Graph and 2-D Optimization Theory and their application for discrete simulation of gas transportation networks and industrial processes with repetitive operations

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften
der Rheinisch-Westfälischen Technischen Hochschule Aachen
zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften
genehmigte Dissertation

vorgelegt von

Diplom-Mathematiker

Dymkou Siarhei Michailovich

aus Minsk, Weißrußland

Berichter: Universitätsprofessor Dr. Gerhard Jank
Universitätsprofessorin Dr. Eva Zerz

Tag der mündlichen Prüfung: 22. September 2006

To my parent Dymkov Michail, Dymkova Raisa
my brother Dymkou Vitali
my wife Irina
and my son Alexei.

Content

Declaration

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university and is less than 40,000 words in length excluding tables, maps, footnotes, bibliographies and appendices. To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgement has been made.

Siarhei Dymkou

October , 2006.

Preface

This report presents the investigation on the research work "Graph and 2-D Optimization Theory and their application for discrete simulation of gas transportation networks and industrial processes with repetitive operations" and gives mainly the theoretical part of as a basis for developing numerical methods. It is conjectured that the corresponding algorithms will be realized as computer programs and next they can be used in practical work.

Chapter 1 presents graph and 2-D theory setting to test a gas distribution network for physical feasibility and their optimization within the predefined ranges of available transport capacity and boundary conditions. The models introduced can be used for problems of flow/pressure control of gas transport network. The aim of the work is the development of a comprehensive optimization theory based on a constructive approach in the graph and 2-D models. The key elements of the proposed optimization method is illustrated by an example. Some aspects of the numerical optimization method for the distributed gas network are discussed, too.

In Chapter 2 we consider a class of discrete 2-D models to study the gas network system. The focus is on the development of a new optimization method based on operator setting. The traditional for optimization theory problems such as existence and uniqueness of optimal control as well the optimal feedback control law are investigated.

Chapter 3 develops the optimization theory for the so-called repetitive processes. These objects arise in the modeling of a lot of industrial processes and they can be used for the perspective planning or learning procedures. The classical approach based on the separation theorem and a new, named constructive, method are developed. These results are illustrated by solving a synthesis problem for the system with simple dynamics.

In Chapter 4 we investigate the links between some classes of linear repetitive processes and delay systems and apply this to analyze system theoretic problems arising such as controllability and optimization of these repetitive processes.

Acknowledgments

I express my sincere thanks to my supervisor Prof. Gerhard Jank, who introduced me to the subject of differential algebraic system theory and repetitive multidimensional processes Theory and encouraged me to work in this area. Special thanks are due to the Head and Administrative Officers of RWTH for encouragement to research in RWTH-Aachen. Also I deeply thankful to my colleagues from Lehrstuhl II für Mathematik for their patience and support over many months. This investigation was supported by the RWTH Aachen, Germany.

Introduction

The past decades, in particular, have seen a continually growing interest in the application of modern mathematical theories to engineering problems. This development is clearly related to the wide variety of applications of both practical and theoretical interests. The progress in computer technologies, the availability of high-speed processors and various programming languages allow nowadays the researchers in different areas of science to investigate and design numerous algorithms to solve physical and engineering phenomena on the computer. However, to construct high precision models of a real process one has to begin with its mathematical description and analysis in order to obtain specific characteristics of the considered problem. This helps to design very efficient numerical methods which can be implemented directly on the computer.

Many technical and information processes in various fields possess identical mathematical structures, that can be described in a common optimization problem form. Such a generalization allows to construct general algorithms to solve a wide class of problems. It is, however, not sufficient to analyze a given problem on a pure theoretical basis. The practical application may pose additional constraints like real-time performance of the system, low-delay requirements or restrictions on computing power or memory. Therefore, it is necessary to develop a discrete model for the solution of the problem which is suitable for computer implementation.

The presented thesis is oriented towards the development of algorithms based on constructive optimization methods in graph and 2-D system theory and their application for solving practical problems arising in the discrete simulation of distributed gas transportation networks and industrial processes with repetitive operations. The need to consider these objects has been stimulated by numerous engineering applications where system models of physical processes yield this class of objects. In spite of a number of papers devoted to this theme and the large popularity of these processes, their investigation still remains a challenging problem.

In this thesis the investigation on the Research work " Graph and 2-D Optimization Theory and their application for discrete simulation of gas transportation networks and industrial processes with repetitive operations". This work gives mainly the theoretical part of this research on the base of that some numerical methods are developed. It is conjectured that the corresponding algorithms will be realized as computer programs and next they can be used in practical work.

The main goals of this work are: (i) to present the state of the art of the optimization theory in the context of gas transportation networks and repetitive processes, (ii) to present the graph and 2-D system setting approach for gas networks modeling, (iii) to design constructive optimization methods in graph and repetitive models, (iv) to state the important system theoretic properties of 2-D and repetitive processes and (v) to illustrate the presented methods on some selected problems of technical interest. These five goals are reflected by Chapter 1 through Chapter 4.

The gas transportation network (GTN) is a well known example of a complex and large scale distributed parameter system of great practical interest [54]. For this reason the modeling approaches, numerical methods and optimization of operating modes of gas transport networks are of permanent interest for researchers. Although in the last decades a number of papers devoted to this theme were published [64], [74], the control of complicated gas networks still remains an actual problem. A general model of a gas transportation network includes a number of nonlinear elements such as pipelines, gasholders, compressor stations and others. A detailed description of the gas transportation through the whole gas network is rather complex. Usually the proposed equations involve a number of variables and can become quite cumbersome. But there exists a need to check quickly the physical feasibility of a gas distribution network within the predefined ranges of available transport capacity and suddenly changing boundary conditions. This can not be done efficient in framework of the full gas network description. Nevertheless, a multistage mathematical formalization is feasible in terms of some generalized parameters and suitable variables. Such formalization is from a simple high level model that can be expanded and developed then into a more complicated representation according to needs.

By this reason in the first two Chapters the mathematical model and the corresponding optimization problem of gas transport networks are introduced on the basis of the constructive approach in the graph and 2-D system setting. Namely, in Chapter 1, the simplest graph model is proposed to express potentially critical flow/pressure values within the given margins of inflows, outflows and setting of active components such as storage capacity gasholders, compressor stations and others in order to satisfy/optimize the demand distributed over different nodes.

The models introduced can be used to test a gas distribution network for physical feasibility and their optimization within the predefined ranges of available transport capacity and boundary conditions. The aim of the chapter is the development of a comprehensive optimization theory based on constructive approach in the graph and 2-D models. The key elements of the proposed optimization method is illustrated by an example. Some aspects of the numerical optimization method for the distributed gas network are discussed, too.

Next, in Chapter 2, the 2-D system setting can be used then for the more sophisticated and detailed description of dynamic processes in gas pipeline units based on partial differential momentum and continuity equations [74]. Also the optimal feedback control problem for 2-D systems is discussed that is of both system theoretic and application interest. Some aspects of the control theory for the multidimensional systems are investigated in [47]. The various optimal control problems for 2-D systems have been considered in [11].

Chapter 3 develops the optimization theory for so-called repetitive processes. These objects arise in the modeling a lot of industrial processes and they can be used for the planning or learning procedures. A multipass process (termed repetitive process in the other literature) is one in which the material involved is processed by a sequence of passes, termed sweeps, of the processing tool. Such systems are characterized by two distinctive features, repetitive operation and dependence of present-pass behaviour on the behaviour of the previous passes. They arise in the modeling a lot of industrial processes such as long-wall coal cutting, metal rolling operations and others. Metal rolling, for example, is an industrial process where deformation of the metal stock takes place between two rolls with parallel axes revolving in opposite directions through a series of passes for successive reductions. An repetitive processes of metal rolling modeling in linearized form can be presented as follows (some details can be found in [72])

$$\frac{d^2 y_k(t)}{dt^2} + \lambda_1 y_k(t) = \lambda_2 \frac{d^2 y_{k-1}(t)}{dt^2} + \lambda_1 y_{k-1}(t) + b u_k(t),\ t \in [0, t^*],\ k \in K = \{1, ..., N\},$$

where $y_k(t)$ and $y_{k-1}(t)$ denote the gauge on the current and previous passes through the rolls; λ_1, λ_2 and b are determined, in fact, by the stiffness of the metal strip and the roll mechanism properties, $u(t)$ can be interpreted as the applied force to the metal strip by the rolls.

Such dynamic systems provide an appropriate mathematical tool for modeling chemical processes, also. In particular, a model of the rectification process of many component mixture in a many-plate column can be represented by the similar model

$$\frac{dx_s(t)}{dt} = V_{s-1}(t)x_{s-1}(t) + V_s(t)x_s(t) - R_s(x_s(t), y_s(t)) + u_{x_s}(t),$$
$$\frac{dy_s(t)}{dt} = L_{s+1}(t)y_{s+1}(t) + L_s(t)y_s(t) + R_s(x_s(t), y_s(t)) + u_{y_s}(t),$$
$$t \in [0, t^*], \quad s \in K \doteq \{1, ..., N\}.$$

Here $x(s,t)$, $y(s,t)$ denote the desired material concentration on s-th plate in the gas and liquid fractions, respectively; L, V and R present the hydrodynamical characteristic of the process under consideration; u_x and u_y are the control material row; K is subset of integers. Some details of the model can be found in [18]. Also problem areas exist where adopting a repetitive process perspective has clear advantages over alternatives. The development of a mature systems theory for these processes has been the subject of considerable research efforts over the past two decades which has resulted in very significant progress on systems theoretic properties. This work is devoted to the optimization theory of some classes of these objects.

The first part of the Chapter 3 uses the classic approach to investigate the traditional optimal control theory problems. It is well known that the separation theorem for convex sets is quite useful approach for studying a wide class of the extremal problems. Here we develop method to establish optimality conditions in the classic form of maximum principal for multipass nonstationary continuous-discrete control system with nonlinear inputs and nonlocal state-phase terminal constraints of general form. The obtained results are traditional for classic optimal control theory. However, their numerical realization is not a trivial task. By this reason in the next sections for the stationary case of the system model and particular case of the constraint and the cost functional we develop the new optimality and sub-optimality conditions that are more suitable for the design of numerical methods and further applications. In contrast to the classic approaches of optimal control theory,in the second part of the Chapter 3 we use the idea of constructive methods reported in [37] and extend this setting to the continuous-discrete case to produce new results and constructive elements of optimization theory for the considered repetitive systems and develop also its relevant basic properties which can be of interest for others purposes, too. It is shown that the obtained optimality and ϵ-optimality conditions are close related to the corresponding classic results of maximum principle and ϵ- maximum principle. The sensitivity analysis and some differential properties of the optimal controls under disturbances are discussed and their application to the optimal synthesis problem is given. It has been conjectured that such setting could be appropriate for development of numerical methods of optimal control problems and related studies on which very little work has yet been reported. The obtained results yield a theoretical background for the design problem of optimal controllers for relevant basic processes. This idea is demonstrated by an illustrative example where the optimal feedback control law is given.

It is already known that repetitive processes can be represented in various dynamical system forms, which can, where appropriate, be used to great effect in the control related analysis of these processes. In the Chapter 4, we investigate further the already known links between some classes of linear repetitive processes and delay systems and apply this to analyze control theory problems arising in controllability and optimal control of these repetitive processes. In particular, so-called characteristic mappings introduced in [37] are used to establish controllability properties criteria. Next, time optimal control problems are considered, where it is well known that the separation theorem for convex sets is a useful approach for studying a wide class of extremal problems. Here we adopt this method to establish optimality conditions in the classic form. The key features of the developed optimization method is illustrated by an example where the time optimal control is calculated for the time delayed differential equation with control input.

1

Mathematical modeling in distributed gas networks

This chapter presents a graph and a 2-D theory setting for a gas distribution network to test their physical feasibility and their optimization within predefined ranges of available transport capacity and boundary conditions. The models introduced can be used for problems of flow/pressure control of gas transport networks. The aim of the work is the development of a comprehensive optimization theory based on a constructive approach in the graph and 2-D models. The key elements of the proposed optimization method is illustrated by an example. Some aspects of the numerical optimization method for the distributed gas network are discussed, too.

1.1 Gas transportation networks background

The gas transportation network (GTN) is a well known example of a complex and large scale distributed parameter system of great practical interest [54]. By this reason the modeling approaches, numerical methods and optimization of operating modes of gas transport networks are of permanent interest for researchers. In the last decades a number of papers devoted to this theme were published (see, for example, [64], [74], [63], [33], [75], [76], [77], [79], [5], [50]) however control of complicated gas networks still remains a challenging problem.

A general model of a gas transportation network includes a number of nonlinear elements such as pipelines, gasholders, compressor stations and others. See Figure 1.1:

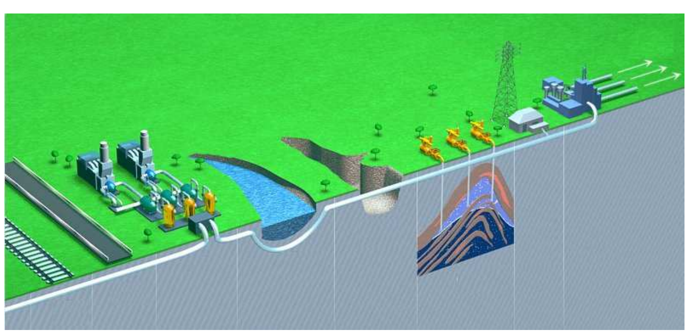

Figure 1.1: Segment of the complex gas network.

A detailed description of the gas transportation through the whole gas network is rather complex.

For the network there are two types of objects: nodes and edges which constitute the network itself. The elements of such types form the required network. The main difference between the elements of these types is that each edge can be connected with no more than two nodes, whereas a nodes in its turn has no limitations with regard to the number of edges connected thereto. The set of edges can be portioned on the different classes such as active and passive, for example, where the compressors/comressor stations exist or not. Usually, the type of nodes consists of the following types of objects:

i) nodes between the edges (external);

ii) nodes along pipes, i.e. internal edge nodes.

The external nodes can be treated as the point where there exist pressure and flow monitoring and control facilities.

Thus, the general gas transportation network can be constructed on the basis of the given elements.

Nevertheless, a detailed description of the gas transportation through the whole gas network is rather complex. Usually the equations obtained in this way involve a number of variables and can become quite cumbersome, and they are used mostly for theoretical studying. But in the ordinary cases there exists a need to check quickly the physical feasibility of a gas distribution network within the predefined ranges of available transport capacity and the sudden changed boundary conditions, which can not be done fast in framework of the full gas network description. In this case a multistage mathematical formalization is feasible in terms of some generalized parameters and suitable variables. Such formalization is from a simple high level model that can be expanded and developed then into a more complicated representation according to needs.

In the presented Chapter this idea of multistage modeling in large scale distributed gas network is realized. Namely, the mathematical models and corresponding optimization problems of gas transport networks are introduced on the basis of the constructive approach in graph and 2-D system setting.

The simplest graph model is proposed for the first stage modeling to express potentially critical flow/pressure values within the given margins of inflows, outflows and setting of active components such as storage capacity gasholders, compressor stations and others in order to satisfy/optimize the demand distributed over different nodes. At the beginning the graph optimization method is developed for the simple case of stationary network. Next we extend this method for nonstationary case when some characteristic properties of the the network elements can be varied as times goes by. In order to consider simultaneity the pressure and flow gas in the transport networks the two-commodity flow graph model is considered, also.

The 2-D system setting can be used then at the second stage modeling for the more sophisticated and detailed description of dynamic processes in gas pipeline units based on partial differential momentum and continuity equations [74]. The various types of the 2-D models are stated due to the applied discretization schemes to the PDE equations used for the gas behavior description. Also the optimal feedback control problem for 2-D systems is discussed that is of both systems theoretic and application interest.

1.2 Graph models for non steady gas transport networks

It was noted above that the detailed description of the gas transportation through the whole gas network is rather complex. Therefore, a multistage mathematical formalization in terms of the corresponding parameters and suitable variables is a natural way to study the key features of the problem to be interesting for the given level.

In this Section the graph optimization method is developed for the simple case of the non steady (stationary) gas network. This method is based mainly on the constructive optimization approach proposed in [38], [37], [39], [40] for the linear programming problems.

1.2.1 Problem formulation

An approach to design the higher level GTN model is based on the assumption that the network consists of several supply points where the gas is injected into the system, several demand points where the gas flows out of system and other intermediate nodes and storage where the gas is rerouted or stockpiled. Pipelines are represented by arcs(edges) linking the nodes. A fragment of GTN network can be graphically illustrated as in Figure (1.2):

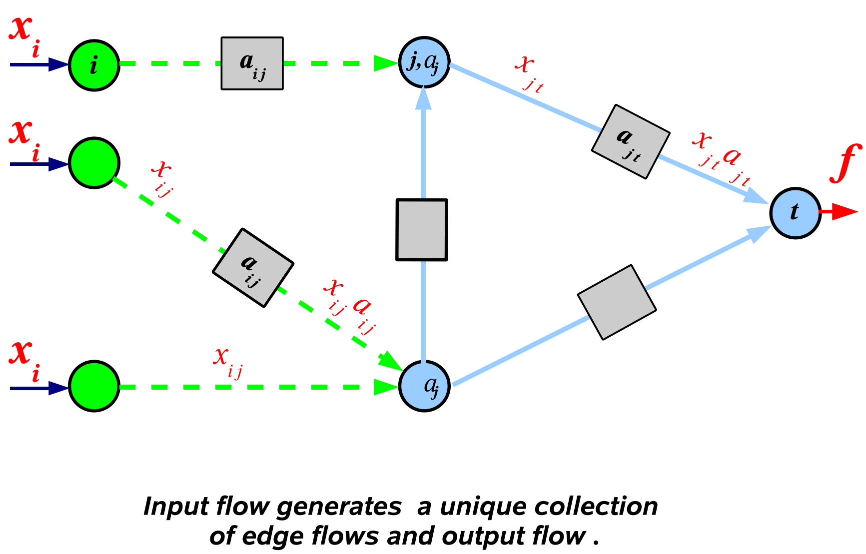

Figure 1.2: Segment of the complex gas network(notations).

In this section we presents the model based on a graph setting which is more attractive for practical implementation at the first stage. In particular, the adaptation of a constructive method of linear programming [37] for the net graph model is presented. This model is treated then as a specific optimal control problem for which ϵ-optimality conditions are given. Such kind of suboptimality conditions are suitable for numerical methods design and present a good tool to realize sensitivity analysis (robustness analysis) of the obtained solution.

We introduce the gas transportation graph network (GTN) model as follows. Let $S \doteq \{I, U\}$ be a stationary net where $I = \{1, 2, \ldots, n\}$ denotes the set of nodes and U denotes the set of edges connecting these nodes. For brevity sake, we suppose that the considered net S has several input sources and one offtake node t, whose output is not used next as input flow. Also, x_i denotes the gas flow in the node i; x_{ij} denotes the gas flow that is transported from the node i to the node j; d^*_{ij}, d_{ij*} denote the upper and lower network throughput gas capacity from the node i to the node j, respectively; $d_i^*,\ d_{i*}$ denote the upper and lower network throughput gas capacity in the node i.

It is convenient to divide the set of nodes I in two subsets $I_\Delta, I_\Sigma, I_\Delta \cap I_\Sigma = \emptyset$ termed as set of input and summation nodes. Hence, the input flow into the gas network is formed by input flows of $x_i, i \in I_\Delta$. The resulting output flow (coming out from final node t)is denoted as f. We assume that each summation node $j \in I_\Sigma$ has several input flows $z_1, z_2, \ldots, z_{q_j}$ and one output flow z such that $\sum_{k=1}^{q_j} z_k + a_j = z$ where a_j denotes the intensity (or available storage capacity) of the node j. In accordance with this partition of nodes, divide the set of edges U in two subsets U_Δ, U_* as follows $U_\Delta = \{(i,j):\ i \in I_\Delta\},\ U_* = U \setminus U_\Delta$. Then for each edge $(i,j) \in U_*$, in addition to the above mentioned throughput capacity $d_{ij}^*,\ d_{ij*}$ of the edges, we introduce another transformation coefficient a_{ij} such that the initial gas flow x_{ij} coming out of node i is transformed into the new gas flow $a_{ij}x_{ij}$ coming into node j. (In the literature, such kind of edges are also termed as active arcs and correspond to pipelines connected to compressor stations). A negative value a_{ij} corresponds to the need of internal consumption. Such assumptions have an obvious physical meaning and make it possible to take into account the real effects of loss and pumping gas in the pipelines.

Let us now briefly describe the problem constraints. For a supply node, the bounds of flow in the network are directly derived from the bounds pre-agreed between the transmission company and the gas producer (on a nominal daily quantity, for example). This implies the following bound constraints on the network inflow:

$$d_{*i} \le x_i \le d_i^*,\ i \in I_\Delta. \tag{1.1}$$

Besides the classical flow balance equations (see (1.6) below) at each node, the nonlinear relation between the flow and the difference of the pressure p_i, p_j at the two ends of each pipe also need to be considered. For the high pressure, this later relation is given by the Weymouth formula for gas (see Osiadacz, 1987) and by the Darcy-Weisbach formula for liquid, and can be expressed as:

$$sign(x_{ij})x_{ij}^2 = C_{ij}^2(p_i^2 - p_j^2), \tag{1.2}$$

where the constants C_{ij} depend on the diameter D_{ij} of the pipe. This allows us to consider the pressure variables and thus to take into account the bounds on pressure. Next we introduce the lower and upper bounds on the pressure p_i at each node $i \in I$. These bounds allow the gas to be delivered, at least, at minimal pressure to the final user and guarantee that the maximal pressure that each producer can provide is not exceeded. This together with relation (1.2) leads to constraints on the throughput capacity of the form:

$$d_{*ij} \le x_{ij} \le d_{ij}^*,\ (i,j) \in U_*. \tag{1.3}$$

In addition, a natural constraint on the transmission capacity of the arches exist and can be joined with (1.3).

Remark 1. *Evidently, there exist more complicated links between pressure and gas flow values that are used for the detailed description of the gas flow in pipes and nodes of the distributed networks. Nevertheless, for the first stage modelling the proposed substitution of the pressure constraints by the generalized flow constraints of (1.3) based on (1.2) are* *feasible, in general. Also, in this Section the so-called two-commodity gas transportation model is introduced. This model gives an ability for simultaneous optimization of pressure and flow variables.*

Standard case is that the cost function presents the common output flow in the form

$$f = \sum_{s \in I_t^-} x_{st} a_{st} + \sum_{j \in I_{\Delta t}} x_j a_{jt} + a_t \to \max \tag{1.4}$$

where

$$I_i^- = \{s \in I_{\sum} : \ (s,i) \in U_*\}, I_{\Delta i} = \{j \in I_\Delta : \ (j,i) \in U\}. \tag{1.5}$$

Finally, the optimization model for gas transportation is given in the following form: maximize the cost function (1.4) over the solution of

$$x_{ij(i)} = \sum_{s \in I_i^-} x_{si} a_{si} + \sum_{j \in I_{\Delta i}} x_j a_{ji} + a_i \tag{1.6}$$

$$d_{*ij} \le x_{ij} \le d_{ij}^*, \ (i,j) \in U_*, d_{*i} \le x_i \le d_i^*, \ i \in I_\Delta \tag{1.7}$$

where $j(i)$ denotes the node connecting with the node i such that $\big(i, j(i)\big) \in U$.

1.2.2 The basic definitions and optimality conditions

Choose some subsets $I_{\Delta supp} \subset I_\Delta, \ I_{\Sigma supp} \subset I_\Sigma \setminus t$. Denote by

$$U_{supp} = U_* \setminus \cup_{i \in I_{\Sigma supp}} (i, j(i)); \ U_{nsupp} = U_* \setminus U_{supp}.$$

Using the given sets, find the solution $\mu_i, \ \ \mu_i \in I_\Sigma$, of the system:

$$\mu_i - a_{ij} \mu_j = 0, \ \ (i,j) \in U_{supp}, \tag{1.8}$$

with the initial condition

$$\mu_t = 1, \ \ \mu_i = 1, \ \ i \in I_{\Sigma supp}. \tag{1.9}$$

Introduce the matrix

$$G_{supp} = \Big(g_{si}, s \in I_{\Sigma supp}, \ i \in I_{\Delta supp} \Big) \tag{1.10}$$

where

$$g_{si} = \sum_{k \in I_{\Delta(i)} \cup I_{\Sigma(s)}} \mu_k a_{ik}, \ s \in I_{\Sigma supp}, \ i \in_{\Delta supp} . \tag{1.11}$$

Denote by $|I|$ the amount of the entries in the integer-valued set I.

Definition 1. *The collection $Q_{supp} = \{I_{\Delta supp}, I_{\Sigma supp}\}$ is called the support of the network S if $|I_{\Delta supp}| = |I_{\Sigma supp}|$ and $\det G_{supp} \neq 0$.*

Introduce the so-called pseudo-flow æ $= \{(x_i, \ i \in I_\Delta); \ x_{ij}, \ (i,j) \in U_*\}$ as a flow which is calculated for the given net S for the given input flow $(x_i, \ i \in I_\Delta)$ that is transformed at the nodes and edges in accordance with given characteristics of the net S, and where the principal constraints (1.7) of the net are omitted.

Remark 2. *The given definition of the support can be equivalently rewritten by the following form.*

The collection of the sets

$$Q_{supp} = \{I_{\Delta supp}, I_{\sum supp}\},\ I_{\Delta supp} \subset I_{\Delta},\ I_{\sum supp} \subset I_{\sum} \setminus t$$

is called the support if in the net S with additional condition $a_i = 0,\ i \in I_{\sum}$ the following equalities

$$x_{ij(i)} = 0, i \in I_{\sum supp};$$
$$x_i = 0, i \in I_{\Delta nsupp},\quad I_{\Delta nsupp} = I_{\Delta} \setminus I_{\Delta supp}$$

holds only for the trivial pseudo-flow æ $= (x_i,\ i \in I_{\Delta};\ x_{ij},\ (i,j) \in U_*)$, but for $\forall i_0 \in I_{\sum supp},\quad \forall j_0 \in I_{\Delta nsupp}$ the equalities

$$x_{ij(i)} = 0, i \in I_{\sum supp};\quad x_i = 0, i \in I_{\Delta nsupp} \setminus j_0$$

or the equalities

$$x_{ij(i)} = 0, i \in I_{\sum supp} \setminus i_0;\quad x_i = 0, i \in I_{\Delta nsupp}$$

are valid for one nontrivial pseudo-flow æ, at least.

Remark 3. *The considered notions extend, in fact, the notions of the so-called support and supporting solution introduced in classic linear programming [39], which, in turn, are an extension of the well known notions of basis and basic solution used in linear programming. In particular, in the simplest case of the LP problem*

$$c^T x \to \max_x,\quad Ax = b,\ x \geq 0,\ x \in \mathbb{R}^n,$$

where A is the given $n \times m$- matrix, c, b are n and m-vectors, the couple of indexes $I_{sup} \subset \{1, 2, ..., n\},\ J_{sup} \subset \{1, 2, ..., m\},\ |I_{sup}| = |J_{sup}|$ is called a support iff the submatrix $A_{sup} \doteq A(I_{sup}, J_{sup})$ of the matrix A is nonsingular $\det A_{sup} \neq 0$. *In this case the equation (rewritten in accordance with index partition)*

$$A_{sup} x_{sup} + A_{nsup} x_{nsup} = b_{sup}$$

possess a nontrivial solution, and, hence, the equation

$$A_{sup} x_{sup} = 0$$

has a unique trivial solution $x_{sup} = 0$.

Next, let z be some admissible flow in the network. Note that this flow can easily be determined by the admissible inputs $x_i, i \in I_{\Delta}$ in accordance with the the given characteristics of the net.

Definition 2. *The pair $\{z, Q_{supp}\}$ that consists of the flow z and the support Q_{supp} of the problem (1.4)—(1.7) is called support flow. The support flow $\{z, Q_{supp}\}$ is called nondegenerate if*

$$d_{i*} < x_i < d_i^*,\ i \in I_{\Delta supp}; d_{ij*} < x_{ij} < d_{ij}^*,\ (i,j) \in U_{supp}.$$

Denote by $c(I_{\Delta supp}) = c_i = \sum_{k \in I_\Delta(i) \cap I_\Sigma(t)} a_{ik}\mu_k, \quad i \in I_{\Delta supp}$ where $\{I_\Sigma(s), U_{supp}(s)\}$ means the connected components of the subset $S_{supp} \doteq \{I_\Sigma, U_{supp}\}$ containing the node $s \in I_{\Sigma_{supp}}$.

For the given support flow $\{z, Q_{supp}\}$ calculate the so-called potentials $y_i, \quad i \in I_{\sum}$:

$$
\begin{gathered}
y_t = -1, \quad y'(I_{\Sigma supp}) = c'(I_{\Delta supp})G^{-1}_{supp}; \\
y_i = y_s\mu_i, \quad i \in I_\Sigma(s) \setminus s, \quad s \in I_{\Sigma supp} \cup t
\end{gathered} \tag{1.12}
$$

and the corresponding estimates:

$$
\begin{gathered}
\Delta_i = \sum_{k \in I_\Delta(i)} a_{ik}y_k, \quad i \in I_{\Delta nsupp}; \\
\Delta_{ij} = -y_i + a_{ij}y_j, \quad (i,j) \in U_{nsupp} = U_* \setminus U_{supp}.
\end{gathered} \tag{1.13}
$$

The obtained estimates are used to formulate the following optimality conditions [37]:

Theorem 1. *The nondegenerate support flow $\{z, Q_{supp}\}$ is optimal for the problem (1.4)—(1.7) if and only if the following conditions are fulfilled*

$$
\begin{gathered}
\Delta_i \geq 0 \quad at \quad x_i = d_{*i}; \quad \Delta_i \leq 0 \quad at \quad x_i = d_i^* \\
\Delta_i = 0 \quad at \quad d_{*i} < x_i < d_i^*, \; i \in I_{\Delta nsupp} \\
\Delta_{ij} \geq 0 \quad at \quad x_{ij} = d_{*ij}, \; \Delta_{ij} \leq 0 \quad at \quad x_{ij} = d_{ij}^*; \\
\Delta_{ij} = 0 \quad at \quad d_{*ij} < x_{ij} < d_{ij}^*, \; (i,j) \in U_{nsupp}.
\end{gathered} \tag{1.14}
$$

1.2.3 Algorithm for optimization method

Usually, in the design of numerical implementation of optimization algorithms we exploit approximate solutions with corresponding error estimation. Hence it is necessary to introduce the 'sub-optimality' concept, as it is often sufficient to stop the numerical computations when a satisfactory accuracy level has been achieved.

Denote by f^0 the optimal criteria function value. Then the suboptimality estimate $\beta(z, Q_{supp}) \doteq f^0 - f$ of the current support flow can be calculated as follows

$$
\begin{gathered}
\beta(z, Q_{supp}) = \\
\sum_{\substack{i \in I_{\Delta nsupp} \\ \Delta_i > 0}} \Delta_i(x_i - d_{*i}) + \sum_{\substack{i \in I_{\Delta nsupp} \\ \Delta_i < 0}} \Delta_i(x_i - d_i^*) + \\
+ \sum_{\substack{(i,j) \in U_{nsupp} \\ \Delta_{ij} > 0}} \Delta_{ij}(x_{ij} - d_{*ij}) + \sum_{\substack{(i,j) \in U_{nsupp} \\ \Delta_{ij} < 0}} \Delta_{ij}(x_{ij} - d_{ij}^*).
\end{gathered} \tag{1.15}
$$

If the given support flow satisfies the optimality criteria or the preassigned suboptimality estimate $\beta(z, Q_{supp}) \leq \epsilon$, then the solution of the our problem stops on this ϵ-optimal flow $z^\epsilon = z$. Otherwise we starting iteration process $\{z, Q_{supp}\} \rightarrow \{\overline{z}, \overline{Q}_{supp}\}$ to improve the current support flow. Each iteration consists of two main parts: changing the flow $z \rightarrow \overline{z}$, and changing the support $Q_{supp} \rightarrow \overline{Q}_{supp}$.

A) *The first part of iteration — changing the flow $z \rightarrow \overline{z}$.*

The new flow is defined as

$$\overline{z} = z + \theta_0 \Delta z \tag{1.16}$$

where Δz denotes the flow improving direction, and θ_0 is the maximal step along direction Δz. The required improving direction is calculated as follows:

$$\Delta z = (\Delta x_i,\ i \in I_\Delta;\ \Delta x_{ij},\ (i,j) \in U_*;\ \Delta f) \tag{1.17}$$

where

$$\begin{gathered} \Delta x_i = d_i^* - x_i \quad at \quad k_i = -1 \\ \Delta x_i = d_{*i} - x_i \quad at \quad k_i = 1, \quad i \in I_{\Delta nsupp} \\ \Delta x_{ij(i)} = d_{ij(i)}^* - x_{ij(i)} \quad at \quad k_i = -1 \\ \Delta x_{ij(i)} = d_{*ij(i)} - x_{ij(i)} \quad at \quad k_i = 1;\ i \in I_{\sum supp} \end{gathered} \tag{1.18}$$

and the vector $k(I_\Delta nsupp) \bigcup I_{\Sigma supp}$ is given as

$$\begin{gathered} k_i = 1 \quad at \quad \Delta_i > 0; \\ k_i = -1 \quad at \quad \Delta_i < 0; \\ k_i = -1 \vee 1 \quad at \quad \Delta_i = 0; \qquad\qquad i \in I_{\Delta nsupp} \\ k_i = 1 \quad at \quad \Delta_{ij(i)} > 0; \\ k_i = -1 \quad at \quad \Delta_{ij(i)} < 0; \\ k_i = -1 \vee 1 \quad at \quad \Delta_{ij(i)} = 0; \qquad\qquad i \in I_{\Sigma supp}. \end{gathered}$$

The remaining components $\Delta x(I_{\Delta supp})$ are given by

$$\Delta x(I_{\Delta supp}) = G_{supp}^{-1} b(I_{\sum supp}) \tag{1.19}$$

where

$$\begin{gathered} b_i = \Delta x_{ij(i)} - \sum_{k \in I_{\sum supp}(i)} \mu_{i(k)} a_{kj(k)} \Delta x_{kj(k)} - \\ - \sum_{j \in I_{\Delta nsupp}} \Delta x_j \sum_{k \in I_\Delta(j) \cap I_\sum(i)} a_{jk} \mu_k,\ i \in I_{\sum supp} \end{gathered} \tag{1.20}$$

and

$$I_{\sum supp}(i) = \{k \in I_{\sum supp} : j(k) \in I_\sum(i)\}. \tag{1.21}$$

The maximal admissible step θ_0 along Δz can be calculated using the standard formulas:

$$\begin{gathered} \theta_0 = \min\{1, \theta_{i_0}, \theta_{i_0 j_0}\} \\ \theta_{i_0} = \min \theta_i,\ i \in I_{\Delta supp}; \theta_i = (d_i^* - x_i)/\Delta x_i \quad at \quad \Delta x_i > 0; \\ \theta_i = (d_{*i} - x_i)/\Delta x_i \quad at \quad \Delta x_i < 0;\ \theta_i = \infty \quad at \quad \Delta x_i = 0, \\ i \in I_{\Delta supp}; \\ \theta_{i_0 j_0} = \min \theta_{ij},\ (i,j) \in U_{supp}; \\ \theta_{ij} = (d_{ij}^* - x_{ij})/\Delta x_{ij} \quad at \quad \Delta x_{ij} > 0; \\ \theta_{ij} = (d_{*ij} - x_{ij})/\Delta x_{ij} \quad at \quad \Delta x_{ij} < 0; \\ \theta_{ij} = \infty \quad at\ \Delta x_{ij} = 0, (i,j) \in U_{supp}. \end{gathered} \tag{1.22}$$

If the new suboptimality estimation $\beta(\overline{z}, Q_{supp}) = (1-\theta_0)\beta(z, Q_{supp})$ is not satisfied then we continue with changing the support $Q_{supp} \rightarrow \overline{Q}_{supp}$.

