad (\text{a.e.)} " , \]  

(5)

\[ "h_1(c,t) \leq -\delta \land h_2(c,t) \leq -\delta \quad (\text{a.e.)} " \]  

(6)

holds, then for every \( a \in S \) there exist \( t_1, t_2 \) such that (1) has a solution on \((t_1, t_2)\).

\[ \text{ii) If the condition of i) is satisfied and if furthermore} \]

\[ |h_0(c,t)| \geq \delta \quad (\text{a.e.)}, \text{ then every solution of (1) is a unisolusion.} \]

\[ \text{iii) If, in addition to the conditions of ii), there exist positive numbers } H_1 \text{ and } H_2 \text{ such that for almost all } t \in (t_1, t_2) \text{ we have} \]

\[ |f(x,t)| \leq H_1 |x| + H_2, \text{ then there exist } t_1, t_2 \text{ such that (1) has a unisolusion } x \text{ on } (t_1, t_2), \text{ with the following property:} \]
If \( t_2 < T_2 \), we have \( x(t) \rightarrow y(t \rightarrow t_2) \) for some \( y \in \mathcal{S} \). If \( t_1 > T_1 \), we have \( x(t) \rightarrow y(t \rightarrow t_1) \) for some \( y \in \mathcal{S} \).

We need to prove a result that is analogous to (7).

**Lemma.** If \( f \in \mathcal{L}(S \times (T_1, T_2) \rightarrow \mathbb{R}^n) \), \( g \in C^1(S \times \mathbb{R}^n) \), \( a \in S \) and \( g(a) = 0 \), and for some \( \delta > 0 \) we have \((\partial g(a)) f(x, t) \geq \delta \) (a.e.), then there exist \( t_1, t_2 \) such that (7) has a unique solution on \((t_1, t_2)\), satisfying

\[ \text{sgn}(g(x(t))) = \text{sgn}(t - T_0) \quad (t_1 < t < t_2). \]  

**Proof.** According to Theorem 1, there exists a unique solution of (7) on some interval \((t_1', t_2')\). Therefore we only have to prove that (7) holds on some subinterval \((t_1, t_2)\) of \((t_1', t_2')\) containing \( T_0 \). Let \( \epsilon_0 \) be a positive number such that \( y \in S \) holds whenever \( |y - a| \leq \epsilon_0 \).

Furthermore, assume that \( T_1 < t_0', t_2' < T \). Then we may define \( M > 0 \) such that for almost all \( t \in (t_1', t_2') \) we have \( |f(y, t)| \leq M \), \( |f(y, t) - f(a, t)| \leq M |y - a| \) for \( |y - a| \leq \epsilon_0 \).

Let \( \epsilon := \min(\epsilon_0, \delta M^{-1}(1 + |\partial g(a)|)^{-1}) \). Then, since \( x \) and \( \partial g \) are continuous there exist \( t_1, t_2 \) with \( t_1 \leq t_0 < t_2 \leq t_2' \) such that \( |x(t) - a| < \epsilon \) and \( |\partial g(x(t)) - \partial g(a)| < \epsilon \) for \( t \in (t_1, t_2) \). Since \( g \) is continuously differentiable (and hence Lipschitz continuous) and \( x \) is absolutely continuous, it follows from ([7], p. 51) that \( t \mapsto g(x(t)) \) is absolutely continuous. Furthermore, for almost all \( t \in (t_1, t_2) \) we have

\[ \frac{d}{dt} g(x(t)) = (\partial g(x(t))) f(x(t), t). \]  

Hence,

\[ g(x(t)) = \int_{T_0}^{t} (\partial g(x(s))) f(x(s), s) ds \quad (t \in (t_1, t_2)). \]  

Now it follows from

\[ (\partial g(x(s))) f(x(s), s) = (\partial g(a)) f(a, s) - \\
- (\partial g(a) - \partial g(x(s))) f(x(s), s) - (\partial g(a))(f(a, s) - f(x(s), s)) \]

that

\[ (\partial g(x(s))) f(x(s), s) \geq \delta - cM - |\partial g(a)| M \geq \delta/2 \quad (a.e.). \]
This yields
\[ g(x(t)) \geq \frac{1}{2}(t - T_0) \quad (t \in (T_0, t_2)) \]
and
\[ g(x(t)) \leq \frac{1}{2}(t - T_0) \quad (t \in (t_1, T_0)) . \]

**Remark 5.** If it is given that the inequality
\[ (\exists \delta > 0) \forall (a, t) \in [T_0, t_2] \]
holds a.e. on \([T_0, t_2]\) instead of \((t_1, t_2)\), we can derive (7) with \((T_0 \leq t < t_2)\) instead of \((t_1 < t < t_2)\) and a similar remark applies if (7a) holds a.e. on \((t_1, T_0)\).

**Proof of i.** If \(a \in G_i\) (\(i = 1\) or \(i = 2\)), the conditions of theorem 1 are satisfied with \(G_1\) instead of \(S\). Therefore we can assume that \(a \in G_0\). Since in that case we have \(a \in G_0^a \cap G_1^a \cap G_2^a\) there exist unisolusions \(x_i\) of the equations
\[ x = \xi_i(x, t), \quad x(T_0) = a \]
on some interval \((t_1, t_2)\) for \(i = 1, 2, 3\). Assume that (3) holds for \(c = a\). Then it follows from the lemma that \(g(x_1(t)) > 0\) \((T_0 < t < t_1)\) and \(g(x_2(t)) < 0\) \((t_1 < t < T_0)\). Therefore the function \(x\), defined by \(x(t) = x_1(t)\) \((T_0 < t < t_2)\), \(x(t) = x_2(t)\) \((t_1 < t < T_0)\) is a solution of (1) on \((t_1, t_2)\). If (6) holds, the argument is analogous. (Here the lemma is applied with \(-g\) in place of \(g\).) This completes the proof of i.

**Proof of ii.** First we show that for every \(a \in S\) there exists a unisolusion of (1) on some interval \((t_1, t_2)\). If \(a \notin G_0\), the existence of a unisolusion follows from theorem 1. Therefore we assume that \(a \in G_0\). Let \(\bar{x}\) be the solution of (1) constructed in the proof of i) and let \((t_1, t_2)\) be the domain of \(\bar{x}\). We may assume that \(t_1 < t_1' < T_0 < t_2' < T_2\) holds. Let us assume that (5) is satisfied for \(c = a\). Then there exists a neighborhood \(O\) of \(a\) such that for almost all \(t \in (t_1', t_2')\) we have \(h_1(c, t) \geq \delta \land h_2(c, t) \geq \delta\) \((c \in O \cap G_0)\). In fact, the function
\[ c = \int_{t_1}^{t_2} h_1(c,t) \, dt \]

is continuous and therefore has constant sign in some neighborhood of \( a \). We may assume that \( \delta \) is bounded and contained in \( S \). Then by definition 2, \( \delta \) is bounded on \( 0 \times (t_1', t_2') \) and hence there exist \( \tau_1, \tau_2 \) with \( t_1' \leq \tau_1 < \tau_2 \leq t_2' \) such that \( \delta(t) = 0 \) (\( t \in (\tau_1, \tau_2) \cap (\tau_3, \tau_4) \)) for all \( t_3, \tau_4 \) and all solutions \( x \) of (1) on \((\tau_3, \tau_4)\). We show that the restriction of \( \bar{x} \) to \((\tau_1, \tau_2)\) is a unique solution. Let \( x \) be an arbitrary solution of (1) on \((\tau_3, \tau_4)\). We may assume that \( \tau_1 = \tau_2 \) and \( t_4 < \tau_2 \) hold, since otherwise we can restrict \( x \) to \((\tau_1, \tau_2) \cap (\tau_3, \tau_4)\), and we prove that \( x(t) = \bar{x}(t) \) for \( \tau_3 < t < \tau_4 \). It is sufficient to show that the conditions

\[ g(x(t)) > 0 \quad (\tau_3 < t < \tau_4), \quad (8) \]

\[ g(x(t)) < 0 \quad (\tau_3 < t < \tau_0) \quad (9) \]

hold. In fact, from (8) and (9) it follows that \( x(t) \) satisfies (1)\( _1 \) a.e. on \([\tau_0, \tau_4]\) and (1)\( _2 \) a.e. on \([\tau_3, \tau_0]\). If we apply Remark 4 to (1)\( _1 \) and (1)\( _2 \) the result follows immediately.

In order to prove (8), suppose that \( x(t) \notin G_2 \) for some \( t \in (\tau_0, \tau_4) \). Let \( \tau_0 := \max \{ t \in \bar{E} \mid x(t) \notin G_2 \} \) and let \( c := x(\tau_0) \). Then \( \bar{x} \) satisfies (1)\( _2 \) (a.e.) on \([\tau_0, \bar{E}]\) (with \( c \) replaced by \( c \)). According to Remark 5 we therefore have \( g(\bar{x}(t)) > 0 \) \((\tau_0 < t < \tau')\) for some \( \tau' \in (\tau_0, \bar{E}) \), and this is a contradiction. It follows that we have \( g(x(t)) > 0 \) \((\tau_3 < t < \tau_4)\). Furthermore, we cannot have \( g(x(t)) = 0 \) on an interval. Otherwise we would have

\[ \left| \frac{d}{dt} g(x(t)) \right| = \left| h_0(x(t),t) \right| \geq \delta \]

a.e. on that interval, and this is impossible. Therefore, if for some \( \bar{E} \in (\tau_3, \tau_4) \) we have \( g(x(\bar{E})) = 0 \), then there exist \( \tau', \tau'' \) with \( \tau_0 < \tau' < \tau'' \leq \bar{E} \) such that \( g(x(t')) = 0 \) and \( g(x(t'')) = 0 \) \((\tau' < t < \tau'')\). Then \( x \) satisfies (1)\( _1 \) a.e. on \((\tau', \tau'')\) and by Remark 5
(with $T_0$ replaced by $t''$) we have $g(x(t)) < 0$ on some interval $(t_0, t'')$. This is a contradiction and hence (8) is proved. The statement (9) can be proved similarly, and so the existence of a unisolusion is established.

We are going to prove that every solution of (1) is a unisolusion. Let $x$ be an arbitrary solution of (1) on $(t_1, t_2)$ and suppose that $x$ is not a unisolusion. Then there exists a solution $\bar{x}$ of (1) on some interval $(\bar{t}_3, \bar{t}_4)$ such that we have $x(\bar{t}_4) = \bar{x}(\bar{t}_4)$ for some $\bar{t}_4 \in (t_3, t_4)$. Since

$$
\emptyset := \{ t \in (t_5, t_6) \mid x(t) = \bar{x}(t) \}
$$

is not empty ($T_0 \notin \emptyset$) and $\emptyset \neq (t_5, t_6)$, there is a boundary point $t_0$ of $\emptyset$ in $(t_5, t_6)$. The foregoing result (with $T_0$ replaced by $t_0$) yields that there exists a unisolusion $x^*$ of (1) on some interval $(t_1, t_2)$ with $T_1 \leq t_1 < t_0 < t_2 \leq T_2$. But then we have $x^*(t) = x(t)$ and $x^*(t) = \bar{x}(t)$ on $(t_1', t_2') \cap (t_5, t_6)$ contradicting the definition of $t_0$. This completes the proof of (ii).

Proof of (iii). Let us first assume that $f$ is bounded on $S \times (T_1, T_2)$ ($|f(x, t)| \leq M_0$, say). Since every solution of (1) is a unisolusion, it follows by Remark 3 that there exists a unisolusion $x$ on a maximal domain $(t_1, t_2)$. We show that $x(t) = \gamma(t + t_2)$ for some $\gamma \in \mathbb{S}$. Indeed, we have for $t > s$:

$$
|\gamma(t) - \gamma(s)| = \left| \int_s^t f(x(t), t) \, dt \right| \leq M_0(t - s).
$$

Hence, if $\{s_n\}$ is a sequence with $s_n + t_2 \to \infty$, then $(\gamma(s_n))$ is a Cauchy sequence which has a limit $y \in \mathbb{S}$. It is easily seen (for example, by mixing two sequences) that $y$ does not depend on the sequence $(s_n)$ and also that $x(t) = \gamma(t + t_2)$. If $y \in \mathbb{S}$ and $t_2 < T_2$, we can apply (ii) with $(s, T_0)$ replaced by $(y, t_2)$. It follows that there exists a unisolusion $\bar{x}$ on some interval $(\bar{t}_1, \bar{t}_2)$ containing $t_2$. Then the function $x^*$ defined by $x^*(t) = x(t)$ ($t \in (t_1, t_2)$),
$x^*(t) = \overline{x}(t)$ (t ∈ [t_2, t_2]) is a unisolvent whose domain is larger than the domain of x and this is a contradiction. In a similar way, we can prove that if $t_2 > T_2$, then we have $x(t) = y(t + t_1)$ for some $y \in \mathcal{S}$.

If we have $|f(x, t)| \leq M_1|\dot{x}| + M_2$ (x ∈ S, $T_1 < t < T_2$), then a solution of (1) on some interval $(t_1, t_2)$ satisfies

$$|x(t)| \leq |a| + \int_{t_0}^{t} (M_1|x(s)| + M_2)ds \leq a + M_2(T_2 - T_0)$$

Together with Gronwall's lemma (see [15] p. 19) this implies

$$|x(t)| \leq (a + M_2(T_2 - T_0)) e^{M_1(T_2 - T_0)} \leq R_2 \quad (T_0 < t < t_2)$$

where $R_2 := (a + M_2(T_2 - T_0)) e^{M_1(T_2 - T_0)}$. Similarly we have $|x(t)| \leq R_2$ ($t_1 < t < T_0$) for every solution x, where $R_1 := (a + M_2(T_1 - T_0)) e^{M_1(T_2 - T_0)}$.

Therefore we can apply the foregoing result, with S replaced by

$$S_1 := \{ y \in \mathcal{S} \mid |y| < 1 + \max(R_1, R_2) \}$$

since on this set x is bounded.

1.5. A theorem for linear control systems

In this section we consider the control system

$$\dot{x}(t) = A(t)x(t) + f(t,u(t)),$$

$$x(0) = a.$$ 

Here $A$ is a continuous $n \times n$ matrix-valued function on $[0, \omega)$ and $f \in C([0, \omega) \times U \rightarrow \mathbb{R}^n)$ where U is a compact set in $\mathbb{R}^m$. The set $\mathcal{U} := L([0, \omega) \rightarrow U)$ is the set of control functions. A trajectory is a solution of (1) (in the sense of Definition 4.1) corresponding to some control $u$, and is denoted by $x_u$. Hence, $x_u \in C([0, \omega) \rightarrow \mathbb{R}^n)$.

Now let $b \in \mathbb{R}^n$, then $u \in \mathcal{U}$ is called admission if there exists $t \geq 0$ such that $x_u(t) = b$. Furthermore, an admission control $u$ is optimal (or time-optimal) if there exists a $\bar{t} \geq 0$ such that
\( x_u(t) = b \), and such that for every \( u \in \mathcal{U} \) and \( t \in [0, \bar{t}) \) we have \( x_u(t,a) = b \). So, the quantity \( \bar{t} \) is the minimum time at which the state variable can reach the point \( b \). Now we have the following fundamental result:

**Theorem.** With the foregoing notation we have

(i) If there exists an admissible control, then there exists an optimal control.

(ii) If \( \tilde{u} \in \mathcal{U} \) is an optimal control with minimal time \( \bar{t} \), then there exists a non-trivial solution \( \psi \in ([0, \bar{t}) \times \mathbb{R}^m) \) of the adjoint equation:

\[
\dot{\psi}(t) = - \psi(t)A(t),
\]

such that

\[
\psi(t)u(t, \tilde{u}(t)) = \max_{v \in \mathcal{U}} \psi(t)f(t,v) \quad \text{(a.e.)}.
\]

The existence theorem (i) is in its general form due to L.W. Neustadt (see [5] for a proof). The necessary condition (ii) is the *maximum principle* of Pontryagin for this special case (see [1] Chapters IV, V; [2] Chapters I, II, III, where a proof of (ii) can be found).

When applying this theorem it is not necessary in general to establish a priori the existence of an admissible control. Rather one tries to find directly admissible controls which satisfy the *maximum principle*. If there exists no admissible control which satisfies the *maximum principle*, then there does not exist an optimal control, and hence (by (i)) no admissible control exists at all. If there do exist admissible controls satisfying the *maximum principle*, one of them has to be an optimal control.
CHAPTER II
THE YO-YO EQUATION

II.1. Derivation of the equation

A simple example of a system with a discontinuous right-hand side is provided by the differential equation satisfied by the idealized yo-yo.

\[ \dot{\varphi} = \frac{mg}{F(t)} \]

\[ \varphi \text{ is the angle between } PM \text{ and the downwards directed vertical line through the point } P. \]

Let \( \mathcal{M} \) be the center of a cross-section of the axle of the yo-yo, let \( R \) be the radius of the yo-yo and let \( P \) he the point where the string is fixed to the axle (see fig. 1). We denote by \( \dot{\varphi} \) the angle between \( PM \) and the downwards directed vertical line through the point \( P \) (that is, the line \( PB \)). Here we agree that \( \dot{\varphi} \) is positive if \( \mathcal{M} \) is to the right of \( PB \) and negative if \( \mathcal{M} \) is to the left of \( PB \). (If \( |\varphi| > \frac{\pi}{2} \), then a piece of the string is wound on the axle.)
An upward-directed vertical force $F(t)$ is acting on the string. The upper end $A$ of the string has the height $h(t)$.

In the following calculations we neglect the horizontal velocity of the yo-yo and we assume that the string is always vertical so that the horizontal movements do not influence the vertical movements.

First suppose that $|\phi| \leq \pi/2$. Then the height of $H$ is $h(t) - \lambda - R \cos \phi(t)$, where $\lambda$ is the length of the string. Applying Newton's law concerning the forces and accelerations in the vertical direction we obtain:

$$F(t) = mg = \frac{m \dot{h}(t)}{R} \sin \phi(t) + R \ddot{s}(t) \cos \phi(t).$$

From the law concerning the relation between moment and impulse we get:

$$RF(t) \sin \phi(t) = -J \dot{s}(t),$$

where $J$ is the moment of inertia of the yo-yo.

If $|\phi| \geq \pi/2$ holds, the corresponding formulas are

$$F(t) = mg = \frac{m \dot{h}(t)}{R} \sin \phi(t) + R \ddot{s} \text{ sgn} \phi,$$

$$RF(t) \text{ sgn} \phi(t) = -J \ddot{s}(t).$$

If $\phi$ is sufficiently large for $|\phi| \leq \pi/2$, the time intervals on which we have $|\phi| \leq \pi/2$ are very small. We will neglect these intervals so that we only have to deal with equations (1) and (2).

We make this more precise in the following way: Let the sets $A$, $B$, $C$ be defined by

$$A := \{t \geq 0 \mid |\phi(t)| \leq \pi/2\},$$

$$B := \{t \geq 0 \mid \phi(t) > \pi/2\},$$

$$C := \{t \geq 0 \mid \phi(t) < -\pi/2\}.$$
Now define a function $\theta$ as follows:

$$
\theta(t) := t - \mu \{ s \in A \mid s \leq t \}.
$$

Hence

$$
\theta(t) = t - \sum_{t_{i-1} < t} (t_i - s_i) \quad (t \in B \cup C)
$$

$$
\theta(t) = s(t) - \sum_{t_{i-1} < t} (t_i - s_i) \quad (t \in A)
$$

where $\mu$ is the Lebesgue-measure and $s(t) := \max\{ s_i \mid s_i < t \}$. We see that $\theta$ is strictly increasing on $B \cup C$ and constant on each interval of $A$. In particular, we have $\theta(s_{i-1}) = \theta(s_i) = \tau_i$. We introduce the new independent variable $\tau := \theta(t)$ on $B \cup C$, and define the functions $\phi$, $\Phi$, $h$, by

$$
\phi(t) = \phi(t) - \tau/2 \quad (t \in B)
$$

$$
= \phi(t) + \tau/2 \quad (t \in C),
$$

$$
\Phi(t) = \Phi(t) \quad (t \in B \cup C),
$$

$$
h(t) = h(t) \quad (t \in B \cup C).
$$

These functions are well-defined and satisfy the equations

$$
\Phi'(\tau) = mg - m \left[ \frac{d^2h}{dt^2} + \tau \frac{d^2\phi}{dt^2} \right] \text{sgn} \psi(\tau),
$$

$$
\Phi''(\tau) \text{sgn} \psi(\tau) = -J \frac{d^2\phi}{dt^2},
$$

for $\tau > 0$ except at the points $\tau_1, \tau_2, \ldots$. It is easily seen that the function $\psi$ is continuous for all $\tau > 0$, if we define $\psi(\tau_i) := 0$ for $i = 1, 2, \ldots$. The behaviour of $d\psi/d\tau$ at $\tau_i$ depends on $F(t)$ ($s_i \leq t \leq s_{i-1}$). In general, $d\psi/d\tau$ has a jump at $\tau_i$ equal to

$$
\delta(\tau_i) = \delta(\tau_i).
$$

It follows from (2) that
II.1

\[ J \hat{\psi}^2(\tau_i) - \hat{\psi}^2(\tau_i) = \pm R \int_{-\pi/2}^{\pi/2} F(\tau) \sin \phi \, d\phi. \]

Therefore, if \( F \) is constant on \([\tau_i, \tau_{i+1}]\), we have \( \hat{\psi}(\tau_i) = \hat{\psi}(\tau_{i+1}) \). We will assume henceforth that on the intervals where \( |\phi| \leq \pi/2 \) holds, \( F \) is constant. Then it follows that \( d\phi/d\tau \) is continuous at the junction points \( \tau_i \). Thus, we have the equation:

\[ \frac{d^2 \psi}{d\tau^2} = -v(\tau) \, \text{sgn} \, \psi(\tau) \tag{3} \]

where \( v(\tau) := J^{-1} R F^{+}(\tau) \).

We can consider \( v \) a control function here. It is obvious that \( v \geq 0 \) must hold. If we impose the further restriction \( v(\tau) \leq M \) for some positive \( M \) on the control function, we are led to optimal control problems of the type studied in this chapter. With this constraint we can normalize (3) by the substitutions:

\[ v = \frac{1}{M} (1 + u) ; \quad \psi = \frac{1}{M} \xi. \]

Writing \( t \) instead of \( \tau \) we then get what we call the yo-yo equation:

\[ \ddot{\xi} + (1 + u) \, \text{sgn} \, \xi = 0 \tag{4} \]

where \( u \) satisfies \( |u(t)| \leq 1 \ \ (t \geq 0) \). In this chapter we pose the somewhat more general restriction \( |u(t)| \leq \alpha \) where \( \alpha \) is some number in \((0,1]\).