B) *The second part of iteration — changing the support* $Q_{supp} \rightarrow \overline{Q}_{supp}$.

For this purpose the duality optimization theory and the associated notions are used [37]. In particular, for the optimization problem (1.4)—(1.7) the dual optimization problem is defined as follows: minimize the function

$$q(y) = -\sum_{i\in I_\Sigma} a_i y_i + \sum_{(s,i)\in U_*} (w_{si}d^*_{si} - v_{si}d_{*si}) + \sum_{i\in I_\Delta}(w_i d^*_i - v_i d_{*i}) \rightarrow \min_{y,w,v} \qquad (1.23)$$

subject to

$$y_t = -1, \quad -y_s + a_{si}y_i + w_{si} - v_{si} = 0, \quad w_{si} \geq 0, \quad v_{si} \geq 0, \quad (s,i) \in U_*; \qquad (1.24)$$
$$\sum_{j\in I_\Delta(i)} y_j a_{ij} + w_i + v_i, \quad w_i \geq 0, \quad v_i \geq 0, \quad i \in I_\Delta.$$

The collection $\lambda = \{y_i, i \in I_\Sigma; w_{si}, v_{si}, (s,i) \in U_*; w_i, v_i, i \in I_\Delta\}$ satisfying to the constraints (1.24) is called the dual plan of (1.23)—(1.24).

Each dual plan λ generates the so-called co-flow by the following formulae:

$$\Delta = (\Delta_i, i \in I_\Delta; \Delta_{ij}, (i,j) \in U_*;$$
$$\Delta_i = \sum_{j\in I_\Delta(i)} y_i a_{ij}, \Delta_{ij} = -y_i + a_{ij}y_j).$$

The dual plan λ is called a conforming dual plan if it satisfies the conditions:

$$w_{si} = 0, v_{si} = \Delta_{si} \quad if \quad \Delta_{si} \geq 0; \qquad (1.25)$$
$$w_{si} = -\Delta_{si}, v_{si} = 0 \quad if \quad \Delta_{si} < 0, (s,i) \in U_*; \qquad (1.26)$$
$$w_i = 0, v_i = \Delta_i \quad if \quad \Delta_i \geq 0; \qquad (1.27)$$
$$w_i = -\Delta_i, v_i = 0, \quad if \quad \Delta_i < 0, \quad i \in I_\Delta. \qquad (1.28)$$

It is easy to see that each conforming dual plan is uniquely determined by the potentials $y = (y_i,\ i \in I_\Sigma)$. Therefore, instead the dual plans the potentials are often used (in this case the remaining dual variables are given by (1.25).

The couple $\{\Delta, Q_{supp}\}$ formed by a co-flow Δ and a support Q_{supp} is called a support co-flow.

Next, let us $\{\Delta, Q_{supp}\}$ is the given support co-flow. For the given support co-flow find the associated pseudo-flow as follows

$$\varkappa_i = d_{*i} \quad if \quad k_i = 1; \varkappa_i = d^*_i \quad if \quad k_i = -1, \quad i \in I_{\Delta nsupp};$$
$$\varkappa_{ij(i)} = d_{*i} \quad if \quad k_i = 1; \ \varkappa_{ij(i)} = d^*_i \quad if \quad k_i = -1, \quad i \in I_{\Sigma supp};$$

The remaining components are defined by

$$\varkappa(I_{\Delta_{supp}}) = -G^{-1}_{supp} b(I_{\Sigma_{supp}})$$

where and the vector b is

$$b_i = \sum_{k\in I_\Sigma(i)} a_k\mu_k - \varkappa_{ij(i)} + \sum_{k\in I_{\Sigma supp}(i)} \mu_{j(k)}a_{kj(k)}\varkappa_{kj(k)} + \sum_{j\in I_{\Delta nsupp}} \varkappa_j \sum_{k\ \ I_\Delta(j)\cap I_\Sigma(i)} a_{jk}\mu_k$$

where

$$I_{\Sigma supp}(i) = \{k \in I_{\Sigma supp} : j(k) \in I_{\Sigma}(i)\}.$$

The constructed pseudo-flow is optimal iff the following conditions

$$\varkappa_i = d_{*i} \quad if \quad \Delta_i > 0$$
$$\varkappa_i = d_i^* \quad if \quad \Delta_i d_{*i} \le \varkappa_i \le d_i^* \quad at \quad \Delta_i = 0, \quad i \in I_{\Delta supp};$$
$$\varkappa_{ij(} = d_{*i} \quad if \ \Delta_{ij} > 0;$$
$$\varkappa_{ij} = d_i^* \quad if \quad \Delta_{ij} d_{*ij} \le \varkappa_{ij} \le d_{ij}^* \quad at \quad \Delta_{ij} = 0, \quad (i,j) \in U_{supp}$$

hold.

Thus, the optimal solution for the problem (1.4)—(1.7) in some cases can be estimated directly on the phase of the dual optimization.

Otherwise, we continue the iteration procedure by the following manner.

From (1.22) follows that two situations are possible: $\theta_0 = \theta_{i_0}$ and $\theta_0 = \theta_{i_0 j_0}$.

(i) In the case $\theta_0 = \theta_{i_0}$ the new support $\overline{Q}_{supp} = \{\overline{I}_{\Delta supp}, \overline{I}_{\sum supp}\}$ is modified as follows

$$\overline{I}_{\Delta supp} = (I_{\Delta supp} \setminus i_0) \cup i_{(\nu)}, \ \overline{I}_{\sum supp} = I_{\sum supp}, \ if \quad i_{(\nu)} \in I_{\Delta nsupp} \cup i_0; \tag{1.29}$$
$$\overline{I}_{\Delta supp} = I_{\Delta supp} \setminus i_0, \overline{I}_{\sum supp} = I_{\sum supp} \setminus i_{(\nu)}, \ if \quad i_{(\nu)} \in I_{\sum supp},$$

where the index $1 \le \nu \le p$ is determined from the inequalities $v_{\nu-1} < 0, \quad v_\nu \ge 0$. Here $v_0 = -|\alpha_{i_0}|$ where the value $|\alpha_{i_0}|$ denotes the maximal value among the numbers $\alpha_i, \ \alpha_{ij}$ calculated on the basis of dual variables as follows

$$\begin{cases} \alpha_i = \varkappa_i - d_{*i}, if \Delta_i > 0 \quad or \quad \Delta_i = 0, \ \varkappa_i \alpha_i = \varkappa_i - d_i^*, if \Delta_i d_i^*, \quad i \in I_{\Delta supp}; \\ \alpha_{ij} = \varkappa_{ij} - d_{*ij}, if \Delta)ij > 0 \quad or \quad \Delta_{ij} = 0, \ \varkappa_{ij}\alpha_{ij} = \varkappa_{ij} - d_{ij}^*, if \Delta)ij d_{ij}^*, \quad (i,j) \in U_{supp}. \end{cases}$$

The other values v_k are given by

$$v_k = v_{k-1} + \Delta v_k, \quad k = \overline{1,p}, \tag{1.30}$$

where

$$\Delta v_k = |\delta_{i_{(k)}}|(d^*_{i_{(k)}} - d_{*i_{(k)}}) \quad if \quad i_{(k)} \in I_{\Delta nsupp} \cup i_0; \tag{1.31}$$
$$\Delta v_k = |\delta_{i_{(k)} j_{(k)}}|(d^*_{i_{(k)} j_{(k)}} - d_{*i_{(k)} j_{(k)}}) \quad if \quad i_{(k)} \in I_{\sum supp},$$
$$j_{(k)} = j(i_{(k)}); \sigma_{i_{(p+1)}} = \infty.$$

Here δ_i, δ_{ij} denote the so-called improving direction for dual co-flow Δ_i, Δ_{ij} (in particular, the estimates obtained in (1.13) are also dual co-flows), and are given as

$$\delta_{ij} = -\Delta y_i + a_{ij}\Delta y_j, \ (i,j) \in U_*; \delta_i = \sum_{j \in I_\Delta(i)} \Delta y_j a_{ij}, \ i \in I_\Delta,$$

where

$$\Delta y_t = 0, \quad \Delta y'(I_{\sum supp}) = -e'_{i_0}(I_{\Delta supp}) G^{-1}_{supp} sign \alpha_{i_0}, \tag{1.32}$$
$$\Delta y_i = \Delta y_s \mu_i, \quad i \in I_{\sum}(s) \setminus s, \quad s \in I_{\sum supp} \cup t.$$

and

$$e_{i_0}(I_{\Delta supp}) = (e_i = 0, \ i \in I_{\Delta supp} \setminus i_0, \ e_{i_0} = 1).$$

Finally, the new indices $i_{(\nu)}$ are rearranged in accordance with the increasing order as

$$\sigma_{i_{(1)}} \leq \sigma_{i_{(2)}} \leq ... \leq \sigma_{i_{(p)}}. \tag{1.33}$$

Here the values σ_i (in fact, they are used in the dual optimization problem to determine the maximal step for dual co-flow along the improving direction) are given as

$$\begin{aligned} \sigma_i = \tfrac{-\Delta_i}{\delta_i}, \quad if \quad \delta_i k_i < 0;\ \sigma_i = \infty \quad otherwise, \quad i \in I_{\Delta nsupp}; \\ \sigma_i = \tfrac{-\Delta_{ij(i)}}{\delta_{ij(i)}}, \quad if \quad \delta_{ij(i)} k_i < 0;\ \sigma_i = \infty \quad otherwise, \quad i \in I_{\sum supp}. \end{aligned} \tag{1.34}$$

ii For the case $\theta_0 = \theta_{i_0 j_0}$, the needed calculation is, in its essence, analogous to the one described above. We give them shortly. Putting $\alpha_{i_0 j_0} = \Delta z_{i_0 j_0}(1-\theta_0)$ and

$$\begin{aligned} &\Delta y_t = 0; \\ &\Delta y(I_{\sum supp}) = -(\sum_{k \in I_\Delta(j) \cap I_{\sum}(s_0, i_0)} \frac{a_{jk}\mu_k}{\mu_{i_0}}, j \in I_{\Delta supp}) G^{-1}_{supp} sign\ \alpha_0; \\ &\Delta y_i = \Delta y_s \mu_i,\ i \in I_{\sum}(s),\ s \in (I_{\sum supp} \cup t) \setminus s_0; \\ &\Delta y_i = \Delta y_{s_0} \mu_i,\ i \in I_{\sum}(s_0) \setminus I_{\sum}(s_0, i_0); \\ &\Delta y_i = (\Delta y_{s_0} + \frac{sign\ \alpha_0}{\mu_{i_0}})\mu_i, i \in I_{\sum}(s_0, i_0). \end{aligned} \tag{1.35}$$

we can rearranged in accordance with the increasing order the numbers (1.34)—(1.33). Putting $v_0 = -|\alpha_{i_0 j_0}|$ and using (1.30) to find the v_k, we determine f the index $1 \leq \nu \leq p$ such that $v_{\nu-1} < 0, \quad v_\nu \geq 0$. Then the new support $\overline{Q}_{supp} = \{\overline{I}_{\Delta supp}, \overline{I}_{\sum supp}\}$ and vector $\overline{k}(\overline{I}_{\Delta nsupp} \cup \overline{I}_{\sum supp})$are given

$$\begin{aligned} &\overline{I}_{\Delta supp} = I_{\Delta supp},\ \overline{I}_{\sum supp} = (I_{\sum supp} \setminus i_{(\nu)}) \bigcup i_0,\ if \quad i_\nu \in I_{\sum supp} \\ &\overline{I}_{\Delta supp} = I_{\Delta supp} \bigcup i_{(\nu)}, \overline{I}_{\sum supp} = I_{\sum supp} \bigcup i_0,\ if \quad i_{(\nu)} \in I_{\Delta nsupp}; \end{aligned} \tag{1.36}$$

and

$$\begin{aligned} &\overline{k}_{i_{(s)}} = -k_{i_{(s)}}, \quad s = \overline{1, \nu - 1}; \\ &\overline{k}_i = k_i,\ i \in (I_{\Delta nsupp} \bigcup i_0 \bigcup I_{\sum supp}) \setminus \bigcup_{s=1}^{\nu} i_{(s)}, \end{aligned} \tag{1.37}$$

where $k_{i_0} = sign \Delta_{i_0 j_0} \quad at \quad \Delta_{i_0 j_0} \neq 0; \quad k_{i_0} = \delta_{i_0 j_0} \quad at\ \Delta_{i_0 j_0} = 0.$

Thus, the iteration of the method is complete.

It is not hard task to verify that the suboptimality estimation of the new support flow $\overline{z}, \overline{Q}_{supp}$ is

$$\beta(\overline{z}, \overline{Q}_{supp}) = (1-\theta_0)\beta(z, Q_{supp}) + v_0 \sigma_{i_{(1)}} + \sum_{k=1}^{\nu-1} v_k(\sigma_{i_{k+1}} - \sigma_{i_{(k)}}) \leq \beta(z, Q_{supp}) \tag{1.38}$$

If $\beta(\overline{z}, \overline{Q}_{supp}) \leq \epsilon$ then stop the calculation of the solution of problem (1.4)-(1.7). Otherwise,if $\beta(\overline{z}, \overline{Q}_{supp}) > \epsilon$ then we continue with the new iteration beginning with the obtained currently support flow $\{\overline{z}, \overline{Q}_{supp}\}$ and the vector $\overline{k}(\overline{I}_{\Delta nsupp} \cup \overline{I}_{\sum supp})$.

Remark 4. *It is obviously that the optimization problem (1.4)–(1.7) can be embedded into the general case of linear programming optimization. But such approach ignores the particularities and special form of the considered problem and, hence, leads to a great computational effort and reduces the computational speed. However, the constructive form of the established optimality and ϵ - optimality condition is suitable for numerical methods and present a good tool to realize sensitivity analysis (robustness analysis) of the obtained solution.*

1.2.4 Numerical Example

In the report [54] of the Pipeline Simulation Interest Group of the gas network of Belgium is presented. This network consists of 20 nodes and 24 pipelines connecting these nodes.

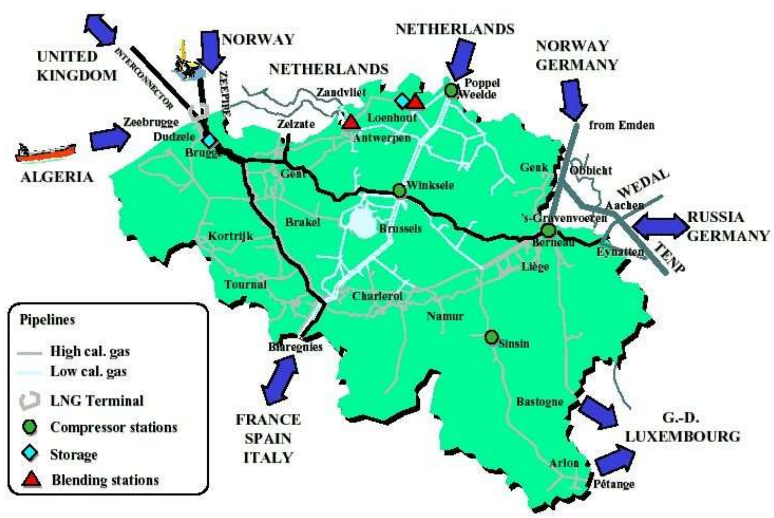

Figure 1.3: Belgian natural gas infrastructure

In order to demonstrate the key moment of the proposed optimization method we give the solution of an illustrative example of a gas network that images (in some sense) the part of the Belgium net with the parameters chosen arbitrary. To simplify our considerations, we restrict ourselves to a network containing 6 nodes and 5 pipelines.

According to our notation we have the net $S = (I, U)$ where $I = \{1, 6, 2, 3, 5, t = (sink\ \ node) = 4\}$ is the collection of nodes,

$U = \{(1, 3), (6, 5), (2, 5), (3, 4), (5, 4)\}$ is the set of edges.

Here the node numbering is cited in the circle together with their intensity a_j for summation nodes, the throughput capacities $d^*_{ij},\ d_{ij*}$ of edges are written under lines, the transformation coefficient a_{ij} of some edges are given in rectangles. Let $a_3 = 1,\ a_4 = a_5 = 0, a_{13} = 3,\ a_{65} = 2,\ a_{25} = 1,\ a_{34} = 2,\ a_{54} = -4;$

$d_{1*} = -1, d^*_1 = 2,\ d_{2*} = -2, d^*_2 = 1,\ d_{6*} = -1, d^*_6 = 2,\ d_{*34} = -2, d^*_{34} = 2, d_{*54} = -1, d^*_{54} = 1$ (see Figure 1.4-1.5). Put $I_\Delta = \{1, 6, 2\},\ I_\Sigma = \{3, 4, 5\},$
$U_\Delta = \{(i, j) \in U : i \in I_\Delta\} = \{(1, 3), (6, 5), (2, 5)\},\ U_* = U \setminus U_\Delta = \{(3, 4), (5, 4)\}.$

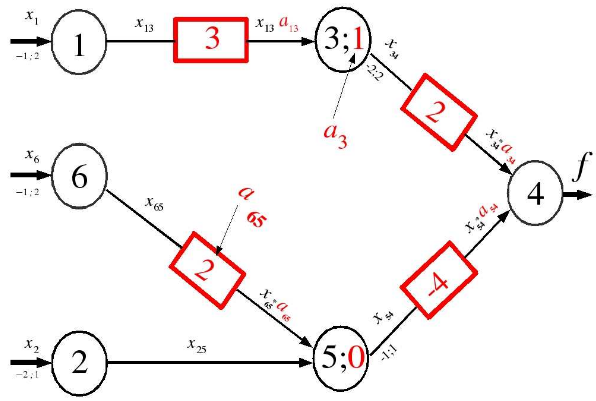

Figure 1.4: Compressors and storages

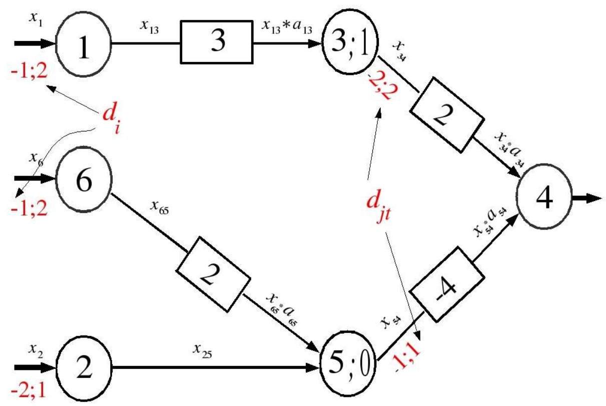

Figure 1.5: Throughput capacities

Note that the negative values for transformation coefficients of a pipeline can be used to formalize the need to accumulate gas in emergency funds, for example.

Below we give the detailed step-by-step procedure to determine the optimal solution. In order to show the most essential and crucial moments of the algorithm we select arbitrary initial data.

1) Select the initial support $Q_{supp} = \{I_{\Delta supp}, I_{\Sigma supp}\} = \{2, 5\}$
where $I_{\Delta supp} = \{2\}, I_{\Sigma supp} = \{5\}$.

Then $I_{\Delta nsupp} = I_\Delta \setminus I_{\Delta supp} = \{1, 6\}$. First, we need to verify that Q_{supp} is a support indeed. For this purpose consider the subnetwork S_{supp} given as follows $S_{supp} = \{I_\Sigma; U_{supp}\} = \{\{3, 4, 5\} - nodes; (3, 4) - edge\}$, where
$U_{supp} = U_* \setminus \bigcup\limits_{i \in I_{\Sigma supp}=5} (i, j(i)) = \{(3, 4), (5, 4) \setminus (5, 4)\} = \{(3, 4)\}$.

Denote also $U_{nsupp} = U_* \setminus U_{supp} = \{(3, 4), (5, 4)\} \setminus \{(3, 4)\} = \{(5, 4)\}$.
It also should be noted that the network $S_{supp} = \{I_\Sigma, U_{supp}\}$ consist of two connected components:

$$\{I_\Sigma(4) = \{3, 4\}, U_{supp}(4) = \{(3, 4)\}\}$$

and

$$\{I_\Sigma(5) = \{5\}, U_{supp}(5) = \{\varnothing\}.$$

Hence $I_\Sigma(5) = \{5\}$ is an isolated node of the subnetwork S_{supp}. In accordance with the

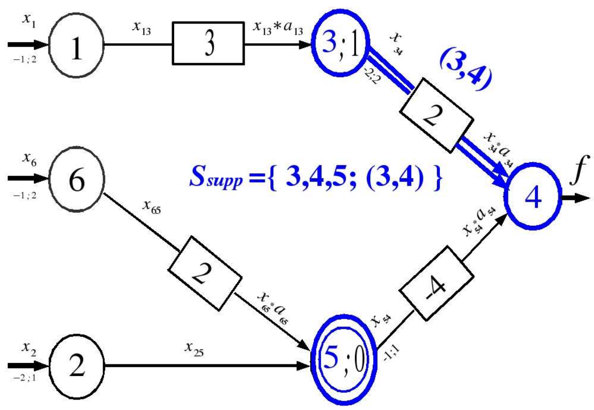

Figure 1.6: Subnetwork S_{supp}

algorithm calculate the coefficients $\mu_i, i \in I_\Sigma = \{3,5\} \bigcup \{t=4\} = \{3,5,4\}$ as follows

$$\begin{aligned} &\mu_t = \mu_4 = 1, \quad \mu_5 = 1 \ \ (i \in I_{\Sigma supp} = \{5\}); \\ &(i,j) \in U_{supp} = (3,4) \Rightarrow \mu_3 - a_{34}\mu_4 = 0 \Rightarrow \mu_3 - 2\mu_4 = 0. \end{aligned} \tag{1.39}$$

Thus, we have

$$\mu_5 = 1, \qquad \mu_4 = 1, \qquad \mu_3 = 2.$$

Next, construct the matrix $G_{supp} = G(I_{\Sigma supp}, I_{\Delta supp}) = (g_{si}, s \in I_{\Sigma supp}, i \in I_{\Delta supp})$. In our case we have $s \in I_{\Sigma supp} = \{5\}$, $i \in I_{\Delta supp} = \{2\}$ and the matrix G_{supp} is (1×1)-matrix with the entry

$$g_{52} = \sum_{k \in I_\Delta(i) \bigcap I_\Sigma(s)} \mu_k a_{2k} = \mu_5 a_{25} = 1 \cdot 1 = 1.$$

Then, in accordance with the support definition, the collection $Q_{supp} = \{2,5\}$ is the support of the network S since:

(i) $|I_{\Delta supp}| = |I_{\Sigma supp}| = 1$, (the notation $|I_{\Delta supp}|$ means the amount of elements in the set $I_{\Delta supp}$, and, hence $|I_{\Delta supp}| = 1$ means that the set $I_{\Delta supp}$ contains one element),
(ii) $detG_{supp} = g_{52} = 1 \neq 0$.

2) Consider now the initial support flow $\{z, Q_{supp}\}$ formed by the flow z corresponding to the initial admissible input flow

that is chosen here randomly, but satisfying the constraints on the throughput capacity $d_i^*,\ d_{i*}, i \in I_\Delta$ and in accordance with Kirchhoff's law such that $z = \{x_i, i \in I_\Delta; x_{ij}, (i,j) \in U_*; f\}$ presents the complete net supporting flow. The symbol f means the resulting output flow of the node $\{t\}$. Namely, we put $x_{input} = \{x_1 = 0, x_6 = 0, x_2 = -1\}$ and then starting this data the remaining values of the net flow we uniquely define in accordance with the given characteristics of the nodes and arcs of the net such that (see Figure 1.7, also) we have

$$\begin{aligned} z = \{x_i, i \in I_\Delta; x_{ij}, (i,j) \in U_*; f\} = \\ \{x_1 = 0, x_6 = 0, x_2 = -1; x_{13} = 0, x_{34} = 0 + 1 = 0, \\ x_{25} = -1, x_{65} = 0, x_{54} = -1, f = 6\}. \end{aligned}$$

3) Verify now the optimality criteria for the given support flow.

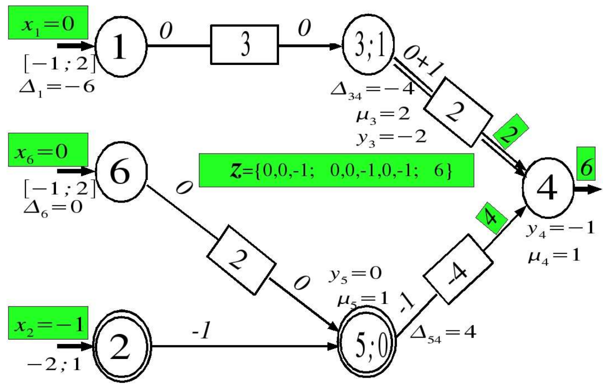

Figure 1.7: Network supporting flow z

At the beginning, for the given support Q_{supp} calculate the vector $c(I_{\Delta supp})$. Since

$$I_\Delta(2)\bigcap I_\Sigma(4) = \{5\}\bigcap\{3,4\} = \varnothing$$

then

$$c(I_{\Delta supp}) = c_2 = \sum_{k\in I_\Delta(2)\bigcap I_\Sigma(4)} \mu_k a_{k2} = 0.$$

Find now the potentials $y_i,\ i\in I_\Sigma$ corresponding to the chosen support Q_{supp} as follows. Then

$$y_t = y_4 = -1,\ y(I_{\Sigma supp}) = y_5 = c_2 g_{52}^{-1} = 0 \to y_5 = 0.$$

Since $s \in I_{\Sigma supp}\bigcup t = \{5,4\}$ then the set $I_\Sigma \setminus s$ is :
a) at $s = 5$ we have $I_\Sigma(5)\setminus\{5\} = \{5\}\setminus\{5\} = \varnothing$;
b) at $s = 4$ we have $I_\Sigma(4)\setminus\{4\} = \{3,4\}\setminus\{4\} = \{3\}$. Hence, for $y_i, i \in I_\Sigma(s)\setminus s,\ s\in I_{\Sigma supp}\bigcup t$ we find:

$$i = I_\Sigma(4)\setminus\{4\} = \{3\}, \to y_3 = y_t\mu_3 = (-1)\cdot 2 = 2.$$

Finally, the required potentials are

$$y_3 = -2, y_4 = -1, y_5 = 0.$$

Next, calculate the following estimations associated with the given support

$$\Delta_i, i\in I_{\Delta nsupp}; \Delta_{ij}, (i,j)\in U_{nsupp},\ \Delta_{ij}, (i,j)\in U_{nsupp}.$$

In our case

$$I_{\Delta nsupp} = I_\Delta \setminus I_{\Delta supp} = \{1,6,2\}\setminus\{2\} = \{1,6\},$$

and

$$\Delta_{ij}, (i,j)\in U_{nsupp} = \{(5,4)\}.$$

Hence, we have

$$\Delta_{54} = -y_5 + a_{54}y_4 = 0 + (-4)\cdot(-1) = 4$$

$$\Delta_6 = \sum_{k\in I_\Delta(6)} a_{6k}y_k = \sum_{k=5} a_{6k}y_k = a_{65}y_5 = 2\cdot 0 = 0,$$

$$\Delta_1 = \sum_{k\in I_\Delta(1)} a_{1k}y_k = \sum_{k=3} a_{1k}y_k = a_{13}y_3 = 3\cdot(-2) = -6.$$

Thus, the asked estimates are

$$\Delta_1 = -6, \Delta_6 = 0, \Delta_{54} = 4.$$

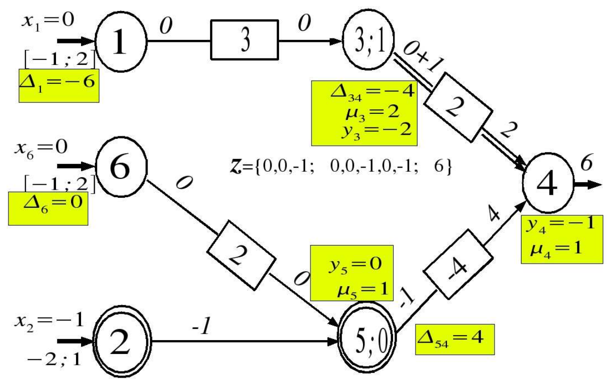

Figure 1.8: Potentials y_i, and corresponding estimations Δ_i

Since $d_{1*} = -1 < x_1 = 0 < d_1^* = 2$, and with the estimate $\Delta_1 = -6 < 0$ for the support flow $\{z, Q_{supp}\}$, reveals that the optimality criteria are not fulfilled. Note that the non-optimality of the current support flow can be determined by other entries, also.

In order to observe the suboptimality estimate of the current support flow $f^o - f \leq \beta$ we find

$$\beta(z, Q_{supp}) = \sum_{\substack{i \in I_{\Delta nsupp} \\ \Delta_i > 0}} \Delta_i(x_i - d_{*i}) + \sum_{\substack{i \in I_{\Delta nsupp} \\ \Delta_i < 0}} \Delta_i(x_i - d_i^*) +$$
$$+ \sum_{\substack{(i,j) \in U_{nsupp} \\ \Delta_{ij} > 0}} \Delta_{ij}(x_{ij} - d_{*ij}) + \sum_{\substack{(i,j) \in U_{nsupp} \\ \Delta_{ij} < 0}} \Delta_{ij}(x_{ij} - d_{ij}^*) =$$
$$= \Delta_1(x_1 - d_1^*) + \Delta_6 \cdot 0 + \Delta_{54}(x_{54} - d_{*54}) =$$
$$-6(0-2) + 0 \cdot 0 + 4(-1-(-1)) = 12 + 4 \cdot 0 = 12.$$

The calculated suboptimality estimates for the considered support flow $\beta(z, Q_{supp})$ shows (see Figure 1.9) that the maximal value f^0 of the flow has not more then

$$f^0 \leq f(x) + \beta(z, Q_{supp}) = 6 + 12 = 18.$$

4) Next realize the iteration $\{z, Q_{supp}\} \to \{\bar{z}, \bar{Q}_{supp}\}$ of the proposed optimization method. This iteration consists in two parts.