Instead of \( F^* \) (or equally \( v \)) we can consider \( h^* \) as a control function. This does not alter the situation, however, because eliminating \( \hat{\psi} \) in (1) and (2) gives

\[ F^* = J \text{Im}(J + M \xi^2)^{-1} (g + h^*). \]

By substituting this relation into (3) we obtain an equation which also can be normalized to (4).

We give two other examples which give rise to equation (4).
Consider a particle which moves on a V-shaped configuration shown in fig. 3. Here the y-axis indicates the vertical direction. The configuration can be translated upwards and downwards. Equivalent to this system is the system at rest, with variable gravitation. Then the gravitation acting on the particle has a vertical direction and equals

\[ F = mg + m \ddot{h} \]

Let \((x,y)\) denote the coordinates of the particle. The component of the force in the direction of the line \( \ell \) towards the origin is \( F \sin \alpha \). The component orthogonal to \( \ell \) is compensated by the normal force on \( \ell \) provided \( F \neq 0 \). Now according to Newton's law, the equation of motion of the particle is given by:

\[ m \ddot{y} = F - mg - m \ddot{h} \]
\( \ddot{x} + F \sin \alpha \cos \alpha = 0 \quad (x > 0) \)

\( \ddot{m} - F \sin \alpha \cos \alpha = 0 \quad (x < 0) \).

Hence,

\( \ddot{x} + F \sin \alpha \cos \alpha \text{ sgn } x = 0 \). \hspace{1cm} (5)

If we pose the condition that \( \ddot{x} \leq M \) holds for some \( M > 0 \) equation (5) can easily be transformed to (4) with \( \lvert u \rvert \leq 1 \) and

\( u = c_1 + c_2 \ddot{x} \) for some constants \( c_1, c_2 \).

**Fig 5.**

**ii) As a final example of a system which gives rise to equation (4),** we mention (without derivation) the electrical relay system given in fig. 5, where the voltage \( v \) acts as a control (see [9] p. 56).

**II.2. Behaviour of the trajectories, summary of Chapter II**

We write equation (1.4) in the form of a system

\( \dot{x} = y \),

\( \dot{y} = -(1 + u) \text{ sgn } x \). \hspace{1cm} (1)

In vector form we have \( \dot{x} = \dot{f}(x,u) \) where \( x := (x,y)' \) and \( \dot{f}(x,u) := (y, -(1 + u) \text{ sgn } x)' \). Control functions of (1) are elements of \( L([0,\infty) \rightarrow [-a,a]) \) where \( a \) is some number in \( (0,1] \). For a given
point \( a \in \mathbb{R}^2 \) and control functions \( u \) we denote by \( \pi = \pi(u(t), a) \) the solution of (1) corresponding to \( u \) with \( \pi(0) = a \). In order to prove the existence and uniqueness of the solution we use theorem (1.4.2, iii). Let the control \( u \) be fixed now. Then we apply theorem (1.4.2, iii) with \( S := \{ x \in \mathbb{R}^2 \mid x \neq 0 \} \), \( \xi := (x, t) \to f(x, u(t)) \) and \( a := (x, t) \). Then the functions \( f_i \) \((i = 0, 1, 2)\) are defined by

\[ f_0(x, t) := (y, 0)' \quad f_1(x, t) := (y, -(1+u))' \quad f_2(x, t) := (y, 1+u)' \]

It follows that \( h_2(c, t) := 2g(c) f_2(c, t) = c_2 \) \((i = 0, 1, 2)\) where \( c = (c_1, c_2) \). We have \( c_1 = 0 \), \( c_2 = 0 \) on \( G_0 \) (since \( 0 \notin S \)). We conclude from theorem (1.4.2, iii) that for every \( T > 0 \) and each \( a \in S \) there exists \( T_0 \leq T \) such that there is a solution of (1) with \( \pi(0) = a \) on \([0, T_0]\) and where either \( T_0 = T \) or \( T_0 = \infty \). However, in this last case we can extend the trajectory by setting \( \pi(t) = 0 \) \((T_0 < t < T)\).

Thus it is easily seen that we obtain a solution of (1) on \([0, T]\).

Since this can be done for every \( T > 0 \) and since the corresponding trajectories (with the same control \( u \) and initial value \( a \)) coincide on common intervals, we see that there exists a solution of (1) on \([0, \infty)\) for every \((a, u)\). It should be remarked that the trajectory given here is not necessarily a unisolution (definition 1.4.2). If \( \pi(t_1) = 0 \) for some \( t_1 > 0 \) we sometimes can extend the trajectory on \([t_1, \infty)\) in a way which differs from the trivial one given above. We will give an example of such a trajectory in section 7. However, if we have \( \pi(t) \neq 0 \) \((t \geq 0)\) then it follows from theorem 1.4.3, iii that \( \pi \) is a unisolution.

We are going to describe now the general behaviour of the trajectories. First suppose that \( 0 < a < 1 \). Let \( u \) be an arbitrary control. In the first quadrant \( x \) is increasing and \( y \) is decreasing \((y \leq -1+u)\) and each trajectory will intersect the positive \( x \)-axis within a finite time interval. In the fourth quadrant, \( y \) is still decreasing, now \( x \) is also decreasing, and the negative \( y \)-axis will be crossed.
In the left half plane the situation is analogous. We see that every trajectory starting at a point \( x = (0,0) \) winds around \( (0,0) \) infinitely often (as we will see later, possibly in a finite time interval). If \( u = 1 \), then for an arbitrary control \( u \) a trajectory \( \mathbf{x} = (x,y) \) will have a non-increasing \( y \) and a strictly increasing \( x \) coordinate in the first quadrant, and a non-increasing \( y \) and a strictly decreasing \( x \) in the fourth quadrant. The behaviour in the second and third quadrant can be obtained again by reflection with respect to the origin. It is possible that \( y \) is constant on a trajectory or on a piece of it. It is even possible that \( x \) is constant on a trajectory (for \( y = 0 \), \( u = 1 \)).

**Remark 1.** The properties of the trajectories given here are intuitively obvious and can be proved rigorously using theorem (1.4.2, iii).

It follows from these considerations that (for an arbitrary \( u \in (0,1) \)) the origin cannot be attained by a trajectory which remains wholly in either the right or the left half plane. Yet, the origin can be reached from each point in the plane by some trajectory as will follow from the results of sections 5 and 6.

System (1) has some symmetry and homogeneity properties which will be important in the sequel. These properties are expressed in the following theorem:

**Theorem 1.** If \( x = (t \mapsto (x(t),y(t)))' \) is a trajectory on \([0,T],\) then the following functions are also trajectories on \([0,T]:\)

i) \( t \mapsto (-x(t),-y(t))' \)

ii) \( t \mapsto (x(T-t),-y(T-t))' \)

Furthermore, if \( p > 0 \), then \( t \mapsto (p^2 x(ot),p^2 y(ot))' \) is a trajectory on
\[ [0, \alpha T], \text{ and if } t_0 \text{ is a real number, then } t \mapsto (x(t_0 + t), y(t_0 + t))' \text{ is a trajectory on } [t_0, t_0 + T]. \]

The theorem can be verified by substituting the functions defined here into (1).

Now we give a summary of the contents of the remaining part of this chapter. We will restrict ourselves to the case \(0 < \alpha < 1\) in sections 3 to 9, and we consider the case \(\alpha = 1\) in section 10. We will investigate several problems in this chapter, and define several optimality criteria which all give rise to different kinds of optimal controls. Since sometimes we have to refer to other sections we must distinguish carefully between several kinds of optimality. Therefore we start every section by stating a certain problem \(P_k\) containing a number of parameters \((\lambda_1, \ldots, \lambda_r)\) and we will call the corresponding optimal controls and trajectories "\(P_k(\lambda_1, \ldots, \lambda_r)\)-optimal".

In section 3 we calculate the control which transfers a point \((0, p)\) on the positive \(y\)-axis in minimal time to a point \((0, -q)\) on the negative \(y\)-axis via the right half-plane. Since sign \(x\) does not change in the right half-plane, the problem is one of the type described in section 1.5. We will see in section 3 that it is possible to attain the point \((0, -q)\) from \((0, p)\) via the right half-plane if and only if \(\beta \leq r < \beta^{-1}\) where \(r := q/p\) and \(\beta := (1 - 1)/(1 + \alpha)\), and that the minimal time is equal to \(T = f(r)p\), where 
\[
f = C(\beta^{1/2}, \beta^{-1}) \in \mathbb{R}^1
\]
is a positive, strictly convex function. By theorem 1 it is evident that the same results hold for the left half-plane (from the negative to the positive \(y\)-axis). The results of section 3 are used in all subsequent sections of this chapter.

**DEFINITION:** If \(p > 0\) and \(N\) is natural, a trajectory \(x\) is called a \(p\)-start, \(N\)-stop trajectory on \([0, T]\) if \(y(0) = p\) and if the function \(x\) has exactly \(N+1\) zeros on \([0, T]\), denoted by \(t_0, t_1, \ldots, t_N\), and
II.2

satisfying $0 = t_0 < t_1 < \ldots < t_N = T$. The numbers $t_k$ will be called alteration times, the points $(0, y(t_k))$ alteration points and the numbers $p_k := (-1)^k y(t_k)$ alteration moduli. Obviously, we have $p_0 = p$, $p_k > 0$ ($k = 0, \ldots, N$). The quantity $T$ is called the final time. If we do not want to specify $p$, we will omit the term "$p$-start".

We will call (see Definition 3.1) an $N$-step, $p$-start trajectory piecewise optimal, if the parts of it between two consecutive intersection points are optimal in the sense of section 3. Therefore, if we denote the ratio $r_{k+1}/p_k$ by $r_k$, then for a piecewise optimal control the time used for the $k$-th step ($k = 0, \ldots, N-1$), is equal to $f(r_k)p_k$ and the final time is given by $\sum_{k=0}^{N-1} f(r_k)p_k$.

In section 4 we determine a $p$-start, $N$-step control with minimal final time. It is shown there that such a control exists and is piecewise optimal. The differential optimization problem is reduced by this result to a discrete optimization problem. In fact, now the problem is: "Find a sequence $(r_0, \ldots, r_{N-1})$ with $0 < r_k < \infty$ $(k = 0, \ldots, N-1)$ such that $\sum_{k=0}^{N-1} f(r_k)p_k$ is minimal, where $\{p_k\}_0^N$ is given by $p_0 := p$, $p_{k+1} := r_kp_k$ ($k = 0, \ldots, N$)". This problem is solved in the following way:

We assume that the problem with $N-1$ steps is already solved. Then we perform one step and for the remaining part we use the optimal sequence of the problem with $N-1$ steps. Finally, we choose the first step in an optimal way. This method is called by R. Bellman the dynamic programming method (see [4]).

In section 5 the problem is to find a control $u$ such that the origin is reached from a point on the positive $y$-axis in minimal time. Since the origin cannot be attained from either the right half-plane or the left half-plane, the desired trajectory has to intersect the $y$-axis infinitely often. (We call such a trajectory a $p$-start, $-\infty$-step trajectory.) It is not directly clear in this case that an optimal control exists. But, it is shown in section 5 that we can restrict our attention to piecewise optimal controls also in this case. Therefore, we now have the following discrete problem:
"Find a sequence \( \{r_0, r_1, \ldots \} \) such that \( r_k \in \left[ a^1, a^{1/2} \right] \) and such that for the sequence \( \{p_0, p_1, \ldots \} \) defined by \( p_0 = p_\star \), \( p_{k+1} = r_k p_k \) \((k = 0, 1, \ldots)\) we have \( p_k \rightarrow 0 \) \((k \rightarrow \infty)\) and \( \sum_{k=0}^{\infty} f(r_k) p_k \) is convergent and minimal". Assuming the existence of an optimal control we can easily find the solution by the same method as in section 4. Afterwards it is proved that the solution is actually an optimal control. It also turns out that the optimal control is unique. (This is also the case for the optimal control of section 4.)

As is shown in section 6, the optimal control of section 5 can conveniently be described graphically in the \((x, y)\)-plane. Then it turns out that we can generalize the result of section 5 to the case of an arbitrary starting point \((a, b)\) in the phase plane instead of a point on the \(y\)-axis as was required in section 5. The graphical solution makes it clear, that the optimal control is a feedback control, that is, \( u \) is given as a function of \( x \) and \( y \) instead of \( t \), \( a \), and \( b \). A similar, though more complicated, graphical solution can be found for the problem treated in section 4.

In section 7 we will consider the problem of finding a control which maximizes \(|x(t)|/t|\). For definiteness we will restrict ourselves to \(p\)-start, \(N\)-step controls. Then the problem can be stated as follows: "Find a \(p\)-start, \(N\)-step control which maximizes the quantity \(p_N/T\), where \(p_N\) is the final alteration modulus and \(T\) the final time". In mechanical terms, the optimal control is the one which maximizes the average acceleration (note that \(p\) has the dimension of velocity). It turns out that the optimal control is strongly related to the one of section 4.

In section 9 we are going to consider the following problem: "Given positive real numbers \(p\), \(q\) and a natural number \(N\), find a \(p\)-start, \(N\)-step control for which we have \(p_N = q\) and such that \(T\) is minimal". We will see that the fact that the endpoint of the trajectory is prescribed, complicates the problem a great deal. The dynamic programming method used so far in this chapter yields a complicated recurrence relation for the solution. A more useful result can be obtained, either by the discrete maximum principle or by means of an
auxiliary problem treated in section 9.

The problem treated in section 9 is the following one: "Given real numbers \( p, \bar{p} \), with \( p > 0 \), and a natural number \( N \), find a \( p \)-step, \( N \)-step control which maximizes the quantity \( p\bar{p} + \bar{p}T \) where \( p\bar{p} \) is the final alteration modulus and \( T \) is the final time". The problem can be solved in the same way as the problem of section 4. It is shown in section 9 that the solution of this problem also yields the solutions of the problems of sections 4, 7, 8.

In section 10 we will discuss the results of this chapter for the case \( n = 1 \).

II.3. A basic optimal control problem for the yo-yo equation

\( P_1(p,q) : \) "Given positive numbers \( p \) and \( q \), find a control \( \bar{u} \) and a number \( \bar{t} \) such that

\[
\begin{align*}
\mathbf{x}(t_1,(0,p)') & = (0,-q)' \\
\mathbf{x}(t,(0,p)') & > 0 \\
\end{align*}
\]

holds for \( u = \bar{u} \) and \( t_i = \bar{t} \) and such that for all controls \( u \) and numbers \( t_i \), satisfying (1) and (2) we have \( t_i \geq \bar{t} \)."

Here, as defined in section 2, \( t \mapsto \mathbf{x}(t,z) \) is the trajectory of system (2.1) corresponding to the control \( u \) and the initial value \( z \). Because of condition (2) we can write (instead of (2.1)):

\[
\begin{align*}
\dot{x} & = y \\
\dot{y} & = -(1 + u) \, .
\end{align*}
\]

Hence, except for condition (2), problem \( P_1 \) is of the type discussed in section 1.5. Therefore, denoting by \( t \mapsto \mathbf{x}(t,z) \) the trajectory of (3) satisfying \( x(0) = z \), we replace \( P_1 \) by the following problem:

\( P'_1(p,q) : \) "Given positive numbers \( p \) and \( q \), find a control \( \bar{u} \) and a number \( \bar{t} \) such that \( \mathbf{x}(\bar{t},(0,p)') = (0,-q)' \), whereas for every control \( u \) and every \( t_i \) with \( \mathbf{x}(t_i,(0,p)') = (0,-q)' \), we have \( t_i \geq \bar{t} \)."
It will turn out afterwards that the solutions of $P_1$ and $P'_1$ coincide. In order to apply the theorem of section 1.5 we write (3) in vector notation:

$$\mathbf{z} = A\mathbf{x} + \mathbf{f}(u)$$

with

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{f}(u) = \begin{bmatrix} 0 \\ -(1+u) \end{bmatrix}.$$ 

If $\bar{u}$ is a $P'_1(p,q)$-optimal control, then according to the theorem there exists a non-trivial solution $\psi = (e,\dot{e})$ of the adjoint equation $\dot{\psi} = -\psi A$ (or $\dot{e} = 0$, $\dot{\psi} = -\psi$) such that

$$\|\psi(\bar{u})\| = \max_{\|v\| \leq 1} \|\psi(v)\|,$$ 

that is, $\bar{u} = \alpha \text{ sign } \psi$ for almost all $t$ with $\psi(t) \neq 0$. Now $\psi$ is linear and not identically zero (since $\dot{\psi}$ is non-trivial), and therefore $\psi$ vanishes at no more than one point. It follows that we may assume

$$\bar{u}(t) = \alpha \text{ sign } \psi(t) \quad \quad (0 \leq t \leq \bar{t})$$

since the values of $\bar{u}$ on a set of measure zero do not influence the trajectory. We conclude that $u(t) = \pm \alpha$, and that there is at most one alteration of sign. On the other hand, if $u$ has this form, it satisfies the necessary conditions of theorem (1.5). We will show that there exist at most two admissible controls satisfying this property. Let us first determine the trajectories in the phase plane corresponding to constant controls $u$. For constant $u$ these trajectories satisfy $y \frac{dy}{dx} + 1 + u = 0$, and hence

$$\frac{1}{2} y^2 + (1 + u)x = C$$

for some $C$. Moreover, for all $C$ there is a trajectory of (3) corresponding to the constant control $u$ for which (4) holds. Let us denote the parabola of the form (4) which intersects the $y$-axis in the points $(0,a)$ and $(0,-a)$ by $\eta(u,a)$. According to the foregoing we are
interested in $\pi(\pm a, p)$ and $\pi(\pm a, q)$. The trajectory corresponding to the optimal control (which will be called the optimal trajectory henceforth) consists of one or two arcs of these parabolas.

Therefore, we have to determine the intersection points of $\pi(a, p)$ and $\pi(-a, q)$ and the ones of $\pi(-a, p)$ and $\pi(a, q)$. But, since $y$ is decreasing ($\dot{y} \leq -1/a$) we can only use intersection points $(\bar{x}, \bar{y})$ with $-q \leq \bar{y} \leq p$, or equivalently, intersection points with $\bar{x} \geq 0$. We distinguish three cases:

i) $q < p$. Here we have to consider the intersections of the parabolas $\pi(a, p)$ and $\pi(-a, q)$. (In fig. 1: $A = (0, p)$, $C = (0, -q)$.)

If there are intersection points they satisfy the equations:

$$
\begin{align*}
\frac{1}{2} y^2 + (1 + a)x &= \frac{1}{2} p^2, \\
\frac{1}{2} y^2 + (1 - a)x &= \frac{1}{2} q^2. 
\end{align*}
$$

(5)

If $(\bar{x}, \bar{y})$ is a solution of (5), we have

$$
\bar{x} = \frac{p^2 - q^2}{4a},
$$

(6)

$$
\bar{y}^2 = \frac{(1 + a)q^2 - (1 - a)p^2}{2a}.
$$
If we put
\[ r := \frac{q}{p}, \quad (7) \]
\[ \beta := \frac{1 - \alpha \beta}{1 - \alpha}, \quad (8) \]
it follows that there is no intersection point if \( r < \beta \). If \( r = \beta \) there is exactly one intersection point (lying on the \( x \)-axis), and if \( r > \beta \), there are two intersection points.

Using the notation of (7) and (8), the intersection points are \((x, \bar{y})\) and \((\bar{x}, \bar{y})\), where
\[ x := \frac{1 - r \beta}{1 - \beta} (1 - r^2)p^2, \quad (9) \]
\[ \bar{y} := -p \left( \frac{r^2 - \beta}{1 - \beta} \right) \]

(note that \( \beta < r < 1 \)).

Thus, we have found two admissible trajectories which satisfy the necessary conditions of theorem (I.3). In order to decide which one of them is an optimal trajectory we have to compare the corresponding numbers \( t_1 \) (see (1), (2)). It is easily seen that the trajectory with intersection point \((x, \bar{y})\) has final time \( t_1 \) which is minimal; indeed, we have
\[ t_1 = t_1^+ := \frac{\bar{y} - \bar{y}}{\alpha} + \frac{\bar{y} - \bar{y}}{\alpha}, \quad (11) \]
whereas the other trajectory has the final time
\[ t_1^- := \frac{p + \tilde{v}}{1 + a} + \frac{q - \tilde{v}}{1 - a} = t_1^+ - \frac{4\alpha \tilde{v}}{1 - a^2} > t_1^+ \]

since \( \tilde{v} < 0 \) and \( 0 < a < 1 \). Thus, the trajectory with intersection point \((\tilde{x}, \tilde{y})\) is the optimal trajectory.

**Remark 1.** The point \((\tilde{x}, \tilde{y})\) is called the **switching point** of the trajectory.

ii) \( q = p \). In this case it is easily seen from figure 1 that we have \( \tilde{u} = a \) on the whole trajectory.

iii) \( q > p \). It can be shown by an argument analogous to the one of case i) that the optimal trajectory now starts on \( \tau(-a, p) \) and ends on \( \tau(a, q) \).

Furthermore, the switching point is the intersection point \((\tilde{x}, \tilde{y})\) of these parabolas with \( \tilde{y} > 0 \). The switching point is given by

\[ \tilde{x} = \frac{1}{4} \frac{1 + 8}{1 - 8} (r^2 - 1)p^2, \]

\[ \tilde{y} = p \left( \frac{1 - 8r^2}{1 - 8} \right)^{\frac{1}{2}}, \] (12)

where it is supposed that \( 1 < r < 8^{-\frac{1}{2}} \). It follows that there are admissible controls if and only if \( r \in J \) where \( J \) is defined by

\[ J := [8^{-\frac{1}{2}}, 8^{-\frac{1}{2}}]. \] (13)

For the same values of \( r \) there exists a unique optimal control \( \bar{u} \), which is of the "bang-bang" type, that is, \( \bar{u} \) assumes only the extremal values \( \pm a \). We calculate the minimal final time, which will be denoted by \( T(p, q) \). Using (7) and (8) it follows that

\[ T(p, q) = \ell(r)p, \] (14)
where
\[ f(r) = \frac{1 + \beta}{2\delta} \left[ r + \beta - \sqrt{(1-\beta)(r^2-\beta)} \right] \quad (\beta^\frac{1}{2} \leq r \leq 1) \quad (15) \]

and
\[ f(r) := \frac{1 + \beta}{2\delta} \left[ 1 + \beta r - \sqrt{(1-\beta)(1-\beta^2)} \right] \quad (1 \leq r \leq \beta^{-\frac{1}{2}}) \quad (16) \]

where (15) follows from (11) and the latter formula can be obtained in a way analogous to the way formula (15) was obtained.

We have found now the solution of problem $P_1'$ and since $x \geq 0$ holds on the $P_1'$-optimal trajectory, it follows that the same control furnishes a solution of $P_1$.

Some properties of $f$, which will be used in the sequel, will be derived now. It is easily seen from Theorem 2.1 that $T(p,q) = T(q,p)$.

According to (7) and (14) this implies
\[ f(r) = r f(1/r) \quad (r \in J) \quad (17) \]

It should be remarked that this symmetry property can be used for deriving (16) from (15). Furthermore, $f$ is continuous on $J$ and twice continuously differentiable on $J$ except at the points $\beta^{-\frac{1}{2}}, 1, \beta^\frac{1}{2}$, and we have:

![Graph](image)
\[ f'(r) = \begin{cases} 
\tau \pm \beta & (\tau \pm \beta) \\
\beta^x < r < 1 & (\beta^x < r < 1) \\
\tau^x & (\tau^x) \\
f'(r) = 1 + \beta & (r + 1) \;
\end{cases} \]

(10)

The function \( f \) is positive and has a unique minimum \( 1 + \beta \) at \( r = 1 \). Furthermore, \( f''(r) > 0 \) on \( (\beta^\frac{1}{2}, 1) \) and \( (1, \beta^{-\frac{1}{2}}) \), and hence \( f \) is strictly convex.

REMARK 2. If one wants to find the control which transfers the state variable from \((0, p)\) to \((0, q)\) via the half-plane \( x < 0 \) in a minimal time, then the optimal control one gets is equal to the \( P_1(p, q) \)-optimal control. The optimal trajectory is obtained by reflecting the \( P_1(p, q) \)-optimal trajectory with respect to the origin. It is obvious then, that there is an admissible control again if and only if \( r := q/p \in J \), and that the minimal final time is given by
\[ T = f(r)p \text{ (see also Theorem 2.1).} \]

REMARK 3. It is clear by the autonomy of the problem, that an optimal control (in the sense of \( P_1(p, q) \)) with \( x(t_0) = (0, p) \) (instead of \( x(0) = (0, p) \)) is obtained from the \( P_1(p, q) \)-optimal control by a translation of time.

REMARK 4. It follows from the uniqueness of the optimal control \( \bar{u} \) that \( \bar{u} \) is strictly optimal; that is, every admissible control which is not equal to \( u \) almost everywhere has a larger final time.

Let \( \mathcal{X} \) be a \( p \)-start, \( N \)-step trajectory (see Definition 2.1), and let \( t_k \) and \( p_k \) for \( k = 0, \ldots, N \) be defined as in Definition 2.1. Then, according to the foregoing, we have \( r_k := p_k \bar{t}_k \in J \), and
\[ t_{k+1} - t_k \geq f(t_k)p_k \quad (k = 0, \ldots, N-1). \]
DEFINITION 1. A p-step, N-step control (trajectory) is called piecewise optimal if on each interval \([t_k, t_{k+1})\) the control (trajectory) is optimal in the sense of \(P_1\).

Hence, for a piecewise optimal control we have \(t_{k+1} - t_k = f(t_k)p_k\). A piecewise optimal control \(u\) is for a given \(p\) completely determined by the sequence \(\{r_0, \ldots, r_{N-1}\}\) where \(r_k := p_{k+1}/p_k\) \((k = 0, \ldots, N-1)\). The sequence \(\Lambda := \{r_0, \ldots, r_{N-1}\}\) is an element of \(J^N\) and is called a strategy (or, if we want to specify \(N\) it is called an "N-step strategy"). The functions \((p, \Lambda) \mapsto p_k(p, \Lambda), (p, \Lambda) \mapsto t_k(p, \Lambda)\) for \(p > 0, \Lambda \in J^N\) and \(k = 0, \ldots, N\) are defined by:

\[ p_0(p, \Lambda) := p, \]
\[ p_{k+1}(p, \Lambda) := r_k p_k(p, \Lambda) \quad (k = 0, \ldots, N-1), \quad (19) \]

and

\[ t_0(p, \Lambda) = 0, \]
\[ t_{k+1}(p, \Lambda) = t_k(p, \Lambda) + f(t_k)p_k(p, \Lambda) \quad (k = 0, \ldots, N-1). \quad (20) \]

Furthermore, for \(N > 1\) the mapping \(\Lambda \mapsto \Lambda'\) of \((J^N - J^{N-1})\) is defined by:

"If \(\Lambda = \{r_0, \ldots, r_{N-1}\}\) then \(\Lambda' = \{t_1, \ldots, t_{N-1}\}\)". \quad (21)

COROLLARY 1. Given the sequence \(\{p_0, \ldots, p_{N-1}\}\) with \(p_{k+1}/p_k \in J\) \((k = 0, \ldots, N-1)\), the piecewise optimal control with alteration moduli \(p_k\) \((k = 0, \ldots, N-1)\) minimizes \(t_N\) among all the \(p\)-step, \(N\)-step controls with the same alteration moduli.

COROLLARY 2. We have the following fundamental equality:

\[ t_N(p, \Lambda) = f(t_0)p + t_{N-1}(r_0p, \Lambda'), \quad (N > 1) \quad (22) \]

where \(\Lambda = \{r_0, \ldots, r_{N-1}\}\). In fact:

\[ t_N(p, \Lambda) = f(t_0)p + \sum_{k=1}^{N-1} f(t_k)p_k \]
II.3.4

and

\[ t_{N-1}(r_0 p, \lambda') = \sum_{k=0}^{N-2} f(r_{k+1}) p_{k+1} \].

REMARK 5. If \( \lambda \) has length 1 (that is, \( \lambda = (r_0) \)), we say that \( \lambda' \) is the empty sequence, and we write \( \lambda' = \emptyset \). We call \( \emptyset \) a 0-step strategy and define \( t_0(p, \emptyset) = 0 \), \( p_0(p, \emptyset) = p \). With this convention (22) is also true for \( N = 1 \).

II.4. Time-optimal \( p \)-start, \( N \)-step control

\( P_2(p, N) : \) "Given a real positive number \( p \) and a natural number \( N \), find a \( p \)-start, \( N \)-step control \( u \) such that the final time \( t_N \) is minimal."

From Corollary 1 of the previous section we have the following result:

THEOREM 1. If \( \tilde{u} \) is \( P_2(p, N) \)-optimal, then \( u \) is piecewise optimal. On the other hand, if \( \tilde{u} \) is piecewise optimal and minimizes the final time among the piecewise optimal controls (with \( p_0 = p \)), then \( \tilde{u} \) is \( P_2(p, N) \)-optimal.

PROOF. If \( \tilde{u} \) is not piecewise optimal, and if \( \widetilde{u} \) is the piecewise optimal control with the same alteration points, then the final time corresponding to \( \tilde{u} \) is less than the one of \( \tilde{u} \). The second part of the theorem is proved similarly.

Theorem 1 reduces \( P_2 \) to a discrete optimization problem:

\( P_2'(p, N) : \) "Find an \( N \)-step strategy \( \lambda \) such that \( t_N(p, \lambda) \) is minimal."

Since \( J^N \) is compact and \( \lambda \mapsto t_N(p, \lambda) \) is continuous, the existence of an optimal strategy is clear. The minimal time, which depends on \( p \) is denoted by \( T_N(p) \). Hence:

\[ T_N(p) := \min\{t_N(p, \lambda) \mid \lambda \in J^N \} . \]
In particular, we have \( T_0(p) = 0 \) for every \( p > 0 \) (see Remark 3.5).
Now let for a given \( N \) and \( p, \tilde{\lambda} = (\bar{\tau}_0, \ldots, \bar{\tau}_{N-1}) \) be an optimal
strategy, so that \( t_N(p, \tilde{\lambda}) = T_N(p) \). According to (3.22) we have:

\[
t_N(p, \tilde{\lambda}) = f(\bar{\tau}_0)p + T_{N-1}(\bar{\tau}_0p, \tilde{\lambda}') .
\]  

(2)

This implies that \( \tilde{\lambda}' \) is an \( F^*_2(\bar{\tau}_0p, N-1) \)-optimal strategy. Otherwise
we would be able to improve on \( t_N(p, \tilde{\lambda}) \). It follows that we have:

\[
T_N(p) = f(\bar{\tau}_0)p + T_{N-1}(\bar{\tau}_0p) .
\]  

(3)

Also, it follows from the optimality of \( \tilde{\lambda} \), that \( f(\bar{\tau}_0)p + T_{N-1}(\bar{\tau}_0p) = 
\min_{r \in J} [f(r)p + T_{N-1}(rp)] \). Therefore we get the following recurrence
relation

\[
T_N(p) = \min_{r \in J} [f(r)p + T_{N-1}(rp)]
\]  

(4)

for \( T_N(p) \), whereas \( \tilde{\lambda} \) satisfies the following properties:

\( i) \) \( \bar{\tau}_0 \) is the value of \( r \) for which the minimum in (4) is
assumed.

(4a)

\( ii) \) \( \tilde{\lambda}' \) is \( F^*_2(\bar{\tau}_0p, N-1) \)-optimal.

(4b)

It follows from (3.19) and (3.20) that \( p \mapsto p_k(p, \lambda) \) and \( p \mapsto t_k(p, \lambda) \)
are homogeneous with degree 1 for \( k = 0, 1, \ldots, N \) and for every \( \lambda \in \mathcal{J}^N \).

This implies that \( p \mapsto T_N(p) \) is homogeneous. In fact, if \( p > 0 \), we have:

\[
T_N(rp) = \min_{\lambda \in \mathcal{J}^N} t_N(rp, \lambda) = \rho \min_{\lambda \in \mathcal{J}^N} t_N(p, \lambda) = \rho T_N(p) .
\]

Therefore, if we introduce

\[
S_k := T_k(1) , \quad (k = 0, \ldots, N)
\]

(5)

we have \( T_N(p) = S_Np \). Now (4) implies

\[
S_N = \mu(S_{N-1}) ,
\]

(6)
II.4

where the function $u \in (\mathbb{R}^1 \rightarrow \mathbb{R}^1)$ is defined by

$$u(s) := \min_{t \in J} \{ f(t) + rs \}.$$  \hfill (7)

Now we derive some properties of the function $u$. Since $f$ is strictly convex, the minimum in (7) is assumed at exactly one value of $r$. This value will be denoted by $h(s)$, so that we have

$$u(s) = f(h(s)) + sh(s).$$  \hfill (8)

![Figure 1](image)

It follows from fig. 1, that we have $h(s) = 1$ $(-1 < s < 0)$, whereas for other values of $s$ the values of the function $h$ is determined by $f'(h(s)) = -s$. Hence

$$h(s) = \frac{1 + \beta + 2s}{\sqrt{(s+1)^2 + 4(s+1)s + 4s^2}} \quad (s > 0),$$

$$h(s) = 1 \quad (-1 < s < 0),$$

$$h(s) = \frac{1 - \beta - 2s}{\sqrt{(s+1)^2 + 4(s+1)s + 4s^2}} \quad (s < -1).$$  \hfill (9)

And by equation (8) this yields
\[ \mu(s) = \begin{cases} \frac{1}{2} [1 + \beta + \sqrt{(\beta+1)^2 + 4(\beta+1)s + 4\beta s^2}] & (s > 0), \\ 1 + \beta + s & (-1-\beta \leq s \leq 0), \\ \frac{1}{2\beta} [1 + \beta - \sqrt{(\beta+1)^2 + 4\beta(\beta+1)s + 4\beta s^2}] & (s < -1-\beta). \end{cases} \] (10)

**Fig 2.**

We have the following asymptotic formulas for \( \mu \):

\[ \mu(s) = s^{1/\delta} + \frac{1}{4} (1+\delta)(1+\delta^{-1}) + O(1/s) \quad (s \to \infty), \] (11)

\[ \mu(s) = s^{-1} + \frac{1}{4} \delta^{-1} (1+\delta) + O(1/s) \quad (s \to -\infty). \]

**Remark 1.** Since these asymptotic formulas are starting terms of a convergent series, the asymptotic formulas of \( h(s) = \mu'(s) \) (see formula (15)) can be derived by formal differentiation. Calculating one more term of the asymptotic series for \( \mu \) and differentiating we obtain

\[ h(s) = s^{1/\delta} + \frac{1}{8} (1-\delta)(1+\delta)^2 s^{-3/2} s^{-2} + O(s^{-3}) \quad (s \to \infty), \]

\[ h(s) = s^{-1} + \frac{1}{8} (1-\delta)(1+\delta)^2 s^{-3/2} s^{-2} + O(s^{-3}) \quad (s \to -\infty). \]

(We will use these formulas in section 9.)

There are exactly two points \( s \) for which \( \mu(s) = s \) holds, namely,
II.4

$s^*$ and $-s^*$ where

$$s^* = \frac{2(1 + \beta)}{1 - \beta} = \frac{2}{\alpha}$$ (12)

Let us return to problem $P'_2$. It follows from (4a) that $\bar{r}_0 = h(S_{N-1})$ and from (4b) $\bar{r}_1 = h(S_{N-2})$, It is easily proved by induction that

$$\bar{r}_k = h(S_{N-k-1}) \quad (k = 0, \ldots, N-1)$$ (13)

holds for a $P'_2(-p,N)$-optimal strategy $\bar{a} = \{r_0, \ldots, r_{N-1}\}$. This provides the solution of problem $P_2$. In particular, it follows that the $P'_2(p,N)$-optimal strategy and hence the $P_2(p,N)$-optimal control is unique. The solution is indicated graphically in fig. 4.

In section 6 we will discuss the graphical constructions in the $(x,y)$-plane. From the fact that $u$ is increasing, $u(s) > s$ ($0 < s < s^*$) and $u(s^*) = s^*$ it follows that

$$0 = S_0 < S_1 < S_2 < \ldots < s^*$$
and

\[ S_N + s^* \quad (N \to \infty) \quad (14) \]

This implies that \( l = \bar{r}_{N-1} > \bar{r}_{N-2} > \ldots > r^* : = h(s^*) \).

REMARK 2. A number of properties of \( \mu \) and \( h \) can be derived using only the following properties of \( f \):

i) \( f \in C(J \to \mathbb{R}^1) \),

ii) for all \( r \in J \) we have \( \frac{1}{r} \in J \),

iii) \( f \) is strictly convex,

iv) \( f(r) = rf(1/r) \quad (r \in J) \).

Again \( \mu \) is defined by (7). Using iii), we see that the minimum in (7) is attained for exactly one value of \( r \). This value of \( r \) is denoted by \( h(s) \). It follows that we have equation (8). Now we have the following results:

LEMMA 1. \( \mu'(s) = h(s) \).

PROOF. For real \( s \) and \( \delta \) we have

\[ \mu(s+\delta) = \min_{r \in [f(r) + rs]} f(h(s)) + (s+\delta)h(s) \leq f(s+\delta) + \delta h(s) \]

and

\[ \mu(s) = \min_{r \in [f(r) + rs]} f(h(s+\delta)) + s \delta h(s) = s \mu(s+\delta) - s \delta h(s+\delta) \]

Hence, if \( \delta \) is positive

\[ h(s+\delta) \leq \frac{\mu(s+\delta) - s\delta}{\delta} \leq h(s) \]

and if \( \delta < 0 \),

\[ h(s) \leq \frac{\mu(s+\delta) - s\delta}{\delta} \leq h(s+\delta) \]

Since \( f \) is strictly convex, the function \( h \) is continuous, hence (14).

From this proof it follows also that \( h \) is non-increasing and hence \( \mu \) is concave.
LEMMA 2. \( u(-u(s)) = s \), \hspace{1cm} (16)
that is, the graph of \( u \) is symmetric with respect to the line \( u = -s \).

PROOF. Let \( v(s) := u(-s) = \min_{t \in J} [f(t) - ts] \). Then we have
\[
v(v(s)) = \min_{t \in J} \max \{f(t) - tf(p) + tsp\} = \min_{t \in J} \max_{p \in J} F(t, p),
\]
where \( F(t, p) := f(t) - tf(p) + tsp \). Since for all \( t \in J \) we have
\[
\max_{p \in J} F(t, p) \geq F(t, 1/r) = s,
\]
it follows that \( v(v(s)) \geq s \). Furthermore, it is well known (see [10]
p. 78-80), since \( F \) is continuous on \( J \times J \), \( t \mapsto F(t, p) \) convex for all \( t \)
and \( p \mapsto F(t, p) \) concave for all \( t \), that
\[
v(v(s)) = \max_{p \in J} \min_{t \in J} F(t, p) \leq \max_{p \in J} F(t, p) = s,
\]
which yields (16).

LEMMA 3. \( h(-u(s)) = 1/h(s) \).

This result is obtained by differentiating (16) and replacing \( s \) by
\(-s\).

II.5. Time optimal null-control

\( P_p(t) \): "Given a real number \( p > 0 \), find the control \( u \) such that if
\( x := (t \mapsto x_0(t, 0, p')) \), we have \( x(T) = 0 \) for a minimal \( T \)."

Since, as we have seen in section 2, the origin cannot be
reached from either the right or the left half-plane, the trajectory \( x \) has to intersect the y-axis infinitely often. We will call such a
trajectory with initial point \((0, p), p > 0, a \) p-start, \( s \)-step tra-
jectory (and the corresponding control a p-start, \( s \)-step control).
The alteration times \((t_0, t_1, \ldots)\) form an infinite sequence and so do
the alteration points \((0, y(t_k))\) and the alteration moduli \((p_k)\)
with \( p_k := (-1)^k y(t_k) \). Again we have \( r_k := p_{k+1}/p_k \in J \) and
\[ t_{k+1} - t_k = f(t_k)p_k \]. If \( \lim_{k \to \infty} t_k =: T \) exists, it is called the final time, otherwise we have \( t_k \to (k \to \infty) \) and then we set \( T = \infty \). Piecewise optimal controls are defined in a similar way as in section 3. For piecewise optimal controls we have \( t_{k+1} - t_k = f(t_k)p_k \). Just as in section 4, it is easily seen that we can restrict our attention to piecewise optimal controls when looking for \( P_3(p) \)-optimal control. Therefore we define an \( \infty \)-step strategy, as an infinite sequence \( (t_k)_{k=0}^{\infty} \) of numbers \( t_k \in J \). The set of \( \infty \)-step strategies will be denoted by \( J^{\infty} \). Furthermore, the functions \( t_0, p_k \) in \( (\mathbb{R}_+ \times J^m \to \mathbb{R}) \) are defined by

\[
P_0(p, \lambda) = p \, ,
\]

\[
P_{k+1}(p, \lambda) = t_k p_k(p, \lambda) \, ,
\]

and

\[
t_0(p, \lambda) = 0 \, ,
\]

\[
t_{k+1}(p, \lambda) = t_k(p, \lambda) + f(t_k)p_k(p, \lambda) \, .
\]

for \( k = 0, 1, \ldots \). Furthermore, the function \( T \in (\mathbb{R}_+ \times J^m \to \mathbb{R}) \) is defined by: If \( \lim_{k \to \infty} t_k(p, \lambda) \) exists, then

\[
T(p, \lambda) = \lim_{k \to \infty} t_k(p, \lambda) \, ,
\]

otherwise (if \( t_k(p, \lambda) \to (k \to \infty) \)), we define \( T(p, \lambda) = \infty \). The problem \( P_3(p) \) is equivalent to

\[ P_3^\infty(p) : \text{"Given a real positive number } p, \text{ find an } \infty \text{-step strategy } \lambda \text{ such that } T(p, \lambda) \text{ is finite and minimal."} \]

We have omitted the condition \( p_k(p, \lambda) \to 0 \) \((k \to \infty)\) in \( P_3^\infty \). But this condition is evidently satisfied if

\[
T(p, \lambda) = \sum_{k=0}^{\infty} f(t_k)p_k(p, \lambda)
\]
is finite, since $f(r) \geq 1 + g$. Examples of strategies $A$ with $T(p, A) = 0$ are easily given. For instance, if $A := (r, r, r, \ldots)$ with $0 \leq r < 1$, we have

$$T(p, A) = p f(r) (1 + r + r^2 + \ldots) = p f(r) / (1 - r).$$

It is not immediately clear that a $P_3$-optimal strategy exists. We will assume for the moment the existence of a $P_3$-optimal strategy for every $p > 0$. Given $p > 0$ we denote this optimal strategy by $\tilde{A} = (\tilde{r}_0, \tilde{r}_1, \ldots)$. Let $\tilde{A}' := (\tilde{r}_1, \tilde{r}_2, \ldots)$. Then in a similar way as in (1.22), or by taking limits in (1.22), we can prove

$$T(p, \tilde{A}) = f(\tilde{r}_0)p + T(\tilde{r}_0, \tilde{A})' . \tag{4}$$

In a similar way as in section 4 it can be proved that $\tilde{A}'$ is a $P_3(\tilde{r}_0, p)$-optimal strategy. If we define

$$\overline{f}(p) := \inf \{ T(p, A) \mid A \in \mathcal{A} \} \quad (p > 0) , \tag{5}$$

then it follows that

$$\overline{f}(p) = f(\tilde{r}_0)p + \overline{T}(\tilde{r}_0, p) = \min_{r \in \mathbb{R}} [f(r)p + \overline{T}(rp)] . \tag{6}$$

Also, $\overline{f}$ is easily seen to be homogeneous. Furthermore, defining

$$S := \overline{f}(1) , \tag{7}$$

it follows from (6) that:

$$S = u(S) , \tag{8}$$

where $u$ is the function defined by (4.7). Since clearly $S > 0$ holds, we see that

$$S = \frac{2}{\alpha} . \tag{9}$$

(cf. (4.12)). Also we have $\tilde{r}_k = r^* (k = 0, 1, \ldots)$ where

$$r^* := h(s^*) = \frac{1 + 3 \alpha}{3 + \alpha} = \frac{2 - \alpha}{2 + \alpha} . \tag{10}$$

This constitutes the solution of problem $P_3$, provided the optimal
control exists. We show the existence of the optimal strategy by proving that \( \tilde{\lambda} := (x^*, r^*, \ldots) \) actually is a \( P_3(p) \)-optimal strategy. We have

\[
T(p, \lambda) = \frac{f(h(s^*))}{1 - r^*} = \frac{f(h(s^*))}{1 - h(s^*)} = \frac{p u(s^*) - s^* h(s^*)}{1 - h(s^*)} = p s^*,
\]

where we have used (4.8) and the fact that \( u(s^*) = s^* \).