A) The first part of the iteration $\{z, Q_{supp}\} \to \{\bar{z}, \bar{Q}_{supp}\}$ consists in changing the flow $z \to \bar{z}$ and given by the following formula

$$\overline{z} = z + \theta_0 \Delta z,$$

where the improvement direction

$$\Delta z = (\Delta x_i,\ i \in I_\Delta;\ \Delta x_{ij},\ (i,j) \in U_*;\ \Delta f)$$

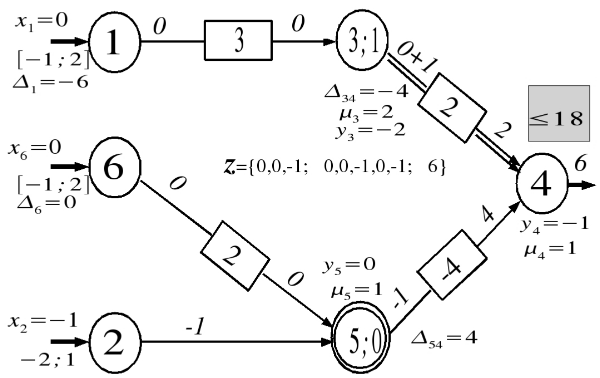

Figure 1.9: Suboptimality estimate $\beta(z, Q_{supp})$ for maximal value of the flow f^0

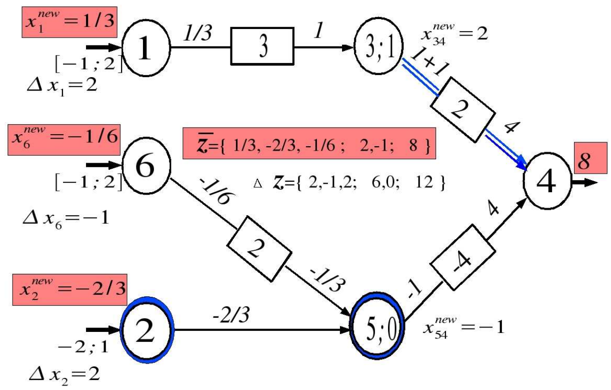

Figure 1.10: New flow $\overline{z}$

are determined as follows

$$\Delta x_i = d_i^* - x_i \quad at \quad k_i = -1; \Delta x_i = d_{*i} - x_i \quad at \quad k_i = 1, \quad i \in I_{\Delta nsupp}$$
$$\Delta x_{ij(i)} = d_{ij(i)}^* - x_{ij(i)} \quad at \quad k_i = -1; \Delta x_{ij(i)} = d_{*ij(i)} - x_{ij(i)} \quad at \quad k_i = 1; \ i \in I_{\sum supp}$$

where the vector $k(I_\Delta nsupp) \bigcup I_{\Sigma supp}$ is defined by the following

$$k_i = 1 \quad at \quad \Delta_i > 0; k_i = -1 \quad at \quad \Delta_i < 0;$$
$$k_i = -1 \vee 1 \quad at \quad \Delta_i = 0 \ i \in I_{\Delta nsupp};$$
$$k_i = 1 \quad at \quad \Delta_{ij(i)} > 0; k_i = -1 \quad at \quad \Delta_{ij(i)} < 0;$$
$$k_i = -1 \vee 1 \quad at \quad \Delta_{ij(i)} = 0; \ i \in I_{\Sigma supp}.$$

In our case $i \in I_{\Delta nsupp} = \{1, 6\}$, $\qquad i \in I_{\Sigma supp} = \{5\}$, and hence

$$\begin{cases} k_1 = -1, \quad since \quad \Delta_1 = -6; \\ k_6 = 1, \quad since \quad \Delta_6 = 0; \\ k_5 = 1, \quad since \quad \Delta_{54} = 4 > 0. \end{cases} \tag{1.40}$$

Calculate now $\Delta x_i, i \in I_{\Delta nsupp}$ and $\Delta x_{ij(i)}, i \in I_{\Sigma supp}$:

$$\Delta x_1 = d_1^* - x_1 = 2 - 0 = 2 \quad since \quad k_1 = -1;$$
$$\Delta x_6 = d_{*6} - x_6 = -1 - 0 = -1 \quad since \quad k_6 = 1;$$
$$\Delta x_{54} = d_{*54} - x_{54} = -1 - (-1) = 0 \quad since \quad k_5 = 1.$$

Further, in order to find the remaining $\Delta x(I_{\Delta supp})$ we need to find the value b_i for the $i \in I_{\Sigma supp} = \{5\}$:

$$b_i = \Delta x_{ij(i)} - \sum_{k \in I_{\Sigma supp}(i)} \mu_{j(k)} a_{kj(k)} \Delta x_{kj(k)} - \\ - \sum_{j \in I_{\Delta nsupp}} \Delta x_j \sum_{k \in I_\Delta(j) \cap I_\Sigma(i)} a_{jk}\mu_k, \ i \in I_{\Sigma supp}$$

where the set $I_{\Sigma supp}(i)$ is determined as

$$I_{\Sigma supp}(i) = \{k \in I_{\Sigma supp} : j(k) \in I_\Sigma(i)\}.$$

Then

$$b_5 = \Delta x_{54} - \sum_{k \in I_{\Sigma supp}(5) = \{5\}} \mu_{j(k)} a_{kj(k)} \Delta x_{kj(k)} - \\ - \sum_{j \in I_{\Delta nsupp} = \{1,6\}} \Delta x_j \sum_{k \in I_\Delta(j) \cap I_\Sigma(i) = \{5\}} a_{jk}\mu_k = \\ = \Delta x_{54} - \mu_4 a_{54} \Delta x_{54} - \Delta x_6 a_{65} \mu_5 = 0 - 1 \cdot (-4) \cdot 0 - (-1) \cdot 2 \cdot 1 = 2.$$

Now we can calculate

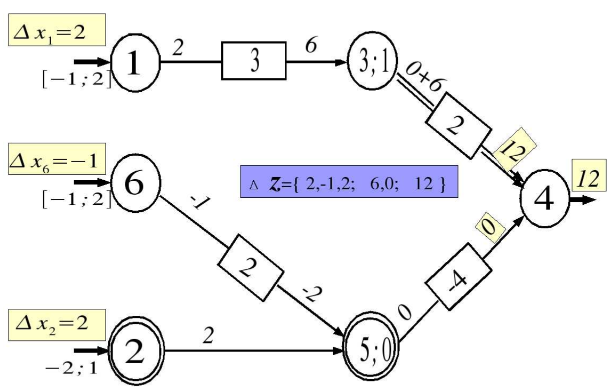

Figure 1.11: Pseudo-flow Δz

$$\Delta x_2 = G_{supp}^{-1} b(I_{\Sigma supp}) = \frac{b_5}{g_{52}} = \frac{2}{1} = 2.$$

Using the input signals $\Delta x_1 = 2, \Delta x_6 = -1, \Delta x_2 = 2$ find the associated pseudo-flow

$$\Delta z = (\Delta x_i, \ i \in I_\Delta; \ \Delta x_{ij}, \ (i,j) \in U_*; \ \Delta f)$$

putting in the net $a_i = 0, \ i \in I_\Sigma$. Note that if the obtained pseudo-flow is admissible and satisfies the optimality conditions [37] then it is required optimal supporting flow. In our case we have $\Delta z = (\Delta x_1 = 2, \Delta x_6 = -1, \Delta x_2 = 2;\ \Delta x_{34} = 2 \cdot 3 + 0 = 6, \Delta x_{54} = 0)$. It is easy to see (see Figure 1.11) the obtained pseudo-flow is not optimal.

Now we determine the maximal available step pace θ_0 along the improvements direction Δz as follows

$$\theta_0 = \min\{1, \theta_{i_0}, \theta_{i_0 j_0}\}$$

where

$$\theta_{i_0} = \min \theta_i,\ i \in I_{\Delta supp}; \theta_i = (d_i^* - x_i)/\Delta x_i \ \ at \ \ \Delta x_i > 0;$$
$$\theta_i = (d_{*i} - x_i)/\Delta x_i \ \ at \ \ \Delta x_i < 0;\ \theta_i = \infty \ \ at \ \ \Delta x_i = 0,\ i \in I_{\Delta supp};$$

and

$$\theta_{i_0 j_0} = \min \theta_{ij},\ (i,j) \in U_{supp};$$
$$\theta_{ij} = (d_{ij}^* - x_{ij})/\Delta x_{ij} \ \ at \ \ \Delta x_{ij} > 0;$$
$$\theta_{ij} = (d_{*ij} - x_{ij})/\Delta x_{ij} \ \ at \ \ \Delta x_{ij} < 0;$$
$$\theta_{ij} = \infty \ \ at\ \Delta x_{ij} = 0, (i,j) \in U_{supp}.$$

Then we have

$$\theta_{i_0} = \min \theta_i,\ i \in I_{\Delta supp} = \{2\} \quad \rightarrow \theta_2 = \frac{d_2^* - x_2}{\Delta x_2} = \frac{1-(-1)}{2} = 1,$$

$$\theta_{i_0 j_0} = \min \theta_{ij},\ (i,j) \in U_{supp} = \{(3,4)\} \quad \rightarrow \theta_{34} = \frac{d_{34}^* - x_{34}}{\Delta x_{34}} = \frac{2-1}{6} = \frac{1}{6}.$$

Hence, the maximal admissible step θ_0 along the Δz is $\theta_{i_0 j_0} = \theta_{34} = \frac{1}{6}$. Therefore, the first part of iteration is completed by the construction of the following new flow

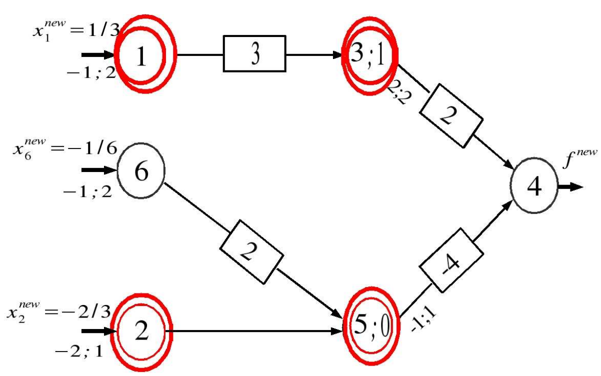

Figure 1.12: New support $\hat{Q}_{supp}$

$$\overline{z} = z + \theta_0 \Delta z = \{\overline{x}_1 = 0 + \tfrac{1}{6} \cdot 2 = \tfrac{1}{3}, \overline{x}_2 = -1 + \tfrac{1}{6} \cdot 2 = -\tfrac{2}{3},$$
$$\overline{x}_6 = 0 + \tfrac{1}{6} \cdot (-1) = -\tfrac{1}{6}; \overline{x}_{34} = 1 + \tfrac{1}{6} \cdot 6 = 2, \overline{x}_{54} = -1 + \tfrac{1}{6} \cdot 0 = -1\}.$$

The realized iteration decreases the suboptimality estimate as follows

$$\overline{\beta}(\overline{z}, Q_{oldsupp}) = (1 - \theta_0)\beta = \frac{5}{6} \cdot 12 = 10,$$

and, hence $f^o \leq f + \overline{\beta} = 6 + 10 = 16$

B) The second part of the iteration (changing support set $Q_{supp} \to \overline{Q}_{supp}$).

For this purpose introduce the dual optimization problem

$$q(y) = -\sum_{i \in I_\Sigma} a_i y_i + \sum_{(s,i) \in U_*} (w_{si} d^*_{si} - v_{si} d_{*si}) + \sum_{i \in I_\Delta} (w_i d^*_i - v_i d_{*i}) \to \min$$

$$y_t = -1, \quad -y_s + a_{si} y_i + w_{si} - v_{si} = 0, \quad w_{si} \geq 0, \quad v_{si} \geq 0, \quad (s,i) \in U_*;$$

$$\sum_{j \in I_\Delta(i)} y_j a_{ij} + w_i + v_i, \quad w_i \geq 0, \quad v_i \geq 0, \quad i \in I_\Delta.$$

The collection $\lambda = \{y_i, i \in I_\Sigma; w_{si}, v_{si}, (s,i) \in U_*; w_i, v_i, i \in I_\Delta\}$ is called a dual plan for the dual problem. Find the variables of this plan $\lambda = \lambda(y, w, v)$ as follows

$$y_4 = -1,$$

$$\begin{cases} (s,i) \in U_* = \{(3,4),(5,4)\} \\ i \in I_\Delta = \{1,2,6\} \\ j \in I_\Delta(i), \quad where \quad i \in I_\Delta \\ I_\Delta(1) = \{3\}, \\ I_\Delta(2) = \{5\}, \\ I_\Delta(6) = \{5\}. \end{cases}$$

$$\Downarrow$$

$$\begin{cases} -y_5 + a_{54} y_4 + w_{54} - v_{54} = 0, \\ -y_3 + a_{34} y_4 + w_{34} - v_{34} = 0, \\ y_3 a_{13} + w_1 - v_1 = 0, \\ y_5 a_{25} + w_2 - v_2 = 0, \\ y_5 a_{65} + w_6 - v_6 = 0, \end{cases} \Rightarrow \begin{cases} -y_5 + (-4)(-1) + w_{54} - v_{54} = 0, \\ -y_3 + 2(-1) + w_{34} - v_{34} = 0, \\ 3y_3 + w_1 - v_1 = 0, \\ y_5 + w_2 - v_2 = 0, \\ 6y_5 + w_6 - v_6 = 0, \end{cases}$$

From the last system we can, finally, calculate the conforming dual plan. Taking into account $w_{si} \geq 0, v_{si} \geq 0, (s,i) \in U_*,\ w_i \geq 0, v_i \geq 0, i \in I_\Delta$)and using the calculated early the potentials, we have

$$\lambda = (y_3 = -2, y_4 = -1, y_5 = 0; w_{34} = 4, v_{34} = 0, w_{54} = 0, v_{54} = 4,$$

$$v_1 = 0, w_1 = 6, v_2 = 0, w_2 = 0, v_6 = 0, w_6 = 0.)$$

Next, using the obtained dual plan λ we find the so-called co-flow as

$$\Delta = (\Delta_i, i \in I_\Delta; \Delta_{ij}, (i,j) \in U_*;\ \Delta_i = \sum_{j \in I_\Delta(i)} y_i a_{ij}, \Delta_{ij} = -y_i + a_{ij} y_j).$$

Thus, in our case we have

$$\Delta = (\Delta_1 = -6, \Delta_2 = 0, \Delta_6 = 0, \Delta_{34} = -4, \Delta_{54} = 4).$$

Improve now the co-flow by the formula

$$\Delta \to \overline{\Delta} = \Delta + \sigma\delta.$$

For this purpose we need to construct pseudo-flow $\varkappa$. Also it should be reminding that vector $k(I_{\Delta nsupp} \bigcup I_{\Sigma supp}) = (\{1,6,5\})$ has the following entries

$$k_1 = -1, \quad k_6 = 1, \quad k_5 = 1.$$

Hence the pseudo-flow is given as

$$\begin{cases} i \in I_{\Delta nsupp} = \{1,6\} \\ \\ i \in I_{\Sigma supp} = \{5\} \end{cases} \Rightarrow \begin{cases} \varkappa_1 \overset{(k_1=-1)}{=} d_1^* = 2, \\ \varkappa_6 \overset{(k_6=1)}{=} d_{*6} = -1, \\ \varkappa_{54} \overset{(k_5=1)}{=} d_{*54} = -1; \end{cases}$$

Since

$$b_i = \sum_{k\in I_\Sigma(i)} a_k\mu_k - \varkappa_{ij(i)} + \sum_{k\in I_{\Sigma supp}(i)} \mu_{j(k)}a_{kj(k)}\varkappa_{kj(k)} + \sum_{j\in I_{\Delta nsupp}} \varkappa_j \sum_{k\ I_\Delta(j)\cap I_\Sigma(i)} a_{jk}\mu_k =$$

$$= \sum_{I_\Sigma(5)=\{5\}} a_5\mu_5 - \varkappa_{54} + \sum_{k\in I_{\Sigma supp}(5)=\{5\}\ or\ \varnothing} \mu_4 a_{54}\varkappa_{54} + \varkappa_1 \sum_{I_\Delta(1)\cap I_\Sigma(5)=\varnothing} (\varnothing) +$$

$$+\varkappa_6 \sum_{I_\Delta(6)\cap I_\Sigma(5)=\{5\}} a_{65}\mu_5 =$$

$$= 0\cdot 1 - (-1) + 1\cdot(-4)\cdot(-1) + \varnothing + (-1)\cdot 2\cdot 1 = 1+4-2 = 3.$$

them $\varkappa_2 = -g_{25}^{-1} = -3$. Thus, the asked pseudo-flow is

$$\varkappa_1 = 2,\ \varkappa_2 = -3,\ \varkappa_6 = -1.$$

Check now the optimality condition for the obtained pseudo-flow. We have $\varkappa_2 = -3 < d_{*2} = -2$ is out of constraint limit. Therefore, $\varkappa_i$ is not optimal flow.

Next calculate the numbers α_i corresponding to the edges $(i,j) \in U_{supp}$ and to node $i \in I_{\Delta supp}$.

$$\begin{cases} \alpha_i = \varkappa_i - d_{*i}, if \Delta_i > 0 \quad or \quad \Delta_i = 0,\ \varkappa_i < d_{*i}, \\ \alpha_i = \varkappa_i - d_i^*, if \Delta_i < 0 \quad or \quad \Delta_i = 0,\ \varkappa_i > d_i^*, \quad i \in I_{\Delta supp}; \\ \\ \alpha_{ij} = \varkappa_{ij} - d_{*ij}, if \Delta)ij > 0 \quad or \quad \Delta_{ij} = 0,\ \varkappa_{ij} < d_{*ij}, \\ \alpha_{ij} = \varkappa_{ij} - d_{ij}^*, if \Delta)ij < 0 \quad or \quad \Delta_{ij} = 0,\ \varkappa_{ij} > d_{ij}^*, \quad (i,j) \in U_{supp}. \end{cases}$$

and find among them the maximal one (by absolutely value)

$$\begin{cases} i \in I_{\Delta supp} = \{2\} \to \alpha_2 \overset{\Delta_2=0}{=} \varkappa_2 - d_{*2} = -3-(-2) = -1 \\ \\ (i,j) \in U_{supp} = \{(3,4)\} \to \alpha_{34} \overset{\Delta_{34}=-4<0}{=} \varkappa_{34} - d_{34}^* = 6-2 = 4. \end{cases}$$

Thus

$$\alpha_0 = \max\{\alpha_2, \alpha_{34}\} = 4 \to (i_0, j_0) = (3,4) \in U_{supp}.$$

In general, the two cases are possible:
(i) $\alpha_0 = \alpha_{i_0}, i_0 \in I_{\Delta supp}$; and (ii) $\alpha_0 = \alpha_{i_0,j_0}, (i_0,j_0) \in U_{supp}$.

We have here the case when $\alpha_0 = \alpha_{i_0,j_0} = (\alpha_{34})$ and, hence Let $(i_0, j_0) \in U_{supp}(s_0), s \in I_{\Sigma supp} \bigcup t$.

Our purpose is to find the new co-flow, on the base of which we can then change the old support set. For this we find first the new potentials as follows

$$\begin{cases} y_t = 0; \\ \Delta y(I_{\Sigma supp}) = -(\sum\limits_{k\in I_\Delta(j)\cap I_\Sigma(s_0,i_0)} a_{jk}\mu_k \setminus \mu_{i_0})G^{-1}_{supp} sign\alpha_0, \\ j \in I_{\Delta supp}; \\ \Delta y_i = \Delta y_s \mu_i, i \in I_\Sigma(s), s \in (I_{\Sigma supp} \cup t) \setminus s_0; \\ \Delta y_i = \Delta y_{s_0}\mu_i, i \in I_\Sigma(s_0) \setminus I_\Sigma(s_0, i_0); \\ \Delta y_i = (\Delta y_{s_0} + \frac{sign\alpha_0}{\mu_{i_0}})\mu_i,\ i \in I_\Sigma(s_0, i_0). \end{cases}$$

and hence

$$\begin{cases} y_4 = 0; \\ \Delta y(I_{\Sigma supp}) = \Delta y_5 = -\sum\limits_{I_\Delta(2)\cap I_\Sigma(4,3)=\varnothing} (\varnothing) = 0; \\ \\ i \in I_\Sigma(s_0, i_0) = I_{\Sigma(3,4)} = \{3\}, \\ \Delta y_3 = (\Delta y_4 + \frac{sign\alpha_0}{\mu_3})\mu_3 = 0 + \frac{1}{2}\cdot 2 = 1. \end{cases}$$

Here $I_\Sigma(s_0, i_0)$ is the set of nodes such that the connected component of the net $\{I_\Sigma(s_0), U_{supp}(s_0) \setminus (i_0, j_0)\}$ contains the node i_0. In our case for $s_0 = 4$ the set $I_\Sigma(s_0, i_0)$ is the set of nodes of the connected component of the net $\{I_\Sigma(4), U_{supp}(4) \setminus \{(3, 4)\}\} = \{3, 4; \varnothing\}$ containing node $i_0 = \{3\}$.

Thus

$$\Delta y_4 = 0, \quad \Delta y_5 = 0, \quad \Delta y_3 = 1.$$

Now we need to find the direction δ along that we can improve the co-flow Δ:

$$\delta_i = 0, i \in I_{\Delta supp};\ \delta_{ij} = 0, (i, j) \in U_{supp} \setminus (i_0, j_0);\ \delta_{i_0 j_0} = -sign\alpha_{i_0 j_0}.$$

and

$$\begin{cases} i \in I_{\Delta supp} = \{2\} \to \delta_2 = 0; \\ \\ i \in I_{\Delta nsupp} = \{1, 6\}, \\ \delta_i = \sum\limits_{j\in I_\Delta(i)} \Delta y_j a_{ij}, i \in I_\Delta; \\ \\ (i, j) \in U_* = \{(5, 4), (3, 4)\}, \\ \delta_{ij} = -\Delta y_i + a_{ij}\Delta y_j; \\ \delta_{i_0 j_0} = -sign\alpha_{i_0 j_0}. \end{cases} \Rightarrow \begin{cases} \delta_2 = 0; \\ \\ \delta_1 = \sum\limits_{j\in I_\Delta(1)=\{3\}} \Delta y_3 a_{13} = 1\cdot 3 = 3, \\ \delta_6 = \sum\limits_{j\in I_\Delta(6)=\{5\}} \Delta y_5 a_{65} = 0\cdot 2 = 0; \\ \\ \delta_{54} = \Delta y_5 + a_{54}\Delta y_4 = 0 + 0 = 0; \\ \\ \delta_{34} = -sign\alpha_{34} = -sign4 = -1. \end{cases}$$

Thus, we obtain

$$\delta_1 = 3, \delta_6 = 0, \delta_2 = 0, \delta_{54} = 0, \delta_{34} = -1.$$

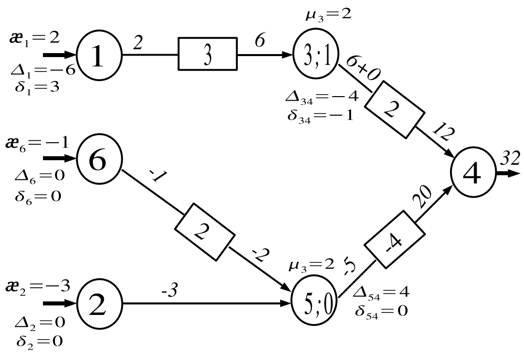

Figure 1.13: Intermediate data

It is convenient to image the obtained data on the Figure (1.13):

Next, we need to find the admissible step pace σ along the obtained improvement direction

$$\begin{gathered}
\sigma_i = \tfrac{\Delta_i}{\delta_i} \quad if \quad \delta_i k_i < 0; \\
\sigma_i = \infty \quad if \quad i \in I_{\Delta nsupp}; \\
\sigma_i = -\tfrac{\Delta_{ij(i)}}{\delta_{ij(i)}} \quad \delta_{ij(i)} k_i < 0; \\
\sigma_i = \infty \quad otherwise, \quad i \in I_{\Sigma supp}; \\
\sigma_{i_0} = |\Delta_{i_0}| \quad at \quad \Delta_{i_0} \alpha_{i_0} > 0; \\
\sigma = \infty \quad at \quad \alpha_{i_0} \Delta_{i_0} \leq 0.
\end{gathered}$$

$$\begin{cases} i \in I_{\Delta nsupp} = \{1, 6\}, \\ \\ i \in I_{\Sigma supp} = \{5\}, \\ \\ i_0 = \{3\}. \end{cases} \Rightarrow \begin{cases} \sigma_1 = -\frac{\Delta_1}{\delta_1} = 2; \\ \sigma_6 = +\infty; \\ \\ \sigma_5 = +\infty; \\ \\ \sigma = +\infty \quad since \quad \Delta_{34} \cdot \alpha_{34} = (-1) \cdot 4 = -16 < 0. \end{cases}$$

Rearrange the calculated values in the required order: $\sigma_{i_{(1)}} \leq \sigma_{i_{(2)}} \leq \sigma_{i_{(2)}}$. Here $i_{(1)} = 1$. In accordance with the duality theory the velocity of decreasing dual cost function for each subintervals of $\sigma_{i_{(1)}} \leq \sigma_{i_{(2)}} \leq \sigma_{i_{(2)}}$ is calculated as follows

$$v_0 = -|\alpha_0| = -|\alpha_{34}| = -4 < 0,$$
$$v_1 = v_0 + \Delta v_1 = -4 + 9 = 5; \Delta v_1 \overset{i_{(1)}=1}{=} |\delta_{i_{(1)}}|(d^*_{i_{(1)}} - d_{*i_{(1)}}) = 3 \cdot (2 - (-1)) = 9.$$

Since $v_0 = -4 < 0$ and $v_1 = 5 > 0$ then $\nu = 1$, and the the maximal admissible step pace is

$$\sigma_0 = \sigma_{i_{(1)}} = \sigma_1 = 2.$$

Finally, we can construct the new co-flow $\hat{\Delta} = \Delta + \sigma_0 \cdot \delta$ and the new support set

$\hat{Q}_{supp} = \{\hat{I}_{\Delta supp}, \hat{I}_{\Sigma supp}\}$ by the following

$$\hat{I}_{\Delta supp} = I_{\Delta supp}, \quad \hat{I}_{\Sigma supp} = (I_{\Sigma supp} \setminus i_{(\nu)}) \bigcup i_0, \quad if \quad i_{(\nu)} \in I_{\Sigma supp};$$
$$\hat{I}_{\Delta supp} = I_{\Delta supp} \bigcup i_{(\nu)}, \quad \hat{I}_{\Sigma supp} = I_{\Sigma supp} \bigcup i_0, \quad if \ \imath_{(\nu)} \in I_{\Delta nsupp};$$
$$\hat{k}_{i_{(s)}} = -k_{i_{(s)}}, s \in \overline{1, \nu - 1}; \hat{k}_i = \quad i \in (I_{\Delta supp} \bigcup I_{\Sigma supp} \bigcup i_0) \setminus \bigcup_{s=1}^{\nu} i_{(s)},$$
$$where$$
$$k_{i_0} = sign \Delta_{i_0 j_0} \quad at \quad \Delta_{i_0 j_0} \neq 0;$$
$$k_{i_0} = \delta_{i_0 j_0}, \quad at \quad \Delta_{i_0 j_0} = 0.$$

Thus

$$\hat{I}_{\Delta supp} = I_{\Delta supp} \cup i_{(\nu)} = \{2\} \cup \{1\} = \{1, 2\},$$
$$\hat{I}_{\Sigma supp} = I_{\Sigma supp} \cup i_0 = \{5\} \cup \{3\} = \{3, 5\}.$$

$$\begin{cases} \hat{\Delta}_1 = \Delta_1 + \sigma_0 \delta_1 = -6 + 2 \cdot 3 = 0, \\ \hat{\Delta}_2 = \Delta_2 + \sigma_0 \delta_2 = 0 + 0 = 0, \\ \hat{\Delta}_6 = \Delta_6 + \sigma_0 \delta_6 = 0 + 0 = 0, \\ \hat{\Delta}_{34} = \Delta_{34} + \sigma_0 \delta_{34} = -4 + 2 \cdot (-1) = -6, \\ \hat{\Delta}_{54} = \Delta_{54} + \sigma_0 \delta_{54} = 4 + 0 = 4. \end{cases}$$

In order to construct the new pseudo-flow $\overline{\varkappa}$, corresponding to the new support co-flow $\{\hat{\Delta}, \hat{Q}_{supp}\}$ we are needed to determine the new numbers $\hat{k}(\hat{I}_{\Delta nsupp} \cup \hat{I}_{\Sigma supp}) = \hat{k}(\{6\} \cup \{3, 5\}) : \hat{k}_6 = 1, \quad \hat{k}_3 = -1, \quad \hat{k}_5 = 1.$

Then the new pseudo-flow $\overline{\varkappa}$ corresponding to the new support co-flow $\{\hat{\Delta}, \hat{Q}_{supp}\}$ are defined as

$$\varkappa_i = d_{*i} \quad if \quad k_i = 1,$$
$$\varkappa_i = d_i^* \quad if \quad k_i = -1, \quad i \in I_{\Delta nsupp};$$
$$\varkappa_{ij(i)} = d_{*i} \quad if \quad k_i = 1,$$
$$\varkappa_{ij(i)} = d_i^* \quad if \quad k_i = -1, \quad i \in I_{Sigmasupp};$$
$$and \quad for \quad b_i :$$
$$b_i = \sum_{k \in I_\Sigma(i)} a_k \mu_k - \varkappa_{ij(i)} +$$
$$+ \sum_{k \in I_{\Sigma supp}(i)} \mu_{j(k)} a_{kj(k)} \varkappa_{kj(k)} + \sum_{j \in I_{\Delta nsupp}} \varkappa_j \sum_{k \ I_\Delta(j) \cap I_\Sigma(i)} a_{jk} \mu_k$$
$$where$$
$$I_{\Sigma supp}(i) = \{k \in I_{\Sigma supp} : j(k) \in I_\Sigma(i)\}.$$

In our case we have

$$\begin{cases} i \in \hat{I}_{\Delta nsupp} = \{6\} \\ \\ i \in \hat{I}_{\Sigma supp} = \{5, 3\} \end{cases} \Rightarrow \begin{cases} \hat{\varkappa}_6 \overset{(\hat{k}_6 = 1)}{=} d_{*6} = -1, \\ \hat{\varkappa}_{34} \overset{(\hat{k}_3 = -1 < 0)}{=} d_{34}^* = 2, \\ \hat{\varkappa}_{54} \overset{(\hat{k}_5 = 1)}{=} d_{*54} = -1; \end{cases}$$

and for b_i:

$$i \in \hat{I}_{\Sigma supp} = \{5, 3\}:$$
$$\hat{b}_3 = a_3 \cdot \hat{\mu}_3 - \hat{\varkappa}_{34} = 1 \cdot 1 - 2 = -1;$$
$$\hat{b}_5 = \sum_{I_\Sigma(5)=\{5\}} a_5 \hat{\mu}_5 - \hat{\varkappa}_{54} + \sum_{\hat{I}_{\Delta nsupp}=\{6\}} \hat{\varkappa}_6 \cdot \sum_{\hat{I}_\Delta(6)\cap \hat{I}_\Sigma(5)=\{5\}} a_{65}\hat{\mu}_5 =$$
$$= 0 \cdot 1 - (-1) + (-1) \cdot 2 \cdot 1 = -1$$

where the new coefficients $\hat{\mu}_i, i \in I_\Sigma$ are

$$\hat{\mu}_4 = \mu_t = 1, \hat{\mu}_i = 1,\ \hat{\mu}_3 = \hat{\mu}_4 = \hat{\mu}_5 = 1.$$

Hence, the reminding $\hat{\varkappa}(\hat{I}_{\Delta supp}) = \hat{G}^{-1}_{supp}\hat{b}(\hat{I}_{\Sigma supp})$ are given

$$g_{31} = \sum_{k\in \hat{I}_\Delta(1)\cap \hat{I}_\Sigma(3)} \hat{\mu}_3 \cdot a_{13} = 3;$$
$$g_{32} = \sum_{k\in \hat{I}_\Delta(2)\cap \hat{I}_\Sigma(3)=\varnothing} (......) = 0;$$
$$g_{51} = \sum_{k\in \hat{I}_\Delta(1)\cap \hat{I}_\Sigma(5)=\varnothing} (......) = 0;$$
$$g_{52} = \sum_{k\in \hat{I}_\Delta(2)\cap \hat{I}_\Sigma(5)} \hat{\mu}_5 \cdot a_{25} = 1.$$

Then

$$\hat{G}_{supp} = \begin{bmatrix} g_{13} & g_{32} \\ g_{51} & g52 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \Rightarrow \hat{G}^{-1}_{supp} = \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & 1 \end{bmatrix}.$$

and the entries of the new pseudo-flow are

$$\begin{bmatrix} \hat{\varkappa}_1 \\ \hat{\varkappa}_2 \end{bmatrix} = -\hat{G}^{-1}_{supp} \cdot \begin{bmatrix} \hat{b}_3 \\ \hat{b}_5 \end{bmatrix} = -\begin{bmatrix} \frac{1}{3} & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ -1 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ 1 \end{bmatrix}.$$

In final, the new pseudo-flow are:

$$\hat{\varkappa}_1 = \frac{1}{3},\ \hat{\varkappa}_2 = 1,\ \hat{\varkappa}_6 = -1,\ \hat{\varkappa}_{34} = 2,\ \hat{\varkappa}_{54} = -1.$$

Check now the obtained pseudo-flow for their optimality conditions of the form

$$\hat{\varkappa}_i = d_{*i} \quad if \quad \hat{\Delta}_i > 0$$
$$\hat{\varkappa}_i = d^*_i \quad if \quad \hat{\Delta}_i < 0;$$
$$d_{*i} \le \hat{\varkappa}_i \le d^*_i \quad at \quad \hat{\Delta}_i = 0, \quad i \in \hat{I}_{\Delta supp};$$

$$\hat{\varkappa}_{ij(} = d_{*i} \quad if \quad \hat{\Delta}_{ij} > 0;$$
$$\hat{\varkappa}_{ij} = d^*_i \quad if \quad \hat{\Delta}_{ij} < 0;$$
$$d_{*ij} \le \hat{\varkappa}_{ij} \le d^*_{ij} \quad at \quad \Delta_{ij} = 0, \quad (i,j) \in \hat{U}_{supp}.$$

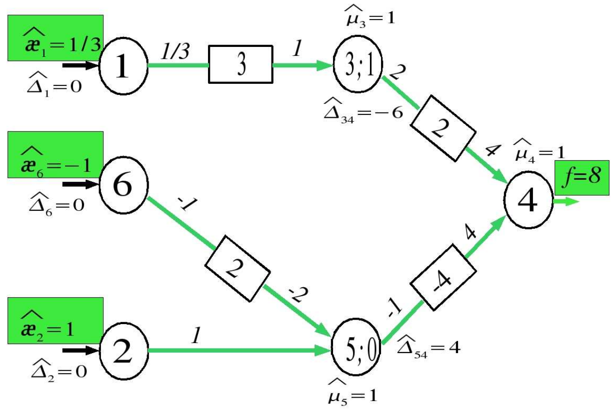

Figure 1.14: Maximal flow in the network

In our case we have:

$$\begin{cases} i \in \hat{I}_{\Delta supp} = \{1, 2\} \\ \hat{U}_{supp} = U^* \setminus \bigcup\limits_{i \in \hat{I}_{\Sigma supp}} (i, j(i)) = \\ = \{(3,4),(5,4)\} \setminus \{(3,4),(5,4)\} = \varnothing \end{cases} \qquad \begin{cases} \hat{\varkappa}_1 = \frac{1}{3} \;\; (i.e. \;\; -1 < \frac{1}{3} < 2) \;\; and \;\; \hat{\Delta}_1 = 0, \\ \hat{\varkappa}_2 = 1 \;\; (i.e. \;\; d_2^* = 1) \;\; and \;\; \hat{\Delta}_2 = 0. \end{cases}$$

The last equations means that optimality criteria condition are fulfilled.

Thus

$$\hat{\varkappa} = z^o = \{x_1 = \frac{1}{3},\; x_2 = 1,\; x_6 = -1,\; x_{34} = 2,\; x_{54} = -1\}$$

is an optimal solution of the considered optimization problem.

Remark 5. *The optimality of the obtained solution can be stated also by the optimality conditions formulated in Theorem 1 applied for the new support set and support flow.*

1.3 Two-commodity flow gas networks

In the previous paragraphs the gas network models are given in the form where the volume of the transported gas was the subject for optimization problem. Nevertheless, there is another crucial characteristic such as flow pressure for which the predefined demands should can be kept, too. It is obvious that an analogous optimization model can be separately introduced for each indicated characteristics. It is of a great interest to optimize several characteristics in frame of an unique mathematical model which is an aim of the present subsection. In particular, in order to consider simultaneously the pressure and flow gas in the transport networks the two-commodity flow graph model is proposed.

1.3.1 Problem formulation

We introduce the two-commodity flow gas transportation graph network model as follows. Let $S \doteq \{I, U\}$ be a stationary net where $I = \{1, 2, \ldots, n\}$ denotes the set of nodes, U means the set of edges. In contrast to the previous section here we assume that through the considered net two kind of signals are running. The signals of such type can be selected from some characteristics of the transported gas. For example, the values of the pressure and volume of the transported gas flow. It is obvious that these signals are transformed by the different physical law due to their transportation through distributed gas network. These effects should be imaged in the mathematical model, in general.

Also, let x_i^1, x_i^2 denotes the gas flow (characterized by two signals x_i^1 and x_i^2) in the node i; x_{ij}^1, x_{ij}^2 denote the gas flow that is transported from the node i to the node j; d_{ij}^{1*}, d_{ij*}^1, d_{ij}^{2*}, d_{ij*}^2 denote the upper and lower bounds for the considered signals throughput from the node i to the node j, respectively; j; d_i^{1*}, d_{i*}^1, d_i^{2*}, d_{i*}^2 denote the upper and lower network bounds for gas in the node i.

The set of nodes I is convenient to divide on two subsets $I_\Sigma, I_\prod$, $I_\Sigma \cap I_\prod = \emptyset$ termed as summations and multiplications nodes respectively. Such partition has a physical meaning.