Suppose that there exists a strategy \( \lambda = (r_0, \lambda_1, \ldots) \) such that \( T(p, \lambda) < s^* p \). We have seen in section 4, that for the \( P_2(p, N) \)-optimal strategy we have \( T_N(p) = s_N^* p + s^* p \) \((N = \infty)\). Hence, there exists \( N_0 \) such that \( s_{N_0}^* p > T(p, \lambda) \). But

\[
T(p, \lambda) = \sum_{k=0}^{\infty} f(r_k)p_k(p, \lambda) \geq \sum_{k=0}^{N_0-1} f(r_k)p_k(p, \lambda) + \sum_{k=N_0}^{\infty} f(r_k)p_k(p, \lambda) \geq T_{N_0}(p),
\]

where \( \lambda^* = (r_0, \ldots, r_{N_0-1}) \in J_{N_0}^0 \), so that we have a contradiction.

So we have obtained the following result: The strategy \( \lambda := (r^*, r^*, \ldots) \) where \( r^* \) is given by (10), is the unique \( P_2(p) \)-optimal strategy for every \( p > 0 \). The \( P_2(p) \)-optimal control \( \tilde{\lambda} \) is the piecewise optimal control with \( p_k(p, \lambda) \) as alteration moduli for \( k = 0, 1, \ldots \). Hence the \( P_2(p) \)-optimal control is unique also.

**Remark.** The existence of a \( P_2'(p) \)-optimal strategy may also be proved (without using the results of \( P_2 \)) by Tychonoff’s theorem, according to which \( J^0 \) is compact, with respect to the Tychonoff topology (see [11] p. 5). Unfortunately, the function \( \lambda \rightarrow T(p, \lambda) \) is not continuous (not even finite) on all of \( J^0 \). Therefore we introduce an auxiliary system in the following way: The function \( g_\delta(t) \) is defined by \( g_\delta(t) := \min(t, 1-\delta) \), where \( \delta \) is some positive number. Then we define functions \( P_{k, \delta} \) and \( t_{k, \delta} \) in \((R^1 \times J^0, R^1)\) by:

\[
P_{0, \delta}(p, \lambda) = p,
\]
\[
P_{k+1, \delta}(p, \lambda) = g(r_k)p_{k, \delta}(p, \lambda),
\]
\[
t_{0, \delta}(p, \lambda) = 0.
\]
\( t_{k+1,\delta}(p,\lambda) = t_k(p,\lambda) \star f(r_k)p_{k,\delta}(p,\lambda) \),

where \( \lambda = (\tau_0, \tau_1, ...) \). Furthermore, we define \( T_\delta \in (R^1 \times J^\omega \rightarrow R^1) \) by

\[
T_\delta(p,\lambda) = \lim_{k \to \infty} t_{k,\delta}(p,\lambda) = \sum_{k=0}^{\infty} f(r_k)p_{k,\delta}(p,\lambda).
\]

\( T_\delta \) is well defined since \( |p_{k,\delta}(p,\lambda)| \leq (1 - \delta)^k \). In the Tychonoff topology, a function \( F \in (J^\omega \rightarrow R^1) \) is continuous, if for any finite sequence \( I = \{i_1, \ldots, i_v\} \) the mapping \( \{\tau_{i_1}, \ldots, \tau_{i_v}\} \mapsto F(\lambda) \) is continuous for each sequence \( \{r_{k_i}\}_{k_i=0}^{v}, k_i \neq I \). Therefore, \( \lambda \rightarrow T_\delta(p,\lambda) \) is easily seen to be continuous, so that there exists \( \bar{\lambda} \in J^\omega \), such that

\[
T_\delta(p,\bar{\lambda}) = \min(T_\delta(p,\lambda) \mid \lambda \in J^\omega) = \bar{T}_\delta(p).
\]

The optimal strategy \( \bar{\lambda} \) can be obtained in the same way as the one for \( P_\delta^*(p) \). It follows that \( \bar{T}_\delta(p) = s_\delta p \), where \( S_\delta \) satisfies

\[
S_\delta = \mu_\delta(S_\delta) \quad \text{and where} \quad \mu_\delta \text{ is given by}
\]

\[
\mu_\delta(s) = \min \{ f(r) + \delta g(r) \} = \min_{r \in J} \{ f(r), 1 + \delta + (1 - \delta) s \}.
\]

Now, if \( \delta > 0 \) is sufficiently small, it follows that \( S_\delta = s^* \) and \( \bar{\lambda} = \{t^*, t^*, \ldots\} \). But we have also \( T(p,\bar{\lambda}) = s^* \). Therefore we have for arbitrary \( \lambda \in J^\omega \):

\[
T(p,\lambda) \geq T_\delta(p,\lambda) \geq T_\delta(p,\bar{\lambda}) = T(p,\bar{\lambda})
\]

which proves the \( P_\delta(p) \)-optimality of \( \bar{\lambda} \).

II.6. Switching curves

In this section we give a description in terms of graphs of the solution found in the previous section. As is shown in section 3, for a \( P_1(p,x,p) \)-optimal control (with \( r > 1 \)) the control assumes the values 0 and \( -\alpha \) consecutively, and the point where \( u \) changes sign, called the switching point, is given by (3.9) and (3.10). As we saw
in the previous section, a $P_3(p)$-optimal trajectory consists of an infinity of pieces, each of which is a $P_1(p, r^p)$-optimal trajectory with a translated initial time. The switching points of these pieces in the right half-plane lie on a curve which is given in parameter representation by

$$x = \frac{2}{(2 + a)^2} \bar{p}^2, \quad y = \frac{2}{2 + a} \bar{p}$$  \hspace{1cm} (1)

where $\bar{p}$ is the parameter (see formulas (3.9), (3.10)). The curve is also given by the equations $x = 2y^2/a^2$, $y < 0$. For the pieces in the left half-plane, the switching points lie on the curve $x = -2y^2/a^2$, $y > 0$. In general, the set of switching points is given by the equation:

$$x = -2a^{-2}y|y|.$$  \hspace{1cm} (2)

This set is called the switching curve and will be denoted by $C$. If we introduce the function $g \in C^1(\mathbb{R}^2 \to \mathbb{R})$ by

$$g(x) = x^2 + 2a^{-2}xy|y|,$$  \hspace{1cm} (3)

then it follows from the foregoing that if $\bar{u}$ is $P_3(p)$-optimal control and $\bar{x}$ is the corresponding trajectory, we have $\bar{u}(t) = a \text{ sgn } g(\bar{x}(t))$ a.e. On the other hand, by theorem (1.4.2) it is easily seen that the system

$$\begin{align*}
\dot{x} &= y \\
\dot{y} &= -(1 + a \text{ sgn } g(x, y)) \text{ sgn } x
\end{align*}$$  \hspace{1cm} (4)

has a unisolution for initial value $\bar{a} \in \mathbb{R}^2$, $\bar{x} \neq 0$, and that this solution can be continued, either for all $t > 0$ or to the boundary of $\mathcal{S} := \{x \in \mathbb{R}^2 \mid \bar{x} \neq 0\}$; that is, until we have $\bar{x} = 0$. 
This implies that trajectories of (4) are $P_3$-optimal trajectories (see fig. 2).

An optimal control which is given as a function of the state variable is called a feedback control. Therefore, by the equation $u = a \text{ sgn } g(x)$ we have given the optimal feedback.

The feedback formulation of the solution of $P_3$ enables us to find the optimal trajectory for a more general problem than $P_3$. In fact, now we can solve the problem of time optimal null-control with arbitrary value $x = (a, b) \in \mathbb{R}^2$ instead of the point $(0, b)$ on the y-axis. For this problem, the optimal control is also given by the feedback $u = a \text{ sgn } g(x)$. 
In order to see this, consider for definiteness a point \( a = (a, b) \) with \( a > 0 \). There exists exactly one real \( p > 0 \) such that the \( P_3(p) \)-optimal trajectory passes through the point \( a \) on the first step. Now the part of this trajectory, starting at \( a \) is the time optimal null-control with initial value \( a \). For, if there exists a control \( \bar{G} \), which transfers \( a \) to the origin in a shorter time interval, then we can construct a trajectory which transfers \( (0,p) \) to the origin in a shorter time; this is done by concatenating the part between \( (0,p) \) and \( (a,b) \) of the \( P_3(p) \)-optimal control and the trajectory corresponding to \( \bar{G} \) from \( a \) to \( 0 \). The same argument shows the optimality of the feedback \( u = \alpha \text{ sgn } g(x) \) for initial values \( a = (a,b) \) with \( a < 0 \).

We find a surprisingly simple result for the switching time \( t_k \) in the time interval \((t_k, t_{k+1})\). In fact, we have

\[
0_k = \frac{1}{2}(t_k + t_{k+1}).
\] (5)

Hence, the control assumes the values \( u = \alpha \) and \( u = -\alpha \) on intervals of equal length in \((t_k, t_{k+1})\). We can obtain formula (5) by substituting \( r = r^* = (2-\alpha)/(2+\alpha) \) in (3.10) and computing the first and second term of the right-hand side of (3.11) with \( q = r^* \). Now we describe an independent method for finding the \( P_3(p) \)-optimal control (in particular formula (5)), without the results of sections 3, 4, 5 and even without the maximum principle. Admittedly, we assume in this derivation that the existence of a \( P_3(p) \)-optimal control has already been proved, but, if we wish to do so, we can establish the existence of an optimal control independently by means of the solution of the problem \( P_3 \) below. Then it is a consequence of Theorem 2.1 that the optimal final time \( T(p) \) depends homogeneously on \( p \), so that we can write \( T(p) = Sp \) for some \( S > 0 \). Let \( x = (x,y) \) be an optimal trajectory and let \( t_1 \) be the smallest value of \( t > 0 \) for which we have \( y(t_1) = 0 \). Then we have \( T(|y(0)|) = t_1 + T(|y(t_1)|) \) and hence \( S = t_1/(|y(0)| - |y(t_1)|) \). Because of the homogeneity of the problem, it is no loss of generality to assume that \( y(0) = 1 \). Then we have \( y(t_1) < 0 \) and hence
\[ s = \frac{t_1}{1 + y(t_1)} = \frac{1}{\tan \phi} \quad (6) \]

(see fig. 4, where \( A = (0,1), \ B = (t_1,1), \ A' = (0,-1), \ B' = (t_1,-1), \ O = (0,0), \ P = (t_1,0), \ C = (0,y(t_1)), \ D = (t_1,y(t_1)). \)  

![Diagram]

Fig. 4.

Furthermore, it follows from \( x(0) = x(t_1) = 0 \) that

\[ \int_{0}^{t_1} y(t)dt = 0 \quad \text{ (7) } \]

Also (according to (2.1)), the function \( y \) satisfies \( |\dot{y} + 1| \leq \sigma. \) We will consider the somewhat more general problem, where \( y \) has to satisfy \( -\sigma \leq \dot{y} \leq -\sigma, \) and \( \sigma, \rho \) are arbitrary numbers satisfying \( \rho > \sigma > 0. \) We show that, if (7) is satisfied, the area of the triangle \( A'B'D \) equals the area of the shaded figure (in fig. 4), that is,

\[ \int_{0}^{t_1} (\ddot{y}(t) - y(t))dt , \]

where \( \ddot{y} \) is the function of which the line segment \( AD \) is the graph.

Denoting the area of a polygon \( XYZ... \) by \( <XYZ...> \) we have
\[ \int_0^{t_1} (\bar{y}(t) - y(t))dt = \int_0^{t_1} y(t)dt = \langle \text{ABC} \rangle = \langle \text{OPCD} \rangle = \]
\[ = \frac{1}{4} \langle \text{ABCD} \rangle = \langle \text{OPCD} \rangle = \frac{1}{4} (\langle \text{ABPD} \rangle - \langle \text{OPCD} \rangle) = \langle \text{A'B'D} \rangle. \]

Since \( \tan \theta = \frac{\langle \text{A'B'D} \rangle}{\langle \text{OP} \rangle} \) we have reduced \( P_3 \) to the following problem:

\[ \mu_3^*: \text{"Find a number } t_1 \text{ and an absolutely continuous function } \]
\[ y \in (\mathbb{R}^1) \rightarrow \mathbb{R}^1, \text{ satisfying } (7), y(0) = 1, \text{ and } -\rho \leq \dot{y} \leq \rho \quad (a.e.), \text{ such that } \]
\[ I(y, t_1) := t_1^2 \int_0^{t_1} (\bar{y}(t) - y(t))dt \]
\[ \text{is maximal, where } \bar{y}(t) := 1 + (y(t_1) - 1)t/t_1.\]

Let \( y \) be a function in some interval \([0, t_1]\) satisfying \( y(0) = 1, \)
\[ -\rho \leq \dot{y} \leq \rho \quad (a.e.), \text{ and } \int_0^{t_1} y(t)dt < t_1. \] We define \( y^\lambda \in (\mathbb{R}^1) \rightarrow \mathbb{R}^1 \)
\[ y^\lambda(t) := 1 + \lambda(y(t) - 1), \quad (8) \]
for \( \lambda > 0. \) Then we have \( y^\lambda(0) = 1, -\rho \leq \dot{y}^\lambda(t) \leq \rho \quad (a.e.). \) Note that
the graph of \( y^\lambda \) is obtained by multiplying the graph of \( y \) by \( \lambda \) with
respect to \((0, 1). \) There exists \( \lambda > 0 \) such that \( \int_0^{t_1} y^\lambda(t)dt = 0. \) In
fact, it follows from
\[ J_\lambda := \int_0^{\lambda t_1} y^\lambda(t)dt = \lambda^2 \left( \int_0^{t_1} y(t)dt \right) \]
that \( J_\lambda \sim \lambda^2(J_1 - t_1) < 0 \quad (\lambda \rightarrow \infty) \) and \( J_\lambda \sim \lambda t_1 > 0 \quad (\lambda = 0). \) Also,
it follows from (9) that \( I(y^\lambda, \lambda t_1) = I(y, t_1). \) This implies that we
can restate \( P_3^* \) as follows:
\( F^x_3 \) : "Find an absolutely continuous function \( y \in ([0,1] \to \mathbb{R}^1) \), satisfying:

\[
y(0) = 1, \quad \int_0^1 y(t) dt < 1, \quad -p \leq \dot{y}(t) \leq -c \quad \text{(a.e.)},
\]

such that

\[
J(y) := \int_0^1 (\dot{y}(t) - y(t)) dt
\]

is maximal, where \( \dot{y}(t) := 1 + t(y(1) - 1) \quad (0 \leq t \leq 1) \).

We will call functions \( y \), satisfying (10) admissible. Let \( y \) be admissible. Because of the restriction \(-p \leq \dot{y} \leq -c\) we have \( y(t) \geq 1- pt \) and \( y(t) \geq y(1) + c(1-t) \). Hence

\[
y(t) \geq y^*(t) := \max\{1- pt, y(1) + c(1-t)\}.
\]

But it is easily seen that also \( y^* \) is admissible and that \( J(y^*) > J(y) \).

It follows that if \( y \) is optimal, then there exists \( \theta \in [0,1] \) such that \( y \) is linear on each of the intervals \([0,\theta]\), \([\theta, 1]\) and such that \( \dot{y} = -p \ (0 < t < \theta) \), \( \dot{y} = -c \ (\theta < t < 1) \). Hence \( y \) is of the form

\[
y(t) = 1 - pt \quad \text{ (} 0 < t < \theta) , \\
y(t) = 1 - pt + \frac{1}{2}(p - c)t^2 \quad \text{ (} \theta < t < 1) .
\]
A short calculation yields that in this case we have

\[ J(y) = \frac{1}{2}(\varphi - y)\theta(1 - 0) \]

so that \( J(y) \) is maximal for \( \vartheta = \frac{1}{2} \). This formula is equivalent to (5).

**Remark 1.** Mechanically the switching point is determined by a fixed ratio of the potential and kinetic energy of the system. Apart from constant factors, the potential energy is given by \( E_p = |x| \) and kinetic energy by \( E_k = y^2 \). Hence for the switching point we have \( E_p = 2\alpha^{-2}E_k \). Furthermore, \( E_p \) decreases (considered as function of \( t \)) at the switching time (i.e. \( dE_p/dt < 0 \)).

We are going to give now an analogous description in terms of graphs of the \( P_3 \)-optimal trajectories. Consider the set of \( P_1(p, xp) \)-trajectories for fixed \( r \in J, r \leq 1 \) and varying \( p > 0 \). The switching points of these trajectories are given by (3.9) and (3.10) and form together a semi-parabola: \( x = \vartheta(r)y^2, y < 0 \), where

\[ \vartheta(r) = \frac{1 + \frac{3}{4} \frac{1 - r^2}{r^2 - \beta}}{2} , \]  

(11)

provided \( r > \beta^{1/2} \). If \( r = \beta^{1/2} \) the switching points are given by the equation \( y = 0 \). If \( r = 1 \), the \( y \)-axis is the switching curve. We will denote this curve together with its mirror image with respect to the origin by \( \Gamma(r) \). The \( P_3(p, N) \)-optimal trajectory is piecewise optimal corresponding to the strategy \( \bar{A} = (T_0, \ldots, T_{N-1}) \) given in (4.13). The switching curve for the piece between the \( k \)-th and the \( (k+1) \)-st intersection point is \( \Gamma(T_k) \). As a result we have a set of \( N \) switching curves \( \Gamma(T_0), \ldots, \Gamma(T_{N-1}) \). We denote the curve \( \Gamma(T_{N-k}) \) by \( \Gamma_k \), so that

\[ \Gamma_k = \Gamma(h(S_{k-1})) \quad (k = 1, 2, \ldots) . \]

We have \( \Gamma_k \rightarrow \Gamma(x^\ast) \quad (k \rightarrow \infty) \) (in an appropriate sense) and \( \Gamma_1 \) is the \( y \)-axis.
At each piece of the $P_2$-optimal trajectory we have to choose one of the curves as a switching curve. In fact, if the trajectory still has to intersect the $y$-axis $\ell$ times (irrespective the numbers of intersections before), then we have to choose $\Gamma_1$ as the switching curve. Exactly in the same way as for time optimal null control, in this case we can find optimal trajectories for the generalized version of $P_2$, where the initial value is arbitrary (instead of on the positive $y$-axis).

II.7. Maximal average acceleration

$P_\alpha(p,N)$: "Given a natural number $N$ and a positive real number $p$, find the $p$-start, $N$-step control which maximizes the ratio $p_N/T$ where $p_N$ is the last alteration modulus and $T$ is the final time."

Just as in the problems $P_2$ and $P_3$ it is easy to show here that we can restrict our attention to piecewise optimal controls. Therefore, equivalent to $P_\alpha$ is the following problem:

$P'_\alpha(p,N)$: "Given $p > 0$, $N$ a natural number, find an $N$-step strategy $\lambda = \{x_0, \ldots, x_{N-1}\}$ such that $v_N^\alpha(p,\lambda) = t_N(p,\lambda)/p_N(p,\lambda)$ is
minimal."

Just as in problem $F'_2$, the existence of an optimal strategy is obvious. It follows from Remark 3.5 that $V(p, \emptyset) = 0$ ($p > 0$).

Let for $N > 1$ and an arbitrary $\hat{\lambda} = \{r_0, \ldots, r_{N-2}\} \in J^N$, the $(N-1)$-step strategy $\hat{\lambda}$ be defined by $\hat{\lambda} = \{r_0, \ldots, r_{N-2}\}$ and for $N = 1$, $\hat{\lambda} = \emptyset$. Then we have for $N > 1$:

$$V_N(p, \hat{\lambda}) = \frac{t_{N-1}(p, \hat{\lambda}) + f(r_{N-1})P_{N-1}(p, \hat{\lambda})}{t_{N-1}P_{N-1}(p, \hat{\lambda})} = \frac{1}{t_{N-1}} \left[ V_{N-1}(p, \hat{\lambda}) + f(r_{N-1}) \right]. \quad (1)$$

It follows from this formula, that if $\hat{\lambda}$ is $F'_N(p, N)$-optimal, $\hat{\lambda}$ is $F'_N(p, N-1)$-optimal. If $\bar{V}_N$ is defined by

$$\bar{V}_N := \min_{\hat{\lambda} \in J^N} \{V_N(p, \hat{\lambda})\}, \quad (2)$$

(so that $\bar{V}_0 = 0$), we infer from (1) that

$$\bar{V}_N = \min_{r \in J} \left( \bar{V}_{N-1} + r \frac{f(r)}{r} \right),$$

whereas $t_{N-1}$ is a value of $r$ for which the minimum in (2) is assumed.

But since $1/r \in J$ if and only if $r \in J$, it follows by the substitution $\rho = 1/r$, that

$$\bar{V}_N = \min_{\rho \in J} \left[ \rho \bar{V}_{N-1} + f(1/\rho) \right] = \min_{\rho \in J} \left[ \rho \bar{V}_{N-1} + f(0) \right] = u(\bar{V}_{N-1}),$$

with $u$ as defined by (4.7). The value of $\rho$ for which the minimum is assumed is $h(\bar{V}_{N-1})$, and hence $t_{N-1} = 1/h(\bar{V}_N)$. Since $\hat{\lambda}$ is $F'_N(p, N-1)$-optimal, we have $t_{N-2} = 1/h(\bar{V}_{N-2})$. It follows by induction that we have

$$\bar{V}_k = u(\bar{V}_{k-1}) \quad (k = 1, \ldots, N), \quad \bar{V}_0 = 0, \quad (3)$$

and the $F'_N(p, N)$-optimal strategy $\hat{\lambda} = \{r_0, \ldots, r_{N-1}\}$ is given by
\[ \bar{r}_k = \frac{1}{h(V_k)} \quad (k = 0, \ldots, N-1). \] (4)

In particular, we see that the optimal strategy (and hence the \( P_3(p, N) \)-optimal control) is unique.

REMARK. It is somewhat surprising that \( \bar{V}_0 = 1 \), but in the case \( N = 1 \) it is easily seen to be correct from fig. 3.5. Note also that \( V_N \) does not depend on \( p \).

It follows from (3) that \( \bar{V}_N = S_N \), in particular, we have
\[ \bar{V}_N \rightarrow a^* (N \rightarrow \infty) \] where the convergence is monotonic.