Each multiplications node $i \in I_\prod$ possesses the following property: it has one input flow $z \doteq (x^1, x^2)$ and several output flows $z_1 = (x_1^1, x_1^2), \ldots, z_{q_i} = (x_{q_j}^1, x_{q_j}^2)$ (see Figure 1.15) .

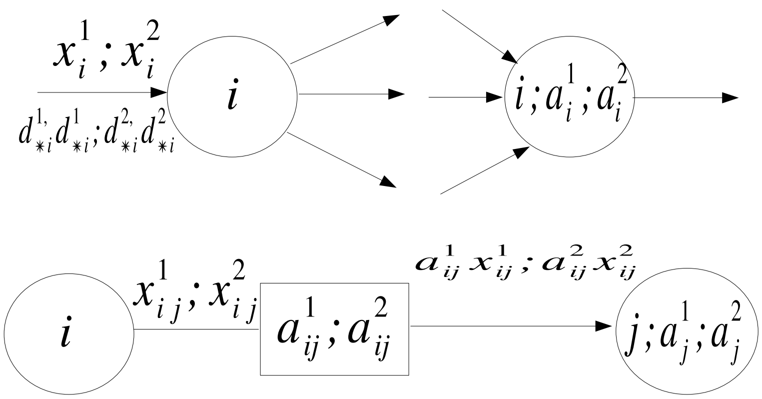

Figure 1.15: The notations of two commodity model

The components of these outputs are characterized by different properties: the first components $(x_1^1, ..., x_{q_i}^1)$ of the outputs $z_1, z_2, \ldots, z_{q_i}$ satisfies the equality $\sum_{k=1}^{q_i} x_k^1 = x^1$ (Kirchhof's law for the gas flow); the second components $(x_1^2, ..., x_{q_i}^2)$ of the outputs $z_1, z_2, \ldots, z_{q_i}$ are the exact copy of the original second component x^2 of the input z. Such kind of effects for the second component are proper for the nodes from that several pipelegs are outcoming and, hence, the values of the pressure are copied in accordance with ordinary physical law. For a multiplication node, the bounds on net inflow can be derived, for example, directly from the agreements between the transmission company and the gas producers.

Each summation nodes $j \in I_\Sigma$ possesses the following property: it has several input flows $\{z_1 = (x_1^1, x_1^2), \ldots, z_{q_j} = (x_{q_j}^1, x_{q_j}^2)\}$ and one output flow $z = (x^1, x^2)$(see Figure 1.15) such that $\sum_{k=1}^{q_j} x_k^1 + a_j^1 = x^1$ and $\max_{k=1,\ldots,q_j} x_k^1 + a_j^2 = x^2$ where two-dimensional vector $a_j = (a_j^1, a_j^2)$ denotes the intensity of the node j. In other words, the vector $a_j = (a_j^1, a_j^2)$ presents the available storage capacity for the first component, and the pressure regulation for outcoming flow produced by compressor/compressor station for the second component. The relation $\max_{k=1,\ldots,q_j} x_k^1$ images the fact that the pressures with which the incoming flows reach the node j should be grown up at this node.

By analogy the corresponding properties of edges can be treated (see Figure 1.15).

Besides the classical flow balance equations at each node, the nonlinear relation between the flow and the difference of the pressure p_i, p_j at the two ends of each pipe also need to be considered. For the high pressure, this later relation is given by the Weymouth formula for gas (see Osiadacz, 1987) and by the Darcy-Weisbach formula for liquid, and can be expressed with help of the introduced components as:

$$sign(x^1_{ij})\left(x^1_{ij}\right)^2 = C^2_{ij}\Big(\left(x^2_i\right)^2 - \left(x^2_j\right)^2\Big), \tag{1.41}$$

where the constants C_{ij} depend on the physical properties of of the pipe.

The resulting outcoming node t is a specific one: it summarizes only the first components signal that will be denote by f, too.

The optimization problem can be formulated in the following form; minimize the cost value function

$$f = \sum_{s \in I^-_t} x^1_{st} a^1_{st} + \sum_{j \in I_{\Delta t}} x^1_j a^1_{jt} + a^1_t \to \max \tag{1.42}$$

subject to

$$x^1_{ij(i)} = \sum_{s \in I^-_i} x^1_{si} a^1 k_{si} + \sum_{j \in I_{\Delta i}} x^1_j a^1_{ji} + a^1_i, \tag{1.43}$$

$$x^2_{ij(i)} = \max\{x^2_{si}, s \in I^-_i; x^2_j a^2_{ji} + a^2_i,\ j \in I_{\Delta i}\} \tag{1.44}$$

$$d^k_{*ij} \le x^k_{ij} \le d^{k*}_{ij},\ (i,j) \in U_*, d^k_{*i} \le x^k_i \le d^{k*}_i,\ i \in I_\Delta \tag{1.45}$$

$$sign(x^1_{ij(i)})\left(x^1_{ij(i)}\right)^2 = C^2_{ij(i)}\Big(\left(x^2_i\right)^2 - \left(x^2_{j(i)}\right)^2\Big),\ (i,j) \in U_*, \tag{1.46}$$

where $j(i)$ denotes the node connecting with the node i such that $\big(i, j(i)\big) \in U$, and

$$I^-_i = \{s \in I_{\sum} :\ (s,i) \in U_*\}, I_{\Delta i} = \{j \in I_\Delta :\ (j,i) \in U\}. \tag{1.47}$$

For the optimization problem (1.42—1.47), by analogy with previous Section a numerical algorithm can be developed. It should be noted that the the nonlinear constraints of (1.46, 1.44) can be linearized in accordance, for example, with the linearization approach proposed in [13].

1.4 Nonstationary gas networks

In the sections above the stationary case of the networks are considered when it is assumed that the basic characteristics of the net are not varied in the time. The temporary effects can be described by the nonstationary graph model setting. A simple model of such type is proposed in this section.

1.4.1 Problem statement

In this Section we consider the simple case of the dynamical network that presents an extension of the stationary net introduced in the previous sections.

Let $S(t) \doteq \{I, U(t)\}$ be a nonstationary network where $I = \{1, 2, \ldots, n\}$ denotes the set of nodes, $U(t)$ means the set of edges which can change with time $t = 0, 1, \ldots, T-1$. Let $Q_i(t)$ denote the gas flow in the node i at the moment t; and $a_i(t)$ the intensity of the node i. Let, furthermore, $Q_{ij}(t)$ denote the gas flow that is transported from the node i to the node j at the moment t. Then

$$w_i(t) = \sum_{j \in I_i^-(U(t))} Q_{ji}(t) - \sum_{j \in I_i^+(U(t))} Q_{ij}(t) \tag{1.48}$$

where

$$\begin{aligned} I_i^-(U(t)) &= \{j \in I|\ (j,i) \in U(t)\}, \\ I_i^+(U(t)) &= \{j \in I|\ (i,j), \in U(t)\} \end{aligned} \tag{1.49}$$

is the gas flow through the node i at the time t. Further, let $c_{ij}(t)$ denote the transport costs for the per unit of the gas flow from node i to the node j; and let $c_i(t)$ be the storage cost per unit of the gas flow in the node i. Denote by $d_{ij}(t)$ the network throughput from the node i to the node j and $d_i(t)$ is the gas capacity in the node i at the moment t.

Finally, the transport model for gas transportation is given in the following form: minimize the cost functional

$$\sum_{t=0}^{T-1} \Big(\sum_{(i,j) \in U(t)} c_{ij}(t) Q_{ij}(t) + \sum_{i \in I} c_i(t+1) Q_i(t+1) \Big) \tag{1.50}$$

over the solution of the equations

$$\begin{aligned} Q_i(t+1) = Q_i(t) - &\sum_{j \in I_i^+(U(t))} Q_{ij}(t) \\ + \sum_{j \in I_i^-(U(t))} & Q_{ji}(t) + a_i(t),\ i \in I \end{aligned} \tag{1.51}$$

subject to constraints

$$0 \le Q_i(t+1) \le d_i(t+1),\ Q_i(0) = q_{i0},\ i \in I \tag{1.52}$$

$$0 \le Q_{ij}(t) \le d_{ij}(t), (i,j) \in U(t), t = 1, \ldots, T \tag{1.53}$$

Note that the equations (1.51) mean the balance conditions for the gas flow in the nodes of the network,where $q_{i0}, i = 1, \ldots, n$ denotes the gas flow at the initial moment. The cost functional describes, in fact, the storage and transmissions costs.

Remark 6. *The given system (1.51) presents, in general, a specific class of $2-D$ control systems. In order to emphasize the possible loss effect in the various pipeline legs we can extend the given model by introducing the so-called loss coefficient $a_{ij}(t)$ of transmission from node i to node j, $(i,j) \in U(t)$.*

1.4.2 Optimality and suboptimality conditions

The given transport problem (1.50)—(1.52) we can rewrite as a specific discrete control problem (see [37]). For this purpose introduce the following new variables. The n-tuple $Q(t) = (Q_1, ...Q_n) = \{Q_i(t), i \in I\}$ marks the state space vector at the moment t; the set $u(t) = \{Q_{ij}(t), (i,j) \in U(t)\}$ is the control vector at the moment t. Also denote

$$d_u(t) = \{d_{ij}(t),\ (i,j) \in U(t)\},\ d_Q(t) = \{d_i(t),\ i \in I\};$$

$$c^*(t) = \{c_{ij}(t) - \sum_{k=t+1}^{T} (c_i(k) - c_j(k)), (i,j) \in U(t)\}.$$

Then the problem (1.50)—(1.52) is rewritten in the following form: Minimize

$$J(u) \triangleq \sum_{t=0}^{T-1} c^{*'}(t)u(t) \to \min \tag{1.54}$$

subject to the constraints

$$Q(t+1) = Q(t) - B(t)u(t) + a(t),\ Q(0) = q_0, \tag{1.55}$$

$$0 \leq u(t) \leq d_u(t)\},\ t = 0, 1, \ldots, T-1, \tag{1.56}$$

$$0 \leq Q(t+1) \leq d_Q(t+1),\ t = 0, 1, \ldots, T-1. \tag{1.57}$$

where the $|I| \times |U(t)|$-dimensional matrix $B(t)$ is generated by the net $S(t) = \{I, U(t)\}$ as

$$B(t) = \{b_k^{ij}(t), k \in I, (i,j) \in U(t)\}$$

where

$$b_k^{ij} = 0 \ \ if \ \ k \neq i, k \neq j;\ b_i^{ij} = 1;\ b_j^{ij} = -1.$$

Remark 7. *The given optimization problem possesses some special properties such as a simple dynamic, the specific input matrix, and a great control vector dimension. These properties produce the main difficulties for numerical implementation. A feasible way to overcome these obstacles is to design special optimization methods for the problem mentioned above.*

In the paper we present an optimization method based on the constructive approach proposed in [37] for the optimal control problem. For brevity denote also by $\hat{T} \triangleq \{0, 1, \ldots, T-1\}$. The control function $u(\cdot) = \{u(t), t \in \hat{T}\}$ is called admissible, if it satisfies the constraints (1.56). Usually, in the processing the numerical implementation of optimal control algorithms we exploit the approximate solutions with corresponding error estimation. Hence it is necessary to introduce the sub-optimality concept as it is often sufficient to stop the numerical computations when a satisfactory accuracy level has been achieved. We call a control function $u^\epsilon(\cdot)$, $\epsilon-$ optimal, if

$$J(u^\epsilon) - J(u^0(t)) \leq \epsilon,$$

where $u^0(\cdot)$ is the optimal control. By analogy with [37] (see, also, Section 1.2.1) introduce the following definition.

Definition 3. *The collection of the sets*

$$G_{supp} = \{\{I_{supp}(t+1), U_{supp}(t)\}, t \in \hat{T}\},$$

$$I_{supp}(t+1) \subset I,\ U_{supp}(t) \subset U(t)$$

is called the support for the problem (1.54)—(1.56) if the following system of equations

$$\begin{aligned} &Q(t+1) = Q(t) - B(t)u(t), Q(0) = 0,\\ &Q_i(t+1) = 0, i \in I \setminus I_*(t+1),\\ &u_{ij}(t) = 0, (i,j) \in U(t) \setminus U_*(t), t \in \hat{T}, \end{aligned}$$

with

$$I_*(t+1) = I_{supp}(t+1), U_*(t) = U_{supp}(t),\ t \in \hat{T},$$

has only the trivial solution $u_{ij}(t) \equiv 0, (i,j) \in U_{supp}(t), t \in \hat{T}$, *but this system has nontrivial solutions* $u_{ij} \neq 0, for(i,j) \in U_*(t), t \in \hat{T}$ *in the case when*

$$\begin{aligned} &I_*(t) = I_{supp}(t), t = 1, \ldots, T, t \neq t_0,\\ &I_*(t_0) = I_{supp}(t_0) \cup i_0; U_*(t) = U_{supp}(t), t \in \hat{T}, \end{aligned}$$

with $i_0 \bar{\in} I_{supp}(t_0)$ *where* t_0 *is arbitrary element from the set* $\hat{T}$.

By analogy to [37] the following theorem can be stated.

Theorem 2. *(ϵ-maximum principle). The control function* $u^{\epsilon}(\cdot)$ *is* ϵ- *optimal in the problem (1.54)—(1.56), if and only if there exist the support* G_{supp} *and the functions* $\epsilon_u(t) \geq 0$ *and* $\epsilon_\chi(t+1) \geq 0, t \in T$ *such that*

$$\sum_{t=0}^{T-1} (\epsilon_u(t) + \epsilon_\chi(t+1)) \leq \epsilon$$

holds and the following quasimaximum control conditions

$$\begin{aligned} &\big(\psi'(t)B(t) - c^{*'}(t)\big)u(t) = \\ \max_{u \in V(t)} &\big(\psi'(t)B(t) - c^{*'}(t)\big)u - \epsilon_u(t), t \in \hat{T} \end{aligned}$$

and the quasimaximum trajectory conditions

$$\lambda'(t)Q(t) = \max_{0 \leq Q \leq d_Q(t)} \lambda'(t)Q - \epsilon_\chi(t), t = 1, \cdots, T$$

are fulfilled.

Here $\psi(t), t \in \hat{T}$ is the solution of the dual(adjoint) system of the form

$$\psi(t-1) = \psi(t) + \lambda(t), t = T-1, T-2, \cdots, 0, \tag{1.58}$$

with initial condition $\psi(T-1) = \lambda(T)$ and $\lambda(t) = (\lambda_i(t),\ i \in I),\ t = 1, \cdots, T$ are the vectors of jumps constructed with help of the support G_{supp} by the following procedure.

For $t = T$ set $\lambda_i(T) = 0, i \in I_{supp}(T)$; the remaining $\lambda_i(T)$ are uniquely determined from the equalities

$$\lambda_i(T) - \lambda_j(T) = c_{ij}(T-1),\ (i,j) \in U_{supp}(T-1) \tag{1.59}$$

For $t = T-1, T-2, \ldots, 1$ the corresponding vectors $\lambda(t)$ are defined analogously: $\lambda_i(t) = 0$, for $i \in I_{supp}(t)$ and the remaining $\lambda_i(t)$ are determined from the equalities

$$\lambda_i(t) - \lambda_j(t) = c_{ij}(t-1) - \sum_{k=t+1}^{T} (\lambda_i(k) - \lambda_j(k)),\ (i,j) \in U_{supp}(t-1). \tag{1.60}$$

Remark 8. *Also the optimality condition in the maximum principle form can be obtained. The distinction of the given result from the discrete analogue of the Pontryagin maximum principle lies in the fact that it gives a constructive procedure to check the optimality and suboptimality of the admissible control. By analogy with the stationary network case of Section 1.2.1 the corresponding iterative method can be developed on the principle of the decreasing suboptimality estimate, i. e. the iteration* $\{(u, G_{supp})\} \to \{(\hat{u}, \hat{G_{supp}})\}$ *is performed in such a way as to achieve a decrease of the suboptimality estimate. These transformations involve, in effect, the duality theory and exploit the* ϵ*-optimality conditions.*

Remark 9. *As it is mentioned above the first (high level) stage modelling can be based on the simultaneously consideration some other crucial characteristics such as the pressure* p *and flow volume gas* Q *in the transport networks, which leads to the two-commodity flow graph model.*

Analogously to the two-commodity stationary networks case considered in the Section 1.2.1 we can introduce nonstationary multi-commodity networks. A fragment of such kind graph model can be presented, for example, by Figure 1.16

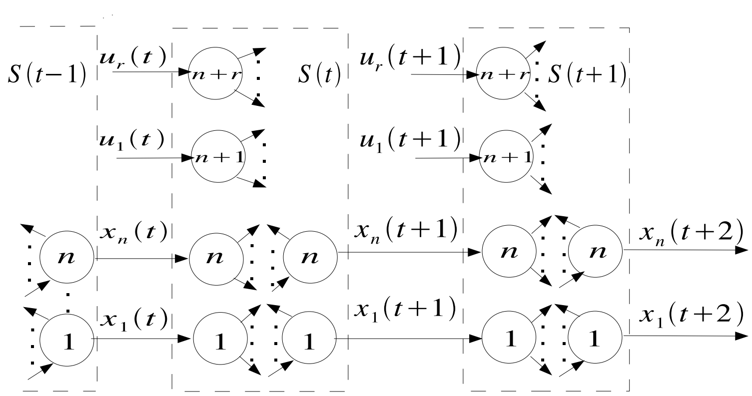

Figure 1.16: Multi-commodity network graph model

The optimization theory corresponding such kind of dynamical multi-commodity networks can be developed,too.

Remark 10. *The mathematical models mentioned in the previous Sections present, generally speaking, a specific class of the so-called multidimensional* $(n-D)$ *systems. Some classes of* $2-D$ *control systems will be studied in the next sections.*

1.5 Gas flow model in a pipeline unit

The main problem for the second-step modeling is to keep (or optimize if it is possible) the preassigned regime $\overline{Q}(x,t), \overline{p}(x,t)$ for the state space parameters of gas pressure p and mass flow Q for each pipeline unit. In particular, this regime can be determined and passed by the high level stage modelling results.

The aim of this section is to use the 2-D control theory setting for studying control problems in gas pipeline units. The state space parameters are gas pressure p and mass flow Q at the points of the pipe. All other physical parameters of the pipe and gas used here are constant at the moment of calculation. For calculating the state space parameters for the turbulent, isothermal gas flow in a long pipeline the following system of non-linear differential equations from the theory of gas dynamics can be used see ([63])

$$\begin{aligned} \frac{\partial Q(\tau,x)}{\partial \tau} &= -S\frac{\partial p(\tau,x)}{\partial x} - \frac{\lambda c^2}{2DS}\frac{Q^2(\tau,x)}{p(x,\tau)}, \\ \frac{\partial p(\tau,x)}{\partial \tau} &= -\frac{c^2}{S}\frac{\partial Q(\tau,x)}{\partial x}. \end{aligned} \tag{1.61}$$

where x denotes the space variable, τ the time variable, S the cross sectional area, D the pipeline diameter, c the isothermal speed of sound and λ the friction factor.

It is known that some important dynamic characteristics of the processes can be evaluated from the linearized model of the processes. The most accurate linear model can be realized in some neighborhood of the known basic regime $\overline{Q}(x,t), \overline{p}(x,t)$ of the considered process. In the section below we give such kind of a linearized model.

1.5.1 Linearization scheme

Let (Q,p) and $(\bar{Q},\bar{p})$ are the current and known state variables for a gas pipeline unit. Therefore, they satisfy the system (1.61):

$$\begin{aligned} \frac{\partial Q(t,x)}{\partial t} &= -S\frac{\partial p(t,x)}{\partial x} - \gamma\frac{Q^2(t,x)}{p(x,t)}, \\ \frac{\partial p(t,x)}{\partial t} &= \alpha\frac{\partial Q(t,x)}{\partial x}. \end{aligned} \tag{1.62}$$

and

$$\begin{aligned} \frac{\partial \bar{Q}(t,x)}{\partial t} &= -S\frac{\partial \bar{p}(t,x)}{\partial x} - \gamma\frac{\bar{Q}^2(t,x)}{\bar{p}(x,t)}, \\ \frac{\partial \bar{p}(t,x)}{\partial t} &= \alpha\frac{\partial \bar{Q}(t,x)}{\partial x}. \end{aligned} \tag{1.63}$$

Let us

$$Q = \bar{Q} + \Delta Q, \;\; p = \bar{p} + \Delta p. \tag{1.64}$$

Substracting from (1.62) the (1.63) yields

$$\begin{aligned} \frac{\partial Q(t,x)}{\partial t} - \frac{\partial \bar{Q}(t,x)}{\partial t} &= -S\frac{\partial p(t,x)}{\partial x} - \gamma\frac{Q^2(t,x)}{p(x,t)} + S\frac{\partial \bar{p}(t,x)}{\partial x} + \gamma\frac{\bar{Q}^2(t,x)}{\bar{p}(x,t)}, \\ \frac{\partial p(t,x)}{\partial t} - \frac{\partial \bar{p}(t,x)}{\partial t} &= \alpha\frac{\partial Q(t,x)}{\partial x} - \alpha\frac{\partial \bar{Q}(t,x)}{\partial x}. \end{aligned}$$

$$\begin{aligned}
\frac{\partial(Q(t,x)-\bar{Q}(t,x))}{\partial t} &= -S\frac{\partial(p(t,x)-\bar{p}(x,t))}{\partial x} - \gamma\left[\frac{Q^2(t,x)}{p(x,t)} - \frac{\bar{Q}^2(t,x)}{\bar{p}(x,t)}\right],\\
\frac{\partial(p(t,x)-\bar{p}(t,x))}{\partial t} &= \alpha\frac{\partial(Q(t,x)-\bar{Q}(t,x))}{\partial x}.
\end{aligned}$$

Noting (1.64) we have

$$\begin{aligned}
\frac{\partial\Delta Q(t,x)}{\partial t} &= -S\frac{\partial(\Delta p(t,x)}{\partial x} - \gamma\left[\frac{(\bar{Q}(t,x)+\Delta Q)^2}{\bar{p}(x,t)+\Delta p(t,x)} - \frac{\bar{Q}^2(t,x)}{\bar{p}(x,t)}\right],\\
\frac{\partial\Delta p(t,x)}{\partial t} &= \alpha\frac{\partial\Delta Q(t,x)}{\partial x}.
\end{aligned}$$

Using the property of geometric progression

$$\frac{1}{1+q} = \frac{1}{1-(-q)} = 1 + (-q) + (-q)^2 + ...$$

we have

$$\begin{aligned}
\left[\frac{(\bar{Q}+\Delta Q)^2}{\bar{p}+\Delta p}\right] &= \frac{(\bar{Q}(1+\frac{\Delta Q}{\bar{Q}}))^2}{\bar{p}(1+\frac{\Delta p}{\bar{p}})} = \frac{\bar{Q}^2}{\bar{p}}\frac{(1+\frac{\Delta Q}{\bar{Q}})^2}{(1+\frac{\Delta p}{\bar{p}})} =\\
&= \frac{\bar{Q}^2}{\bar{p}}\left(1+2\frac{\Delta Q}{\bar{Q}} + (\frac{\Delta Q}{\bar{Q}})^2\right)\left[1-\frac{\Delta p}{\bar{p}} + (\frac{\Delta p}{\bar{p}})^2 + ...\right]\\
&\approx \frac{\bar{Q}^2}{\bar{p}}\left(1+2\frac{\Delta Q}{\bar{Q}} - \frac{\Delta p}{\bar{p}} + (more\quad higher\quad order\ )\right)
\end{aligned}$$

Then the linearized model for disturbances $(\Delta Q, \Delta p)$ are defined as

$$\begin{aligned}
\frac{\partial\Delta Q}{\partial t} &= -S\frac{\partial(\Delta p}{\partial x} - \gamma\left[\frac{\bar{Q}^2}{\bar{p}}(1+2\frac{\Delta Q}{\bar{Q}} - \frac{\Delta p}{\bar{p}}) - \frac{\bar{Q}^2}{\bar{p}}\right],\\
\frac{\partial\Delta p}{\partial t} &= \alpha\frac{\partial\Delta Q}{\partial x}.
\end{aligned}$$

$$\Downarrow$$

$$\begin{aligned}
\frac{\partial\Delta Q}{\partial t} &= -S\frac{\partial(\Delta p}{\partial x} - \gamma\frac{\bar{Q}^2}{\bar{p}}\left[1+2\frac{\Delta Q}{\bar{Q}} - \frac{\Delta p}{\bar{p}} - 1\right],\\
\frac{\partial\Delta p}{\partial t} &= \alpha\frac{\partial\Delta Q}{\partial x}.
\end{aligned}$$

$$\Downarrow$$

$$\begin{aligned}
\frac{\partial\Delta Q}{\partial t} &= -S\frac{\partial\Delta p}{\partial x} - 2\gamma\frac{\bar{Q}^2}{\bar{p}\bar{Q}}\Delta Q - \gamma\frac{\bar{Q}^2}{\bar{p}^2}\Delta p,\\
\frac{\partial\Delta p}{\partial t} &= \alpha\frac{\partial\Delta Q}{\partial x}.
\end{aligned}$$

Introducing the new variables (we can say about new local coordinates $\Delta Q \to Q,\ \Delta p \to p$) we can present the linearized model in the neighbourhood of the known function $(\bar{Q}, \bar{p})$ in the following form

$$\frac{\partial Q}{\partial t} = -S\frac{\partial p}{\partial x} - \rho Q - \beta p, \tag{1.65}$$

$$\frac{\partial p}{\partial t} = \alpha \frac{\partial Q}{\partial x}.$$

where

$$\rho = 2\gamma \frac{\bar{Q}^2}{\bar{p}\bar{Q}},\ \beta = \gamma \frac{\bar{Q}^2}{\bar{p}^2},\ \gamma = \frac{\lambda c^2}{2DS},\ \alpha = -\frac{c^2}{S}.$$

1.5.2 On the link of 2-D discrete models with gas network system

The main object of study in this section is to present a class of the discrete $2-D$ systems that are derived under suitable discretization of differential equations describing the transported gas by pipelines. As it is shown above the linearized model in the neighborhood of the indicated regime $(\bar{Q}, \bar{p})$ has the following form

$$\frac{\partial Q}{\partial \tau} = -S\frac{\partial p}{\partial x} - \rho Q - \beta p, \qquad \frac{\partial p}{\partial \tau} = \alpha \frac{\partial Q}{\partial x}. \tag{1.66}$$

In order to obtain the wanted discrete model, we use the classical discretization scheme for the linear partial differential equations of (1.66).

Introduce the following combined discretization scheme for the partial derivatives with steps h_1, h_2, respectively

$$\frac{\partial Q(t,x)}{\partial t} = \frac{Q(t+h_1,x) - Q(t,x)}{h_1}, \tag{1.67}$$

$$\frac{\partial Q(t,x)}{\partial x} = \frac{Q(t,x+h_2) - Q(t,x-h_2)}{2h_2}, \tag{1.68}$$

$$\frac{\partial p(t,x)}{\partial t} = \frac{p(t+h_1,x) - p(t,x)}{h_1}, \tag{1.69}$$

$$\frac{\partial p(t,x)}{\partial t} = \frac{p(t,x+h_2) - p(t,x-h_2)}{2h_2}. \tag{1.70}$$

Replacing these derivatives in the system (1.66) gives

$$\frac{Q(t+h_1,x) - Q(t,x)}{h_1} = -S\frac{p(t,x+h_2) - p(t,x-h_2)}{2h_2} - \rho Q(t,x) - \beta p(t,x),$$

$$\frac{p(t+h_1,x) - p(t,x)}{h_1} = \alpha \frac{Q(t,x+h_2) - Q(t,x-h_2)}{2h_2} \tag{1.71}$$

and we have

$$Q(t+h_1,x) = Q(t,x) - \frac{h_1}{2h_2}S\Big(p(t,x+h_2) - p(t,x-h_2)\Big) - h_1\rho Q(t,x) - h_1\beta p(t,x)$$

$$p(t+h_1,x) = p(t,x) - \alpha\frac{h_1}{2h_2}\Big(Q(t,x+h_2) - Q(t,x-h_2)\Big).$$

Thus, the discrete values $Q(k_1h_1, k_2h_2)$ and $p(k_1h_1, k_2h_2)$ of the function $Q(x,t)$ and $p(x,t)$ calculated in the nodes of integer lattice $\{(k_1h_1, k_2h_2)\}$ satisfy the following equations

$$\begin{aligned} Q((k_1+1)h_1, k_2h_2) &= Q(k_1h_1, k_2h_2) - \frac{h_1}{2h_2} S\Big(p(k_1h_1, (k_2+1)h_2) - p(k_1h_1, (k_2-1)h_2)\Big) \\ &\quad - h_1\rho Q(k_1h_1, k_2h_2) - h_1\beta p(k_1h_1, k_2h_2), \\ p((k_1+1)h_1, k_2h_2) &= p(k_1h_1, k_2h_2) + \frac{h_1}{2h_2}\alpha\Big(Q(k_1h_1, (k_2+1)h_2) - Q(k_1h_1, (k_2-1)h_2)\Big). \end{aligned} \tag{1.72}$$

Introduce the following notations:

$$x_1(t,s) = Q(th_1, sh_2),\ \ x_2(t,s) = p(th_1, sh_2)$$

where t, s are integers.

Hence, the system (1.72) can be rewritten as follows

$$x_1(t+1,s) = x_1(t,s) - \frac{h_1 S}{2h_2}\Big(x_2(t,s+1) - x_2(t,s-1)\Big) - h_1\rho x_1(t,s) - \beta x_2(t,s),$$

$$x_2(t+1,s) = x_2(t,s) + \frac{h_1\alpha}{2h_2}\Big(x_1(t,s+1) - x_1(t,s-1)\Big).$$

The matrix form of this system is

$$\begin{aligned} \begin{bmatrix} x_1(t+1,s) \\ x_2(t+1,s) \end{bmatrix} = \begin{bmatrix} 1-\rho & -\beta \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1(t,s) \\ x_2(t,s) \end{bmatrix} + \begin{bmatrix} 0 & -\frac{h_1S}{2h_2} \\ \frac{h_1\alpha}{2h_2} & 0 \end{bmatrix} \begin{bmatrix} x_1(t,s+1) \\ x_2(t,s+1) \end{bmatrix} \\ + \begin{bmatrix} 0 & -\frac{h_1S}{2h_2} \\ -\frac{h_1\alpha}{2h_2} & 0 \end{bmatrix} \begin{bmatrix} x_1(t,s-1) \\ x_2(t,s-1) \end{bmatrix} \end{aligned} \tag{1.73}$$

and finally

$$x(t+1,s) = A_0x(t,s) + A_1x(t,s+1) + A_2x(t,s-1) \tag{1.74}$$

where

$$x(t,s) = \begin{bmatrix} x_1(t,s) \\ x_2(t,s) \end{bmatrix};$$

$$A_0 = \begin{bmatrix} 1-\rho & -\beta \\ 0 & 1 \end{bmatrix};\ A_1 = \begin{bmatrix} 0 & -\frac{h_1S}{2h_2} \\ \frac{h_1\alpha}{2h_2} & 0 \end{bmatrix};\ A_2 = \begin{bmatrix} 0 & -\frac{h_1S}{2h_2} \\ -\frac{h_1\alpha}{2h_2} & 0 \end{bmatrix};$$

Remark 11. *The considered model is a discrete version of the gas transport network problem along a single pipe.*

Next, in order to obtain the discrete $2-D$ system with control parameters, we define the part of the initial data as a control parameter. In particular, it is of interest to determine an optimal control programm for gas pressure and gas flow at the pipe. It is natural to assume that the gas regulation is realized at the incoming node of the pipe. For brevity we assume that this incoming node is the the beginning of the pipe where there

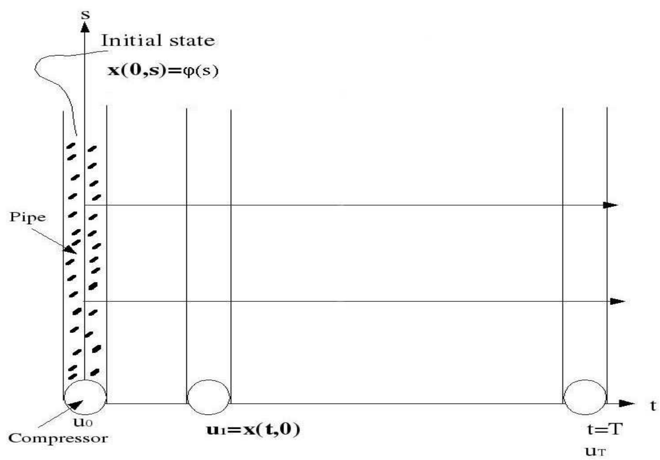

Figure 1.17: Initial data

exists a compressor for which it is necessary to set a regime how to "pump in - pump out" with time. This means that the initial data $x(t,0) = u_t, \quad t = 0, 1, 2, ..., T$ are treated as control parameters.

In final, the optimization problem is formulated as follows: to minimize the cost function

$$J(u) = \sum_{t=1}^{T} \Big[\sum_{s\in\ Z_+} \big(Qx(t,s), x(t,s)\big) + \big(Ru(t), u(t)\big)\Big],$$

over the solution of $2-D$ system of the form

$$x(t+1,s) = A_0 x(t,s) + A_1 x(t,s+1) + A_2 x(t,s-1) \tag{1.75}$$

with initial and boundary condition:

$$\begin{aligned} x(0,s) &= \varphi(s), & s \in \mathbb{Z}_+ \setminus 0 \\ x(t,0) &= \psi(t) = u(t), & t = 0, 1, 2, ...T \end{aligned} \tag{1.76}$$

where $\mathbb{Z}_+$ is the set of nonnegative integers, $Q \geq 0$ and $R > 0$ are given matrixes.

Here the quadratic cost function summarizes the total losses generated by deviation of the current on the basic regime.

Note that the optimization problem (1.75) can be rewritten in more general form as follows: minimize the cost functional

$$J(u) = \sum_{t=1}^{T} \Big[\sum_{s\in\ Z} \big(Qx(t,s), x(t,s)\big) + \big(Ru(t), u(t)\big)\Big],$$

over the solution of the system

$$x(t+1,s) = \sum_{i=-N}^{N} A_i x(t,s+i) + Bu(t,s) \tag{1.77}$$

with initial and boundary condition:

$$\begin{aligned} x(0,s) &= \varphi(s), \qquad s \in \mathbb{Z}, \;\; \varphi \in l_2(E) \\ x(t,j) &= \psi(t,j) = u(t), \quad t = 0,1,2,...,T; \quad j = 0,-1,-2,...,-N. \end{aligned}$$

where N is some integer, $u(t,s)$ is the control function. The parameters $u(t,s)$ can be interpreted as the controlled factors: gas pressure and gas flow at the pre-assigned points of the pipe (or the points to be determined) needed to keep the desired regime. For example, in the case when the points of the pipe are fixed, the corresponding elements of the matrix B can be put zero. Here the assumption $s \in \mathbb{Z}_+$ images the fact that upon discrete approximation the amount of discrete values can be huge. Further, keeping in mind an ability to apply this approach for optimization of gas networks that is composed by some collection of the pipes we admit the multiple shifts $x(t,s+i),\ i = -N,\dots,N$ in the system. Here the boundary condition $\psi(t,j), \quad t = 0,\dots,T; \quad j = 0,\dots,-N$ can be explained as known data from "previous" site of the pipe which is connected to our site of the pipe.

Remark 12. *Note that some others discretization schemes lead to another models of $2-D$ system where the proper defined shift operator should be modified by suitable manner.*

2

2-D optimization theory

The model obtained in the previous Chapter gives a good motivation to start the investigation of a class of the two-dimensional control systems given in this Chapter. The main feature of the model involved is that we first consider the discrete variables of two kinds: one of them runs the finite set in contrast to second that takes their values from the infinite set $\mathbb{Z}_+$. This fact can be illustrated by restricted shape of spatial variable (the finite length of the the gas pipeline, for example). And from second hand, the temporary variables for longtime duration the discretization of which leads usually to a huge amount of discrete values. Such consideration possesses some positivity: finiteness of the discrete variable allows us to obtain the exact optimal solution, and the infinite case, if it is necessary, can be realized as a limit case of the obtained solution. Moreover, such consideration gives an ability to construct the optimal feedback control law in a simple form.

Moreover, there exist a number of other technical processes that can be represented by suitable two-dimensional discrete (2-D) systems which gives a good mathematical tools for their analysis and investigation of their structural properties. This interest is clearly related to the wide variety of applications of both practical and/or theoretical interest. The key unique feature of an nD system is that the plant or process dynamics depend on more than one indeterminate and hence information is propagated in many independent directions. A key point is that the applications areas for nD systems theory can be found within the general disciplines of circuits, control and signal processing (and many others). For a representative cross-section of these see, for example, Levi B.C., Adams M.B. and Willsky A.S. (1990), Pratt (1982), Bose N.(2000), K.Galkowski and J. Wood (2001), E.Zerz (2000).

Some aspects of the control theory (controllability, observability, stabilizability) for the multidimensional systems are investigated in the papers of Kaczorek T. (1987), G. Jank (1999), Klamka J. (1994), Gaishun I. (1996), Zerz E. (2004). Some optimal control problems are considered by Givone D. and Roesser R. (1993), Bisiacco M. and Fornasini E. (1990), G.Jank (2002), S.Dymkou (2003) and others.

The aim of the Chapter 2 is to give a strong mathematical background for the $2-D$ control problem optimization.

2.1 Some basic notations and definitions

In this section we present a short resume for the basic notions used in this chapter.

First, we ill exploit often the notion of the inner product which is a generalization of the dot product. In a vector space, it is a way to multiply vector together, with the result of this multiplication being a scalar.

Let u, v and w be vectors in the real vector space H and α be a scalar from the field $\mathbb{R}$.

Definition 4. *The mapping* $(\cdot,\cdot) : H \times H \to \mathbb{R}$ *is called inner product if the following properties are fulfilled:*
1)$(u+v, w) = (u, w) + (v, w)$;
2)$(\alpha v, w) = \alpha(v, w)$;
3)$(v, w) = (w, v)$;
4)$(v, v) \geqslant 0$ *for any* $v \in H$, *and* $(v, v) = 0$ *if and only if* $v = 0$.

This definition also applies to an abstract vector space over any field. When given a complex vector space, the third property is usually replaced by $(v, w) = \overline{(w, v)}$ where $\bar{z}$ refers to complex conjugation.

Definition 5. *A Hilbert space is a vector space* H *with an inner product* (f, g) *such that the norm defined by* $| f |= (f, f)^{1/2}$ *turns* H *into a complete metric space.*

Definition 6. *A space* X *is called finite-dimensional (n-dimensional) if in* X *exist a finite basis (basis from n-elements)*

Theorem 3. *For any bounded functional* f *on Hilbert space* H *there exists a unique element* $u \in H$ *such that* $f(x) = (x, u)$ *and* $|f| = |u|$

Definition 7. *Let* H_1, H_2 *are Hilbert spaces and* $\mathcal{A} : H_1 \to H_2$ *be linear operator. We will say that operator* $\mathcal{A}^* : H_2 \to H_1$ *adjoint to operator* $\mathcal{A}$ *iff*

$$(\mathcal{A}x, y)_{H_2} = (x, \mathcal{A}^* y)_{H_1}$$

holds for any $x \in H_1$ *and* $y \in H_2$.

Definition 8. *Let* $|\cdot|_1$ *and* $|\cdot|_2$ *the given norms in a vector space* H. *We will called them equivalent, if exist constants* $c_1 > 0$ *and* $c_2 > 0$ *such that* $c_1\|x|_1 \leq |x|_2 \leq c_2|x|_2$ *for all* $x \in X$.

Let E and V be finite-dimensional Hilbert spaces, the inner product in which defined by the same symbol $(\cdot,\cdot)$. Let $l^2(E)$ and $l^2(V)$ denote the Hilbert spaces of the square summarised sequences in E and V respectively, i.e. the spaces of the functions $\varphi : \mathbb{Z} \to E$ and $\varphi : \mathbb{Z} \to V$ ($\mathbb{Z}$ is the set of integers) such that $\|\varphi\| = \sum\limits_{s\in\mathbb{Z}} | \varphi(s) |^2 < \infty$ where $| \cdot |$ are the norms in E and V, and $\| \cdot \|$ denotes the norm in the space $l^2(E)$ (or $l^2(v)$) that is defined by the associated inner products. The inner product in $l^2(E)$ is defined usually as $(\varphi, \psi)_2 = \sum\limits_{s\in\mathbb{Z}} (\varphi(s), \psi(s))$.

2.2 System model and preliminary notions

Let E and V be finite-dimensional Hilbert spaces, the inner product in which defined by the same symbol $(\cdot,\cdot)$. Consider the linear maps $A_i : E \to E$, $i = -N, -N+1, ..., 0, ...N$ and the linear operator B acting from V into E.

The main object is the two-dimensional $(2-D)$ discrete system described by

$$x(t+1, s) = \sum_{i=-N}^{N} A_i(t, s+i) + Bu(t, s), t = 0, 1...;\ s \in \mathbb{Z}, \tag{2.1}$$

where $u(t,s)$ is the V-valued function.

Clearly, for any function $\varphi : \mathbb{Z} \to E$ the equation (2.1) admits a unique solution $x(t,s)$ that satisfies the initial condition

$$x(0,s) = \varphi(s), \;\; s \in \mathbb{Z}. \tag{2.2}$$

Remark 13. *The parameters $u(t,s)$ can be interpreted as the controlled factors: gas pressure and gas flow at the pre-assigned points of the pipe (or the points to be determined) needed to keep the desired regime. The assumption $s \in \mathbb{Z}_+$ images the fact that upon discrete approximation the amount of discrete values of the state variable x corresponding long leg pipeline can be huge. Further, keeping in mind an ability to apply this approach for optimization of gas networks that is composed by some collection of the pipes we admit the multiple shifts $x(t,s+i), \;\; i = -N,\dots,N$ in the system.*

For the linear maps $A_i : E \to E, \; i = -N, -N+1, ..., 0, ...N$ define the operator $\mathcal{A} : l^2(E) \to l^2(E)$ as follows

$$(\mathcal{A}\varphi)(s) = \sum_{i=-N}^{N} A_i \varphi(s+i), \;\; s \in \mathbb{Z}. \tag{2.3}$$

It is easy to establish that the operator is bounded and the norm $\parallel \mathcal{A} \parallel$ satisfies the following inequality

$$\parallel \mathcal{A} \parallel^2 \leq 2N \sum_{i=-N}^{N} \mid A_i \mid^2,$$

where $\mid A_i \mid$ is the norm of the operator A_i, conforming with the norm $\mid \cdot \mid$ in E.

A conjugate operator for the operator $\mathcal{A}$ is the map $\mathcal{A}^* : l^2(E) \to l^2(E)$ defined as follows

$$(\mathcal{A}^*\psi)(s) = \sum_{i=-N}^{N} A_i^* \psi(s-i), \;\; s \in \mathbb{Z} \tag{2.4}$$

where A_i^* is the conjugate operator for the operator A_i.

Using the obtained conjugate operator (2.4) define the conjugate equation for the equation (2.1) in the following form

$$z(t,s) = \sum_{i=-N}^{N} A_i^* z(t+1, s-i) + g(t,s), \tag{2.5}$$

where $z(t,s)$ is a unknown function.

It can be shown that if the maps u and g are equal zero then the equalities

$$(x(t,s), z(t,s)) = const, \;\; \forall \; (t,s)$$

are hold for any solutions $x(t,s), z(t,s)$ of the equations (2.1) and (2.5).

Next it is suitable the equations (2.1) and (2.5) represent in an operator form. For this purpose we assume that the sequences $s \to u(t,s), s \to g(t,s)$ are to be square summarized for each fixed $t, \;\; t \geq 0$.

Let $y_t, \; \omega_t, \; \psi_t$ are the elements of the $l^2(E)$ space, v_t is the element of the $l^2(V)$ space defined as follows

$$(y_t)(s) = x(t,s), (\omega_t)(s) = z(t,s), (\psi_t)(s) = g(t,s), (v_t)(s) = u(t,s), \;\; t \geq 0, s \in \mathbb{Z}. \tag{2.6}$$

Now, the equations (2.1) and (2.5) can be represented as follows

$$y_{t+1} = \mathcal{A} y_t + \mathcal{B} v_t,$$
$$\omega_t = \mathcal{A}^* \omega_{t+1} + \psi_t, \tag{2.7}$$

where $\mathcal{B}$ is the linear operator from $l^2(V)$ into $l^2(E)$ defined by $(\mathcal{B}\varphi)(s) = B\varphi(s)$, $s \in \mathbb{Z}$.

The obtained operator form is used next to prove the solvability for optimization problem below.

2.3 Linear quadratic optimization in the strip

Let $T > 1$ be a given integer. In this section we consider the control system (2.1) defined in the strip

$$t \in \{0, 1, \dots, T\}, \ \ s \in \mathbb{Z} \tag{2.8}$$

with the initial condition $x(0, s) = \varphi(s), s \in \mathbb{Z}$, where $\varphi \in l^2(E)$.

The function $u(t, s)$ is called an admissible control if the sequence $s \to u(t, s)$ is square summable for each fixed $t \in \{0, 1, \dots, T\}$.

For the given initial function $\varphi \in l_2(E)$ and the admissible control function $u \in \mathcal{B}(V)$ the function $x : T \times \mathbb{Z} \to V$ such that $x(t, s) \in l^2(E)$ for each fixed $t \in \{0, 1, \dots, T\}$ is called the solution of (2.1), if it satisfies the equation (2.1) and the initial condition $x(0, s) = \varphi(s)$, $s \in \mathbb{Z}$.

The optimization problem is to find the admissible control function $u^0(t, s)$ that minimizes the following cost functional

$$J(u) = \sum_{t=1}^{T} \sum_{s \in \mathbb{Z}} \left[\big(Qx(t, s), x(t, s)\big) + \big(Ru(t-1, s), u(t-1, s)\big) \right], \tag{2.9}$$

where $x(t, s)$ is the solution of (2.1) in the strip (2.8) corresponding to the initial data $x(0, s) = \varphi(s)$, $s \in \mathbb{Z}$ and control u. Here $Q : E \to E$, $R : V \to V$ are self-adjoint operators such that $Q \geq 0$ and $R > 0$.

The following theorem is true.

Theorem 4. *The optimization problem (2.1), (2.9) is solvable.*

Proof. For this purpose we represent the system (2.1) in the operator form

$$y_{t+1} = \mathcal{A} y_t + \mathcal{B} v_t, \ t \in \{0, 1, \dots, T\}, \tag{2.10}$$

where $\mathcal{A}$ is a linear operator from $l_2(E)$ to $l_2(E)$ and $\mathcal{B}$ is another linear operator from $l_2(V)$ to $l_2(E)$ that are defined by analogy with making early. Denote by $\mathcal{B}_T(E)$ and $\mathcal{B}_T(V)$ the spaces of the functions defined on the set $\{0, 1, \dots, T\}$ with values in the spaces $l_2(E)$, $l_2(V)$, respectively, such that

$$\mathcal{B}_T(E) = \big(l_2(E)\big)^{T+1}, \ \mathcal{B}_T(V) = \big(l_2(V)\big)^{T+1}.$$

Also, let $L : \mathcal{B}_T(V) \to \mathcal{B}_T^0(E)$ be the mapping given as follows

$$(L\gamma)_t = \mathcal{B}\gamma_{t-1} + \mathcal{A}\mathcal{B}\gamma_{t-2} + \dots + \mathcal{A}^{t-1}\mathcal{B}\gamma_0, t > 0, (L\gamma)_0 = 0, \tag{2.11}$$

where $\gamma = \big(\gamma_0, \gamma_0, \dots, \gamma_T\big) \in \mathcal{B}_T(E)$, $\mathcal{B}_T^0(E)$ denotes the subspace of $\mathcal{B}_T(E)$ containing the functions with zero values at $t = 0$.

The task now is to prove that the solution of (2.10) can be presented in the form

$$y = Lv + w. \tag{2.12}$$

Indeed, we have step-by-step

$$\begin{aligned}
t =0:&\quad y_0 = x(0,s) = (Lv+\omega)_0 = \varphi(s);\\
t =1:&\quad y_1 = x(1,s) = (Lv+\omega)_1 = \mathcal{B}v_0 + \mathcal{A}\varphi;\\
t =2:&\quad y_2 = x(2,s) = (Lv+\omega)_2 = \mathcal{B}v_1 + \mathcal{A}\mathcal{B}v_0 + \mathcal{A}^2\varphi;\\
\dots&\\
t =T:&\quad y_T = x(T,s) = (Lv+\omega)_T = \mathcal{B}v_{T-1} + \mathcal{A}\mathcal{B}v_{T-2} + \dots + \mathcal{A}^{T-1}\mathcal{B}v_0 + \mathcal{A}^T\varphi
\end{aligned}$$

which together (2.11) proves (2.12).

Now rewrite the cost functional (2.9) in the operator form as

$$\begin{aligned}
\big(Qx,x\big) + \big(Ru,u\big) = \big(\mathcal{Q}y,y\big) + \big(\mathcal{R}v,v\big) = &\\
\big(\mathcal{Q}(Lv+\omega),(Lv+\omega)\big) + \big(\mathcal{R}v,v\big) = &\\
\big(\mathcal{Q}Lv,Lv\big) + \big(\mathcal{Q}Lv,\omega\big) + \big(\mathcal{Q}\omega,Lv\big) + \big(\mathcal{Q}\omega,\omega\big) + \big(\mathcal{R}v,v\big) = &\\
\big(L^*\mathcal{Q}Lv,v\big) + \big(\mathcal{Q}Lv,\omega\big) + \big(L^*\mathcal{Q}\omega,v\big) + \big(\mathcal{Q}\omega,\omega\big) + \big(\mathcal{R}v,v\big) = &\\
\big((\mathcal{R}+L^*\mathcal{Q}L)v,v\big) + \big(\mathcal{Q}Lv,\omega\big) + \big(L^*\mathcal{Q}\omega,v\big) + \big(\mathcal{Q}\omega,\omega\big) = &\\
\big((\mathcal{R}+L^*\mathcal{Q}L)v,v\big) + 2\big(L^*\mathcal{Q}\omega,v\big) + \big(\mathcal{Q}\omega,\omega\big)&
\end{aligned}$$

Thus

$$J(v) = ((\mathcal{R}+L^*\mathcal{Q}L)v,v)_{\mathcal{B}} + 2(L^*\mathcal{Q}\omega,v)_{\mathcal{B}} + (\mathcal{Q}\omega,\omega)_{\mathcal{B}}, \tag{2.13}$$

where $\omega = \big(\varphi, \mathcal{A}\varphi, \dots, (\mathcal{A})^T\varphi\big) \in \mathcal{B}_T(E)$. Here the symbol $(\cdot,\cdot)$ means the inner product in the Hilbert space $\mathcal{B}_T(E)$ (or in the space $\mathcal{B}_T(V)$) defined as usually $(a,b)_{\mathcal{B}} = \sum_{i=0}^{T}(a_i,b_i)_E$. The operators of $\mathcal{R} : \mathcal{B}_T(V) \to \mathcal{B}_T(V)$ and $\mathcal{Q} : \mathcal{B}_T(E) \to \mathcal{B}_T(E)$ are given by obvious manner

$$\big(\mathcal{R}u\big)(t,s) = Ru(t-1,s),\ t \neq 0,\ \big(\mathcal{Q}x\big)(t,s) = Qx(t,s),\ t = 0,\dots,T,\ s \in Z_+.$$

Note that since $\mathcal{G} \geq 0$, $\mathcal{R} > 0$ then the operator $\mathcal{R} + L^*\mathcal{Q}L$ is inverted.

In order to find the minimum of (2.13) we calculate its Freshet derivative

$$\frac{\partial J(v)}{\partial v} = (\mathcal{R}+L^*\mathcal{Q}L)v + 2\mathcal{L}^*\mathcal{Q}\omega = 0.$$

It is obviously that this equation has the solution of

$$v = -(\mathcal{R}+L^*\mathcal{Q}L)^{-1}L^*\mathcal{Q}\omega$$

Put

$$v^0 = -(\mathcal{R}+L^*\mathcal{Q}L)^{-1}L^*\mathcal{Q}\omega. \tag{2.14}$$

In order to prove the optimality of the obtained solution we need to check the following inequality $J(v) - J(v^0) \geq 0$ for any $v \in \mathcal{B}_T(V)$. The validity of the inequity means that v^0 is optimal control in the initial problem.

Denoting $\Pi = (\mathcal{R} + L^*\mathcal{Q}L)$ we have

$$J(v) - J(v^0) = (\Pi v, v) + 2(L^*\mathcal{Q}\omega, v) - (L^*\mathcal{Q}\omega, \Pi^{-1}L^*\mathcal{Q}\omega) =$$
$$(\Pi v + \Pi\Pi^{-1}L^*\mathcal{Q}\omega, v) + (L^*\mathcal{Q}\omega, v + \Pi^{-1}L^*\mathcal{Q}\omega) =$$
$$(\Pi(v - v^0), v) + (L^*\mathcal{Q}\omega, v - v^0) = (\Pi(v - v^0), v) - (\Pi v^0, v - v^0) =$$
$$(\Pi(v - v^0), v - v^0) > 0$$

since accordingly the inequalities $Q \geq 0,\ R > 0$ the operator Π positive and, hence, is inverted.

Since $J(v) - J(v^0) = ((\mathcal{R} + L^*\mathcal{Q}L)(v - v_0), (v - v_0))_{\mathcal{B}} > 0$ for any $v \in \mathcal{B}_T(V), v \neq v_0$ then v^0 is a unique optimal solution for the problem (2.1),(2.9). The theorem is proved.

Thus, the optimization problem (2.1), (2.9) is solvable, and optimal control is given by the formula (2.14). Nevertheless, the obtained formula (2.14) is not suitable for applications since the inverting operator procedure presents a nontrivial problem, in general. By this reason we propose another way to find the optimal control function $u^0(t, s)$. This approach is based on the duality theory that can be developed for the considered $2 - D$ control systems.

The following theorem is true.

Theorem 5. *The boundary problem*

$$x(t+1, s) = \sum_{i=-N}^{N} A_i(t, s+i) - BR^{-1}B^*z(t, s),\ \ (t, s) \in \{0, \ldots, T\} \times \mathbb{Z} \tag{2.15}$$

$$z(t, s) = \sum_{i=-N}^{N} A_i^* z(t+1, s-i) + Qx(t+1, s),\ \ (t, s) \in \{0, \ldots, T\} \times \mathbb{Z} \tag{2.16}$$

with the conditions

$$x(0, s) = \varphi(s),\ \ z(T, s) = 0,\ \ s \in \mathbb{Z}. \tag{2.17}$$

is solvable in the space $l_2(V)$.

Proof. Denote by y_t, w_t the elements of the space $l_2(E)$ for which

$$(y_t)(s) = x(t, s),\ \ (w_t)(s) = z(t, s)\ s \in \mathbb{Z},\ \ t \in \{0, \ldots, T\}.$$

Then the problem (2.15)—(2.17) can be rewritten in the operator form

$$\begin{gathered} y_{t+1} = \mathcal{A}y_t - \mathcal{B}\mathcal{R}^{-1}\mathcal{B}^*w_t, \quad y_0 = \varphi, \\ w_t = \mathcal{A}^*w_{t+1} + \mathcal{Q}y_{t+1}, \quad w_T = 0. \end{gathered} \tag{2.18}$$

In the equation (2.10) put $y_0 = \varphi$ and $v_t = v_t^0$, where v_t^0 are elements of space $l_2(V)$ defined by formula (2.14). Next, from equation (2.10) we can determine the function $y_t^0,\ t \in \{0, \ldots, T\}$. Indeed, step by step procedure applied to the following equation

$$y_{t+1} = \mathcal{A}y_t + \mathcal{B}v_t = \mathcal{A}y_t - \mathcal{B}(\mathcal{R} + L^*\mathcal{Q}L)^{-1}L^*\mathcal{Q}w_t, \quad t \in \{0, 1, \ldots, T\} \tag{2.19}$$

we have

for $t=0:\ y_1 = \mathcal{A}y_0 + \mathcal{B}v_0 = \mathcal{A}\varphi - \mathcal{B}(\mathcal{R}+L^*\mathcal{Q}L)^{-1}L^*\mathcal{Q}w_0;$

for $t=1:\ y_2 = \mathcal{A}y_1 + \mathcal{B}v_1 = \mathcal{A}(\mathcal{A}\varphi-$

$$- \mathcal{B}(\mathcal{R}+L^*\mathcal{Q}L)^{-1}L^*\mathcal{Q}w_0) - \mathcal{B}(\mathcal{R}+L^*\mathcal{Q}L)^{-1}L^*\mathcal{Q}w_1 = \mathcal{A}^2\varphi - \hat{\Pi}(\mathcal{A}w_0 + w_1).$$

To simplify we denote $\hat{\Pi} = \mathcal{B}(\mathcal{R}+L^*\mathcal{Q}L)^{-1}L^*\mathcal{Q}$. Then continuing

$$\text{for } t=2:\ y_3 = \mathcal{A}\varphi - \hat{\Pi}(\mathcal{A}^2w_0 + \mathcal{A}w_1 + w_2);$$

. . . .

$$\text{for } y_t = \mathcal{A}^t\varphi - \hat{\Pi}\sum_{i=0}^{t-1}\mathcal{A}^i w_{t-1-i}$$

The obtained above formulae

$$v^0 = -(\mathcal{R}+L^*\mathcal{Q}L)^{-1}L^*\mathcal{Q}\omega_0$$

can be rewritten as

$$v^0 = -\mathcal{R}^{-1}\mathcal{B}^*\omega_0$$

and hence

$$\hat{\Pi} = -\mathcal{B}\mathcal{R}^{-1}\mathcal{B}^*$$

Therefore the obtained elements y_t can be written as

$$y_t = \mathcal{A}^t\varphi - \sum_{i=0}^{t-1}\mathcal{A}^i\mathcal{B}\mathcal{R}^{-1}\mathcal{B}^*w_{t-1-i},\ t \in \{0,\dots,T\}$$

Substituting this function into the second equation of the system (2.18) and using the boundary condition $w_T = 0$, we have

$$w_t^0 = \sum_{i=0}^{T-t}\mathcal{A}^{*i}\mathcal{Q}y_{t+i}^0, \quad t \in \{0,\dots,T\}. \tag{2.20}$$

It easy to obtain the following equalities

$t=T-1: \quad w_{T-1} = \mathcal{A}^*w_T + \mathcal{Q}y_{T-1} = \mathcal{Q}y_{T-1}$

$t=T-2: \quad w_{T-2} = \mathcal{A}^*w_{T-1} + \mathcal{Q}y_{T-2} = \mathcal{A}^*\mathcal{Q}y_{T-1} + \mathcal{Q}y_{T-2}$

$t=T-3: \quad w_{T-3} = \mathcal{A}^*w_{T-2} + \mathcal{Q}y_{T-3} = \mathcal{A}^*(\mathcal{A}^*w_{T-1} + \mathcal{Q}y_{T-2}) + \mathcal{Q}y_{T-3} =$

$\mathcal{A}^{*2}\mathcal{Q}y_{T-1} + \mathcal{A}^*\mathcal{Q}y_{T-2} + \mathcal{Q}y_{T-3}$

. . .

$$w_t^0 = \sum_{i=0}^{T-t}\mathcal{A}^{*i}\mathcal{Q}y_{t+i}^0, \quad t \in \{0,\dots,T\}.$$

These equalities prove that the couple of the functions $(y_t^0, w_t^0),\ t \in \{0,\dots,T\}$ satisfies the second equation of the system (2.18) and condition $y_0^0 = \varphi$, $w_T^0 = 0$.

For the proof of the theorem it is sufficient now to show that

$$v_t^0 = -\mathcal{R}^{-1}\mathcal{B}^* w_t^0, \quad t \in \{0, \dots, T\}.$$

Multiplying both sides (2.14) on $(\mathcal{R} + L^*\mathcal{Q}L)$ we have

$$(\mathcal{R} + L^*\mathcal{Q}L)v^0 = -(\mathcal{R} + L^*\mathcal{Q}L)(\mathcal{R} + L^*\mathcal{Q}L)^{-1}L^*\mathcal{Q}\omega$$

that is equivalent to the following equalities

$$\mathcal{R}v^0 + L^*\mathcal{Q}Lv^0 + L^*\mathcal{Q}w = 0$$

$$y^0 = Lv^0 + \omega \implies Lv^0 = y_0 - \omega,$$

$$\mathcal{R}v^0 + L^*\mathcal{Q}(y^0 - \omega) = -L^*\mathcal{Q}\omega$$

such that $\mathcal{R}v^0 = -L^*\mathcal{Q}y^0$ and, finally $v^0 = -\mathcal{R}^{-1}L^*\mathcal{Q}y^0$.

Find now for the operator L its conjugate operator $L^* : \mathcal{B}_T^0(V) \to \mathcal{B}_T(E)$ using the following reasons

$(Lf, \gamma)_{\mathcal{B}_\mathcal{T}(\mathcal{V})} = (f, L^*\gamma)_{\mathcal{B}_\mathcal{T}(\mathcal{V})};$
$(Lf, \gamma) = ((Lf)_0, \gamma_0) + ((Lf)_1, \gamma_1) + \cdots + ((Lf)_T, \gamma_T) =$
$(0, \gamma_0) + (\mathcal{B}f_0, \gamma_1) + (\mathcal{A}\mathcal{B}f_0 + \mathcal{B}f_1, \gamma_2) + (\mathcal{A}^2\mathcal{B}f_0 + \mathcal{A}\mathcal{B}f_1 + \mathcal{B}f_2, \gamma_3) + \cdots =$
$(0, \gamma_0) + (\mathcal{B}f_0, \gamma_1) + (\mathcal{A}\mathcal{B}f_0, \gamma_2) + (\mathcal{B}f_1, \gamma_2) + (\mathcal{A}^2\mathcal{B}f_0, \gamma_3) + (\mathcal{A}\mathcal{B}f_1, \gamma_3) + (\mathcal{B}f_2, \gamma_3) + \cdots =$
$(0, \gamma_0) + (f_0, \mathcal{B}^*\gamma_1) + (f_0, \mathcal{A}^*\mathcal{B}^*\gamma_2) + (f_1, \mathcal{B}^*\gamma_2) + (f_0, \mathcal{A}^{*2}\mathcal{B}^*\gamma_3) + (f_1, \mathcal{A}^*\mathcal{B}^*\gamma_3) + (f_2, \mathcal{B}^*\gamma_3)$
$+ \cdots = grouping \;\; elements =$
$(0, \gamma_0) + (f_0, \mathcal{B}^*\gamma_1 + \mathcal{A}^*\mathcal{B}^*\gamma_2 + \mathcal{A}^{*2}\mathcal{B}^*\gamma_3 + \cdots) +$
$(f_1, \mathcal{B}^*\gamma_2 + \mathcal{A}^*\mathcal{B}^*\gamma_3 + \cdots) + (f_2, \mathcal{B}^*\gamma_3 + \cdots) + \cdots$

Thus, we have the following formula for the L^*:

$$(L^*\beta)_t = \mathcal{B}^*\beta_{t+1} + \mathcal{B}^*\mathcal{A}^*\beta_{t+2} + ... + \mathcal{B}^*\mathcal{A}^{*T-t-1}\beta_T, \;\; (L^*\beta)_T = 0. \tag{2.21}$$

Therefore, from (2.20) we have

$$(\mathcal{B}^*w^0)_t = \sum_{i=0}^{T-t-1} \mathcal{B}^*\mathcal{A}^{*i}\mathcal{Q}y^0_{t+i+1} = (L^*\mathcal{Q}y^0)_t \;\; t \in \{0, \dots, T\}. \tag{2.22}$$

This yields that $v^0 = -\mathcal{R}^{-1}(L^*\mathcal{Q}y^0) = -\mathcal{R}^{-1}\mathcal{B}^*w^0$. Theorem is proved.

Theorem 6. *Optimal control problem (2.1),(2.9) has a unique solution , which is defined by formula*

$$u^0(t, s) = -R^{-1}B^*z(t, s), \;\; t \in \{0, \dots, T\}, \;\; s \in \mathbb{Z},$$

where $z(t, s)$ is given by (2.15)—(2.17).

Proof. The uniqueness of the optimal control was established before. Let $x(t, s), z(t, s)$, $t \in \{0, \dots, T\}, s \in \mathbb{Z}_+$ is a solution of the system (2.15)—(2.17). Consider the following function

$$\hat{u}(t, s) = -R^{-1}B^*z(t, s), \;\; t \in \{0, \dots, T\}, \;\; s \in \mathbb{Z}_+.$$

Taking into account the introduced notations the systems (2.15)—(2.17) we can rewrite as

$$y_{t+1} = \mathcal{A}y_t - \mathcal{B}\mathcal{R}^{-1}\mathcal{B}^* w_t,\ w_t = \mathcal{A}^* w_{t+1} + \mathcal{Q}y_{t+1},\ y_0 = \varphi,\ w_T = 0.$$

Then it follows immediately that

$$y_t = \mathcal{A}^t\varphi - \sum_{i=0}^{t-1}\mathcal{A}^i\mathcal{B}\mathcal{R}^{-1}\mathcal{B}^* w_{t-1-i},\ t \in \{0,\dots,T\}$$

$$w_t = \sum_{i=0}^{T-t}\mathcal{A}^{*i}\mathcal{Q}y_{t+i}, \qquad \hat{v}_t = \mathcal{R}^{-1}\mathcal{B}^* w_t, \tag{2.23}$$

where $(\hat{v}_t)(s) = \hat{u}(t,s)$. According (2.22), (2.21), the element $\hat{v} = (\hat{v}_0,\dots,\hat{v}_T)$ from the space $\mathcal{B}_T(V)$ can be presented in the form

$$\hat{v} = -\mathcal{R}^{-1}L^*\mathcal{Q}y,\ \text{where}\ y = (y_0,...,y_T) \in \mathcal{B}_T(E).$$

Then

$$\mathcal{R}\hat{v} = -L^*\mathcal{Q}y.$$

From other hand, from the first equality of (2.18) we can find that

$$y = \omega - L\mathcal{R}^{-1}\mathcal{B}^* wa \text{ and } y = \omega + L\hat{v},$$

where $w = (w_0,...,w_1) \in \mathcal{B}_T(E)$. Therefore

$$\mathcal{R}\hat{v} = -L^*\mathcal{Q}y + L^*\mathcal{Q}\omega - L^*\mathcal{Q}\omega = L^*\mathcal{Q}(\omega - y) - L^*\mathcal{Q}\omega = -L^*\mathcal{Q}L\hat{v} - L^*\mathcal{Q}\omega$$

and

$$\hat{v} = -(\mathcal{R} + L^*\mathcal{Q}L)^{-1}L^*\mathcal{Q}\omega.$$

Thus, the element $\hat{v}$ coincides with element v^0 defined by formula (2.14). Hence, $\hat{u}(t,s) = u^0(t,s) = -\mathcal{R}^{-1}\mathcal{B}^* z(t,s)$ is optimal control problem. The theorem proved.

2.3.1 Optimal control in feedback form

In this paragraph we wish to find another presentation for optimal solution.

Namely, we would like to find the operators $\mathcal{P}_t,\ \ t \in \{0,1,...,T\}$ such that the optimal control v_t^0 is determined as resulting action of some operator $\mathcal{P}_t$ acting on trajectory such that the following equality

$$v_t^0 = \mathcal{P}_t y_t^0, \quad t \in \{0,\dots,T\}$$

holds.

Such kind problem statement is traditional for automation theory and engineering reasons.

Denote by $\mathcal{P}_t : l_2(E) \to l_2(E)$, $t \in \{0,\dots,T\}$ some collection of the linear bounded operators, satisfying the following condition $\mathcal{P}_T = 0$.

Let $v^0 \in \mathcal{B}_T(V)$ is an optimal control in the initial problem (2.1)—(2.9), and $y^0 \in \mathcal{B}_T(E)$ is the corresponding solution of the (2.10).

The problem is to find the operators $\mathcal{P}_t$, $t \in \{0,\dots,T\}$ such that

$$v_t^0 = -\mathcal{R}^{-1}\mathcal{B}^*\mathcal{P}_t y_t^0, \quad t \in \{0,\dots,T\}. \tag{2.24}$$

Here $v^0 = (v_0^0, v_1^0, v_2^0, \cdots, v_T^0)$, $y^0 = (y_0^0, y_1^0, y_2^0, \cdots, y_T^0)$, and the additional term of $-\mathcal{R}^{-1}\mathcal{B}^*$ we take for convenience sake. We call the formulated problem as a control problem with the feedback control for the system (2.1), (2.9).