If we use the \( N \)-step strategy
\[ \Lambda^* = \left\{ 1/r^*, 1/r^*, \ldots \right\}, \] (5)

with \( r^* \) as defined in (5.10), we have \( V_k(p, \Lambda^*) = s^* \) for \( k = 1, \ldots, N \). Although this strategy is not optimal we see that \( V_N(p, \Lambda^*) - \bar{V}_N \rightarrow 0 \) \((N \rightarrow \infty)\). Therefore we will call this strategy "quasi-optimal". The advantage of the quasi-optimal strategy is that the corresponding piecewise optimal control is simple to describe. We shall do this in a graphical way just as in the previous section. The switching curve corresponding to (5) is given by (3.12) with \( r = 1/r^* \) and where \( p \) is a parameter. Eliminating \( p \) we obtain
\[ x = 2y|y|/a \] (6)

for the switching curve (see fig. 1). For the actual optimal control we have a number of switching curves similar to the ones of the \( P_2 \)-optimal controls. In this case we get for the set of switching curves:
\[ x = \delta(r_k)y|y| \quad (k = 0, 1, \ldots), \] (7)

where \( \delta \) is defined by
\[ \delta(r) = \frac{1 + 3r^2 - 1}{4(1 - \delta r^2)} \quad (1 \leq r < \delta^{-1}) \]

and \( \bar{r}_k \) by (4) (see fig. 2).
REMARK 1. Unlike the situation of section 5 we cannot use this feedback control for optimal trajectories starting at an arbitrary point of the plane.

REMARK 2. We are now in the position to give the example (promised in section 2) of a situation where the trajectories are not unique. For that aim we use the feedback control indicated in fig. 1: that is, \( u = \alpha \text{sgn} \ g(x) \), where \( g(x) = x^2 - 2y|y|/a \). If we reverse the time in the system

\[
\dot{x} = y, \quad \dot{y} = - (1 + \alpha \text{sgn} \ g(x))x,
\]

then the trajectory attains the origin in a finite time interval \( T \). Therefore, the solution of (8) starting at \( \bar{x} \) are not unisolutions.

II.8. Time optimal \( p \)-start, \( N \)-step control with fixed endpoint

\( F_5(p,q,N) : \) "Given real positive numbers \( p \) and \( q \) and a natural number \( N \), determine the \( p \)-start, \( N \)-step control with final alteration modulus \( p_N = q \) and minimal final time \( t_N \)."
Again, we can confine ourselves to piecewise optimal controls. Hence equivalent to $P_5$ we have:

$$P_5^I(p,q,N): \text{"Given real numbers } p,q > 0 \text{ and a natural number } N, \text{ find the } N \text{-step strategy } \lambda \text{ such that } p_N(p,\lambda) = q \text{ and } t_N(p,\lambda) \text{ is minimal, where the functions } p_N, \text{ and } t_N \text{ are defined in (3.19) and (3.20)."}$$

In order that there exists $\lambda \in J^N$ for which $p_N(p,\lambda) = q$ holds, it is necessary and sufficient that $q/p \in J_N$, where

$$J_N := [\delta^{N/2}, \delta^{-N/2}] .$$

Since $J^N$ is compact and $\lambda = p_N(p,\lambda), \lambda = t_N(p,\lambda)$ are continuous, it is obvious that there exists an optimal strategy whenever the condition $q/p \in J_N$ is satisfied. We will denote the $P_5(p,q,N)$-optimal time by $T_N(p,q)$. Hence

$$T_N(p,q) := \min \{t_N(p,\lambda) \mid \lambda \in J^N \land p_N(\lambda) = q\} .$$

Here the minimum of an empty set is defined to be infinity. Hence $T_N(p,q) = \infty$ if $q/p \notin J_N$. Let $\lambda = (\bar{r}_0, \ldots, \bar{r}_{N-1})$ be a $P_5^I(p,q,N)$-optimal strategy. By (3.22) we have

$$t_N(p,\lambda) = f(\bar{r}_0)p + t_N-1(\bar{r}_0,\lambda') .$$

Since $p_N(p,\lambda) = q$, it follows that $p_{N-1}(\bar{r}_0,\lambda') = q$. Therefore, it is easily seen that $\lambda'$ is $P_5^I(\bar{r}_0,p,q,N-1)$-optimal. It follows then that $T_N(p,q) = f(\bar{r}_0)p + T_N(\bar{r}_0,p,q)$ and because of the optimality of $\lambda$ this yields the equation

$$T_N(p,q) = \min_{r \in J} \{f(r)p + T_{N-1}(rp,q)\} .$$

Properly speaking the minimum should be over those $r \in J$ for which $q/(rp) \in J_{N-1}$ holds. However, as $T_N(p,q) = (q/p \notin J_N)$, we can use the simpler formulation (4). Now it is easily seen that for $p > 0$
we have $T_N(p,q) = q T_N(p,q)$; that is, $T_N$ is homogeneous. Therefore, if we set

$$y_N(q) = T_N(1,q) \quad (5)$$

we obtain the following recurrence relation for $S_N(q)$:

$$S_N(q) = \min_{r \in \mathcal{J}} \left\{ f(r) + q s_{N-1}(r/q) \right\}$$

(q < 1) \quad (6)

$$S_1(q) = f(q) \quad (q \in \mathcal{J})$$

(q < 1) \quad (6)

From this recurrence relation the functions $S_N(q)$ can be calculated successively for $N = 1, 2, \ldots$. However, these calculations cannot be performed analytically, but only by a numerical method. We shall not go into this subject here, since we will obtain a much more tractable result in the next section.

We conclude this section with some simple properties of $S_N$. At first, it is easily shown by induction that $S_N$ is continuous. Furthermore, we have

$$S_0(1/q) = q^{-1} S_N(q) \quad (7)$$

This is a consequence of the relation $T_N(p,q) = T_N(q,p)$. The latter equality is an easy consequence of Theorem 2.1. By means of (7) we can write instead of the first equation of (6):

$$S_N(q) = \min_{r \in \mathcal{J}} \left\{ f(r) + q s_{N-1}(r/q) \right\} \quad (8)$$

Furthermore, it follows from section 4 that $\min_{q \in \mathcal{J}_N} S_N(q) = S_N$ (defined by (4.6)) and that the minimum is assumed for $q = p_N(1,\bar{\lambda})$, where $\bar{\lambda} := (\bar{r}_0, \ldots, \bar{r}_{N-1})$ and $\bar{r}_k (k = 0, \ldots, N-1)$ is defined in (4.13).

More properties of $S_N$ will be obtained in section 9. It will be shown there that $S_N$ is continuously differentiable in the interior of $\mathcal{J}_N$ for $N \geq 2$, $S_N$ is strictly convex, and that the graph of $S_N$ has
vertical slope at the endpoints of $J_N$. Furthermore, it will be shown that $S_N(q) = s^*(1 + q) (N + w)$ for every $q > 0$, where $s^*$ is given by (4.12).

II.9. An auxiliary optimization problem

$P_6(p,\rho,N)$: "Given $p > 0$, $\rho$ real and $N$ natural, find a $p$-start, $N$-step control for which $s_N := t_N + \rho p_N$ is minimal, where $t_N$ is the final time and $p_N$ the final alteration modulus."

Note that $P_6(p,0,N) = P_2(p,N)$. Again we have to consider only piecewise optimal control. Hence:

$P_6'(p,\rho,N)$: "Given $p > 0$, $\rho$ real, $N$ natural, find a $N$-step strategy $\lambda$ which minimizes $s_N(p,\rho,\lambda) := t_N(p,\lambda) + \rho p_N(p,\lambda)".

The existence of an optimal strategy is clear. From (3.22) it follows that for $\lambda = (r_0,\ldots, r_{N-1}) \in J^N$ we have:

$$\theta_N(p,\rho,\lambda) = \epsilon(r_0) p + \theta_{N-1}(r_0 p,\rho,\lambda). \tag{1}$$

If we introduce $\theta_N \in (R^1 \times R^1 \to R^1)$ by

$$\theta_N(p,\rho) := \min_{\lambda \in J^N} \theta_N(p,\rho,\lambda), \tag{2}$$

then the relation

$$\theta_N(p,\rho) = \min_{r \in J} [\epsilon(r) p + \theta_{N-1}(r p,\rho)] \tag{3}$$

can be derived from (1) in a way analogous to the way (4.4) is derived from (4.1). Now it is easily seen that $p = \theta_N(p,\rho)$ is homogeneous of degree 1. Hence, if we define $\theta_N \in (R^1 \to R^1)$ by

$$\theta_N(1,\rho) := \theta_N(1,\rho), \tag{4}$$

we have $\theta_N(p,\rho) = p \theta_N(\rho)$. It follows then that
\[ \phi_N(\sigma) = \nu(\phi_{N-1}(\sigma)), \]
\[ \phi_0(\rho) = \rho, \]

with \( \nu \) as defined in (4.7). By this recurrence relation \( \phi_n(\rho) \) can be calculated successively. Furthermore, \( \bar{T}_k \) can be found from the equalities
\[ \bar{T}_k = \omega_{N-1-k}(\rho) \quad (k = 0, \ldots, N-1) \]

where the functions \( \omega_k : (R^1 + J) \) are defined by
\[ \omega_k(\rho) = h(\phi_k(\rho)) \quad (k = 0, \ldots, N-1) \]

and \( h \) is defined in (4.9).

We see that the \( P_0^*(\rho, p, N) \)-optimal strategy (and hence the \( P_0(\rho, p, N) \)-optimal control) is unique. The optimal strategy as well as \( \phi_N \) depends continuously on \( \rho \). Furthermore, \( \phi_k \) is strictly increasing, \( \omega_k \) is non-increasing and \( \omega_k \) is strictly decreasing except
at the points \( \rho \) for which \( \Phi_k(\rho) \in [-1,-\beta,0] \). We will denote \( p_N(\rho,\Lambda) \) by \( \tau_N(\rho) \). Hence

\[
\tau_N(\rho) = p_N(1,\Lambda). 
\] (7a)

Then \( \tau_N \) is continuous. Furthermore, since \( \tau_N(\rho) = \omega_{N-1}(\rho) \cdots \omega_0(\rho) \) and since \(-1 \leq \Phi_k(\rho) \leq 0\) holds for at most one \( k \), it follows that \( \tau_N \) is strictly decreasing if \( N \geq 2 \).

**Remark 1.** Note that, since \( \pi_1(\rho) = \omega_0(\rho) = h(\rho) \), the function \( \tau_1 \) is not strictly decreasing (see fig. 4.2).

Finally it is obvious that we have

\[
\Phi_k(\rho) \to -\infty \quad (\rho \to -\infty), \quad \Phi_k(\rho) \to +\infty \quad (\rho \to +\infty), 
\] (8a)

\[
\omega_k(\rho) \to -1 \quad (\rho \to -\infty), \quad \omega_k(\rho) \to 1 \quad (\rho \to +\infty), 
\] (8b)

\[
\tau_N(\rho) \to +\infty \quad (\rho \to -\infty), \quad \tau_N(\rho) \to +\infty \quad (\rho \to +\infty). 
\] (8c)

Problem \( P_e \) is related to problems \( P_2, P_4, P_5 \). We will give a geometric interpretation of these problems which will make this relationship clearer. Because of homogeneity it is no loss of generality to assume that \( \rho = 1 \). We introduce the **attainable set**

\[
\Pi_{N} := \{ p_N(1,\Lambda), \tau_N(1,\Lambda) \mid \Lambda \in J^N \}. 
\] (9)
We have the following relations:
\[ J_N = \{ q \mid \exists c (q,t) \in \Pi_N \} , \]  
\[ S_N = \min \{ t \mid \exists q (q,t) \in \Pi_N \} , \]
\[ S_N(q) = \min \{ t \mid (q,t) \in \Pi_N \} , \]
\[ \phi_N(q) = \min \{ t + pq \mid (q,t) \in \Pi_N \} . \]

Here \( S_N \) is defined in (4.1), \( S_N(q) \) in (8.2) and \( J_N \) by (8.1).

Now we can solve the problems \( P_2, P_4, P_5 \) using the solutions of \( P_6 \).

I. As we have observed already, \( P_2(p,N) \) is equivalent to \( P_6(p,0,N) \).

II. We shall show that \( -\tilde{V}_N \) (defined in (7.2)) is the unique value of \( \rho \) for which \( \theta_N(\rho) = 0 \) holds, and that the \( P_6(1,N) \)-optimal strategy is the \( P_6(1,\rho,N) \)-optimal strategy with this value of \( \rho \). The existence and uniqueness of the zero of \( \theta_N \) is a consequence of (8a) and the monotonicity of \( \theta_N \). Let us denote the zero of \( \theta_N \) by \( \tilde{\rho} \) and let \( (\tilde{q}, \tilde{l}) \in \Pi_N \) be a point for which \( \theta_N(\tilde{l}) = \tilde{t} - \tilde{p} \tilde{q} \) (see formula (13)). Then for all \( (q,t) \in \Pi_N \) we have \( \tilde{t} - \tilde{p} \tilde{q} \leq t - p q \). Since \( \tilde{t} = \tilde{p} \tilde{q} = 0 \) it follows that \( \rho = -\tilde{t}/\tilde{q} \), and \( t + pq \geq 0 \) for all \( (t,q) \in \Pi_N \). Hence, \( t/q \geq -\rho = -\tilde{t}/\tilde{q} \) \( \forall (t,q) \in \Pi_N \), and this implies \( \tilde{V}_N = \tilde{v} \). Therefore, we have the following result for \( P_6 \): "The \( P_6 \)-optimal strategy \( \tilde{\theta} = (\tilde{r}_0, \ldots, \tilde{r}_{N-1}) \) satisfies the equations \( \tilde{r}_k = h(\phi_{N-k-1}) \), where \( \phi_k \) is determined by \( \phi_0 = 0, \phi_{k+1} = u(\phi_k) \) \( (k = N-1, \ldots, 0) \)." Note that \( \phi_k \) is uniquely determined by these relations, since \( u \) is strictly increasing.

REMARK 2. An argument, similar to the foregoing one, is used in [12].

Now we will prove that the result obtained here coincides with the one given in section 7. For that aim we express \( \phi_k \) in terms of \( \phi_{k+1} \) by means of the formula \( \mu(-u(-s)) = s \) (see (4.16)). Then it follows that \( \phi_k = \mu(-\phi_{k+1}) \) \( (k = 0,1,\ldots,N-1) \). Furthermore, introducing the notation \( \nu_k = -\phi_{N-k} \), we obtain \( \nu_0 = 0, \nu_{N-k} = h(\phi_{N-k}) \)
(k = N-1, ..., 0), \[ t_k = \frac{1}{\mu(v_k)} \]. Finally, using formula (4.17) we have: \( v_0 = 0, v_{k+1} = \mu(v_k) \) (k = 0, ..., N-1), \[ t_k = \frac{1}{\mu(v_k)} \], and this result is equivalent to (7.3) and (7.4).

III. Consider problem \( F_S(1, q, N) \) for \( N \geq 2 \). (Note that \( F_S(1, q, 1) \) is equivalent to \( F_1(1, q) \)). If \( q = \beta^N/2 \) or \( q = \beta^{-N/2} \) there exists only one N-step strategy \( \lambda \) for which \( p_{N,1}(1, \lambda) = q \) holds. This is necessarily the optimal strategy. We assume henceforth that \( \beta^{N/2} < q < \beta^{-N/2} \).

Then we will show that there exists a unique \( F_S^*(1, q, N) \)-optimal strategy and that this strategy is \( F_S^*(1, \rho, N) \)-optimal, where \( \rho \) is uniquely determined by the equation

\[
\tau_N(\rho) = q. \tag{14}
\]

The existence of a solution of (14) is a consequence of (8c) and the continuity of \( \tau_N \). The solution is unique since \( \tau_N \) is strictly decreasing. We denote the solution \( \rho \) of (14) by \( R_N(q) \). Hence:

\[
R_N(q) = \rho \iff \tau_N(\rho) = q \quad (\beta^{N/2} < q < \beta^{-N/2}) \tag{15}
\]

It follows that \( R_N \) is continuous and strictly decreasing. For a given \( q \), let the quantity \( \tilde{\rho} \) be defined by \( \tilde{\rho} = R_N(q) \) and let \( \tilde{\lambda} \) denote the \( F_S^*(1, \tilde{\rho}, N) \)-optimal strategy. Then \( p_{N,1}(1, \tilde{\lambda}) = \tau_N(\tilde{\rho}) = q \) (see (7a)).

Also, if \( \lambda = \tilde{\lambda} \) is an N-step strategy with \( p_{N,1}(1, \lambda) = q \), we have

\[
t_N(1, \lambda) = \theta_N(1, \tilde{\lambda}) - \tilde{\rho} \rho_{N,1}(1, \tilde{\lambda}) = \tau_N(\tilde{\rho}) - \tilde{\rho} q < \nu_N(1, \tilde{\lambda}) - \tilde{\rho} q = t_N(\rho, \lambda). \]

This proves that \( \lambda = \tilde{\lambda} \) is the unique \( F_S^*(1, q, N) \)-optimal strategy.

REMARK 3. It follows from the foregoing that we have the following formula

\[
\phi_N(R_N(q)) = \phi_N(q) + q R_N(q) \tag{15a}
\]

In order to find the \( F_S^*(1, q, N) \)-optimal strategy we have to solve equation (14). It seems impossible to find an explicit expression for
the solution. However, since $\tau_N$ decreases, the equation is easily solved by a numerical method (e.g. regula falsi).

Let us now derive some properties of $S_N(q)$ for $N \geq 2$. (Note that $S_1(q) = f(q)$.) Let

$$q_1, q_2 \in \text{int}(I_N) := \{ q \mid b^{N/2} < q < b^{-N/2} \}.$$ 

Since $(q_1, S_N(q_1)) \in I_N$, we have by (13): $S_N(q_1) + \rho q_1 \geq \theta_N(\rho)$ for every real $\rho$. Specializing to $\rho = R_N(q_2)$ and using (15a) for $q = q_2$ we obtain

$$S_N(q_1) \geq \Phi_N(R_N(q_2)) - q_2 R_N(q_2) + (q_2 - q_1) R_N(q_2) = S_N(q_2) + (q_2 - q_1) R_N(q_2).$$

Hence

$$S_N(q_1) - S_N(q_2) \geq -R_N(q_2) (q_1 - q_2).$$

Interchanging the role of $q_1$ and $q_2$ and combining the two results we obtain under the assumption $q_2 > q_1$:

$$-R_N(q_1) \leq \frac{S_N(q_2) - S_N(q_1)}{q_2 - q_1} \leq -R_N(q_2).$$

By the continuity of $R_N$ this implies

$$S'_N(q) = -R_N(q). \tag{16}$$

As a consequence we see that $S_N$ is strictly convex.

REMARK 4. It is surprising that $S_N$ is continuously differentiable for $N \geq 2$, whereas $S_1 - f$ is not differentiable at $q = 1$. This circumstance is due to the fact that $\tau_N$ is strictly decreasing for $N \geq 2$ but not strictly decreasing if $N = 1$ (see Remark 1).

Now we will derive some properties of $\Phi_N$. Since $\nu \in C^1$ it follows from (5) that $\Phi_N \in C^1 (k = 0, \ldots, N)$ and from $\nu'(s) = h(s)$ (formula (4.15)) we infer
\[ \Phi_k^*(\rho) = 1, \quad \Phi_{k+1}(\rho) = \omega_k(\rho) \Phi_k^*(\rho). \]  
\text{(17)}

Especially it follows that
\[ \Phi_k^*(\rho) = \pi_N(\rho). \]  
\text{(18)}

Since for \( N \geq 2 \) the function \( \pi_N \) is strictly decreasing we conclude that \( \Phi_N \) is strictly concave. Furthermore, the following relation can be derived from (12) and (13):
\[ \Phi_N(\rho) = \min_{q \in J_N} (S_N(q) + \rho q). \]  
\text{(19)}

In fact, if we define \( \delta_N(\rho) := \min_{q \in J_N} (S_N(q) + \rho q) \), we have
\[ \Phi_N(\rho) = \min \{ t + \rho q \mid (q, t) \in \Pi_N \} \leq S_N(q) + \rho q \]

for all \( q \in J_N \), since \( (q, S_N(q)) \in \Pi_N \). Hence, \( \Phi_N(\rho) \leq \delta_N(\rho) \). On the other hand, if \( (\tilde{q}, \tilde{t}) \in \Pi_N \) satisfies \( \tilde{t} + \rho \tilde{q} = \delta_N(\rho) \), we have
\[ \Phi_N(\rho) = \tilde{t} + \rho \tilde{q} \geq S_N(\tilde{q}) + \rho \tilde{q} \geq \delta_N(\rho). \]

This proves (19).

Since \( \delta_N \) is continuous and strictly convex, and since \( S_N(q) = qS_N(1/q) \) (see formula (8.7)), we can use Remark (4.2) to obtain
\[ \Phi_N(-\delta_N(-\rho)) = \rho, \]  
\text{(20)}

(that is, the graph of \( \Phi_N \) is symmetric with respect to the line \( \Phi = \rho \)), and furthermore
\[ \pi_N(-\delta_N(\rho)) = 1/\pi_N(\rho). \]  
\text{(21)}

REMARK 5. Equation (20) can also be derived by induction from (4.16). In fact, \( \phi_1 = \mu \) and \( \Phi_1(\mu) = \mu(\Phi_{N-1}(\mu)) = \phi_{N-1}(\mu(\rho)) \). Hence (4.16) is equivalent to (20) for \( N = 1 \). Assuming that (20) is proved with \( N-1 \) instead of \( N \), we have
\[ \Phi_N(-\delta_N(-\rho)) = \mu(\Phi_{N-1}(-\delta_{N-1}(\mu(-\rho))) = \mu(-\mu(-\rho)) = \rho. \]
We now derive some asymptotic formulas for the functions discussed in this section.