Theorem 7. *If the feedback control problem (2.24) is solvable, then the operators* $\mathcal{P}_t$ *satisfy the following system of equations*

$$\mathcal{P}_{t-1} + \big(\mathcal{Q} + \mathcal{A}^*\mathcal{P}_t\big)\mathcal{B}\mathcal{R}^{-1}\mathcal{B}^*\mathcal{P}_{t-1} = \big(\mathcal{Q} + \mathcal{A}^*\mathcal{P}_t\big)\mathcal{A}, \tag{2.25}$$

with the boundary conditions of

$$\mathcal{P}_T = 0,\ t \in \{0, \dots, T\}. \tag{2.26}$$

and the optimal trajectory are defined as a solution of the following Cauchy problem

$$y_{t+1} = (\mathcal{A} - \mathcal{B}\mathcal{R}^{-1}\mathcal{B}^*\mathcal{P}_t)y_t,\ t \in \{0, \dots, T\}, \quad y_0 = \varphi. \tag{2.27}$$

Moreover, the optimal cost function value is $J^0 = (\mathcal{P}_0\varphi, \mathcal{A}\varphi)$.

Proof. Let the feedback control problem of (2.24) is solvable. Then the function y^0 satisfies the condition (2.27). To show this it is sufficient into (2.10) to put $v^0 = -\mathcal{R}^{-1}\mathcal{B}^*\mathcal{P}_t y_t^0$:

$$y_{t+1} = \mathcal{A}y_t + \mathcal{B}v_t = \mathcal{A}y_t - \mathcal{B}\mathcal{R}^{-1}\mathcal{B}^*\mathcal{P}_t y_t = (\mathcal{A} - \mathcal{B}\mathcal{R}^{-1}\mathcal{B}^*\mathcal{P}_t)y_t$$

But from other side, the solution of (2.27) we can represented as

$$y_t^0 = F_{t-1}(F_{t-2}...(F_s y_s^0))\ \forall\ t > s \geq 0,\ \text{where } F_t = \mathcal{A} - \mathcal{B}\mathcal{R}^{-1}\mathcal{B}^*\mathcal{P}_t.$$

$y_{t+1} = F_t y_t$ with initial conditions $y_s = y_s^0$, where $s > 0$-$\forall index$.

Substituting this into the $v^0 = -\mathcal{R}^{-1}L^*\mathcal{Q}y^0$ which proves the required equality of Theorem 6. Now show how the operator L^*operates on element Qy^0:

$$\begin{aligned} L^*(Qy^0)_t = \mathcal{B}^*(Qy^0)_{t+1} + \cdots + \mathcal{B}^*\mathcal{A}^{*(T-t-1)}(Qy^0)_T = \\ \mathcal{B}^*[(Qy^0)_{t+1} + \mathcal{A}^*(Qy^0)_{t+2} + \cdots + \mathcal{A}^{*(T-t-1)}(Qy^0)_T]. \end{aligned}$$

Therefore

$$\begin{aligned} -\mathcal{R}^{-1}\mathcal{B}^*\mathcal{P}_t y_t^0 = v^0 = & \\ & -\mathcal{R}^{-1}L^*\mathcal{Q}y^0 = \\ & -\mathcal{R}^{-1}\mathcal{B}^*[(Qy^0)_{t+1} + \mathcal{A}^*(Qy^0)_{t+2} + \cdots + \mathcal{A}^{*(T-t-1)}(Qy^0)_T]. \end{aligned}$$

We will take into account that, solutions y_{t+1}^0 we can present through index s of the system: $y_{t+1} = F_t y_t,\ \ y_s = F_s y_s$

$$\begin{cases} y_{t+1}^0 & = F_t y_t^0; \\ y_{t+2}^0 & = F_{t+1}(F_t y_t^0); \\ & \cdots; \\ y_T^0 & = F_{T-1}(F_{T-2}(\cdots F_t y_t^0)) \end{cases}$$

Then we have

$$\mathcal{P}_t y_t^0 = QF_t y_t^0 + \mathcal{A}^*QF_{t+1}F_t y_t + \cdots + \mathcal{A}^{*(T-t-1)}QF_{T-1}F_{T-2}\cdots F_t y_t$$

or

$$\mathcal{P}_t y_t^0 = (QF_t + \mathcal{A}^*QF_{t+1}F_t + \cdots + \mathcal{A}^{*(T-t-1)}QF_{T-1}F_{T-2}\cdots F_t)y_t$$

Hence, the needed operators $\mathcal{P}_t$ satisfy

$$\mathcal{P}_{t-1} = \mathcal{Q}F_{t-1} + \mathcal{A}^*\mathcal{Q}F_tF_{t-1} + \ldots + \mathcal{A}^{*T-t}\mathcal{Q}F_{T-1}\ldots F_{t-1},\ \mathcal{P}_T = 0,$$

and this we can rewrite as the following recurrent formulas

$$\mathcal{P}_{t-1} + \big(\mathcal{Q} + \mathcal{A}^*\mathcal{P}_t\big)\mathcal{B}\mathcal{R}^{-1}\mathcal{B}^*\mathcal{P}_{t-1} = \big(\mathcal{Q} + \mathcal{A}^*\mathcal{P}_t\big)\mathcal{A},\quad t = 1,\ldots,T,\ \mathcal{P}_T = 0.$$

Let $\mathcal{P}_t^0$ is a solution of (2.26). Since the following formulas are true

$$(\mathcal{P}_{t-1}^0 y_{t-1}^0, \mathcal{A}y_{t-1}^0) - (\mathcal{P}_t y_t^0, \mathcal{A}y_t^0) = (\mathcal{P}_{t-1}^0 y_{t-1}^0, \mathcal{A}y_{t-1}^0) -$$

$$-(\mathcal{A}^*\mathcal{P}_t^0(\mathcal{A} - \mathcal{B}\mathcal{R}^{-1}\mathcal{B}^*\mathcal{P}_{t-1}^0)y_{t-1}^0, y_t^0) = (\mathcal{P}_{t-1}^0 y_{t-1}^0, \mathcal{A}y_{t-1}^0) - (\mathcal{P}_{t-1}^0 y_{t-1}^0, y_{t-1}^0)$$

$$+(\mathcal{Q}(\mathcal{A} - \mathcal{B}\mathcal{R}^{-1}\mathcal{B}^*\mathcal{P}_{t-1}^0)y_{t-1}^0, y_t^0) = (\mathcal{P}_{t-1}^0 y_{t-1}^0, \mathcal{B}\mathcal{R}^{-1}\mathcal{B}^*\mathcal{P}_{t-1}^0)y_{t-1}^0) +$$

$$+(\mathcal{Q}y_t^0, y_t^0) = (\mathcal{R}v_{t-1}^0, v_{t-1}^0) + (\mathcal{Q}y_t^0, y_t^0),$$

then

$$J(v^0) = \sum_{t=1}^T (\mathcal{Q}y_t^0, y_t^0) + (\mathcal{R}v_{t-1}^0, v_{t-1}^0) = \sum_{t=1}^T \Big[(\mathcal{P}_{t-1}^0 y_{t-1}^0, \mathcal{A}y_{t-1}^0)$$
$$-(\mathcal{P}_t^0 y_t^0, \mathcal{A}y_{t-1}^0)\Big] = (\mathcal{P}_0 y_0^0, \mathcal{A}y_0^0) = (\mathcal{P}_0^0\varphi, \mathcal{A}\varphi).$$

Theorem is proved.

2.4 Optimal control via boundary data

The results developed in the previous section can also be extended to other classes of discrete $2-D$ models for control problems optimization appeared in the complex gas network model.

As it was shown in the Chapter 1 for some cases the part of the initial data can be treated as a control parameter for discrete $2-D$ system. In particular, it is of interest to determine an optimal control programm for gas pressure and gas flow at the pipe when the gas regulation in the time is feasible at a fixed node of the pipe.

In particular, an interesting problem is optimization due to the boundary condition $x(t,0) = \psi(t)$ which is given above in (1.75)—(1.76). In other words, we consider the following system

$$x(t+1,s) = A_0x(t,s) + A_1x(t,s+1) + A_2x(t,s-1) \tag{2.28}$$

with initial and boundary condition:

$$\begin{aligned} x(0,s) &= \varphi(s), \qquad s \in \mathbb{Z}_+ \setminus \{0\} \\ x(t,0) &= \psi(t) = u_t, \quad t = 0,1,2,...,T-1 \end{aligned}$$

The optimization problem is to minimize the cost functional of the form

$$J(u) = \sum_{t=1}^T \Big[\sum_{s\in\mathbb{Z}_+} \big(Qx(t,s), x(t,s)\big) + \big(Ru(t), u(t)\big)\Big], \tag{2.29}$$

where $Q \geq 0$ and $R > 0$. We keep here the notations of the Section 2.3 of E, V for the finite dimensional spaces and corresponding operators. In fact, this is equivalent to spaces $\mathbb{R}^n$ and $\mathbb{R}^m$ when some bases are chosen in E and V, respectively.

2.4.1 Embedding to the general system case

The considered problem can be reduced to the optimal control problem with control input in the right hand system side which is investigated in the Section 2.3. To realize this reduction rewrite (2.28)—(2.29) in operator form. In order to rewrite correctly the boundary control data by control input in the right hand side of the system, note first

$$\begin{aligned} x(1,0) =& u_1; \\ x(1,1) =& A_0x(0,1) + A_1x(0,2) + A_2x(0,0) = A_0\varphi(1) + A_1\varphi(2) + A_2u_0; \\ x(1,2) =& A_0x(0,2) + A_1x(0,3) + A_2x(0,1) = A_0\varphi(2) + A_1\varphi(3) + A_2\varphi(1); \\ & \ldots\ldots\ldots\ldots\ldots \\ x(2,0) =& u_2; \\ x(2,1) =& A_0x(1,1) + A_1x(1,2) + A_2x(1,0) = A_0x(1,1) + A_1x(1,2) + A_2u_1; \end{aligned} \tag{2.30}$$

The obtained recurrent formulas lead to the following definitions of the operators $\mathcal{A} : l_2(E) \to l_2(E)$ and $\mathcal{B} : V \to l_2(E)$ (which are needed later) as follows

$$\mathcal{A} :(\xi_1, \xi_2, \xi_3, \ldots) \longrightarrow (A_0\xi_1 + A_1\xi_2 + A_2 \cdot 0,\ \ A_0\xi_2 + A_1\xi_3 + A_2\xi_1, A_0\xi_3 + A_1\xi_4 + A_2\xi_2, \ldots)$$

$$\mathcal{B} :u \longrightarrow (A_2u, 0, 0, 0, \ldots, 0, \ldots). \tag{2.31}$$

Then the equalities (2.30) can be rewritten as the ordinary discrete system of the form

$$y_{t+1} = \mathcal{A}y_t + \mathcal{B}v_t,\ \ t = 0, 1, \ldots, T-1 \tag{2.32}$$

where

$$y_t = \{x(t,0), x(t,1), \ldots\},\ \ v_t = u_t,\ \ t = 0, 1, \ldots, T-1 \tag{2.33}$$

and $\mathcal{A}$, $\mathcal{B}$ are the linear operators defined above. Denote by $\mathcal{B}_T(E)$ and $\mathcal{B}_T(V)$ the spaces of the functions defined on the set $\{0, 1, \ldots, T-1\}$ with values in the spaces $l_2(E)$ and V, respectively, such that $\mathcal{B}_T(E) = \big(l_2(E)\big)^T$, $\mathcal{B}_T(V) = \big(V\big)^T$.

Next we wish to rewrite the system by operator equality. By this reason we introduce the following notations. Let $\mathcal{L} : \mathcal{B}_{T+1}(V) \to \mathcal{B}^0_{T+1}(E)$ is the operator of the form

$$(\mathcal{L}\gamma)_t = \mathcal{B}\gamma_{t-1} + \mathcal{A}\mathcal{B}\gamma_{t-2} + ... + \mathcal{A}^{t-1}\mathcal{B}\gamma_0, t > 0, (\mathcal{L}\gamma)_0 = 0,$$

where $\gamma = \big(\gamma_0, \gamma_0, \ldots, \gamma_T\big) \in \mathcal{B}_{T+1}(E)$, $\mathcal{B}^0_T(E)$ denotes the subspace of $\mathcal{B}_T(E)$ containing the functions with zero values at $t = 0$, and $\omega = \{\varphi, \mathcal{A}\varphi, \mathcal{A}^2\varphi, ..., \mathcal{A}^{T-1}\varphi\} \in \mathcal{B}_{T+1}(E)$.

Thus, in this case the solution of (2.32) can be presented in the general operator form as

$$y = \mathcal{L}v + \omega$$

the coordinates of which are

$$y_t = (\mathcal{L}v + \omega)_t = \mathcal{A}^t\varphi + \mathcal{A}^{t-1}\mathcal{B}v_0 + ... + \mathcal{A}\mathcal{B}v_{t-2} + \mathcal{B}v_{t-1}, \tag{2.34}$$

where

$$v = (u_0, u_1, u_2, ..., u_{T-1}) \in \mathcal{B}^T,\ \ y = (y_0, y_1, y_2, ..., y_T) \in \mathcal{B}^{T+1} = \underbrace{(l^2, l^2, ..., l^2)}_{(T+1)-times}.$$

It is easy to check the validity of the obtained formula. Indeed, we have

$$
\begin{aligned}
&y_0 = \{x(0,s), s \in \mathbb{Z}_+\} = (Lv+\omega)(0) = 0 + \omega(0) = \varphi \\
&y_1 = \{x(1,s), s \in \mathbb{Z}_+\} = (\mathcal{L}v)(1) + (\omega)(1) = \mathcal{A}\varphi + \mathcal{B}v_0 \\
&y_2 = \{x(2,s), s \in \mathbb{Z}_+\} = (\mathcal{L}v)(2) + (\omega)_2 = \mathcal{B}v_1 + \mathcal{A}\mathcal{B}v_0 + \mathcal{A}^2\varphi = \\
&= \begin{bmatrix} A_2v_1 \\ 0 \\ 0 \\ \dots\dots\dots\dots\dots \end{bmatrix} + \mathcal{A} \begin{bmatrix} A_2v_0 \\ 0 \\ 0 \\ \dots\dots\dots\dots \end{bmatrix} + \mathcal{A} \begin{bmatrix} A_0\varphi_1 + A_1\varphi_2 + A_2\cdot 0 \\ A_0\varphi_2 + A_1\varphi_3 + A_2\varphi_1 \\ \dots\dots\dots \end{bmatrix}
\end{aligned}
\tag{2.35}
$$

which coincides with the calculation in (2.30).

Since

$$
\begin{aligned}
\big(Qx,x\big) + \big(Ru,u\big) &= \big(\mathcal{Q}y,y\big) + \big(\mathcal{R}v,v\big) = \\
&\big(\mathcal{Q}(\mathcal{L}v+\omega),(\mathcal{L}v+\omega)\big) + \big(\mathcal{R}v,v\big) = \\
&\big(\mathcal{Q}\mathcal{L}v,\mathcal{L}v\big) + \big(\mathcal{Q}\mathcal{L}v,\omega\big) + \big(\mathcal{Q}\omega,\mathcal{L}v\big) + \big(\mathcal{Q}\omega,\omega\big) + \big(\mathcal{R}v,v\big) = \\
&\big(\mathcal{L}^*\mathcal{Q}\mathcal{L}v,v\big) + \big(\mathcal{Q}\mathcal{L}v,\omega\big) + \big(\mathcal{L}^*\mathcal{Q}\omega,v\big) + \big(\mathcal{Q}\omega,\omega\big) + \big(\mathcal{R}v,v\big) = \\
&\big((\mathcal{R}+\mathcal{L}^*\mathcal{Q}\mathcal{L})v,v\big) + \big(\mathcal{Q}\mathcal{L}v,\omega\big) + \big(\mathcal{L}^*\mathcal{Q}\omega,v\big) + \big(\mathcal{Q}\omega,\omega\big) = \\
&\big((\mathcal{R}+\mathcal{L}^*\mathcal{Q}\mathcal{L})v,v\big) + 2\big(\mathcal{L}^*\mathcal{Q}\omega,v\big) + \big(\mathcal{Q}\omega,\omega\big)
\end{aligned}
$$

then the cost functional (2.29) can be represented in the operator form as

$$
J(v) = ((\mathcal{R}+\mathcal{L}^*\mathcal{Q}\mathcal{L})v,v)_{\mathcal{B}} + 2(\mathcal{L}^*\mathcal{Q}\omega,v)_{\mathcal{B}} + (\mathcal{Q}\omega,\omega)_{\mathcal{B}}, \tag{2.36}
$$

where $\omega = \big(\varphi, \mathcal{A}\varphi, \dots, (\mathcal{A})^T\varphi\big) \in \mathcal{B}_T(E)$.

Here the symbol $(\cdot,\cdot)$ means the inner product in the Hilbert space $\mathcal{B}_T(E)$ (or in the space $\mathcal{B}_T(V)$) defined as $(a,b)_{\mathcal{B}} = \sum\limits_{i=0}^{T}(a_i,b_i)_E$. The operators of $\mathcal{R} : \mathcal{B}_T(V) \to \mathcal{B}_T(V)$ and $\mathcal{Q} : \mathcal{B}_T(E) \to \mathcal{B}_T(E)$ are given obviously:

$$
\big(\mathcal{R}u\big)(t,s) = Ru(t,s),\ t \neq 0,\ \big(\mathcal{Q}x\big)(t,s) = Qx(t,s),\ t = 0,\dots,T,\ s \in \mathbb{Z}_+.
$$

Since $\mathcal{G} \geq 0$, $\mathcal{R} > 0$ then the operator $\mathcal{R} + \mathcal{L}^*\mathcal{Q}\mathcal{L}$ is inverted.

The minimum of (2.36) satisfies to the equation

$$
\frac{\partial J(v)}{\partial v} = (\mathcal{R}+\mathcal{L}^*\mathcal{Q}\mathcal{L})v + 2\mathcal{L}^*\mathcal{Q}\omega = 0
$$

Hence

$$
v^0 = -(\mathcal{R}+\mathcal{L}^*\mathcal{Q}\mathcal{L})^{-1}\mathcal{L}^*\mathcal{Q}\omega. \tag{2.37}
$$

Next we can check inequality $J(v) - J(v^0) \geq 0$. If it is true then v^0 is optimal control in the initial problem. Denoting $\Pi = (\mathcal{R}+\mathcal{L}^*\mathcal{Q}\mathcal{L})$, we have

$$
\begin{aligned}
J(v) - J(v^0) &= (\Pi v,v) + 2(\mathcal{L}^*\mathcal{Q}\omega,v) - (\mathcal{L}^*\mathcal{Q}\omega,\Pi^{-1}\mathcal{L}^*\mathcal{Q}\omega) = \\
&(\Pi v + \Pi\Pi^{-1}\mathcal{L}^*\mathcal{Q}\omega,v) + (\mathcal{L}^*\mathcal{Q}\omega,v+\Pi^{-1}\mathcal{L}^*\mathcal{Q}\omega) = \\
&(\Pi(v-v^0),v) + (\mathcal{L}^*\mathcal{Q}\omega,v-v^0) = (\Pi(v-v^0),v) - (\Pi v^0,v-v^0) = \\
&(\Pi(v-v^0),v-v^0).
\end{aligned}
$$

Since $\mathcal{G} \geq 0$, $\mathcal{R} > 0$ then $J(v) - J(v^0) = ((\mathcal{R} + \mathcal{L}^*\mathcal{Q}\mathcal{L})(v - v_0), (v - v_0))_{\mathcal{B}} > 0$ for any $v \in \mathcal{B}_T(V), v \neq v_0$. This means that v^0 is a unique optimal solution for the problem (2.28),(2.29). Thus the optimization problem (2.28),(2.29) is solvable and optimal control given by the formula (2.37).

But the obtained formula is not suitable for calculation, since the search of the invert operator is not trivial task. By this reason in the next Section for the obtained function we consider another presentation by the so-called adjoint variables.

2.4.2 Conjugate system

First we give some general facts. As well known for the linear equation given in some Hilbert space X in the form

$$Hx = f \tag{2.38}$$

their conjugate system is defined as

$$H^*z = g \tag{2.39}$$

where H and H^* are prime and their adjoint operators acting in X.

The equalities obtained in the Section above are

$$y_{t+1} = \mathcal{A}y_t + \mathcal{B}v_t,\ t \in \{0, 1, \ldots, T-1\},\ y_0 = \varphi$$

Rewriting the last as $y_{t+1} - \mathcal{A}y_t = \mathcal{B}v_t$ we have the operator equation $Hy = f$ where the operator $H : \mathcal{B}^{T+1} \Rightarrow \mathcal{B}^T$ is given by $(Hy)(t) = y_{t+1} - \mathcal{A}y_t$ and $f = \mathcal{B}v$.

It easy to see that:

$$(Hy, z)_{\mathcal{B}^T} = (y_1 - \mathcal{A}y_0, z_0) + (y_2 - \mathcal{A}y_1, z_1) + ... + (y_T - \mathcal{A}y_{T-1}, z_T) =$$

$$= (y_0, -\mathcal{A}^*z_0) + (y_1, z_0 - \mathcal{A} * z_1) + ... + (y_{T-1}, z_{T-2} - \mathcal{A}^*z_{T-1}) + (y_T, z_T) = (y, H^*z)_{\mathcal{B}^{T+1}}$$

In other words the adjoint operator $H^* : \mathcal{B}^T \Rightarrow \mathcal{B}^{T+1}$ is defined as follows

$$\begin{aligned}
(H^*z)_0 =& - \mathcal{A}^*z_0\\
(H^*z)_s =& z_{s-1} - \mathcal{A}^*z_s, \quad s = 1, ...T-1\\
(H^*z)_T =& z_T
\end{aligned}$$

Then for the considered case the adjoint equation of (2.39) is

$$\begin{aligned}
-\mathcal{A}^*z_0 =& g_0,\\
z_0 - \mathcal{A}^*z_1 =& g_1,\\
&..................\\
z_{T-2} - \mathcal{A}^*z_{T-1} =& g_{T-1},\\
z_T =& g_T.
\end{aligned} \tag{2.40}$$

In order to find the required adjoint operator $\mathcal{A}^* : l^2 \to l^2$ for the operator $\mathcal{A}$ that is given as

$$\mathcal{A} : (\xi_1, \xi_2, \xi_3, \ldots) \longrightarrow (A_0\xi_1 + A_1\xi_2 + A_2 \cdot 0, A_0\xi_2 + A_1\xi_3 + A_2\xi_1, A_0\xi_3 + A_1\xi_4 + A_2\xi_2, \ldots)$$

we calculate the inner product

$$(\mathcal{A}\xi,\eta)_{l^2} = (A_0\xi_1 + A_2\xi_2,\eta_1) + (A_0\xi_2 + A_1\xi_1 + A_2\xi_3,\eta_2) + ... =$$
$$= (\xi_1, A_0^*\eta_1 + A_1^*\eta_2) + (\xi_2, A_2^*\eta_1 + A_0^*\eta_2 + A_1^*\eta_3) + ... = (\xi,\mathcal{A}^*\eta)_{l^2}$$

Hence

$$\mathcal{A}^* : (\eta_1,\eta_2,\eta_3,...,) \to (A_0^*\eta_1 + A_1^*\eta_2, A_2^*\eta_1 + A_0^*\eta_2 + A_1^*\eta_3,...)$$

To determine correctly the adjoint system in the coordinate form we rewrite in details the relations 2.40). Then for the element z_0 we have

$$z_0 = (z(0,0), z(0,1),..., z(0,s),...) \Rightarrow -\mathcal{A}^*z_0 = g_0 \Rightarrow -\mathcal{A}^*z_0(s) = g_0(s)$$

such that

$$\begin{aligned} -A_0^*z(0,0) - A_1^*z(0,1) &= g(0,0)\\ -A_0^*z(0,1) - A_1^*z(0,2) - A_2^*z(0,0) &= g(0,1)\\ -A_0^*z(0,s) - A_1^*z(0,s+1) - A_2^*z(0,s-1) &= g(0,s)\ \forall\ s > 1. \end{aligned}$$

Analogously, for the element z_1 we have

$$z_1 = (z(1,0), z(1,1),..., z(1,s),...) \Rightarrow z_0 - \mathcal{A}^*z_1 = g_1 \Rightarrow z_0(s) - \mathcal{A}^*z_1(s) = g_1(s)$$

such that

$$\begin{aligned} &z(0,0) - A_0^*z(1,0) - A_1^*z(1,1) = g(1,0)\\ &z(0,1) - A_0^*z(1,1) - A_1^*z(1,2) - A_2^*z(1,0) = g(1,1)\\ &z(0,s) - A_0^*z(1,s) - A_1^*z(1,s+1) - A_2^*z(1,s-1) = g(1,s) \end{aligned}$$

And for the arbitrary element z_t:

$$z_t = (z(t,0), z(t,1),..., z(t,s),...) \Rightarrow z_{t-1} - \mathcal{A}^*z_t = g_t \Rightarrow z_{t-1}(s) - \mathcal{A}^*z_t(s) = g_t(s)$$

such that

$$\begin{aligned} &z(t-1,0) - A_0^*z(t,0) - A_1^*z(t,1) = g(t,0)\\ &z(t-1,1) - A_0^*z(t,1) - A_1^*z(t,2) - A_2^*z(t,0) = g(t,1)\\ &z(t-1,s) - A_0^*z(t,s) - A_1^*z(t,s+1) - A_2^*z(t,s-1) = g(t,s) \end{aligned}$$

where $t \in \{1,\dots,T\}$.

Summarizing the obtained above we have the asked adjoint system

$$\begin{aligned} &z(t,s) = A_0^*z(t+1,s) + A_1^*z(t+1,s+1) + A_2^*z(t+1,s-1) + g(t+1,s)\\ &z(t,0) = A_0^*z(t+1,0) + A_1^*z(t+1,1) + g(t+1,0),\\ &z(T-1,s) = g(T,s), \quad t \in \{0,1,\dots,T-1\},\ s = 1,2,... \end{aligned} \tag{2.41}$$

2.4.3 Boundary optimal control

The aim of this paragraph is to obtain the representation of $u^o = (u_0^o, u_1^o, u_2^o, ..., u_{T-1}^o)$ by means of adjoint variable z. The following result is hold.

Theorem 8. *The optimal control u^0 of the problem (2.28)-(2.29) is given as*

$$u_t^0 = -R^{-1}A_2^* z^0(t,0), \ \ t = 0, 1, \dots, T-1, \tag{2.42}$$

where $z(t,0)$ is determined by the following system of equations

$$z(t,s) = A_0^* z(t+1,s) + A_1^* z(t+1,s+1) + A_2^* z(t+1,s-1) + Qx(t+1,s), \ \ s \in \mathbb{Z},$$

$$\begin{aligned} x(t+1,s) = A_0 x(t,s) + A_1 x(t,s+1) + A_2 x(t,s-1) - A_2 R^{-1} A_2^* z(t,0), \\ t = 0, 1, \dots, T-1, \end{aligned} \tag{2.43}$$

with the boundary conditions

$$x(0,s) = \varphi(s), \ \ z(T,s) = 0. \tag{2.44}$$

Proof. From (2.37) follows that the optimization problem (2.28)—(2.29) is solvable and optimal control given by the formula

$$v^0 = -(\mathcal{R} + \mathcal{L}^*\mathcal{Q}\mathcal{L})^{-1}\mathcal{L}^*\mathcal{Q}\omega.$$

Multiplying the left-hand side on $(\mathcal{R} + \mathcal{L}^*\mathcal{Q}\mathcal{L})$ and the representation for y as

$$y = \mathcal{L}v + \omega$$

gives

$$v^0 = -\mathcal{R}^{-}\mathcal{L}^*\mathcal{Q}y^0. \tag{2.45}$$

Here the operator $\mathcal{L}^*$ is conjugate operator to $\mathcal{L}$ is presented as

$$\mathcal{L}^* = \mathcal{B}^*\Lambda.$$

The conjugate operator $\mathcal{B}^* : l_2(E) \to V$ can be defined as follows

$$\begin{aligned} (\mathcal{B}u, \eta)_{l_2(E)} =& (A_2 u, v_1) + (0, v_2) + \ldots + (0, v_n) + \ldots = \\ =& (u, A_2^* v_1) + (0, v_2) + \ldots + (0, v_n) + \ldots = \\ =& (u, A_2^* v_1 + 0 \cdot v_2 + \ldots + 0 \cdot v_n + \ldots). \end{aligned}$$

Hence

$$\mathcal{B}^* : (v_1, v_2, \ldots) \to A_2^* v_1. \tag{2.46}$$

It is easy to show that the operator Λ is given as follows

$$(\Lambda\beta)_t = \beta_{t+1} + \mathcal{A}^*\beta_{t+2} + \mathcal{A}^{*2}\beta_{t+3} + ..., \ \ \ \beta_0 = 0, \ \ t = 0, 1, \ldots, T-1 \tag{2.47}$$

where the conjugate operator $\mathcal{A}^*$ is given as

$$\mathcal{A}^* : (\eta_1, \eta_2, \eta_3, ...,) \to (A_0^*\eta_1 + A_1^*\eta_2, A_2^*\eta_1 + A_0^*\eta_2 + A_1^*\eta_3, ...).$$

Thus, we have

$$v_t^0 = -\mathcal{R}^{-}\mathcal{B}^*(\mathcal{Q}y_{t+1}^0 + \mathcal{A}^*\mathcal{Q}y_{t+2}^0 + ...), \ \ t = 0, 1, \ldots, T-1 \tag{2.48}$$

Put

$$z_t = (\mathcal{Q}y^0_{t+1} + \mathcal{A}^*\mathcal{Q}y^0_{t+2} + ...),\ t = 0, 1, \ldots, T-1 \tag{2.49}$$

It is easy to see that the function z satisfies the following equation

$$z_t = \mathcal{A}^* z_{t+1} + \mathcal{Q}y^0_{t+1},\ t = 0, 1, \ldots, T-1 \tag{2.50}$$

which is called the conjugate ones to the equation (2.32).

Since the equation (2.28)and the equation (2.32) is solvable for any admissible control v, then the conjugate system (2.50) is solvable, too. Hence, the equation (2.50) has a unique solution z^0 for any y^0. It is easy to check that the equation (2.50) can be transformed to the required form of (2.43). Then from (2.48) follows

$$v^0_t = -R^{-1}B^* z^0_t.$$

In accordance with (2.46 we have $u^0_t = -R^{-1}A^*_2 z^0(t,0),\ t = 0, 1, \ldots, T-1$ which proves (2.42). The proof of (2.43) can be done similar to the Theorem 6.

2.5 Optimal control in infinite case

In this section we consider the problem (2.1)—(2.9) in the case $T \to \infty$. The solution of this problem will be given as a limit procedure based on the results of the previous section. Let $l^2_2(V)$ be the space of all sequences $u : Z_+ \times Z \to V$ such that $\sum_{(t,s)} \mid u(t,s) \mid^2_E < \infty$. For this reason the solution of the equation (1) we shall consider on the space $l^2_2(E)$, also.

It can be shown that for every admissible control $u \in l^2_2(V)$ and initial state $x(0,s) = \varphi(s)$, $\varphi \in l^2_2(E)$ there exists a unique solution iff $r(\mathcal{A}) < 1$, where $r(\mathcal{A})$ denotes the spectral radius of $\mathcal{A}$.

For some assumptions on the operators A_i (on the spectrum of the pencil for A_i) the following theorem is proved.

Theorem 9. . *Let $r(\mathcal{A}) < 1$ and the following conditions hold*

$$\begin{gathered} |A_{-N} + |A_{-N+1}| + \ldots + |A_{N-1} + |A_N| + |B|^2 < 1, \\ |R| < 1 - (|A_{-N} + |A_{-N+1}| + \ldots + |A_{N-1} + |A_N|)^2/(1 - |B|^2). \end{gathered} \tag{2.51}$$

Then the optimal control for the problem (2.1), (2.9) ($T = \infty$) can be presented in the form

$$u^0 = -\mathcal{R}^{-1}\mathcal{B}{*}Px^0(t), t \in Z_+,$$

where $x^0(t), t \in Z_+$ is a unique solution of the equation

$$x(t+1) = (\mathcal{A} - \mathcal{B}\mathcal{R}^{-1}\mathcal{B}^*P)x(t), x(0) = \varphi.$$

Here the linear bounded operator $P : l^2(E) \to l^2(E)$ is given by the equation

$$P = (\mathcal{R} + \mathcal{A}^*P)(\mathcal{A} - \mathcal{B}\mathcal{R}^{-1}\mathcal{B}^*P).$$

Moreover, the minimal value

$$J^0 = (P\varphi, \mathcal{A}\varphi).$$

Proof. Let $N > 1$ is a fixed integer and P_t, $t = 0, 1, ..., N$ are the solution for the operator equation (2.25)-(2.26). Hence the re-numbered operators $\tilde{P}_t = P_{N-t}$, $t = 0, 1, ..., N$ satisfy the equation

$$\tilde{P}_t + (\mathcal{G} + \mathcal{A}\tilde{P}_{t-1})\mathcal{B}\mathcal{R}^{-1}\mathcal{B}^*\tilde{P}_t = (\mathcal{G} + \mathcal{A}^*\tilde{P}_{t-1})\mathcal{A}, \ \tilde{P}_0 = 0, \tag{2.52}$$

the solutions of which are not depend on the integer N. Now we study the solvability of the given equation in detail. It can be shown that if on the each t-th stage the following condition

$$\|\mathcal{G}\| + \|\mathcal{A}^*\|\|\tilde{P}_{t-1}\| < 1, \ \|\mathcal{B}\mathcal{R}^{-1}\mathcal{B}^*\| < 1$$

is guaranteed, then the unique bounded solution $\tilde{P}_t$ exists for the equation (2.52). Moreover, from (2.52) the following estimate follows

$$\|\tilde{P}_t\| \leq \frac{(\|\mathcal{G} + \mathcal{A}^*\tilde{P}_{t-1}\|)\|\mathcal{A}\|}{(1 - \|\mathcal{G} + \mathcal{A}^*\tilde{P}_{t-1}\|\|\mathcal{B}\mathcal{R}^{-1}\mathcal{B}^*\|)} \leq \frac{\|\mathcal{A}\|}{(1 - \|\mathcal{B}\mathcal{R}^{-1}\mathcal{B}^*)}.$$

Hence, in order to guarantee the solvability of equation (2.52) for the next $(t-1)$-th stage it is sufficient to keep the condition $\|(\mathcal{G} + \mathcal{A}^*\tilde{P}_t)\mathcal{B}\mathcal{R}^{-1}\mathcal{B}^*\| < 1$. The previous inequality yields then that this will be true if $\|\mathcal{G}\| + \|\mathcal{A}\|^2(1 - \|\mathcal{B}\mathcal{R}^{-1}\mathcal{B}^*\|)^{-1} < 1$. Early we shown that $\| \mathcal{A} \|^2 \leq 2N\sum_{i=-N}^{N} \mid A_i \mid^2$ and $\|\mathcal{G}\| \leq \|Q\|$. Unite the given inequality we see that the conditions given in the theorem guarantee the solvability of the operator equation (2.52) for any $t = 0, 1, ...$

Further, the theorem 7 gives that the minimal cost value for each fixed integer N is equal $(\tilde{P}_N x_0, \mathcal{A}x_0)$. Let $N_2 > N_1$. Then for any admissible control $u \in \mathcal{B}([0, N_2], l^2(E))$ and initial data $x \in l_2(E)$ we have

$$\sum_{t=1}^{N_2} \big[(\mathcal{G}x_t, x_t) + (\mathcal{R}u_t, u_t)\big] \geq \sum_{t=1}^{N_1} \big[(\mathcal{G}x_t, x_t) + (\mathcal{R}u_t, u_t)\big] \geq \min_{u \in \mathcal{B}_{N_1}(E)} J(u) \geq 0. \tag{2.53}$$

Hence, $(\tilde{P}_{N_2}x, \mathcal{A}x) \geq (\tilde{P}_{N_1}x, \mathcal{A}x)$ for any $x \in l^2(E)$.

Let $\mathcal{J}_\infty(x)$ denotes the minimal value for the cost functional in optimization problem (2.1),(2.9) with initial data $x \in l^2(E)$ $(N = \infty)$. By analogy with the previous section we can shown that the optimal control in this case is

$$u^0 = -\big(\mathcal{R} + L^*\mathcal{G}L\big)^{-1}L^*\mathcal{G}w, \ \text{where} \quad w = (x, \mathcal{A}x, \mathcal{A}^2x, ...).$$

In addition, from (2.14) it follows that $J_\infty(x) = (\mathcal{P}w, w)$, where $\mathcal{P}$ is the linear operator in $\mathcal{B}(E)$ given by the formula

$$\mathcal{P} = \mathcal{G} - \mathcal{G}L\big(\mathcal{R} + L^*\mathcal{G}L\big)^{-1}L^*\mathcal{G}L$$

(others operators involved were defined early). Using (2.53) we have that for any $x \in l^2(E)$ the following inequalities are fulfilled

$$0 \leq J_\infty(x) = (\mathcal{P}w, w\big) \leq \|\mathcal{P}\|(w, w) \leq C\|\mathcal{P}\|(x, x),$$

where the constant $C \doteq 1/(1 - \|\mathcal{A}\|) > 0$. Moreover, for any integer N the following inequalities are true

$$J_\infty(x) = \min_{u \in \mathcal{B}(E)} J(u, x) = \sum_{t=1}^{\infty} \big[(\tilde{\mathcal{G}}x_t^0, x_t^0) + (\mathcal{R}u_t^0, u_t^0)\big] \geq$$

$$\geq \sum_{t=1}^{N}[(\tilde{\mathcal{G}}x_t^0, x_t^0) + (\mathcal{R}u_t^0, u^0(_t)] \geq \min_{u\in\mathcal{B}_N(V)} J(u) = (\tilde{P}_N x, \mathcal{A}x).$$

Let $0 \leq N_1 < N_2 < \ldots$ be some increasing integer sequence. Then

$$0 \leq (\tilde{P}_{N_1}x, \mathcal{A}x) \leq (\tilde{P}_{N_2}x, \mathcal{A}x) \leq \ldots \leq J_\infty(x) \leq C(x,x), \tag{2.54}$$

where the constant $C > 0$ was given above. This means that $\{\mathcal{A}^*\tilde{P}_{N_i}\}$ is a nondecreasing bounded up sequence of nonnegative selfadjoint operators. In this case the Banach-Steinhause theorem states that this operator sequence has a strong nonnegative operator limit T, that is

$$\lim_{i\to\infty} \mathcal{A}^*\tilde{P}_{N_i}x = Tx \qquad \forall \quad x \in l^2(E).$$

Since $r(\mathcal{A}) < 1$ then the operator $\mathcal{A}^*$ is invertible. Then from (2.54) it follows that the sequence $\tilde{P}_{N_i}$ is convergent, also. Let $\lim\limits_{i\to\infty} \tilde{P}_{N_i}x = Px$. We shown early that $J_\infty(x) \geq (\tilde{P}_N x, \mathcal{A}x)$ for all $x \in l^2(E)$ and any N. Taking limit at $N \to \infty$, we get $J_\infty(x) \geq (Px, \mathcal{A}x)$. Verify now that the cost functional J_∞ takes the value $(Px, \mathcal{A}x)$ on the control function $u^* = -\mathcal{R}^{-1}\mathcal{B}^*Px$. This will means that the control function $u = u^*$ is optimal. By analogy with the theorem 7 it can be shown that the control function $u_t^*,\ t \in \mathbb{Z}_+$, produces the solution $x_t^*,\ t \in \mathbb{Z}_+$ for the equation

$$x_{t+1} = (\mathcal{A} - \mathcal{B}\mathcal{R}^{-1}\mathcal{B}^*P)x_t,\ x_0 = x,\ t \in \mathbb{Z}_+.$$

This solution satisfies the following equality

$$(Px_t^*, \mathcal{A}x_t^*) - (Px_{t+1}^*, \mathcal{A}x_{t+1}^*) = (\mathcal{G}x_{t+1}^*, x_{t+1}^*) + (\mathcal{R}u_t^*, u_t^*).$$

Then

$$J(u^*) = \sum_{t=0}^{\infty}\left[\left(\mathcal{G}x_t^*, x_t^*\right) + \left(\mathcal{R}u_t^*, u_t^*\right)\right] = (Px, \mathcal{A}x) + \lim_{t\to\infty}(Px_t^*, \mathcal{A}x_t^*).$$

Since $x^* \in \mathcal{B}(E)$, then $\|x_t^*\|_{l^2(E)} \to 0,\ t \to \infty$. This proves the required equality $J_\infty(x) = J(u^*) = (Px, \mathcal{A}x)$. The proof is completed.

Another form of the optimal solution for the problem (2.28), (2.29) is given by the following theorem.

Theorem 10. *The Fourier transform*

$$\mathcal{U}_t(\omega) = \sum_{s=-\infty}^{\infty} u^0(t,s)e^{-is\omega},$$

$\omega \in [0, 2\pi]$ *of the optimal control* $u^0(t,s)$ *for the problem (2.28), (2.29)* ($T = \infty$ *can be presented as follows*

$$\mathcal{U}_t(\omega) = K(\omega)X_t(\omega), \tag{2.55}$$

where

$$K(\omega) = -[R + B'P(\omega)B]^{-1}B'P(\omega)A(\omega),\ A(\omega) = \sum_{k=-N}^{N} e^{ik\omega}A_k. \tag{2.56}$$

Here $P(\omega)$, $\omega \in [0, 2\pi]$ *satisfies the following operator equation*

$$P(\omega) = Q + A^* P(\omega)A(\omega) - A^* P(\omega)B[R + B^{'} P(\omega)B]^{-1}B^{'} P(\omega)A(\omega), \tag{2.57}$$

$X_t(\omega)$ *is the Fourier transform of the optimal trajectory* $x^0(t, s)$. *Moreover, the minimal value of the cost functional is*

$$I(u^0) = \frac{1}{2\pi}\int_0^{2\pi} (X_0(\omega), P(\omega)X_0(\omega))d\omega \tag{2.58}$$

Proof. Applying the discrete Fourier transformation to the equation (2.28) with respect to the variable s:

$$X_t(\omega) = \sum_{s\in\mathbb{Z}} x^0(t, s)e^{-is\omega},\ \omega \in [0, 2\pi]$$

leads to the system of

$$X_{t+1}(\omega) = A(\omega)X_t(\omega) + BU_t(\omega),\ A(\omega) = \sum_{k=-N}^{N} e^{ik\omega} A_k, \omega \in [0, 2\pi]$$

In accordance with the Parseval's identity the cost functional can be represented as follows

$$J(u) = \frac{1}{2\pi}\sum_{t\in Z_+}\int_0^{2\pi} (X_t(\omega), QX_t(\omega)X_T(\omega)) + (U_t(\omega), RU_t(\omega))d\omega.$$

Let $P(\omega)$, $\omega \in [0, 2\pi]$ be an arbitrary collection of nonnegative operators such from E to E such that $\int_0^{2\pi} \|P(\omega)\|d\omega < \infty$. The following identity is true

$$0 = (P(\omega)X_0(\omega), X_0(\omega) - \sum_{t\in Z_+}(P(\omega)X_t(\omega), X_t(\omega))+$$

$$+\sum_{t\in Z_+}(P(\omega)(A(\omega)P(\omega)X_t(\omega) + BU_t(\omega)), (A(\omega)P(\omega)X_t(\omega) + BU_t(\omega)).$$

Integrating this identity on $\omega \in [0, 2\pi]$ and adding then the result to J, we obtain

$$J(u^0) = \frac{1}{2\pi}\int_0^{2\pi} [(P(\omega)X_0(\omega), X_0(\omega)) + \sum_{t\in Z_+}(QX_t(\omega), X_t(\omega))-$$

$$-(P(\omega)X_t(\omega), X_t(\omega)) + (RU_t(\omega), U_t(\omega)) + (P(\omega)A(\omega)X_t(\omega), A(\omega)X_t(\omega))+$$

$$+(P(\omega)BU_t(\omega), BU_t(\omega)) + (P(\omega)A(\omega)X_t(\omega), A(\omega)X_t(\omega), A(\omega)X_t(\omega))]d\omega.$$

Adding and subtracting in the obtained expression the following term

$$(P(\omega)A(\omega)X_t(\omega), B[R + B^{'} P(\omega)B]^{-1}B^{'} P(\omega)A(\omega)X_t(\omega)), t \in Z_+,$$

we find that

$$J(u) = \frac{1}{2\pi}\int_0^{2\pi} [(P(\omega)X_0(\omega), X_0(\omega)) + \sum_{t\in Z_+}[(F(\omega)X_t(\omega), X_t(\omega))+$$

$$+((R+B^{'}P(\omega)B)V_t(\omega),V_t(\omega))]d\omega.$$

Here

$$F(\omega)=Q-P(\omega)+A^*(\omega)P(\omega)A(\omega)-$$
$$-A^*(\omega)P(\omega)B[R+B^{'}P(\omega)B]^{-1}B^{'}P(\omega)A(\omega),$$
$$V_t(\omega)=U_t(\omega)+[R+B^{'}P(\omega)B]^{-1}B'P(\omega)A(\omega)X_t(\omega).$$

Note that the needed operators are invertible since $P(\omega)$ is nonnegative and R is positive operators. The second term in J is not depend on control function since $X_0(\omega)=\sum\limits_{s\in\mathbb{Z}}\varphi(s)e^{-is\omega}$, $\omega\in[0,2\pi]$. Choose now the required operators $P(\omega)$ such that the following condition $F(\omega)=0$ holds. Then the cost functional can be rewritten as

$$J(u)=\frac{1}{2\pi}\int_0^{2\pi}[(P(\omega)X_0(\omega),X_0(\omega))+\sum_{t=0}^{+\infty}([R+B'P(\omega)B]^{-1}V_t(\omega),V_t(\omega))d\omega. \quad (2.59)$$

It is obviously that the minimal value for (2.59) is

$$J(u^0)=\frac{1}{2\pi}\int_0^{2\pi}[(P(\omega)X_0(\omega),X_0(\omega))d\omega,$$

which is feasible iff $V_t(\omega)=0$. In other words, this is possible iff $U_t(\omega)=K(\omega)X_t(\omega)$. Thus the required representation for the optimal control law and the function $K(\omega)$ and $P(\omega)$ are obtained. The theorem is proved.

3

Differential Linear Repetitive Processes

Differential repetitive processes are a distinct class of continuous-discrete 2D linear systems of both systems theoretic and applications interest. The feature which makes them distinct from other classes of such systems is the fact that information propagation in one of the two independent directions only occurs over a finite interval. Applications areas include iterative learning control and iterative solution algorithms for classes of dynamic nonlinear optimal control problems based on the maximum principle, and the modeling of numerous industrial processes such as metal rolling, long-wall cutting etc. Repetitive processes share also similar mathematical model with recently extensive developing and important from the practical point of view, so-called spatially interconnected systems, and hence these are the possible application area for them. In particular, some models of the distributed gas networks can be treated in the terms of repetitive processes, also.

The first part of this chapter uses the classic approach to investigate the traditional optimal control theory problems for the repetitive dynamics model. It is well known that the separation theorem for convex sets is quite useful approach for studying a wide class of extreme problems. Here we develop this method to establish optimality conditions in the classic form of maximum principle for multipass nonstationary continuous-discrete control system with nonlinear inputs and nonlocal state-phase terminal constraints of general form. The obtained results are traditional for classic optimal control theory. However, their numerical realization is not a trivial task. By this reason in the next sections for the stationary case of the system model and particular case of the constraint and the cost functional we develop the new optimality and sub-optimality conditions that are more suitable for the design of numerical methods and further applications. In contrast to the classic approaches of optimal control theory,in the second part in this chapter we use the idea of constructive methods reported in [37] and extend this setting to the continuous-discrete case to produce new results and constructive elements of optimization theory for the considered repetitive systems and develop also its relevant basic properties which can be of interest for others purposes, too. It is shown that the obtained optimality and ϵ-optimality conditions are close related to the corresponding classic results of maximum principle and ϵ- maximum principle. The sensitivity analysis and some differential properties of the optimal controls under disturbances are discussed and their application to the optimal synthesis problem is given. The obtained results yield a theoretical background for the design problem of optimal controllers for relevant basic processes. The end goal of the research programme for which this research forms part of the output is the development of numerically reliable algorithms for the synthesis of optimization based control schemes for these processes. Some areas for short to medium term further research are also briefly discussed.

3.1 Background and preliminaries

The essential unique characteristic of a repetitive (termed multipass in the early literature) process can be illustrated by considering machining operations where the material or workpiece involved is processed by a sequence of sweeps, termed passes, of the processing tool. Assuming that the pass length t^* (i.e. the duration of a pass of the processing tool), which is finite by definition, has a constant value for each pass. Then in a repetitive process the output vector, or pass profile, $y_k(t)$, $t \in [0, t^*]$, (t being the independent spatial or temporal variable) produced on pass k acts as a forcing function on the next pass and hence contributes to the dynamics of the new pass profile $y_{k+1}(t)$, $t \in [0, t^*]$, $k \geq 0$.

The dynamics of such processes in the time and space invariant case(see [66]) ares defined over $0 \leq t \leq t^*$, $k \geq 0$, by a state space model of the form

$$\begin{aligned} \dot{x}_{k+1}(t) =& \hat{A}x_{k+1}(t) + \hat{B}u_{k+1}(t) + \hat{B}_0 y_k(t) \\ y_{k+1}(t) =& \hat{C}x_{k+1}(t) + \hat{D}u_{k+1}(t) + \hat{D}_0 y_k(t) \end{aligned} \tag{3.1}$$

To complete the process description, it is necessary to specify the boundary conditions, i. e. the state initial vector on each pass and the initial pass profile. Here no loss of generality arises from assuming these to be of the form $x_{k+1}(0) = d_{k+1}$, $k \geq 0$, and $y_0(t) = f(t)$, where d_{k+1} is an $n \times 1$ vector of known constant entries and $f(t)$ is an $m \times 1$ vector whose entries are known functions of t over $0 \leq t \leq t^*$. Industrial examples (see, for example, [7]) include long-wall coal cutting and metal rolling operations.

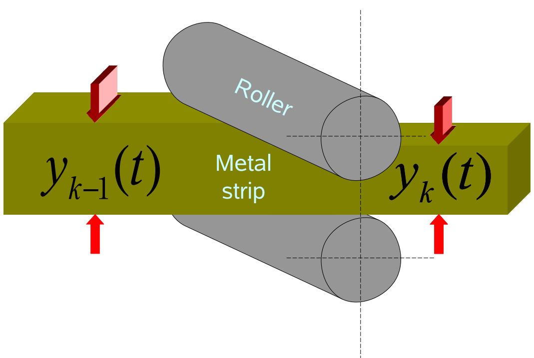

Figure 3.1: Metal rolling

The simulation of these processes in the simplest cases lead to the mathematical models of the following form (see, for example, the monographs by N. Bose, E.Rogers, etc)

$$\frac{d^2 y_k(t)}{dt^2} + \lambda_1 y_k(t) = \lambda_2 \frac{d^2 y_{k-1}(t)}{dt^2} + \lambda_1 y_{k-1}(t) + b u_k(t), \; t \in [0, t^*], \; k \in K = \{1, ..., N\}, \tag{3.2}$$

where $y_k(t)$ and $y_{k-1}(t)$ denote the gauge on the current and previous passes through the rolls; λ_1, λ_2 and b are determined by the stiffness of the metal strip and the roll mechanism properties, $u(t)$ can be interpreted as the applied force to the metal strip by the rolls.

A model of the rectification process of a many component mixture in a many-plate column can be represented by

$$\frac{dx_s(t)}{dt} = V_{s-1}(t)x_{s-1}(t) + V_s(t)x_s(t) - R_s(x_s(t), y_s(t)) + u_{x_s}(t), \tag{3.3}$$

$$\frac{dy_s(t)}{dt} = L_{s+1}(t)y_{s+1}(t) + L_s(t)y_s(t) + R_s(x_s(t), y_s(t)) + u_{y_s}(t), \tag{3.4}$$

$$t \in [0, t^*], \quad s \in K \doteq \{1, ..., N\}. \tag{3.5}$$

Here $x(s,t)$, $y(s,t)$ denote the desired material concentration on s-th plate in the gas and liquid fractions, respectively; L, V and R present the hydrodynamical characteristic of the process under consideration; u_x and u_y are the control material row; K is subset of integers. Details of the model can be found in [18].

Also problem areas exist where adopting a repetitive process setting for analysis has clear advantages over alternatives. This is especially true for classes of iterative learning control schemes (see, for example, [2]) and of iterative solution algorithms for classes of dynamic nonlinear optimal control problems based on the maximum principle (see, for example, [68]). Repetitive processes share also similar mathematical model with recently extensive developing and important from the practical point of view, so-called spatially interconnected systems, see e.g. [17], and hence these are the possible application area for them. In particular, a wide class of the distributed gas networks can be treated in the terms of repetitive processes.

The basic unique control problem for repetitive processes is that the output sequence of pass profiles generated can contain oscillations that increase in amplitude in the pass to pass direction (i.e. in the k direction in the notation for variables used here). Early approaches to stability analysis and controller design for (linear single-input single-output) repetitive processes and, in particular, long-wall coal cutting were based on first converting the process dynamics into those of an equivalent infinite-length single-pass process [32]. This, for example, resulted in a scalar differential/algebraic system to which standard scalar inverse-Nyquist stability criteria were then applied. In general, however, it was soon established that this approach to analysis (and controller design) would, except in a few very restrictive special cases, lead to incorrect conclusions [65]. The basic reason for this is that such an approach effectively neglects their finite pass length repeatable nature together and also the effects of resetting the initial conditions before the start of each new pass. This, in turn, led to the development of a rigorous stability theory based on an abstract model of the process dynamics in a Banach space setting which can be applied to all examples with linear dynamics and a constant pass length [71].

Given a suitable stability theory, it is a natural progression to consider the structure of control schemes for these processes and the development of suitable controller design/synthesis tools. In this latter respect, one obvious way to proceed is to minimize a suitably defined cost function.

In the first part of the paper, we consider the more general, non-stationary case. Note, that the repetitive processes are inherently two-dimensional, and hence non-stationarity can reflect to the from pass to pass parameter variability, and to the along the pass parameter variability. Here we study the particular subclass of repetitive processes, where the pass profile vector is exactly the same as the state one [8]. Hence, the second equation of (3.1) is neglected and both dynamics, i.e. along the pass and from pass to pass are encountered by one equation, similarly as in the Fornasini Marchesini model. Also, the input influence to the state dynamics is nonlinear however additive.

It is well known that the separation theorem [39, 9, 55] for convex sets is an useful method for studying a wide area of optimization problems and hence we apply this method to establish optimality conditions in the form of the maximum principle for a differential non-stationary repetitive process with linear state dynamics and an additive nonlinear term to account for the input to the process and non-local state-space terminal constraints of a general form. The results obtained here in that wayare of the significant theoretical value, however are not very well adopted for computation (and hence for application to numerical examples). The on-going work are devoted to achieve more numerically tractable results.

In the remainder of the paper, we focus on the stationary in both the directions case. We develop the new optimality and sub-optimality conditions that are more suitable from the numerical calculation point of view. These conditions extend the constructive methods of [37] to repetitive processes and it is shown that the resulting optimality and ϵ-optimality conditions are closely related to the corresponding standard, one-dimensional results in the form of the maximum and ϵ- maximum principles. Also, the sensitivity analysis of the resulting optimal controls is undertaken and some relevant properties are established. To illustrate the use of these new results, a numerical example is detailed.

3.2 Notation and Model Definition

Assume that $T = [0, t^*]$ is a given interval of values of the continuous independent variable $t \in T$ ruling the along the pass dynamics and $K = \{1, 2, ..., N\}$, $N < +\infty$ be a set of values of the discrete variable $k \in K$ ruling the dynamics from pass to pass direction. This last assumption can be made as in practice, a repetitive process will only ever complete a finite number of passes. Also introduce the control and state vectors as $u_k(t) \in \mathbb{R}^r$ and $x_k(t) \in \mathbb{R}^n$ respectively. Then the non-stationary repetitive processes can be described as

$$\frac{dx_k(t)}{dt} = A(t)x_k(t) + D(t)x_{k-1}(t) + b_k(u_k(t), t), \quad k \in K, \ t \in T \tag{3.1}$$

with the boundary conditions of the form

$$x_k(0) = \alpha(k), k \in K, \quad x_0(t) = \beta(t), \quad t \in T \tag{3.2}$$

where the $n \times n$ matrix functions $A(t)$ and $D(t)$ and the $n \times 1$ function $\beta(t)$ are measurable and integrable on T, the function $b : K \times U \times T \to \mathbb{R}^n$ is continuous with respect to $(u, t) \in U \times T$ for each fixed $k \in K$, $\alpha(k)$ is an $n \times 1$ vector of known constant entries. Note that the last nonlinear term represents the additive but non-linear input signal influence to the process dynamics. What is interesting this influence is pass number variable, hence non-stationary from pass to pass. Also, the model matrices are t-dependent and hence the process is non-stationary along the pass.

This model can be easily extended to the one when the pass profile and state vectors are decoupled and the pass profile dynamics can be a vector valued function of the state dynamics.

Now, it is to define the class of available and admissible input signals for the above model.

Definition 9. *We say that the function $u : K \times T \to \mathbb{R}^r$ is available for (3.1) if it is measurable with respect to t for each fixed $k \in K$, and satisfies the constraint $u_k(t) \in U$,*

$k \in K$, for almost all $t \in T$, where U is a given compact set from $\mathbb{R}^r$. Also the function $x : K \times T \to \mathbb{R}^n$ is a solution of (3.1) corresponding to the given available control $u_k(t)$ if it is absolutely continuous with respect to $t \in T$ for each fixed $k \in K$ and satisfies (3.1) for almost all $t \in T$ and each $k \in K$.

We denote the set of available controls by $U(\cdot)$ and use $M_i,\ M_i \subset \mathbb{R}^n,\ i = 1, 2, \ldots, l$ to denote given compact convex sets.

Definition 10. *The available control $u_k(t)$ is said to be admissible for the process (3.1) if the corresponding solution $x_k(t) = x_k(t, \alpha, \beta, u)$ of (3.1) and (3.2) satisfies*

$$x_N(\tau_i) \in M_i,\ i = 1, 2, ..., l \tag{3.3}$$

where $0 = \tau_0 < \tau_1 < \tau_2 < ... < \tau_l = t^*$ are specified elements of T.

The optimal control problem considered in this paper can now be stated as:
Minimize a cost function of the form

$$J(u) = \varphi(x_N(\tau_1), x_N(\tau_2), ..., x_N(\tau_l)) \tag{3.4}$$

for processes described by (3.1) and (3.2) in the class of admissible controls $u_k(t) \in U(\cdot)$ where the function $\varphi : \mathbb{R}^{nl} \to \mathbb{R}$ is assumed to be convex. It is easy to see that these conditions guarantee the existence and uniqueness of an absolutely continuous solution of (3.1) and(3.2) for any available control $u_k(t)$. To guarantee the existence of an optimal control, throughout this paper we assume that the set of admissible controls is non-empty.

3.2.1 Reachability set and its properties

To solve (3.1) and (3.2) we define the $n \times n$ matrix function $\Phi(\tau, t)$ which solves the following differential equation

$$\frac{d\Phi(\tau, t)}{d\tau} = A(\tau)\Phi(\tau, t), \quad \Phi(t, t) = I_n \tag{3.5}$$

where I_n denotes the $n \times n$ identity matrix. Also it well known, see, for example, [46] that the entries in the matrix $\Phi(\tau, t)$ are absolutely continuous functions defined on the set $T \times T$. Therefore, there exists a constant $0 < C < \infty$ such that $\|\Phi(\tau, t)\| \leq C$ for any $(\tau, t) \in T \times T$, where $\| \cdot \|$ denotes any matrix norm. Further, we use $H^p(0, t^*)$, where $p > 0$ is an integer, to denote the set of all functions $f : (0, t^*) \to \mathbb{R}^n$, which are absolutely continuous on each closed sub-interval $[\alpha, \beta]$ from the interval $(0, t_1^*)$ and have almost everywhere integrable derivatives of order up to p on $(0, t^*)$. Also it can be shown that $H^p(0, t^*)$ is a Banach space with the norm $\|f\|_H = \sum_{i=0}^{p} \|f^{(i)}\|_{L_1}$ and the following inclusions $H^p(0, t^*) \subset C^p(0, t^*) \subset L_1(0, t^*)$ hold, where $C^p(0, t^*)$ denotes the space of $n \times 1$ vector functions which are continuously differentiable on $(0, t^*)$ up to order p, and $L_1(0, t^*)$ the space of $n \times 1$ vector valued functions which are integrable on $(0, t^*)$.

Now define the mapping $\mathcal{P} : L_1(0, t^*) \to H^1(0, t^*)$ as

$$(\mathcal{P}f)(\tau) = \int_0^\tau \Phi(\tau, t)D(t)f(t)dt, \quad \tau \in (0, t^*). \tag{3.6}$$

and its power composition $\mathcal{P}^k : H^{k-1}(0,t^*) \to H^k(0,t^*)$ as $(\mathcal{P}^k f)(\tau) = \mathcal{P}(\mathcal{P}^{k-1} f)(\tau)$, $\tau \in (0,t^*)$. Also define the mapping $Q : L_1(0,t^*) \to H^1(0,t^*)$ by

$$(Qf)(\tau) = \int_0^\tau \Phi(\tau,t) f(t)dt, \quad \tau \in (0,t^*). \tag{3.7}$$

For the given available control $u \in U(\cdot)$ the corresponding solution of (3.1) and (3.2) at $t = \tau_j$ on pass $k = N$ can now be written in the form

$$x_N(\tau_j) = \Phi(\tau_j,0)\alpha(N) + \sum_{i=1}^{N-1} (\mathcal{P}^i \Phi(\cdot,0))(\tau_j)\alpha(N-i) + (\mathcal{P}^N \beta)(\tau_j) +$$
$$\sum_{i=1}^{N-1} (\mathcal{P}^i Q b_{N-i}(u_{N-i},\cdot))(\tau_j) + \int_0^{\tau_j} \Phi(\tau_j,t) b_N(u_N(t),t)dt, \ N > 1, \ j = 1,2,\ldots,l, \tag{3.8}$$

where $\Phi(\cdot,\tau)$ denotes the projection of the function $\Phi(t,\tau)$ with the second variable fixed to some $\tau \in T$.

Next, introduce $c = (c_1, c_2, \ldots, c_l)^T \in \mathbb{R}^{nl}$, where

$$c_j = \Phi(\tau_j,0)\alpha(N) + \sum_{i=1}^{N-1} (\mathcal{P}^i \Phi(\cdot,0))(\tau_j)\alpha(N-i) + (\mathcal{P}^N \beta)(\tau_j), \ j = 1,2,\ldots,l. \tag{3.9}$$

and the mapping $S : U(\cdot) \to \mathbb{R}^{nl}$ as $Su = (S_1 u, S_2 u, \ldots, S_l u)^T$ where

$$S_j u = \sum_{i=1}^{N-1} (\mathcal{P}^i Q b_{N-i}(u_{N-i},\cdot))(\tau_j) + \int_0^{\tau_j} \Phi(\tau_j,t) b_N(u_N(t),t)dt, \ j = 1,2,\ldots,l. \tag{3.10}$$

Then we can state the following basic problem whose solution is to be used for solving the optimal control problem.

Problem A

Find necessary and sufficient conditions for

$$z = c + Su \tag{3.11}$$

to hold, subject to

$$z \in M, \ \varphi(z) \le \delta, \ z \in \mathbb{R}^{nl}, \ u \in U(\cdot) \tag{3.12}$$

where $M = M_1 \times M_2 \times \ldots \times M_l \subset \mathbb{R}^{nl}$, and δ is a fixed number from $\mathbb{R}$

To solve Problem (A), introduce first the following sets

$$\mathcal{R} = \{z \in \mathbb{R}^{nl}, \ z = c + Su, \ u \in U(\cdot)\}, \qquad K(\delta) = \{z \in \mathbb{R}^{nl}, \ z \in M, \ \varphi(z) \le \delta\}. \tag{3.13}$$

Then it is easy to see that the necessary and sufficient condition for Problem (A), to have a solution is $\mathcal{R} \cap K(\delta) \neq \emptyset$. In what follows, we establish the analytical form of this geometric criteria which is based on the separation theorem for convex sets.

Consider first the problem of obtaining the required properties of the sets $\mathcal{R}$ and $K(\delta)$. The main technical difficulties here are related to the convexity and closeness of the set $\mathcal{R}$

which must be established in order to apply the separation theorem. To overcome them we extend known results for $1D$ systems (see, for example, [9]) to the repetitive process case.

First define the set

$$Z = \left\{ z = (z_1, \ldots, z_l) \in \mathbb{R}^{nl} : z_j = \int_0^{\tau_j} f(v(t), t)dt, \quad v \in V(\cdot), \quad j = 1, 2, \ldots, l \right\}, \tag{3.14}$$

where τ_j are given points such that $0 < \tau_1 < \tau_2 < \ldots < \tau_l = t^*$, and $V(\cdot)$ is the set of all measurable functions $v : T \to \mathbb{R}^r$ such that $v(t) \in U$ for almost all $t \in T$, and the function $f : U \times T \to \mathbb{R}^n$ is continuous. Then the response formulas (3.8) and (3.10) show that the required properties of the set $\mathcal{R}$ can be established by studying analogous properties for the set Z.

Now we have the following results to be further used for solving the main problem.

Lemma 1. *Let $g : U \to R^n$ be a continuous function. Then*

$$Y = \left\{ z \in R^{nk} : z_j = \int_0^{\tau_j} g(v(t))dt, \quad v \in V(\cdot), \quad j = 1, 2, ..., k \right\} \tag{3.15}$$

is a convex set.

Proof. The proof is based on the constructing the corresponding combination function by pressing of the given functions along the interval T. Indeed, if z^1 and z^2 are some points of the set Y with corresponding functions $v_1(t)$ and $v_2(t)$ from $V(\cdot)$, then the desired function $v_\alpha \in V(\cdot)$ corresponding to the point $z^\alpha = \alpha_1 z_1^1 + \alpha_2 z^2$, $\alpha_1 + \alpha_2 = 1$, $\alpha_1 \geq 0$, $\alpha_2 \geq 0$, is constructed as follows :

$$v_\alpha(t) = \begin{cases} v_1(t/\alpha_1), & 0 \leq t < \alpha_1\tau_1, \\ v_2((t - \alpha_1\tau_1)/\alpha_2), & \alpha_1\tau_1 \leq t < \tau_1, \\ \ldots & \ldots \\ v_1((t - \alpha_2\tau_{k-1})/\alpha_1), & \tau_{k-1} \leq t < \tau_{k-1} + \alpha_1(\tau_k - \tau_{k-1}), \\ v_2((t - \alpha_1\tau_k)/\alpha_2), & \tau_{k-1} + \alpha_1(\tau_k - \tau_{k-1}) \leq t \leq \tau_k. \end{cases}$$

Then

$$\int_0^{\tau_i} g(v_\alpha(t))dt = \int_0^{\tau_1} g(v_\alpha(t))dt + ... + \int_{\tau_{i-1}}^{\tau_i} g(v_\alpha(t))dt =$$

$$= \int_0^{\alpha_1\tau_1} g(v_1(\frac{t}{\alpha_1}))dt + \int_{\alpha_1\tau_1}^{\tau_1} g(v_2(\frac{t - \alpha_1\tau_1}{\alpha_2}))dt + ...$$

$$+ \int_{\tau_{i-1}}^{\tau_{i-1}+\alpha_1(\tau_i-\tau_{i-1})} g(v_1(\frac{t - \alpha_2\tau_{i-1}}{\alpha_1}))dt + \int_{\tau_{i-1}+\alpha_1(\tau_i-\tau_{i-1})}^{\tau_i} g(v_2(\frac{t - \alpha_1\tau_i}{\alpha_2}))dt =$$

$$= \alpha_1 \int_0^{\tau_1} g(v_1(t))dt + \alpha_2 \int_0^{\tau_1} g(v_2(t))dt + ... + \alpha_1 \int_{\tau i-1}^{\tau_i} g(v_1(t))dt + \alpha_2 \int_{\tau i-1}^{\tau_i} g(v_2(t))dt =$$

$$= \alpha_1 \int_0^{\tau_i} g(v_1(t))dt + \alpha_2 \int_0^{\tau_i} g(v_2(t))dt = \alpha_1 z_i^1 + \alpha_2 z_i^2 = z_i^\alpha, \; i = 1, ..., k.$$

Thus for each z^α there exists an available function $v^\alpha(\cdot) \in U(\cdot)$ generating this vector. The proof is complete.

Lemma 2. *Let $f : U \times T \to \mathbb{R}^n$ be a continuous function. Then for any measurable function $v(\cdot) \in V(\cdot)$ and for a given number $\varepsilon > 0$ $\exists$ a partition of the interval T by points $0 = s_0 < s_1 < \ldots < s_m = t^*$ such that*

$$\sum_{j=0}^{m-1} \int_{s_j}^{s_{j+1}} \|f(v(t), \tau_j) - f(v(t), s_j)\| dt < \varepsilon \tag{3.16}$$

holds for any τ_j satisfying $s_j \le \tau_j \le s_{j+1}$, $j = 0, \ldots, m$.