A. Suppose that $N$ is fixed. Then we derive asymptotic formulas for $\varphi_N(\rho)$, $\omega_N(\rho)$, and $\eta_N(\rho)$ for $\rho \to +\infty$ (a first estimate is given in (8)). In (4.11) we have seen that $\nu(\varepsilon) = \varepsilon^{\frac{1}{2}}s + a_1 + O(s^{-1})$ ($s \to \infty$), where

$$a_1 := \frac{1}{2}(1 + \beta)(1 + \beta^{-\frac{1}{2}}) = f(\varepsilon^\frac{1}{2}) .$$

It is easily shown by induction that this implies

$$a_N := \frac{1 - \beta^{N/2}}{1 - \beta^{-\frac{1}{2}}} , a_1 = S_N(\beta^{N/2}) ,$$

and similarly

$$a_N := \frac{1 - \beta^{-N/2}}{1 - \beta^{-\frac{1}{2}}} , a_1 = S_N(\beta^{-N/2}) ,$$

where

$$b_N := \frac{\beta^{-N/2} - 1}{1 - \beta^{-\frac{1}{2}}} , b_1 = S_N(\beta^{-N/2}) ,$$

and

$$b_1 := (2\beta^{-1})(1 + \beta)(1 + \beta^{-\frac{1}{2}}) = f(\varepsilon^{-\frac{1}{2}}) .$$

The asymptotic formulas for $\omega_k(\rho)$ follow from remark (4.1): We have

$$\omega_1(\varepsilon) = \varepsilon^{\frac{1}{2}}s + c_2 + O(s^{-\frac{3}{2}})$$

where

$$c_2 := \frac{1}{3} (1 - \beta)(1 + \beta) \beta^{-3/2} .$$

Therefore the following formulas obtain for $\omega_k(\rho)$:
\[ w_k(\rho) = \gamma^\frac{1}{2} + c\gamma^{-k}\rho^{-2} + \mathcal{O}(\rho^{-3}) \quad (\rho \to \infty), \] (28)

and similarly
\[ w_k(\rho) = \gamma^\frac{1}{2} - c\gamma^k\rho^{-2} + \mathcal{O}(\rho^{-3}) \quad (\rho \to \infty). \] (29)

Finally, from \( w_N(\rho) = \cdots w_1(\rho) \) it follows that
\[ v_N(\rho) = \gamma^{N/2} + d_N\rho^{-2} + \mathcal{O}(\rho^{-3}) \quad (\rho \to \infty), \] (30)

and
\[ v_N(\rho) = \gamma^{-N/2} - d_N\rho^{-2} + \mathcal{O}(\rho^{-3}) \quad (\rho \to \infty), \] (31)

where
\[ d_N := \frac{1}{8}(1 + \beta)^2\beta^{-1-N/2}(1 - \beta^N). \] (32)

B. Again, we assume that \( N \) is fixed and now we derive formulas for \( S_N(q) \) and \( R_N(q) \) for \( q = \beta^{N/2} \) (\( R_N \) is defined by (15a) and \( S_N \) is defined by (8.5)). Denoting \( (q - \beta^{N/2})/\varepsilon \) by \( \varepsilon \) and \( R_N(q) \) by \( \varepsilon \) we have
\[ q = v_N(\rho) = \gamma^{N/2} + d_N\rho^{-2} + \mathcal{O}(\rho^{-3}) \quad (q \to \beta^{N/2}), \]

since \( \rho \to \infty \) if \( q \to \beta^{N/2} \). Hence, \( \varepsilon = \rho^{-2} = \mathcal{O}(\rho^{-3}) \) \((\varepsilon \to 0)\), which yields \( \rho = \varepsilon^{-\frac{1}{2}} + \mathcal{O}(1) \) \((\varepsilon \to 0)\). Thus, we have found
\[ R_N(q) = d_N^{-\frac{1}{2}}(q - \beta^{N/2})^{-\frac{1}{2}} + \mathcal{O}(1) \quad (q + \beta^{N/2}), \] (33)

and analogously
\[ R_N(q) = -d_N^{-\frac{1}{2}}(\beta^{-N/2} - q)^{-\frac{1}{2}} + \mathcal{O}(1) \quad (q + \beta^{-N/2}). \] (34)

Formulas for \( S_N(q) \) can be obtained by quadrature using (23) and (25). This yields
\[ S_N(q) = a_N - 2d_N^{-\frac{1}{2}}(q - \beta^{N/2})^{\frac{1}{2}} + \mathcal{O}(q - \beta^{N/2}) \quad (q + \beta^{N/2}), \] (35)

and
\[ S_N(q) = b_N - 2d_N \left( \beta^{N/2} - q \right)^{1/2} + O(\beta^{-N/2} - q) \quad (q + \beta^{-N/2}) . \] (36)

These formulas imply in particular that the graph of \( S_N \) has vertical slope at the endpoints.

C. Now let \( \phi \in (s^*, s^*) \) be fixed, let \( N \to \infty \), and consider \( \Phi_N(\phi), \omega_N(\phi), \tau_N(\phi) \) (\( s^* \) is defined by (4.12)). Then we get (cf. [13] Ch. 8)

\[ \Phi_N(\phi) = s^* + O((r^*)^N) \quad (N \to \infty) , \] (37)

\[ \omega_N(\phi) = r^* + O((r^*)^N) \quad (N \to \infty) , \] (38)

\[ \tau_N(\phi) = C_N(r^*)^N(1 + O((r^*)^N)) \quad (N \to \infty) . \] (39)

Here \( r^* \) is defined by (5.10): \( r^* = h(s^*) = \mu'(s^*) \) and \( C_N \) is some constant not depending on \( N \). The last formula is obtained from

\[
\log \tau_N(\phi) = \sum_{k=0}^{N-1} \log \omega_k(\phi) = N \log r^* + \sum_{k=0}^{N-1} O((r^*)^k) = \\
= N \log r^* + C_1 + O((r^*)^N)
\]
where \( C_1 \) does not depend on \( N \).

D. Let \( q > 0 \) be fixed and \( N \to \infty \), and consider \( R_N(q), S_N(q) \). We first observe that \( -s^* < R_N(q) < b_N(R_N(q)) < s^* \) holds for sufficiently large \( N \). In fact, if \( R_N(q) \leq -s^* \), then \( \omega_N(R_N(q)) \leq -s^* \) (\( k = 0, \ldots, N-1 \)), and hence \( \omega_k(R_N(q)) \geq (r^*)^{-k} (k = 0, \ldots, N-1) \). But then we have \( q = \tau_N(R_N(q)) \leq (r^*)^{-N} \) which is impossible for sufficiently large \( N \). In a similar way we get a contradiction from the assumption that \( R_N(q) \geq s^* \). Hence, \( -s^* < R_N(q) < s^* \). But then it follows from the monotonicity of \( \mu \) and from \( \mu(s^*) = s^* \), that \( R_N(q) < a_N(R_N(q)) \leq s^* \).

Furthermore, we will show that

\[ R_N(q) \to -s^* \quad (N \to \infty) , \] (40)

and
Suppose that (40) does not hold. Then there exists \( \delta > 0 \) with \( \delta < 2s^* \) and an increasing sequence \( (N_1, N_2, \ldots) \) such that \( R_{N_k}(q) \geq -s^* + \delta \) \((k = 1, 2, \ldots)\). But then we have
\[
q = \tau_{N_k}(-s^* + \delta) \rightarrow 0 \quad (k \rightarrow \infty)
\]
(see (39)) which is a contradiction. We also have \( \Phi_N(R_N(q)) = s^* \). If this were not the case, there would be \( \delta > 0 \) and strictly increasing sequence \( (N_1, N_2, \ldots) \) such that \( \Phi_N(R_N(q)) \leq s^* - \delta \). But then it would follow from (21) that
\[
q = \tau_{N_k}(-s^* + \delta) \geq 1/\tau_{N_k}(-s^* + \delta) = \infty \quad (k \rightarrow \infty)
\]
which is again contradictory. Formula (41) now follows from (15b).

REMARK 6. Formula (41) can be made intuitive in the following way. If for fixed \( q \) and very large \( N \) one has to find an \( N \)-step trajectory \( x_u \) with \( x_u(r_N) = q \), the final time \( r_N \) will be reasonable small if \( |x_u(t)| \) is small on a large number of steps. Therefore the optimal trajectory will have the following nature: First we go from \((0,1)\) to a point very close to the origin. By the result of section 5 we can do this in a time interval of approximate length \( s^* \). Having done this, we go back to \((0,q)\). This takes us as much time as from \((0,q)\) to \((0,0)\), which can be done in a time interval of approximate length \( qs^* \) (solution of problem \( P_3 \) with reversed time). Therefore the total time is about \( s^*(1 + q) \).

REMARK 7. By the symmetry properties of the problem it is possible to calculate \( S_N(1) \) and \( R_N(1) \) explicitly. In fact, we have: If \( N = 2K \), then
\[
R_N(1) = -\Phi_K(0), \quad S_N(1) = 2\Phi_K(0).
\]
If \( N = 2K + 1 \), then
\[
R_N(1) = -\Phi_K(0), \quad S_N(1) = 2\Phi_K(0).
\]
where \( \sigma = \frac{1}{2}(1 + \tilde{\rho}) \).

**Proof.** Let \( \tilde{\sigma} := R_N(1) \), hence \( \mu(\tilde{\sigma}) = \frac{1}{2}. \) Because of (21) we have
\[
\pi_N(-\pi_N(\tilde{\sigma})) = 1, \quad \text{so} \quad -\pi_N(\tilde{\sigma}) \quad \text{is also a solution of (14). But the solution of (14) is unique and therefore we must have}
\]
\[
\sigma_N(\tilde{\sigma}) = -\tilde{\sigma}.
\] (44)

Since \( \sigma_N \) is strictly increasing, \( \tilde{\rho} \) is uniquely determined by (44).

Furthermore, by (15a) we have
\[
S_N(1) = 2\tilde{\sigma} = -2R_N(1).
\] (45)

If \( N \) is even, say \( N = 2K \), we have
\[
\phi_K(\mu(\tilde{\sigma})) = -\mu(\tilde{\sigma}) = -\phi_K(\mu(\tilde{\sigma}))
\]
(see (20)), which implies \( \phi_K(\tilde{\rho}) = -\phi_K(\tilde{\rho}) \), and hence \( \phi_K(\tilde{\rho}) = 0. \)

Again by (20) we obtain \( \tilde{\rho} = -\phi_K(-\phi_K(\tilde{\rho})) = -\phi_K(0) \). Together with (45) this implies (42).

If \( N \) is odd, say \( N = 2K + 1 \), define \( \sigma := -\phi_K(\tilde{\sigma}) \). Then we have
\[
\tilde{\rho} = -\phi_K(-\phi_K(\tilde{\rho})) = -\phi_K(\sigma).
\]
and furthermore
\[
\phi_K(\mu(\sigma)) = \phi_{2K+1}(\sigma) = \phi_{2K+1}(\tilde{\rho}) = -\tilde{\rho} = \phi_K(\sigma).
\]

Hence \( \mu(\sigma) = \sigma \). It is easily seen that this implies \( \sigma = \frac{1}{2}(1 + \tilde{\rho}) \)
(see formula (4.10) and fig. 4.3). This proves (43).

Note that the relation (45) also can be proved by differentiating (8.7) with respect to \( q \), using (16) and setting \( q = 1. \)

Finally we remark that the formulas (42) and (43) can be derived from Theorem 2.1 and the observation that for the \( F_k(1, N) \)-optimal strategy \( \tilde{N} \) we have \( \tilde{N}(T-t) = \tilde{N}(t) \).
II.10. The case $a = 1$

We consider in this section the problems $P_1, \ldots, P_6$ for the case in which $a = 1$.

$P_1(p,q)$: The analysis here is very similar to the one of section 3. Therefore we will indicate only the results, with special emphasis on the points where these results deviate from the ones in which $a < 1$. Just as in section 3, the optimal control is unique and of the "bang-bang" type, taking on only the values $\pm 1$. But the parabolas $\bar{v}(-1,p)$ and $\bar{v}(1,q)$ (for the notation see section 3) degenerate to straight lines parallel to the $x$-axis. If $q = rp$, the optimal control exists for $r \in J := (0,\infty)$. The formulas for the switching point $(\bar{x}, \bar{y})$ and the function $f(r)$ can be obtained by taking the limit $a \to 1$ (or $b \to 0$) or by direct calculation.

Thus, we obtain: If $r \neq 1$, we have $(\bar{x}, \bar{y}) = (\{1-r^2\}p^2, -rp)$ and $f(r) = (4r(1+r^2))$. If $r = 1$, then $(\bar{x}, \bar{y}) = (\{1-r^2\}p^2, rp)$ and $f(r) = \log(1+r)^2$. The minimum time here is also given by $T(p,q) = f(r)p$. The function $f$ is strictly convex and $f(r)/r \to \infty$ ($r \to 0$), $f(r) \to \infty$ ($r \to \infty$).

The concepts "$p$-start, $N$-step" and "piecewise optimal" and the functions $P_N, P_N$ are defined like in section 3. For problems $P_2, \ldots, P_6$ we can restrict ourselves to piecewise optimal controls. However, since $J$ is no longer compact, the existence of optimal strategies is not obvious. Let us start with problem $P_6$. We will...
proceed just as in section 5. First we assume the existence of an 
optimal strategy. Defining \( \theta_{N}, \theta_{N}, \phi_{N}, \omega_{N}, \pi_{N} \) just as in section 9, 
with \( u \) and \( h \) defined as in section 4 (it is easily seen that these 
functions exist), we can introduce the strategy \( \tilde{\lambda}_{N} = (\tilde{r}_{0}, \ldots, \tilde{r}_{N-1}) \), 
with \( \tilde{r}_{k} \) as defined in (9.6). We show by induction that \( \tilde{\lambda}_{N} \) actually 
constitutes the unique \( P_{0}(p, \omega, N) \)-optimal strategy. For \( N = 1 \), we have 
\[
\theta_{1}(\rho) = \min_{r \in J} [f(r) + r \rho] = f(\tilde{r}_{0}) + \tilde{r}_{0} \rho \leq f(r) + rp 
\]
for all \( r \in J \) with equality only if \( r = \tilde{r}_{0} \). Hence, 
\[
\xi_{1}(p, \rho, \tilde{\lambda}_{1}) = \xi_{1}(p, \tilde{\lambda}_{1}) + \rho p_{1}(p, \tilde{\lambda}_{1}) = (f(\tilde{r}_{0}) + \tilde{r}_{0} \rho)p \leq (f(r_{0}) + r_{0} \rho)p = \theta_{1}(p, \rho, \Lambda) 
\]
for all \( \Lambda = (r_{0}) \in J_{1} \), with equality only if \( \Lambda = \tilde{\lambda}_{1} \). Now consider an 
arbitrary \( N \), and assume that \( \tilde{\lambda}_{N-1} \) has been shown to be \( P_{0}(p, \omega, N-1) \)- 
optimal for all \( p > 0 \). Then we observe from (9.6) that if \( \tilde{\lambda}_{N} = 
(\tilde{r}_{0}, \ldots, \tilde{r}_{N}) \), the strategy \( \tilde{\lambda}_{N} = (\tilde{r}_{0}, \ldots, \tilde{r}_{N}) \) equals \( \tilde{\lambda}_{N-1} \). This im-
plies: 
\[
\theta_{N}(p, \rho, \tilde{\lambda}_{N}) = f(\tilde{r}_{0})p + \tilde{r}_{0} \theta_{N-1}(p, \rho, \tilde{\lambda}_{N-1}) \leq 
(f(r)p + r \theta_{N-1}(p, \rho, \Lambda)) \leq f(r)p + r \theta_{N-1}(p, \rho, \Lambda) 
\]
for all \( r \in J \) and all \( \Lambda \in J_{N-1} \), with equality only if \( r = \tilde{r}_{0} \) and 
\( \Lambda = \tilde{\lambda}_{N-1} \). The first inequality follows from the definition of \( \tilde{r}_{0} \) 
(since \( \theta_{N-1}(p, \rho, \tilde{\lambda}_{N-1}) = \theta_{N-1}(p, \rho) = p \theta_{N}(p) \)), and the second in-
equality is a consequence of the strict optimality of \( \tilde{\lambda}_{N-1} \). Hence, \( \tilde{\lambda}_{N} \) is 
\( P_{0}(p, \omega, N) \)-optimal.

Problem \( P_{0} \) is a special case of \( P_{0} \). As we saw in section 9 the 
solutions of problems \( P_{0} \) and \( P_{0} \) can be constructed from the one of 
\( P_{0} \) without presupposing the existence of \( P_{0} \)- and \( P_{0} \)-optimal strate-
gies. Problem \( P_{3} \) can be dealt with exactly as in section 5.

We give some formulas for \( a = 1 \):
\[ u(s) = \begin{cases} \frac{1}{4}(1 + \sqrt{1 + 4s}) & (s > 0), \\ 1 + s & (-1 \leq s \leq 0), \\ -s(1 + s) & (s < -1), \end{cases} \]

and

\[ h(s) = \begin{cases} (1 + 4s)^{-\frac{1}{2}} & (s > 0), \\ 1 & (-1 \leq s \leq 0), \\ -1 - 2s & (s < -1). \end{cases} \]

Furthermore we have \( a_n = 2, \ t_n = \frac{1}{3} \) and the switching curve of problem \( P_{\rho} \) is given by \( x = 2y|y| \).

The asymptotic formulas given in section 9 have to be modified here. After some calculations one finds

\[ S_n(q) = a_n q - b_n N + O(1) \quad (q \to 0), \]

\[ S_n(q) = c_n q - d_n N + O(q) \quad (q \to -\infty), \]

where

\[ a_n = (2^N - 1)^{-1} \cdot 2^{-N/(2^N - 1)}, \]

\[ b_n = (2^N - 1)^{-1}, \]

\[ c_n = (2^N - 1)^{-1} \cdot 2^{N/(2^N - 1)}, \]

\[ d_n = 2^N (2^N - 1)^{-1}. \]

Furthermore, formula (9.41):

\[ S_n(q) = 2(1 + q) \quad (N = \infty) \]

is also valid in this case.
CHAPTER III
APPLICATION OF THE DISCRETE MAXIMUM PRINCIPLE

III.1. The discrete maximum principle

Although the dynamic programming method used in the previous chapter is an elementary method, it usually results in a functional-recurrence relation which is not easily solved explicitly. The special character of the problems of the previous chapter which made it possible to apply successfully the dynamic programming method, is the homogeneous dependence on p of the functions $p_k$ and $q_k$ defined in (3.19), (3.20). But even for this simple situation we have seen in section II.8 an instance of the difficulties one encounters in the dynamic programming method. The equation (II.8.6) is a typical example of such a functional-recurrence relation. Fortunately, we could solve the problem of section 8 by an auxiliary problem discussed in section 9. However, for more general situations such an artifice usually is not possible, and we are left with an equation of the type (II.8.6).

In this chapter we discuss a different method for solving discrete optimal control problems which is based on a rather complicated theorem, called the discrete maximum principle. This theorem is the discrete analogon of the well-known maximum principle of Pontryagin for continuous (that is, differential) systems (sometimes called the continuous maximum principle), a special case of which we have met in section 1.5. In its general form the discrete maximum principle is due to H. W.altik (see [3]). A simpler proof can be found in [14].

DEFINITION 1. If N is a natural number, then

$$\hat{N} = \{0, \ldots, N-1\}.$$ (1)
III.1

The set of natural numbers is denoted by Nt.

DEFINITION 2. If N \in Nt, U \subset R^m for some positive integer m, then

\begin{align*}
H(R^n \times U \times \hat{N} \rightarrow R^p) := &\ \\
:= \{f \in C(R^n \times U \times \hat{N} \rightarrow R^p) \mid (\pi \mapsto f(\pi,v,k)) \in C^2(R^n \rightarrow R^p) \}
(v \in U, k \in \hat{N}). &
\end{align*}

Now suppose we are given a vector a \in R^n, a set U \subset R^m, an

element N \in Nt and the following functions

\begin{align*}
\alpha) &
\quad f \in H(R^n \times U \times \hat{N} \rightarrow R^p), \\
\beta) &
\quad g \in C^2(R^n \rightarrow R^p), \\
\gamma) &
\quad w \in C^2(R^n \rightarrow R^1).
\end{align*}

Then we will call the elements of \hat{N} \rightarrow U controls. For a given control

u, the function \pi_u \in (\hat{N} \rightarrow R^n), defined by

\begin{align*}
\pi_u(0) = a, \\
\pi_u(k+1) = f(\pi_u(k), u(k), k) \quad (k \in \hat{N}),
\end{align*}

is called the response of u. A control u is called admissible if

\[ g(\pi_u(N)) = 0. \]

The set of admissible controls is denoted by \hat{N}:

\[ \hat{N} := \{u \in \hat{N} \rightarrow U \mid g(\pi_u(N)) = 0\}. \]

An admissible control \hat{u} is called optimal if for every u \in \hat{N} we have

w(\pi_{u}(N)) \geq w(\pi_{\hat{u}}(N)). \]

Hence, an optimal control minimizes w(\pi_{u}(N)) under the restriction g(\pi_{u}(N)) = 0. With these definitions we have the following theorem:

THEOREM 1 (Halkin). Let \hat{u} be an optimal control and let \pi_{\hat{u}} := \pi_{u,\hat{u}} be the corresponding response. If for every \pi \in R^n, k \in \hat{N} the set

\[ f(\pi, u, k) := \{f(\pi, v, k) \mid v \in U\} \]

is convex and if
\[
\text{rank } \begin{bmatrix} \delta g(\tilde{x}(N)) \\ \delta w(\tilde{x}(N)) \end{bmatrix} = p + 1, \tag{4}
\]

then there exists a non-trivial \( \psi \in (\tilde{N}+1 - \mathbb{R}_+^m) \) with the properties

i) \( \psi(k) = \psi(k+1) \tilde{f}(\tilde{x}(k), \tilde{u}(k), k) \quad (k \in \tilde{N}) \),

ii) \( \psi(N) = \lambda \tilde{w}(\tilde{x}(N)) + \mu \delta g(\tilde{x}(N)) \) for some \( \lambda \geq 0, \mu \in \mathbb{R}_+^m \),

iii) \( \psi(k+1) \tilde{f}(\tilde{x}(k), \tilde{u}(k), k) = \min_{v \in U} \psi(k+1) \tilde{f}(\tilde{x}(k), v, k) \quad (k \in \tilde{N}) \).

REMARK 1. In [3] the additional condition \( \tilde{\delta} \tilde{f}(\tilde{x}(k), \tilde{u}(k), k) \) is non-singular" is posed, but as is shown in [14] this condition is superfluous. Note that also we have formulated the problem as a minimum problem, whereas in [3] the problem is considered where a function has to be maximized. Of course these problems are equivalent.

We will derive a modified version of Theorem 1 which will be useful for the problems discussed in this chapter. Suppose that we are given functions \( \tilde{f} \) and \( g \) just as in a) and b), but that instead of \( \tilde{w} \) we have a function \( f_0 \in H(\tilde{N}^0 \times U \times \tilde{N} \rightarrow \mathbb{R}) \). We define the response \( x_u \) and the set of admissible controls \( U \) as before, but now we call \( \tilde{u} \in U \text{ optimal if} \)

\[
\sum_{k=0}^{N-1} f_0(x_u(k), \tilde{u}(k), k) \leq \sum_{k=0}^{N-1} f_0(x_u(k), u(k), k) \tag{5}
\]

holds for all \( u \in U \). Then we have:

THEOREM 2. Let \( \tilde{u} \) be an optimal control, and let \( \tilde{x} := x_{\tilde{u}} \) be the response of \( u \). If for every \( x \in \mathbb{R}^n, k \in \tilde{N} \), the set

\[
V(x,k) := \{(z_0,z') \in \mathbb{R}^{n+1} \mid \exists z = f(x,v,k) \land z_0 \geq f_0(x,v,k) \land v \in U \}
\]

is convex, and if \( \text{rank } 3g(\tilde{x}(N)) = p \), then there exists \( \lambda \geq 0 \) and \( \tilde{u} \in (\tilde{N}+1 - \mathbb{R}_+^m) \) with the properties:
i) \((\lambda, \psi(N)) = 0\),

ii) \(\psi(k) = \psi(k+1) + \lambda f_0(x(k), u(k), k) + \lambda \delta \phi_0(x(k), u(k), k)\) \(k \in \bar{N}\),

iii) \(\psi(N) = \mu g(x(N))\) for some \(\mu \in \mathbb{R}^n\),

iv) \(H(\tilde{u}(k), k) = \min_{v \in \tilde{U}} H(v, k)\) \(k \in \bar{N}\),

where \(H \subseteq (\mathbb{R} \times \mathbb{R} \times \mathbb{R}^n)\) is defined by

\[
H(v, k) := \psi(k+1) f(x(k), v, k) + \lambda f_0(x(k), v, k).
\]

PROOF. We introduce the following notation:

\[
\underline{K} := \begin{bmatrix} x_0 \\ x \end{bmatrix} \subseteq \mathbb{R}^{n+1}, \text{ where } x_0 \in \mathbb{R}^1, x \in \mathbb{R}^n;
\]

\[
\underline{u} := \begin{bmatrix} u_0 \\ u \end{bmatrix} \subseteq \mathbb{R}^{m+1}, \text{ where } u_0 \in \mathbb{R}, u \in \mathbb{R}^m;
\]

\[
\underline{U} := \{ u \in \mathbb{R}^{m+1} \mid u \in \bar{U} \wedge u_0 \geq 0 \};
\]

\[
\bar{U} := \{ u \in \bar{U} \mid g(x_0(N)) = 0 \}.
\]

\(f \subseteq (\mathbb{R}^{n+1} \times \bar{U} \times \bar{U} \times \mathbb{R}^{n+1})\) is defined by

\[
f(x, u, k) := \begin{bmatrix} x_0 + f_0(x, u, k) + u_0 \\ f(x, u, k) \end{bmatrix};
\]

\[
\underline{w}(x) := x_0; \quad \underline{a} := \begin{bmatrix} 0 \\ a \end{bmatrix}; \quad \underline{g}(x) := g(x).
\]

Then we consider the optimal control problem of theorem 1, with

\(n+1, m+1, \underline{K}, \underline{U}, \underline{f}, \underline{g}, \underline{w}, \underline{a}, \underline{w}, \underline{a}, \underline{U}, \underline{g}\) instead of \(n, m, x, u, f, g, w, a, U, G\). It is easily seen that \(\underline{f}, \underline{g}, \underline{w}\) satisfy the conditions \(\alpha, \beta, \gamma\) and \(\psi\). Also we observe that \(\underline{u} := \begin{bmatrix} 0 \\ u \end{bmatrix}\) is an optimal control. In fact, if we denote \(x_0\) by \(\bar{x}\), then \(g(\bar{x}(N)) = g(\bar{x}(N)) = 0\). And for
every \( u = \begin{pmatrix} u_0 \\ u \end{pmatrix} \in \mathcal{U} \) we have

\[
\mathcal{V}(\tilde{u}(N)) = \sum_{k=0}^{N-1} \left( f_0(\tilde{x}(k),\tilde{u}(k),k) + u_0(k) \right) \geq \\
= \sum_{k=0}^{N-1} f_0(\tilde{x}(k),\tilde{u}(k),k) = \mathcal{V}(\tilde{u}(N)),
\]

since \( \tilde{u} \) is optimal and \( u_0(k) \geq 0 \). Furthermore, we see that the conditions of Theorem 1 are satisfied. In fact

\[
\text{rank} \begin{pmatrix} \tilde{g}(\tilde{x}(N)) \\ \tilde{g}(\tilde{x}(N)) \end{pmatrix} = \text{rank} \begin{pmatrix} 0 & \tilde{g}(\tilde{x}(N)) \\ 1 & 0 \end{pmatrix} = p + 1,
\]

and \( f(\tilde{x},u,k) = \begin{pmatrix} x_0 \\ 0 \end{pmatrix} + \tilde{V}(x,k) \) is convex for every \( (\tilde{x},k) \in \mathbb{R}^{n+1} \times \tilde{N} \).

It follows then from Theorem 1 that there exists a non-trivial \( \tilde{x} \in (\tilde{N}^+ \times \mathbb{R}_T^p) \), which (if we set \( \tilde{x} = (\tilde{x},\psi) \)) satisfies:

1) \( \psi_0(k) = \psi_0(k+1) \),
   \( \psi(k) = \psi((k+1)) + f(\tilde{x}(k),\tilde{u}(k),k) + \psi_0(k+1) \mathcal{X} f_0(\tilde{x}(k),\tilde{u}(k),k) \quad (k \in \tilde{N}) \);

2) \( \psi_0(N) = \lambda, \psi(N) = \nu g(\tilde{x}(N)) \) for some \( \lambda \geq 0, \nu \in \mathbb{R}^p_+ \);

3) \( \psi(k+1) f(\tilde{x}(k),\tilde{u}(k),k) + \psi_0(k+1) f_0(\tilde{x}(k),\tilde{u}(k),k) \)
   \( = \min_{\nabla \in U} \psi(k+1) f(\tilde{x}(k),\nabla,k) + \psi_0(k+1) f_0(\tilde{x}(k),\nabla,k) \),

and this yields the result.

**Remark 2.** If there is no restriction on the final value of \( x_u \) (that is, if \( g \) does not appear in theorem 2, and hence all \( u \in (\tilde{N} + \mathcal{U}) \) are admissible), then \( p = 0 \) and the rank condition is satisfied. In this case \( i1) \) has to be interpreted as \( \psi(N) = 0 \). It follows then from \( i) \) that \( \lambda > 0 \). Since \( (\lambda,\psi) \) appears homogeneous in theorem 2 we may assume in this case that \( \lambda = 1 \).
III.2. Application of the maximum principle to problems of Chapter II

We apply Theorem (1.2) to the problems $F_2'$ and $F_3'$ treated in chapter II.

First consider problem $F_2'$. In the notation used in the formulation of Theorem (1.2) we have as a special case of the recurrence (1.2):

$$
x_u(0) = p > 0 ,

x_u(k + 1) = u(k)x_u(k) \quad (k \in \hat{N}) ,
$$

where $x$ and $u$ are scalar valued ($n = m = 1$) and $U = J = \{ \frac{1}{2}, i \}$. The function $u$ should not be confused with the control function $u$ of problem $F_2$. Here $u$ is a function defined on $\hat{N}$ instead of on an interval, and we have $u(k) = r_k$ ($k = 0, ..., N+1$), with $r_k$ as defined in problem $F_2'$. Furthermore we have $f_y(x, u, k) = f(u)x$. Since $f$ is convex and $\delta g(x(N)) = 0$ (cf. remark (1.2)), we can apply Theorem (1.2) and set $\lambda = 1$ in order to find necessary conditions for the optimal control: If $\tilde{u}$ is optimal then exists $\psi \in (\hat{N} + 1) \to \hat{R}^1$ such that

1) $\psi(k) = \psi(k+1)\tilde{u}(k) + f(\tilde{u}(k))$ \quad $(k \in \hat{N})$,

2) $\psi(N) = 0$,

3) $\psi(k+1)\tilde{u}(k) + f(\tilde{u}(k))\tilde{x}(k) = 

   \min_{\nu \in \hat{U}} \{ \psi(k+1)\tilde{x}(k) + f(\nu)\tilde{x}(k) \}$ \quad $(k \in \hat{N})$.

Since $p > 0$ and $u(k) > 0$ $(k \in \hat{N})$, we see that $x(k) > 0$ $(k \in \hat{N}+1)$. Hence, if we use the function $u$ defined in (II.4.7), we have:

1) $\psi(N) = 0$, $\psi(k) = u(\psi(k+1))$ \quad $(k \in \hat{N})$;

2) $\psi(k+1)\tilde{u}(k) + f(\tilde{u}(k)) = u(\psi(k+1))$.

It follows from ii) that $\tilde{u}(k) = h(\psi(k+1))$ where $h$ is the function defined in (II.4.9). Put $\psi(k) = S_{\psi-k}$ $(k \in \hat{N})$, where $S_k$ is defined
by (II.4.3); then we have obtained the same result as in section (II.4) if we have shown that

\[ T_N(a) := \sum_{k=0}^{N-1} f(\bar{u}(k)) \bar{x}(k) = \psi(0)a. \]

But this is an immediate consequence of the equalities

\[ \psi(k) \bar{u}(k) = (\psi(k+1) \bar{u}(k) + f(\bar{u}(k))) \bar{x}(k) = \psi(k+1) \bar{x}(k+1) + f(\bar{u}(k)) \bar{x}(k) \]

and

\[ \psi(N) = 0, \quad \psi(0) = a. \]

Problem \( P_5 \) differs from problem \( P_2 \) only by the condition \( x(N) = q \), where \( b \) is some positive number. Hence, we have \( g(x) = x - q \) (in the notation of Theorem (1.2)). Since rank \( 2g(x) = 1 \), we can apply Theorem (1.2) again. Let us assume for simplicity that \( p = 1 \). This is no loss of generality because of the homogeneity of the problem. Furthermore, just as in section (II.8) we observe that there exists an admissible control if and only if \( b \in J := [\beta^N/2, \beta^{N/2}] \). Also, as we have seen in section (II.9), we may assume that \( \beta = \beta^{N/2} < \beta < \beta^{N/2} \), since otherwise the admissible control is unique. Now it follows from Theorem (1.2) that if \( \bar{u} \) is optimal, there exist \( \lambda \geq 0, \quad \psi \in \mathbb{R}_{1} \) such that

i) \( \psi(k) = \psi(k+1) \bar{u}(k) + \lambda f(\bar{u}(k)) \) \( (k \in \mathbb{N}) \),

ii) \( \psi(N) = \rho \) for some \( \rho \in \mathbb{R} \),

iii) \( \psi(k+1) \bar{u}(k) + \lambda f(\bar{u}(k)) = \min_{v \in J} \{ \psi(k+1) v + \lambda f(v) \} \) \( (k \in \mathbb{N}) \),

iv) \( (\lambda, \rho) \neq 0 \).

It follows from these relations that \( \lambda > 0 \), since otherwise \( \text{sgn} \psi(k) = \text{sgn} \rho \) \( (k \in \mathbb{N}) \) and hence (by iii)) \( \bar{u}(k) = \delta^k \) \( (k \in \mathbb{N}) \) if \( \rho > 0 \) and \( \bar{u}(k) = \delta^k \) \( (k \in \mathbb{N}) \) if \( \rho < 0 \). But this would imply \( b = \beta^{2N/2} \), contradicting our original assumption. Hence, we may assume that \( \lambda = 1 \) and we obtain

i) \( \psi(k) = \psi(k+1) \) \( (k \in \mathbb{N}) \), \( \psi(0) = \rho \).
III.2.3

ii) \( \bar{u}(k) = b(\psi(k+1)) \) \((k \in \mathbb{N})\),

and this result coincides with the one obtained in section II.9, setting \( \psi(k) = \tilde{s}_{N-k}(\bar{u}) \), \( \bar{u}(k) = \tilde{r}_k = \omega_{N-k-1}^0(\bar{u}). \) That is, just as in section II.9 the problem \( P_2 \) is solved if we have found the value of \( r \) for which we have \( x_0 = b \), where \( x \) is defined by i) and ii). The fact that

\[
S_N(b) := \{ \tilde{u}(\tilde{u}(k))\} \tilde{r}(k) = \psi(0) - \rho b
\]

can be proved in a way analogous to the one used in the case of problem \( P_2 \).

III.3. A controlled system with Coulomb friction

Consider the following control system:

\[
\dot{\bar{x}} = \text{sgn } \bar{x} + x - u,
\]

with the functions \( u \in L([0,\infty) \times [-1,1]) \) as controls. We rewrite (1) as

\[
\dot{\bar{x}} = y,
\]

\[
\dot{y} = -x - \text{sgn } y + u.
\]

or in vector form \( \dot{x} = f(x,u) \) where \( x = (x,y)' \) and \( f(x,u) = (y, -x - \text{sgn } y + u)' \). We can apply Theorem (I.4.2) to this system in the region \( S := \{ x \in \mathbb{R}^2 \mid |x| > 2 \land y = 0 \} \). According to the theorem there exists for every control \( u \) and every \( x \in S \) a solution of (2) on either \([0,\infty) \) or \([0,\infty) \) for some \( T_0 > 0 \), where in the last case we have \( x(t) \to x_0 (t + T_0) \) for some \( x_0 \in S = \{ x \in \mathbb{R}^2 \mid |x| \leq 2 \land y = 0 \} \). Whether the solution can be extended outside of \( S \) depends on \( u \). For simplicity we will restrict ourselves to trajectories which are in \( S \).

It follows from (2) that for a trajectory \( x \) the function \( t \mapsto x(t) \) is increasing in the upper half-plane and decreasing in the half-plane \( y > 0 \). Furthermore, it follows from the symmetry of the problem that \( x_u(t,s) = x_{-u}(t,-s) \) where \( t \mapsto x_u(t,s) \) is the
III.3

trajectory of (2) corresponding to the control \( u \) and with initial condition \( g(0) = a \).

For a given natural number \( N \) and a real number \( a > 2 \) we will call a trajectory and the corresponding control an a-start, N-step if the trajectory has initial value \((-a,0)\) and intersects the x-axis \( N \) times for \( t = t_i \) (i = 1,...,N), where \( t_0 := 0 < t_1 < \ldots < t_N \). The points \( (x(t_k),0) \) (\( k = 0,...,N \)) are called alteration points and the numbers \( a_k := (-1)^{k+1} x(t_k) \) are called alteration moduli. We will require always that \( a_k > 2 \) (\( k = 1,...,N \)) holds. The numbers \( t_k \) are called alteration times.

The problem we are going to consider is the following one:

\[ \Psi(a,b,N) : \text{"For fiven numbers } a, b > 2 \text{ and natural number } N \text{ find an } \]

a-start, N-step control which satisfies \( a_N = b \) and such that \( t_N \) (called the final time) is minimal".

In this section we restrict ourselves to the case \( N = 1 \). In this case the trajectory has to satisfy \( x(0) = (-a,0) \), \( x(T) = (b,0) \) for some \( T > 0 \), and \( y(t) > 0 \) (\( 0 < t < T \)). We will call a control admissible if there exists a trajectory satisfying these conditions. It is easily seen that there exists no admissible control if \( b \notin [a-4,a] \). In fact, it follows from

\( \frac{d}{dt} (x^2 + y^2) = -|y| + uy \geq 0 \)  \( (3) \)

that \( b \leq a \), and it follows from

\( \frac{d}{dt} [(x^2+y^2)^2] = (1+u)y \geq 0 \)

that \( b \geq a-4 \) for every trajectory. Also, it is clear from this argument that for \( b = a \) the control \( u = 1 \) is the only admissible control, and for \( b = a-4 \) the only admissible control is \( u = -1 \). We can restate problem \( \Psi(a,b,1) \) now as follows:

\[ \Psi(a,b) : \text{"Given numbers } a, b \text{ satisfying } b > 2 \text{ and } a-4 \leq b \leq a, \text{ find } \]

a control \( u \), such that for the corresponding trajectory of the system
\begin{align*}
\dot{x} &= y, \\
\dot{y} &= -x - 1 + u, \\
\end{align*}

satisfying \( x(0) = (-a, 0) \) we have \( x(T) = (b, 0) \) with minimal \( T' \).

We have ignored the condition \( y > 0 \ (0 < t < T) \) for the moment, since it will turn out to be satisfied automatically. We write (4) in vector notation \( \dot{x} = Ax + f(u) \), with

\[
A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad f(u) = \begin{bmatrix} 0 \\ u-1 \end{bmatrix},
\]

and we apply Theorem (1.5.1). Let \( \bar{u} \) be an optimal control. Then \( \bar{u} \) satisfies the equation

\[
\bar{u}(t) \dot{u}(\bar{u}(t)) = \max_{|v| \leq 1} \bar{u}(t) \dot{u}(v)
\]
a.e. for some non-trivial solution \( u \) of the adjoint equation

\[
\dot{\psi} = -HA \quad (\text{or } \dot{\psi}_1 = \psi_2, \quad \dot{\psi}_2 = -\psi_1).
\]

It follows that \( \psi_2 \) is of the form \( \psi_2(t) = r \cos(t + \phi) \) for some \( r, \phi \) with \( r > 0 \). Hence, \( \bar{u}(t) = \text{sgn } \psi_2(t) \)
a.e. and we may assume that \( \bar{u}(t) = \text{sgn } \psi_2(t) \) holds for all \( t \), since the values of \( \bar{u} \) on a set of measure zero do not influence the trajectory. As a consequence, \( \bar{u} \) assumes only the values \( \pm 1 \) and there is a time interval \( \tau \) between two consecutive switching times (if there is more than one switch). We show that there is at most one switch in \((0, T)\) by proving the existence of an admissible control with final time \( T \leq \tau \).

If \( b = a \) or \( b = a-4 \), it follows from the foregoing that \( \bar{u} \) is constant \( \pm 1 \) and this yields \( T = \tau \) in both cases. We assume henceforth that \( a-4 < b < a \).

The trajectories corresponding to the control \( u = 1 \) are

\begin{align*}
\dot{x} &= r \cos(t + \phi) \\
\dot{y} &= r \sin(t + \phi)
\end{align*}

with constant \( r, \phi \) satisfying \( r > 0 \). In the phase plane these trajectories are determined by the equation
\[ x^2 + y^2 = c^2, \ y > 0. \] \hspace{1cm} (7)

Let us denote the curve given in (7) by \( \gamma_+(c) \).

The trajectories corresponding to the control \( u = -1 \) are given by

\[
\begin{align*}
x &= -2 - r \cos(t + \phi), \\
y &= r \sin(t + \phi).
\end{align*}
\] \hspace{1cm} (8)

In the phase plane we have

\[(x - 2)^2 + y^2 = (c + 2)^2, \ y > 0\] \hspace{1cm} (9)

The curve given by (9) will be denoted by \( \gamma_-(c) \). Since \( x^2 + y^2 \) is strictly decreasing on a trajectory given by (8) (compare (3)) and since \( a-4 < b < a \) holds, there exists exactly one intersection point \( Q \) of \( \gamma_-(a) \) and \( \gamma_+(b) \), and by similar reasoning there is exactly one intersection point \( F \) of \( \gamma_+(a) \) and \( \gamma_-(b) \) (see fig. 1).

Let us introduce the following notation:

\[ A := (-a,0), \quad B := (b,0), \quad C := (-2,0), \quad O := (0,0). \]

Furthermore, we denote by \( t_p, t_Q \) the intersection times in \( P, Q \). Then it follows from formula (6) that \( t_Q = \angle ACQ \) and it follows from (7) that \( T - t_Q = \angle BQC \). Hence, \( T > \pi \) if we use the control \( u = -1 \) \( (0 < t < t_Q) \), \( u = 1 \) \( (t_Q < t < T) \). On the other hand, if we use the control \( u = 1 \) \( (0 < t < t_p) \), \( u = -1 \) \( (t_p < t < T) \), then \( t_p = \angle AOP \) and \( T - t_p = \angle PCB \). Hence \( T = \pi - \gamma \), where \( \gamma := \angle OPC > 0 \). This
control is therefore the optimal one. If we denote the coordinates of \( P \) by \((x, y)\), it follows by the cosine rule (see fig. 2) that

\[
\cos \gamma = \frac{a^2 + (b+2)^2 - 4}{2a(b+2)}.
\]

Furthermore, we have \( \cos \gamma = -\cos \gamma \) \((0 < T < \tau)\) and hence

\[
T = \arccos \frac{a^2 - (b+2)^2}{2a(b+2)}.
\]

This result, being derived for \( a-\frac{4}{4} < b < a \), is also valid if \( b = a \) and \( b = a-\frac{4}{4} \), since in these cases \( T = \tau \).

We sum up the result obtained for \( Q(a, b, l) \)

i) There exists an admissible control if and only if \( a-\frac{4}{4} \leq b \leq a \).

ii) For the same values of \((a, b)\) there is a unique optimal control.

iii) The minimal time is given by (10).

iv) The optimal control is of the bang-bang type, that is, it only assumes the values \( \pm 1 \).

v) If \( a-\frac{4}{4} < b < a \), the optimal control has exactly one switch on some \( t_p \in (0, T) \) and we have \( u(t) = 1 \) \((0 < t < t_p)\), \( u(t) = -1 \) \((t_p < t < T)\).

vi) If \( b = a \) or \( b = a-\frac{4}{4} \), there is no switching point. If \( b = a \), then we have \( u(t) = 1 \) \((0 < t < T)\), and if \( b = a-\frac{4}{4} \), then \( u = -1 \) \((0 < t < T)\).

Furthermore, it is easily seen that similar results can be obtained for the case that the trajectory starts at the point \( x(0) = (a, 0) \) and ends at \( x(T) = (-b, 0) \), whereas \( y(t) < 0 \) \((0 < t < T)\) holds. Also, it is obvious by the autonomy of the problem that a similar result is obtained if we start at an arbitrary time \( t_0 \) instead of \( t = 0 \). Of course, if \( t_1 \) is the final time, then we have \( t_1 - t_0 = T \) for the optimal control where \( T \) is given by (10).

In all these cases we will call the optimal control "\( Q(a, b, l) \)-optimal".
III.4. The case of more alteration points

We consider now problem \( Q(a,b,N) \) for \( N > 1 \). We will call a control \((s,k)\) piecwise optimal if it is an a-start, \( N \)-step control with alteration moduli \( s_0 = a, s_1, \ldots, s_N \) and such that it is
\( Q(a_0, a_{k+1}, 0) \)-optimal on the step between the alteration points
\((s_a, 0)\) and \((s_{a+1}, 0)\) for \( k = 0, 1, \ldots, N-1 \). Then as in the previous chapter we observe that we can restrict ourselves to piecewise optimal controls. Therefore, with a change of notation which is closer to the notation used in the theory of section 1, we can restate \( Q(a,b,N) \) as follows:

\[ Q'(a,b,N) : \text{"Given numbers } a, b > 2 \text{ and a natural number } N, \text{ find the sequence } x(0), x(1), \ldots, x(N) \text{ with } x(0) = a, x(N) = b \text{ and } x(k) = 4 \leq x(k+1) \leq x(k) \quad (k \in \mathbb{R}) \text{ and such that } \]
\[ \sum_{k=0}^{N-1} T(x(k), x(k+1)) \text{ is minimal."} \]

Here the function \( T \) is defined by
\[
T(p, q) = \arccos \frac{4 - p^2 - (q+2)^2}{2pq+2} \tag{1}
\]
(compare formula (3.10)). We introduce the following notation:
\[ v(k) := x(k+1) - x(k) + 2, \quad T(p, p+2) =: S(p, r). \]

Then we can state the problem as follows:

\[ Q''(a,b,N) : \text{"Given real numbers } a, b > 2 \text{ and a natural number } N, \text{ find a strategy } v \in (\mathbb{N} \to [-2,2]) \text{ such that for the function } x \in \left( \mathbb{N} \to \mathbb{R}^2 \right), \text{ defined by } \]
\[ x(0) = a, \]
\[ x(k+1) = x(k) + v(k) + 2 \quad (k \in \mathbb{R}), \tag{2} \]

we have \( x(N) = b \) and such that \( \sum_{k=0}^{N-1} S(x(k), v(k)) \) is minimal, where \( S \in \left( (2,\infty) \times [-2,2] \to \mathbb{R}^3 \right) \) is defined by
\[ S(x,v) := f(x,x+v) , \]

\[ f(x,z) := \arccos(-g(x,z)) , \]

\[ g(x,z) := (x^2 + z^2 - 4)/(2xz) . \]

It is obvious that there exists an optimal strategy if and only if \[ a - 4N \leq b \leq a \] holds. In the cases \( b = a \) or \( b = a - 4N \) the only admissible (and hence optimal) strategies are \( v(x) = 2 \) (\( k \in \mathbb{N} \)) and \( v(k) = -2 \) (\( k \in \mathbb{N} \)) respectively. We will assume henceforth that \( a - 4N < b < a \) holds.

In order to solve \( Q^n \) we apply Theorem (1.2). First, we must show that the set \( \{(x,v-2,S(x,v)+w) \mid |v| \leq 2 \wedge w \geq 0\} \) is convex for every \( x \geq 2 \), that is, we have to prove that \( v = S(x,v) \) is convex. To achieve that aim, and also for further calculations, we give formulas for the derivatives of \( S, f \) and \( g \):

\[ \frac{\partial g(x,z)}{\partial x} = \frac{4 - z^2 + x^2}{2x^2} , \]

\[ \frac{\partial g(x,z)}{\partial z} = \frac{4 - x^2 + z^2}{2xz} , \]

\[ \frac{\partial^2 g(x,z)}{\partial z^2} = \frac{z^2 - 4}{xz^3} , \]

\[ \frac{\partial f(x,z)}{\partial x} = (1 - g^2(x,z))^{-1/2} \frac{\partial f(x,z)}{\partial x} , \]

\[ \frac{\partial f(x,z)}{\partial z} = (1 - g^2(x,z))^{-1/2} \frac{\partial f(x,z)}{\partial z} , \]

\[ \frac{\partial^2 f(x,z)}{\partial z^2} = (1 - g^2(x,z))^{-3/2} . \]

\[ \cdot [(1 - g^2(x,z)) \frac{\partial g(x,z)}{\partial z} + g(x,z) (\frac{\partial g(x,z)}{\partial z})^2] . \]

**Remark.** Here \( \frac{\partial^2 f}{\partial z^2} \) is defined by \( \frac{\partial^2 f}{\partial z^2}(x) := \frac{\partial}{\partial z}(\frac{\partial f}{\partial z}(x)) \).

The following inequalities hold for \( x \geq 2, |x-z| \leq 2; \)
with equality only if \(|x-z| = 2\); furthermore, \(\partial^2 g(x,z) > 0\), and hence \(\partial^2 g(x,z) > 0\) (\(|x-z| < 2\)). Since \(\partial^2 g(x,v) = \partial^2 g(x,x+v)\), we have shown that \(v \mapsto S(x,v)\) is strictly convex for all \(x > 2\).

Now we give some more properties which will be used in the sequel. We have \(\partial g(x,x+2) = 2/x > 0\) and \(\partial g(x,x-2) = -2/x < 0\). This implies that \(\partial g(x,z) \to (z \mapsto x+2)\) and \(\partial g(x,z) \to = (z \mapsto x-2)\).

Now we are ready to apply Theorem (1.2). Let \(\tilde{v}\) be an optimal strategy and let \(\tilde{x}\) denote the corresponding sequence (according to (2)). Then there exists \(\lambda \geq 0\) and \(\psi \in (\bar{\mathcal{N}}^+ + \mathbb{R}^1)\) such that

\[
\begin{align*}
\psi(k) &= \psi(k+1) + \lambda \partial_x S(\tilde{x}(k), \tilde{v}(k)) \quad (k \in \tilde{N}), \\
\psi(N) &= \text{u for some real u}, \\
\psi(k+1)(\tilde{x}(k) + \tilde{v}(k) - 2) + \lambda S(\tilde{x}(k), \tilde{v}(k)) &= \\
&= \min \left[ \psi(k+1)(\tilde{x}(k) + \text{v} - 2) + \lambda S(\tilde{x}(k), \text{v}) \right], \\
&= \text{min}_2 \psi_2. \\
\end{align*}
\]

First, we remark that it follows from (2) that \(x(k) \geq x(k+1)\) (\(k \in \tilde{N}\)). Hence, \(\tilde{x}(k) + \tilde{v}(k) - 2 > 0\) (\(k \in \tilde{N}\). Now it is easily seen that we must have \(\lambda > 0\). In fact, if \(\lambda = 0\) holds, then \(\psi(k) = \psi\) (\(k \in \bar{\mathcal{N}}^+\)) and hence \(\tilde{v}(k) = 2 \text{sgn} \mu\) (\(k \in \tilde{N}\). In that case we have either \(\tilde{x}(N) = a\) or \(\tilde{x}(N) = a = 4N\), but this contradicts the assumption \(a - 4N < b < a\). Because of homogeneity we may set \(\lambda = 1\). Then we get, instead of i) and iii):

\[
\begin{align*}
i') \psi(k) &= \psi(k+1) + \partial_x S(\tilde{x}(k), \tilde{v}(k)) \quad (k \in \tilde{N}), \\
n) \psi(k+1)(\tilde{x}(k) + \tilde{v}(k)) + S(\tilde{x}(k), \tilde{v}(k)) &= \min \left[ \psi(k+1)\text{v} + S(\tilde{x}(k), \text{v}) \right], \\
&= \text{min}_2 \psi_2. \\
\end{align*}
\]

Since \(\psi \mapsto S(x,v)\) is strictly convex for \(x > 2\) it follows that the function \(v \mapsto \psi(\text{v} + S(x,v))\) assumes its minimum for every real \(\text{v}\) and \(x > 2\) in exactly one point. Let us denote this value of \(\text{v}\) by \(N(\psi,x)\).
III.4

Then we can replace iii)’ by

\[ \tilde{y}(k) = N(y(k+1), \bar{x}(k)) \quad (k \in \mathbb{N}) . \]

Because of the asymptotic behaviour of \( y \mapsto \partial_y S(x, y) \) for \( y \rightarrow \pm 2 \) and because of the differentiability of \( S \), iii)” is equivalent to

\[ \psi(k+1) = - \partial_y S(x(k), \bar{y}(k)) . \]  \( (13) \)

The optimal strategy can therefore be obtained by solving the following boundary value problem:

"Find functions \( \psi \in (\mathbb{N}^+ \times \mathbb{R}^1), \bar{x} \in (\mathbb{N}^+ \times \mathbb{R}^1), \bar{y} \in (\mathbb{N}^+ \times [-2,2]) \) satisfying (12), i)’’, iii)” (or (13)) and \( \bar{x}(N) = b. " \)

If there is more than one solution of this problem, we have to select the one which minimizes \( \sum_{k \in \mathbb{N}} S(x(k), \bar{y}(k)) \).

In order to find the function \( M \), one has to solve (13) which can be reduced to a cubic equation. By means of a simple artifice we can modify the boundary value problem in such a way that the cubic equation can be avoided and we only have to solve a quadratic equation at each step. For that purpose we introduce the function \( z \in (\mathbb{N}^+ \times \mathbb{R}^1) \) by

\[ z(k) = \tilde{x}(k) + \tilde{y}(k) \quad (k \in \mathbb{N}) , \]  \( (14) \)

and replace (1) and (13) by

\[ \tilde{x}(k+1) = z(k) - 2 , \]  \( (15) \)

\[ \psi(k+1) = - \partial_z f(\tilde{x}(k), z(k)) , \]  \( (16) \)

and i)’ by

\[ \psi(k) = \psi(k+1) + \partial_{\tilde{x}} f(\tilde{x}(k), z(k)) + \partial_z f(\tilde{x}(k), z(k)) , \]

which (using (16)) reduces to

\[ \psi(k) = \partial_{\tilde{x}} f(\tilde{x}(k), z(k)) . \]  \( (17) \)

We can solve for \( z(k) \) in (17). In fact, the equation
\[ \partial_x \xi(x,z) = \psi, \quad z > 0 \] (18)

in \( z \), is equivalent to

\[ (\partial_x \xi(x,z))^2 = \psi^2 (1 - g^2(x,z)), \quad \text{sgn } \psi = \text{sgn } \partial_x g(x,z), \quad z > 0 \] (19)

or

\[ (4 - z^2 + x^2)^2 = x^2 \psi^2 (4x^2 z^2 - (x^2 + z^2 - 4)^2). \] (20)

With the substitution \( z^2 - 4 = \zeta \), equation (20) is equivalent to

\[ (x^2 - \zeta)^2 = x^2 \psi^2 (16x^2 - (x^2 - \zeta)^2) \]

which is a linear equation for \((x^2 - \zeta)^2\). Hence,

\[ (x^2 - \zeta)^2 = 16x^4 \psi^2 (1 + x^2 \psi^2)^{-1}. \]

This yields

\[ z^2 = x^2 + 4 + 4x^2 \psi (1 + x^2 \psi^2)^{-\frac{1}{2}}. \]

Since \( \text{sgn } \partial_x g(x,z) = \text{sgn } (4 - z^2 + x^2) \) for positive \( z \), we see by (19) that (18) is equivalent to:

\[ z = [x^2 + 4 - 4x^2 \psi (1 + x^2 \psi^2)^{-\frac{1}{2}}]^\frac{1}{2}. \] (21)

Therefore we can restate the boundary value problem as follows:

"Given numbers \( a, b \), satisfying \( b > 2, a - 4N < b < a \), find a real number \( \theta \) such that, if the functions \( \psi \in (S^1, \mathbb{R}^1), \quad \bar{\xi} \in (S^1, \mathbb{R}^1), \quad z \in (\mathbb{R}^1, \mathbb{R}^1) \) are defined by \( \psi(0) = \theta, \bar{\xi}(0) = a \) and

\[ a(k) = [\bar{\xi}^2(k) + 4 - 4\bar{\xi}^2(k)\psi(k)(1 + \bar{\xi}^2(k)\psi(k))]^{-\frac{1}{2}} \]

\[ \bar{\xi}(k+1) = z(k) - 2 \]

\[ \psi(k+1) = (a(k))^{-\frac{1}{2}} (a(k) - 4)[16a^2(k) - (a^2(k) - z^2(k) - 4)^2]^{-\frac{1}{2}} \]

for \( k \in \mathbb{R} \), then we have \( \bar{\xi}(N) = b. \)"
REFERENCES


SAMENVATTING

In dit proefschrift wordt een methode behandeld voor de bepaaling van optimale besturingen (optimal controls) voor systemen die beschreven worden door differentiaalvergelijkingen van de vorm \( \dot{x} = f(x,u) \), waarbij de functie \( f \) discontinuïteiten heeft voor sommige waarden van \( x \). Hierbij is \( u \) de besturingsfunctie en \( t \) de onafhankelijke variabele. Als er geen discontinuïteiten aanwezig zijn kunnen zulke optimale besturingen worden gevonden met behulp van het maximumprincipe van Pontryagin (zie [2]). Als \( f \) discontinuïteiten heeft is het maximumprincipe niet meer geldig. De methode gevolgd in dit proefschrift kan als volgt worden omschreven:

We verdelen de banen (dit zijn de oplossingskrommen van de differentiaalvergelijking) in stukken waarop \( f \) wel continu is. Op deze stukken passen we het maximumprincipe toe. Een baan die optimaal is op elk van die stukken noemen we stuksgewijs optimaal. Om nu een besturing te vinden die optimaal is voor de hele kromme moeten we de "stukken" aan elkaar passen. Dit resulteert in een discreet optimaliseringsprobleem. Een probleem van dit type kan worden opgelost door middel van dynamische programmering of met behulp van het discrete maximumprincipe, een discrete versie van het maximumprincipe van Pontryagin.

In dit proefschrift wordt geen algemene theorie behandeld. De methode wordt toegepast op enkele voorbeelden, die in detail worden uitgewerkt. Onder optimale besturing wordt steeds tijdoptimaal besturing verstaan, hoewel de methode ook kan worden gebruikt bij algemener optimaliteitscriteria.

Omdat van differentiaalvergelijkingen met discontinu rechterlijk de existentie van oplossingen niet uit klassieke stellingen hierover kan worden afgeleid, wordt in hoofdstuk I een criterium voor de existentie van oplossingen van zulke differentiaalvergelijkingen
gegeven.

In hoofdstuk II wordt een eenvoudige vergelijking besproken (de zogenaamde jojo-vergelijking) en er worden een aantal optimale besturingsproblemen opgelost voor deze vergelijking. De "discrete gedeelten" van deze problemen worden opgelost door middel van dynamische programmering.

In hoofdstuk III wordt het discrete maximumprijsprincipe gebruikt om de discrete gedeelten van de problemen van hoofdstuk II op te lossen. Verder wordt in hoofdstuk III de optimale besparing van een mechanisch systeem met coulomb-wrijving besproken.
CURRICULUM VITAE

STELLINGEN

I
Een rij reële getallen \( \{a_n\}_{n=0}^{\infty} \) heeft de eigenschap dat \( \sum a_n x_n^2 > 0 \)
geldt voor elke niet-triviale eindige rij \( \{x_n\}_{n=0}^{N} \) met \( \sum x_n = 0 \), dan en
slechts dan als voldaan is aan de volgende voorwaarden:

i) Hoogstens één \( a_n \) is niet-positief.

ii) Als er een \( a_n < 0 \) is, dan is \( \sum a_n^{-1} \) convergent en niet-positief.

II
De kubus \( \{ x \in \mathbb{R}^n \mid |x_i| \leq 1 \ (i = 1, \ldots, n) \} \) en het hypervlak
\( \{ x \in \mathbb{R}^n \mid (a,x) = c \} \) met \( a \in \mathbb{R}^n \), \( a \neq 0 \), hebben een niet-lege
doorsnede \( D \) als \( \sum |a_i| \geq |c| \). Het punt \( y \in D \) met de kleinste (euclidische)
afstand tot de oorsprong wordt beschreven door \( y_i = \text{sat}(\lambda a_i) \)
\( (i = 1, \ldots, n) \), waarbij sat de oneven functie is met sat \( a = \min(1, a) \)
\( (a \geq 0) \). De waarde van \( \lambda \) kan op een eenvoudige grafische manier
worden bepaald uit de vergelijking \( (a,y) = c \).

III
Een cel in het platte vlak is een \( 1 \times 1 \)-vierkant met hoekpunten in
roosterpunten \( (i,j) \) met \( i,j \) geheel. Een kleuring is een afbeelding van de collectie
der cellen in een verzameling \( S \) (de verzameling der kleuren). Een kleuring heet \( n \times m \)-invariant als voor
elke kleur \( s \in S \) het volgende geldt:

Het aantal cellen met de kleur \( s \) in een \( n \times m \)-rechthoek van
cellen is onafhankelijk van de positie van de rechthoek.
Er geldt de volgende eigenschap: Als \( n \) en \( m \) onderling ondeelbaar
zijn, dan is een \( n \times m \)-invariante kleuring een \( n \times 1 \)- of een \( m \times 1 \)-
invariante kleuring.

Literatuur: M.L.J. Hautus en D.A. Klarner, "Invariant colorings for
boxes" (in voorbereiding).
IV

Zij \( X \) een genormeerde lineaire ruimte en \( F \) een begrense afbeelding van \( X \) in \( X \) (d.w.z. \( F(M) \) is begrensd als \( M \subset X \) begrensd is). Dan bestaat er een rij \( \{\lambda_1, \lambda_2, \ldots\} \) van positieve getallen zodanig dat elke rij \( \{x_n^r\} \) in \( X \) die voldoet aan \( x_{n+1}^r = F(\lambda_n x_n^r) \) \((n = 1, 2, \ldots)\), begrensd is.

V

De oplossing van de differentiaalvergelijking \( y'(x) = -2y(x) \) \((x > 0)\), met beginwaarde \( y(0) = 1 \) wordt benaderd door het Euler–Cauchy polygoon waarvan de hoekpunten gegeven worden door \( x_n = nh \) \((n = 0, 1, \ldots)\), \( y_0 = 1, y_{n+1} = y_n - 2h \cdot y_n^2 \) \((n = 0, 1, \ldots)\). Hierbij is \( h \) een positief getal. Laat \( \xi \) de waarde van \( x \) zijn, waarvoor het polygoon de \( x \)-as snijdt. Dan geldt:

\[
\xi = 1 + \frac{1}{2} \log h + O(h) \quad (h \to 0).
\]

VI

Laten \( \phi(x,y) \) en \( f(x,y) \) analytische functies zijn in een omgeving van de origine \((0,0)\) in \( \mathbb{C}^1 \times \mathbb{C}^1 \) met waarden in \( \mathbb{C}^1 \) resp. \( \mathbb{C}^2 \). Laat de formule reeks \( u(x) = \sum_{k=1}^{\infty} a_k x^k \) formueel voldoen aan de differentiaalvergelijking

\[
\phi(x,y)y'(x) = f(x,y).
\]

Als \( \phi(x,u(x)) = \sum_{k=0}^{\infty} b_k x^k \), dan wordt \( \omega \) gedefinieerd door \( \omega := \min(k \mid b_k \neq 0) \). (Als \( b_k = 0 \) voor alle \( k \) geldt, dan is \( \omega = \infty \).)

Er gelden dan de volgende eigenschappen:

i) Als \( \omega = 0 \) of \( \omega = 1 \), dan heeft \( u \) een positieve convergente strook en is derhalve een oplossing van (1) in een omgeving van \( 0 \).

ii) Als \( \omega < \infty \), dan bestaat er een \( \theta > 0 \), zodat voor elke sector \( S \) in het complexe \( x \)-vlak met openingshoek kleiner dan \( \theta \) een oplossing van (1) bestaat, waarvoor \( u(x) \) de asymptotische reeks
is voor $x > 0$, $x \in S$.


J. Malquist, "Sur l'étude analytique des solutions d'un système d'Équations différentielles dans le voisinage d'un point singulier d'indétermination II", Acta Math. 74 (1941), 1-64.

VII


VIII

Zij $A$ een $n \times n$-matrix en $B$ een $n \times m$-matrix. Dan heet het paar $(A,B)$ bestuurbaar als de $n \times nm$-matrix $(B,A B, \ldots, A^{n-1} B)$ de rang $n$ heeft. Er geldt de volgende eigenschap: $(A,B)$ is bestuurbaar dan en slechts dan als voor elke eigenwaarde $\lambda$ van $A$ geldt: $\text{rang}(A - \lambda I, B) = n$.


IX

Zij $S$ een niet-lege verzameling van complexe getallen, $A$ een $n \times n$-matrix and $B$ een $n \times m$-matrix. Dan bestaat er een matrix $D$ zodanig dat de eigenwaarden van $A + DB$ in $S$ liggen, dan en slechts dan als $\text{rang}(A - \lambda I, B) = n$ voor elke eigenwaarde $\lambda$ van $A$ buiten $S$ geldt.

Literatuur: M.L.J. Hautus, "Controllability, observability and stabilizability of linear autonomous systems" (in voorbereiding).

X

De voorwaarde, die door R.E. Kalman noodzakelijk en voldoende wordt genoemd voor het bestaan van een "asymptotic state estimator" voor
een lineair constant systeem met één output-variable, is wel voldoen-
de maar niet noodzakelijk.

Literatuur: R.E. Kalman, P.L. Falb, M.A. Arbib, "Topics in mathe-

XI

Zij \( \mathcal{N}_n \) de collectie der niet-lege compacte deelverzamelingen van \( \mathbb{R}^n \).

Op \( \mathcal{N}_n \) wordt een afstand gedefinieerd door

\[
\rho(S,T) := \max \{ \max_{x \in S} d(x,T), \max_{y \in T} d(y,S) \}.
\]

Hierbij is

\[
d(x,A) := \min_{x \in A} |x - a|, \quad (x \in \mathbb{R}^n, A \in \mathcal{N}_n).
\]

Laat \( V \) een absoluut continue afbeelding van \([0,1]\) in \( \mathcal{N}_n \) zijn (d.w.z. voor alle \( \epsilon > 0 \) bestaat er een \( \delta > 0 \), zodat voor alle rijen

\( s_1, t_1, s_2, t_2, \ldots, s_n, t_n \) met \( 0 \leq s_1 < t_1 \leq \ldots \leq s_n < t_n \leq 1 \) met

\( \Sigma(t_i - s_i) < \delta \) geldt \( \Sigma d(V(t_i), V(s_i)) < \epsilon \). Dan bestaat er een abs-
oluut continue afbeelding \( v \) van \([0,1]\) in \( \mathbb{R}^n \) met \( v(t) \in V(t) \)

\( 0 \leq t \leq 1 \).

Literatuur: H. Hermes, "Existence and properties of solutions of

\( \dot{x} \in R(t,x) \)". Adv. in Diff. and Int. Eq. SIAM, Philadelphia, 1969.

Eindhoven, 29 juni 1970

M.L.J. Hautus