Proof. This is based on the so-called C-property of measurable functions [61]. Since the function $f(u,t)$ is continuous on the compact set $U \times [t_0, t_1]$ then there exists a number $M > 0$ such that $\|f(u,t)\| \le M$ for each (u,t). Now the c-property of measurable functions yields that $\forall\ \varepsilon > 0$ *exists* a continuous function $u^*(t)$ defined on $[t_0, t_1]$ such that

$$mes\ (E = \{t \in [t_0, t_1] : u(t) \neq u^*(t)\}) \le \varepsilon / 6M.$$

Since $f(u,t)$ is continuous on $U \times [t_0, t_1]$ then the function $f(u^*(t), \tau)$ is continuous also on the set $[t_0, t_1] \times [t_0, t_1]$ and, hence, it is uniformly continuous. It means that there exists a partion of the interval $[t_0, t_1]$ by points

$$t_0 = s_0 < s_1 < ... < s_{m-1} < s_m = t_1, \tag{3.17}$$

such that

$$|f(u^*(t), \tau_j) - f(u^*(t), s_j| \le \varepsilon / 3(t_1 - t_0) \quad \text{for} \quad s_j \le \tau_j \le s_{j+1},$$

and each j, $1 \le j \le m-1$. Hence

$$\sum_{j=0}^{m-1} \int_{s_j}^{s_{j+1}} |f(u^*(t), \tau_j) - f(u^*(t), s_j| \le \frac{\varepsilon}{3(t_1 - t_0)} \sum_{j=0}^{m-1} (s_{j+1} - s_j) = \frac{\varepsilon}{3}. \tag{3.18}$$

Put $E_i = E \bigcap [s_j - s_{j+1}]$. The sum in (3.18) we separate into two parts $\sum_1$ and $\sum_2$. Namely, let the set $\sum_1$ includes the terms containing the integrals around the intervals which have not a joint points with the set E, and the set $\sum_2$ includes the remaining terms. Then

$$\sum_{j=0}^{m-1} \int_{s_j}^{s_{j+1}} |f(u(t), \tau) - f(u(t), s_j| dt = \sum_1 \int_{s_j}^{s_{j+1}} |f(u(t), \tau) -$$

$$- f(u(t), s_j| dt + \sum_2 \int_{[s_j, s_{j+1} \setminus E_j} |f(u(t), \tau) - f(u(t), s_j| dt +$$

$$+ \sum_2 \int_{E_j} |f(u(t), \tau) - f(u(t), s_j| dt \le \varepsilon/3 + \varepsilon/3 + mes\ E \cdot 2M = \varepsilon.$$

The Lemma is proved.

Lemma 3. *Let $f : U \times T \to \mathbb{R}^n$ be a continuous function. Then the closure $\overline{Z}$ of the set Z of (3.14) is convex.*

Proof. Let us assume that it is not true. Then there exists the points $z^1,\, z^2 \in \bar{Z}$ and numbers $\lambda_1,\, \lambda_2 > 0,\, \lambda_1 + \lambda_2 = 1$ such that $z^\lambda - \lambda_1 z^1 + \lambda_2 z^2 \notin \bar{Z}$. Hence, there exists a number $\varepsilon > 0$ such that

$$\|z - z^\lambda\| \geq 0 \;\forall\; z \in \bar{Z}. \tag{3.19}$$

(To clarify the definiteness we assume that $\|\cdot\|$ is the Euclid norm). Since $z^i \in \bar{Z},\, i = 1,2$ then there exist the sequences $\left\{z^i(n)\right\} = \left\{(z_1^i(n), ..., z_k^i(n)\right\},\, i = 1,2;\, n = 1,2,...,$ from Z which are convergent to some points $z^i,\, i = 1,2$. Now we fix the number n for which

$$\|z^i - z^i(n)\| \;<\; \frac{\varepsilon}{(1+2\sqrt{k})}, \quad i = \overline{1,2} \tag{3.20}$$

is fulfilled. Since $z^i(n) \in Z$ then there are an available controls $u_i(\cdot) \in U(\cdot),\, i = 1,2$ such that

$$z_j^i(n) = \int\limits_{t_0}^{\tau_j} f(u_i(t), t)dt, \quad i = 1,2, \quad j = \overline{1,k}. \tag{3.21}$$

Now Lemma 2 yields that for the controls $u_i(\cdot),\; i = 1,2$ for the given number $\alpha = \varepsilon/(1+2\sqrt{k})$ there are the partitions of intervals $[t_0, \tau_1], ..., [\tau_{k-1}, \tau_k]$ such that the estimation (3.16) is valid for the functions $u_i(\cdot),\, i = 1,2$ on the each interval $[t_0, \tau_1], ..., [\tau_{k-1}, \tau_k]$, where ε is replaced by $\varepsilon/(1+2\sqrt{k})$. We suppose that the chain

$$t_0 = s_0 < s_1 < ... < s_{m_1} = \tau_1 < s_{m_1+1} < ... < s_{m_j} = \tau_j < s_{m_{j+1}} < ... < s_{m_k} = \tau_k = t_1 \tag{3.22}$$

includes the all points of partitions of intervals $[t_0, \tau_1], ..., [\tau_{k-1}, \tau_k]$ constructed for the control functions $u_1(\cdot)$ and $u_2(\cdot)$ for the given number α. Since adding new points is not influence to the estimate (3.16) obtained in the Lemma 2 then

$$\sum_{j=0}^{m_l-1} \int\limits_{s_j}^{s_{j+1}} \|f(u_i(t), \tau) - f(u_i(t), s_j\|dt < \frac{\varepsilon}{1+2\sqrt{k}}, \quad i = 1,2, \quad l = \overline{1,k}, \quad s_j \leq \tau \leq s_{j+1} \tag{3.23}$$

Put $\tilde{z}^i = \tilde{z}_1^i, ..., z_k^i) \in R^{nk},\, i = 1,2$ where

$$\tilde{z}_l^i = \sum_{j=0}^{m_l-1} \int\limits_{s_j}^{s_{j+1}} f(u_i(t), s_j)dt, \quad i = 1,2, \quad l = \overline{1,k}. \tag{3.24}$$

Then (3.21),(3.23) yield

$$\|z^i(n) - \tilde{z}^i\|_{R^{nk}} = \left(\sum_{l=1}^{k} \|z_l^i(n) - \tilde{z}_l^i\|_{R^n}^2\right)^{1/2} = \left(\sum_{l=1}^{k} \| \sum_{j=0}^{m_l-1} \int\limits_{s_j}^{s_{j+1}} [f(u_i(t), t) -\right.$$

$$-f(u_i(t),s_j)]dt\|^2\Bigg)^{1/2} \leq \left(\sum_{l=1}^{k}\left(\frac{\varepsilon}{1+2\sqrt{k}}\right)^2\right)^{1/2} = \frac{\varepsilon\sqrt{k}}{1+2\sqrt{k}}, \quad i=1,2.$$

Now we deternine the available control defined on the interval $[t_0, t_1]$ as follows

$$u_\lambda(t) = u_\lambda^{(j)}(t), \quad t \in [s_j, s_{j+1}), \quad j = \overline{0, m_k - 1},$$

where

$$u_\lambda^{(j)}(t) = \begin{cases} u_1\left(\frac{t-\lambda_2 s_j}{\lambda_1}\right), & s_j \leq t < s_j + \lambda_1(s_{j+1} - s_j), \\ u_2\left(\frac{t-\lambda_1 s_{j+1}}{\lambda_2}\right), & s_j + \lambda_1(s_{j+1} - s_j) \leq t < s_{j+1}. \end{cases} \tag{3.25}$$

It easy to see that

$$\int_{s_j}^{s_{j+1}} f(u_\lambda^{(j)}(t), s_j)dt = \lambda_1 \int_{s_j}^{s_{j+1}} f(u_1(t), s_j)dt + \lambda_2 \int_{s_j}^{s_{j+1}} f(u_2(t), s_j)dt, \quad j = \overline{0, m_{k-1}}.$$

Summing these inequalities with respect to j and using (3.24) yield

$$\sum_{j=0}^{m_l-1} \int_{s_j}^{s_{j+1}} f(u_\lambda^{(j)}(t), s_j)dt = \lambda_1 \tilde{z}_l^1 + \lambda_2 \tilde{z}_l^2, \quad l = \overline{1, k}. \tag{3.26}$$

Set $\tilde{z}^\lambda = \lambda_1 \tilde{z}^1 + \lambda_2 \tilde{z}^2$ and determine the vector $\hat{z}^\lambda = (\hat{z}_1^\lambda, ..., \hat{z}_k^\lambda)$ where

$$\hat{z}_j^\lambda = \int_{t_0}^{\tau_j} f(u_\lambda(t), t)dt, \quad j = \overline{1, k}. \tag{3.27}$$

It is obvious that $\hat{z}^\lambda \in Z$. Now we evaluate the spacing between z^λ and $\hat{z}^\lambda$:

$$\|z^\lambda - \hat{z}^\lambda\| \leq \|z^\lambda - \hat{z}^\lambda\| + \|\tilde{z}^\lambda - \hat{z}^\lambda\| = \|\lambda_1(z^1 - \tilde{z}^1) + \lambda_2(z^2 - \tilde{z}^2)\| +$$

$$+\|\hat{z}^\lambda - \hat{z}^\lambda\| \leq \lambda_1\|z^1 - z^1(n)\| + \lambda_1\|z^1(n) - \tilde{z}^1\| + \lambda_2\|z^2 - z^2(n)\| +$$

$$+\lambda_2\|z^2(n) - \tilde{z}^1\| + \|\tilde{z}^\lambda - \hat{z}^\lambda\| < \lambda_1\frac{\varepsilon}{1+2\sqrt{k}} + \lambda_1\frac{\varepsilon\sqrt{k}}{1+2\sqrt{k}} +$$

$$+\lambda_2\frac{\varepsilon}{1+2\sqrt{k}} + \lambda_2\frac{\varepsilon\sqrt{k}}{1+2\sqrt{k}} + \|\tilde{z}^\lambda - \hat{z}^\lambda\| = \frac{\varepsilon(1+\sqrt{k})}{1+2\sqrt{k}} + \|\tilde{z}^\lambda - \hat{z}^\lambda\|.$$

Then we should estimate the value of $\|\tilde{z}^\lambda - \hat{z}^\lambda\|$. From the above relations we have

$$\|\tilde{z}^\lambda - \hat{z}^\lambda\| = \left(\sum_{l=1}^{k}\|\tilde{z}_l^\lambda - \hat{z}_l^\lambda\|^2\right)^{1/2} = \Bigg(\sum_{l=1}^{k}\|\lambda_1\sum_{j=0}^{m_l-1}\int_{s_j}^{s_{j+1}} f(u_1(t), s_j)dt +$$

$$+\lambda_2\sum_{j=0}^{m_l-1}\int_{s_j}^{s_{j+1}} f(u_2(t), s_j)dt - \sum_{j=0}^{m_l-1}\int_{s_j}^{s_j+\lambda_1(s_{j+1}-s_j} f(u_\lambda^{(j)}(t), t)dt -$$

$$-\sum_{j=0}^{m_l-1}\int\limits_{s_j+\lambda_1(s_{j+1}-s_j)}^{s_{j+1}} f(u_\lambda^{(j)}(t),t)dt\|^2\Big)^{1/2} \leq \Big(\sum_{l=1}^{k}\Big[\lambda_1\sum_{j=0}^{m_l-1}\int\limits_{s_j}^{s_{j+1}}\|f(u_1(t),s_j)-$$

$$-f(u_1(t),\lambda_1 t+\lambda_2 s_j)\|dt+\lambda_2\sum_{j=0}^{m_l-1}\int\limits_{s_j}^{s_{j+1}}\|f(u_2(t),s_j)-$$

$$-f(u_2(t),\lambda_1 s_{j+1}+\lambda_2 t)\|dt\Big]^2\Big)^{1/2} \leq \frac{\varepsilon\sqrt{k}}{1+2\sqrt{k}}.$$

Thus

$$\|z^\lambda-\hat{z}^\lambda\| < \frac{\varepsilon(1+\sqrt{k})}{1+2\sqrt{k}}+\frac{\varepsilon\sqrt{k}}{1+2\sqrt{k}}=\varepsilon.$$

This inequality contradicts to the initial preposition (3.19). The lemma is proved.

Remark 14. *Convexity of $\overline{Z}$ is guaranteed by the presence of the integral terms in Z. This fact, known as hidden convexity, is an important property of continuous time control systems which follows, in general, from the Lyapunov theorem on the convexity of the range of an integral operator acting on vector measures. This result is often used, see, for example, [9, 55], to prove the convexity of the reachability set for control systems which are linear in the state variables.*

Formulas (3.8), (3.10) state that each integral expression in $\mathcal{R}$ contains an available control $u_s(t)$ with a fixed single value of the discrete variable s and, therefore, is independent of the others. Hence, to prove that $\mathcal{R}$ is a closed set it is sufficient to show that a set formed by controls with some fixed value of the discrete variable k, $k=1,\ldots,N$ is closed. The simplest case is often to consider $k=N$ and then the set to be studied has the following form

$$\mathcal{R}_N=\big\{z\in\mathbb{R}^{nl}: z_j=a_j+L_jv,\quad v(\cdot)\in V(\cdot),\quad j=1,2,\ldots,l\big\}. \tag{3.28}$$

Here $a_j=\Phi(\tau_j,0)\alpha(N)$, and the mappings L_j defined on the set $V(\cdot)$ are given by

$$L_jv=\int\limits_0^{\tau_j}\Phi(\tau_j,t)g(v(t),t))dt$$

where here $g(v(t),t)$ denotes the function $b_N(v_N(t),t)$, $t\in T$.

Lemma 4. *The set $\mathcal{R}_N$ defined by (3.28) is closed.*

Proof. Suppose that the vector sequence $\{z^n\}=\{(z_1^n,\ldots,z_l^n)^T\}\in\mathcal{R}_N$ converges to a point $z^*=(z_1^*,\ldots,z_l^*)^T\in\mathbb{R}^{nl}$. Then there exists a sequence $\{v^n(\cdot)\}$ of functions from $V(\cdot)$ such that $z_j^n=a_j+L_jv^n$, $j=1,\ldots,l$ and we show that there exists a function $v^*(t)$, $t\in T$ from $V(\cdot)$ such that $z_j^*=c_j+L_jv^*$, $j=1,\ldots,l$.

Consider the set $R(\alpha_N,0)=\{y\in\mathbb{R}^n: y=a_1+L_1v,\ v\in V(\cdot)\}$. Then it is easy to see that $R(\alpha_N,0)$ is the reachability set at $t=\tau_1$ for the following system

$$\dot{y}(t)=A(t)y(t)+g(v(t),t),\quad y(0)=\alpha(N),\quad v\in V(\cdot),\quad t\in T \tag{3.29}$$

Also it is well known, see, for example, [55], that $R(\alpha_N,0)$ is a closed set. Hence, for the sequence $\{z_1^n\}\to z_1^*$, $n\to\infty$, $z_1^n\in R(\alpha_N,0)$, $n=1,2,\ldots$ there exists a function

$v^1 \in V(\cdot)$ such that $z_1^* = a_1 + L_1 v^1$. Now introduce the sequence $\tilde{z}_2^n = \tilde{a}_2 + \tilde{L}_2 v^n$, where $\tilde{a}_2 = \Phi(\tau_2, \tau_1) z_1^*$ and $\tilde{L}_2 v^n = \int\limits_{\tau_1}^{\tau_2} \Phi(\tau_2, t) g(v^n(t), t))dt$, i.e. $\tilde{z}_2^n$ is the solution of the system (3.29) corresponding to the function $v^n(t)$ and initial condition $y(\tau_1) = z_1^*$, where $\tilde{z}_2^n$ and $v^n(t)$ are restricted to the interval $[\tau_1, \tau_2]$. Next, we show that $\tilde{z}_2^n \to z_2^*$.

It is known [46] that the fundamental matrix $\Phi(\tau, t)$ satisfies $\Phi(\tau, s)\Phi(s, t) = \Phi(\tau, t)$, $0 \le \tau < s < t \le t^*$, and the Cauchy response formula now yields

$$\begin{aligned} z_2^n &= \Phi(\tau_2, 0)\alpha(N) + \int\limits_0^{\tau_2} \Phi(\tau_2, t) g(v^n(t), t)dt = \Phi(\tau_2, \tau_1)\Big[\Phi(\tau_1, 0)\alpha(N) + \\ &\int\limits_0^{\tau_1} \Phi(\tau_1, t) g(v^n(t), t)dt\Big] + \int\limits_{\tau_1}^{\tau_2} \Phi(\tau_2, t) g(v^n(t), t)dt = \\ &\Phi(\tau_2, \tau_1) z_1^n + \int\limits_{\tau_1}^{\tau_2} \Phi(\tau_2, t) g(v^n(t), t)dt. \end{aligned}$$

Then

$$\tilde{z}_2^n = \Phi(\tau_2, \tau_1) z_1^* + \int\limits_{\tau_1}^{\tau_2} \Phi(\tau_2, t) g(v^n(t), t)dt.$$

Therefore

$$\|\tilde{z}_2^n - z_2^*\| \le \|\tilde{z}_2^n - z_2^n\| + \|z_2^n - z_2^*\| \le C\|z_1^n - z_1^*\| + \|z_2^n - z_2^*\|,$$

where $C = \|\Phi(\tau_2, \tau_1)\| < \infty$ is a constant. Since $z_1^n \to z_1^*, z_2^n \to z_2^*$, it follows immediately from the last inequality that also $\tilde{z}_2^n \to z_2^*$.

Introduce the set

$$R(z_1^*, \tau_1) = \{y \in \mathbb{R}^n : y = \tilde{a}_2 + \tilde{L}_2 v,\ v \in V(\cdot)\}. \tag{3.30}$$

Then it is obvious that $R(z_1^*, \tau_1)$ is the reachability set at $t = \tau_2$ for the system (3.29) restricted to the interval $[\tau_1, \tau_2]$ with initial condition $y(\tau_1) = z_1^*$. As shown above, $R(z_1^*, \tau_1)$ is a closed set. Therefore for the sequence $\tilde{z}_2^n \to z_2^*$, $n \to \infty$ such that $\tilde{z}_2^n \in R(z_1^*, \tau_1)$, there exists a function $v^2(t)$, $\tau_1 \le t \le \tau_2$, $v^2 \in V(\cdot)$, such that $z_2^* = \tilde{a}_2 + \tilde{L}_2 v^2$.

In an analogous way, it can be established that on every interval $[\tau_j, \tau_{j+1}]$, there exists a function $v^{j+1} \in V(\cdot)$, $j = 1, \ldots, l-1$, such that $z_{j+1}^* = \tilde{a}_{j+1} + \tilde{L}_{j+1} v^{j+1}$, where

$$\tilde{a}_{j+1} = \Phi(\tau_{j+1}, \tau_j) z_j^*, \quad \tilde{L}_{j+1} v = \int\limits_{\tau_j}^{\tau_{j+1}} \Phi(\tau_{j+1}, t) g(v(t), t))dt$$

Finally, we define on $T = [0, t^*]$ the function

$$v^*(t) = \begin{cases} v_1(t), & 0 \le t < \tau_1, \\ v_2(t), & \tau_1 \le t < \tau_2, \\ \ldots & \ldots \\ v^l(t), & \tau_{l-1} \le t \le t^* \end{cases}$$

where clearly $v^* \in V(\cdot)$. Also, it follows immediately from $z_j^* = \tilde{a}_j + \tilde{L}_j v^j$:

$$\tilde{a}_j + \tilde{L}_j v^j = \Phi(\tau_j, \tau_{j-1}) z_{j-1}^* + \int_{\tau_{j-1}}^{\tau_j} \Phi(\tau_j, t) g(v^j(t), t) dt = \Phi(\tau_j, \tau_{j-1}) \Big[\Phi(\tau_{j-1}, \tau_{j-2}) z_{j-2}^*$$

$$+ \int_{\tau_{j-2}}^{\tau_{j-1}} \Phi(\tau_{j-1}, t) g(v^{j-1}(t), t) dt \Big] + \int_{\tau_{j-1}}^{\tau_j} \Phi(\tau_j, t) g(v^j(t), t) dt = \Phi(\tau_j, \tau_{j-2}) z_{j-2}^*$$

$$+ \int_{\tau_{j-2}}^{\tau_{j-1}} \Phi(\tau_j, t) g(v^{j-1}(t), t) dt + \int_{\tau_{j-1}}^{\tau_j} \Phi(\tau_j, t) g(v^j(t), t)) dt = \cdots = \Phi(\tau_j, 0) \alpha(N)$$

$$+ \int_0^{\tau_1} \Phi(\tau_j, t) g(v^1(t), t) dt + \int_{\tau_1}^{\tau_2} \Phi(\tau_j, t) g(v^2(t), t) dt + \cdots + \int_{\tau_{j-1}}^{\tau_j} \Phi(\tau_j, t) g(v^j(t), t) dt$$

$$= \Phi(\tau_j, 0) \alpha(N) + \int_{\tau_0}^{\tau_j} \Phi(\tau_j, t) g(v^*(t), t) dt = a_j + L_j v^*, \quad j = 1, \ldots, l,$$

that $v^*(t)$ is the required function. Hence $z^* \in \mathcal{R}_N$, i. e. $\mathcal{R}_N$ is a closed set and the proof is complete.

■

Note. In the cases when $k \neq N$, the additional terms in the formulas for a_j and L_j in the set $\mathcal{R}_k$ do not change the essence of given proof.

At this stage, we have established that $\mathcal{R}$ and $K(\delta)$ are closed and convex sets and the next result gives the solution of *Problem* (A), where the inner product of vectors g and f from $\mathbb{R}^{nl}$ is denoted by $g^T f$.

Theorem 11. *Problem* (A) *has a solution if, and only if,*

$$\max_{\|g\|_{\mathbb{R}^{nl}}=1} \{ g^T c - \max_{z \in K(\delta)} g^T z + \min_{u \in U(\cdot)} g^T S u \} \leq 0 \tag{3.31}$$

holds.

Proof. *Sufficiency.* Suppose that the condition of (3.31) is valid, but Problem (A) has no solution. Then, $\mathcal{R} \cap K(\delta) = \emptyset$ and the separation theorem for convex sets yields that there exists a nontrivial vector $g \in \mathbb{R}^{nl}$, $\|g\| = 1$ such that

$$\min_{z \in \mathcal{R}} g^T z > \max_{z \in K(\delta)} g^T z. \tag{3.32}$$

Hence

$$g^T c - \max_{z \in K(\delta)} g^T z + \min_{u \in U(\cdot)} g^T S u > 0 \tag{3.33}$$

which contradicts (3.31).

Necessity. Suppose that Problem (A) has a solution. Then there exist $\bar{u}$ and $\bar{z}$ satisfying (3.11)–(3.12) such that $g^T c + g^T S \bar{u} = g^T \bar{z}$ holds for each $g \in \mathbb{R}^{nl}$. Taking the maximum and minimum respectively of the two terms in this last expression now yields

$$g^T c - \max_{z \in K(\delta)} g^T z + \min_{u \in U(\cdot)} g^T S u \leq 0, \tag{3.34}$$

as required and the proof is complete.

■

3.2.2 Optimality conditions

In this sub-section we use the results of the previous sub-section to establish the maximum principle for the optimal control problem (3.1)–(3.4).

Introduce the function $\Lambda : \mathbb{R} \to \mathbb{R}$ as

$$\Lambda(\delta) = \max_{\|g\|_{\mathbb{R}^{nl}}=1} \big\{ g^T c - \max_{z \in K(\delta)} g^T z + \max_{u \in U(\cdot)} g^T S u \big\}. \tag{3.35}$$

where it can be shown that $\Lambda : \mathbb{R} \to \mathbb{R}$ defined by (3.35) is a non increasing continuous function. Hence the optimal value of the performance index (3.4) can be characterized as follows.

Theorem 12. *The control $u^0 \in U(\cdot)$ is the optimal solution of the problem defined by (3.1)–(3.4) if, and only if, $\delta^0 := J(u^0)$ is the smallest root of the equation $\Lambda(\delta) = 0$.*

Proof. *Necessity.* Let $u^0 \in U(\cdot)$ be an optimal control of the problem (3.1)–(3.4). Then u^0 is the solution of Problem (A) with $\delta^0 := J(u^0)$. Therefore, Theorem 11 yields that $\Lambda(\delta^0) \leq 0$.

Suppose now that $\Lambda(\delta^0) < 0$. Then since $\Lambda(\delta)$ is a continuous and monotone function, there exists a number $\bar{\delta}$ such that $\bar{\delta} < \delta^0$ and $\Lambda(\bar{\delta}) \leq 0$. Hence, Theorem 11 yields that Problem (A) has a solution with $\delta = \bar{\delta}$ since otherwise there would be an available control $\bar{u} \in U(\cdot)$ and a vector $\bar{z} \in M$ satisfying (3.11)–(3.12) in the case when $\delta = \bar{\delta}$. Hence, $J(\bar{u}) < J(u^0)$, which contradicts the optimality of the control u^0 and therefore $\Lambda(\delta^0) = 0$. Finally, the fact that δ^0 is the smallest root of the equation $\Lambda(\delta) = 0$ can be proved as above.

Sufficiency. Let $u^0 \in U(\cdot)$ be a control function such that δ^0 is the smallest root of $\Lambda(\delta) = 0$. Suppose also that $u^0(t)$ is not an optimal solution of the problem (3.1)–(3.4). Then there exists an available control function $\bar{u} \in U(\cdot)$ and a vector $\bar{z} \in M$ such that $c - \bar{z} + S\bar{u} = 0$ and $J(\bar{u}) < J(u^0)$ holds. This establishes that Problem (A) has a solution for $\bar{\delta} = J(\bar{u})$, and hence $\Lambda(\bar{\delta}) \leq 0$.

Conversely, since the function $\Lambda(\delta)$ is monotone $\Lambda(\bar{\delta}) \geq \Lambda(J(u^0)) = 0$, which contradicts the assertion that δ^0 is the smallest root. Hence u^0 is an optimal control and the proof is complete.

■

Now let $g^0 = (g_1^0, ..., g_l^0)^T \in \mathbb{R}^{nl}$ be a maximizing vector for $\Lambda(\delta^0)$ and on the interval $T = [0, t^*]$ we introduce the following function $\lambda : \mathbb{R} \to \mathbb{R}^m$

$$\lambda(t) = \sum_{i=j+1}^{l} (g_i^0)^T \Phi(\tau_i, t), \quad \tau_j \leq t < \tau_{j+1}, \quad j = 0, \ldots, l-1. \tag{3.36}$$

Then it is a simple task to verify that the function $\lambda(t)$ satisfies

$$\frac{d\lambda(t)}{dt} = -\lambda^T(t)A(t), \quad \lambda(\tau_j - 0) - \lambda(\tau_j + 0) = g_j^0, \quad j = 1, \ldots, l-1. \tag{3.37}$$

and the optimality conditions for (3.1)–(3.4) are given by the following theorem.

Theorem 13. *If the number δ^0 is the smallest root of the equation $\Lambda(\delta) = 0$, then there exists an optimal control $u_k^0(t)$, $k \in K$, $t \in T$ for the problem (3.1)–(3.4) such that $J(u^0) = \delta^0$ and for almost all $t \in T$*

$$\psi_k^T(t) b_{N-k+1}\big(u_{N-k+1}^0(t), t\big) = \min_{v \in U} \psi_k^T(t) b_{N-k+1}(v, t), \tag{3.38}$$

holds for all $k \in K$. Here the function $\psi : K \times T \to \mathbb{R}^n$ is given by

$$\psi_k(t) = \int_0^t \psi_{k-1}^T(\tau) D(\tau)\Phi(\tau, t) d\tau, \quad \psi_1(t) = \lambda(t), \quad k \in K, \tag{3.39}$$

where the function $\lambda(t)$ is given by (3.37).

Proof. Since $\Lambda(\delta^0) = 0$, Theorem 12 yields that Problem (A) has a solution for $\delta = \delta^0$. This implies that there exists an available control $u^0 \in U(\cdot)$ and a vector $z^0 \in M$ satisfying (3.11)–(3.12). Hence $\varphi(z^0) = J(u^0) \leq \delta^0$ and the assumption $J(u^0) < \delta^0$ leads to a contradiction with the assumption that δ^0 is the smallest root of the equation $\Lambda(\delta) = 0$. Therefore, $J(u^0) = \delta^0$, and, consequently, u^0 is optimal control for (3.1)–(3.4).

The function $u_k^0(t)$, $k \in K$, $t \in T$ satisfies

$$(g^0)^T S u^0 = \min_{u \in U(\cdot)} (g^0)^T S u. \tag{3.40}$$

and if we assume that $(g^0)^T S u^0 > \min_{u \in U(\cdot)} (g^0)^T S u$, then

$$\Lambda(\delta^0) < (g^0)^T c - (g^0)^T z^0 + (g^0)^T S u = 0, \tag{3.41}$$

which is impossible since δ^0 is a root of $\Lambda(\delta) = 0$. Finally, to establish the desired optimality condition (3.38) we employ (3.40). For ease of notation the function $b_k(u_k(t), t),\ t \in T$ is subsequently denoted by $b_u(k)$. Then

$$\min_{u \in U(\cdot)} (g^0)^T S u = \min_{u \in U(\cdot)} \sum_{j=1}^{l} (g_j^0)^T \left(\sum_{i=1}^{N-1} \mathcal{P}^i Q b_u(N-i)(\tau_j) + \int_0^{\tau_j} \Phi(\tau_j, t) b_N(u_N(t), t) dt \right) =$$

$$= \min_{u \in U(\cdot)} \Bigg\{ \int_0^{\tau_1} \Big[(g_1^0)^T \Phi(\tau_1, t) + \cdots + (g_l^0)^T \Phi(\tau_l, t) \Big] b_N(u_N(t), t) dt + \int_{\tau_1}^{\tau_2} \Big[(g_2^0)^T \Phi(\tau_2, t) + \cdots$$

$$+ (g_l^0)^T \Phi(\tau_l, t) \Big] b_N(u_N(t), t) dt + \cdots + \int_{\tau_{l-1}}^{\tau_l} (g_l^0)^T \Phi(\tau_l, t) b_N(u_N(t), t) dt + \cdots$$

$$+ \int_0^{\tau_1} \Big[(g_1^0)^T \Phi(\tau_1, t) + \cdots + (g_l^0)^T \Phi(\tau_l, t) \Big] D(t) \int_0^t \Phi(t, s) b_{N-1}(u_{N-1}(t), t) ds dt$$

$$+ \int_{\tau_1}^{\tau_2} \Big[(g_2^0)^T \Phi(\tau_2, t) + \cdots + (g_l^0)^T \Phi(\tau_l, t) \Big] D(t) \int_0^t \Phi(t, s) b_{N-1}(u_{N-1}(t), t) ds dt + \cdots$$

$$+ \int_{\tau_{l-1}}^{\tau_l} (g_l^0)^T \Phi(\tau_l, t) D(t) \int_0^t \Phi(t, s) b_{N-1}(u_{N-1}(t), t) ds dt + \cdots + \int_0^{\tau_1} \Big[(g_1^0)^T \Phi(\tau_1, t) + \cdots$$

$$+ (g_l^0)^T \Phi(\tau_l, t) \Big] D(t) \mathcal{P}^{N-1} Q b_u(1)(t) dt + \cdots + \int_{\tau_{l-1}}^{\tau_l} (g_l^0)^T \Phi(\tau_l, t) D(t) \mathcal{P}^{N-1} Q b_u(1)(t) dt \Bigg\}$$

$$= \min_{u \in U(\cdot)} \Big\{ \psi_1^T(t) b_N(u_N(t), t) + \cdots + \psi_N^T(t) b_1(u_1(t), t) \Big\} = \sum_{k \in K} \min_{v \in U} \psi_k^T(t) b_{N-k+1}(v, t).$$

which yields (3.38) and the proof is complete.

■

The analysis just completed gives the optimal control solution in the standard maximum principle form which can be very difficult for numerical computations as required in applications. Hence, in the remainder of the paper we proceed to develop new optimality and sub-optimality conditions which are more suitable for numerical purposes. However, as the first step we limit our attention to the stationary case.

3.3 Stationary Differential Linear Repetitive Processes

In this Section, the process (3.1) is assumed to be stationary, and the pass constraints (3.3) and the cost function (3.4) have a special form as detailed below. Also the solutions here are, in effect, developed by extending the constructive methods approach developed in [37] to the repetitive process setting.

The processes considered in this section are described in $\mathbb{R}^n$ by the following linear matrix differential equation

$$\frac{dx_k(t)}{dt} = Ax_k(t) + Dx_{k-1}(t) + bu_k(t), \quad k \in K = \{1, \cdots, N\},\ t \in T = [0, t^*] \quad (3.42)$$

with boundary conditions

$$x_k(0) = \alpha_k, \quad k \in K, \quad x_0(t) = f(t), \quad t \in T, \quad (3.43)$$

and a pass end, or terminal, constraint of the form

$$H_k x_k(t^*) = g_k, \qquad k \in K, \quad (3.44)$$

Here $b,\ \alpha_k$ are specified $n \times 1$ vectors and $A, D, H_k,\ k \in K$ are constant matrices of compatible dimensions. In addition, we assume that the matrix A has simple eigenvalues $\lambda_j,\ 1 \le i \le n$, and that it is a stable matrix in the sense that $Re\ \lambda_i < 0,\ 1 \le i \le n$.

Definition 11. *For every $k \in K$ the piecewise continuous function $u_k : T \to \mathbb{R}$ is termed an admissible control for pass k if it satisfies*

$$|u_k(t)| \le 1,\ t \in T. \quad (3.45)$$

The optimization problem is to find the admissible controls $u_1(t), ..., u_N(t)$ such that the corresponding solution of the system (3.42)–(3.44) maximizes the following cost function

$$J(u) = \sum_{k \in K} p_k^T x_k(t^*), \quad (3.46)$$

where $p_k,\ k = 1, \ldots, n$ are given $n \times 1$ vectors.

3.3.1 Optimality conditions for supporting control functions.

First, note that the solution of the form (3.42)–(3.43) (with no terminal conditions of (3.44)) can be written as follows

$$x_k(t) = \sum_{j=1}^{k} K_j(t)\alpha_{k+1-j} + \int_0^t K_k(t-\tau)Df(\tau)d\tau + \sum_{j=1}^{k} \int_0^t K_j(t-\tau)bu_{k+1-j}(\tau)d\tau \quad (3.47)$$

where the $K_i(t)$ are the solutions of the following $n \times n$ matrix differential equations

$$\dot{K}_1(t) = AK_1(t), \quad \dot{K}_i(t) = AK_i(t) + DK_{i-1}(t), \quad i = 2, \ldots, N, \quad (3.48)$$

with initial conditions

$$K_1(0) = E, \quad K_i = 0, \quad i = 2, \ldots, N. \quad (3.49)$$

Also it is easy to show that these solutions have the following properties

$$
\begin{aligned}
K_j(t-\sigma) &= \int_{\sigma}^{t} K_{j-k}(t-\tau)DK_k(\tau-\sigma)d\tau, \qquad 0 \le \sigma < t \le t^*, \quad k = 1,\dots,j-1;\\
K_j(t-\sigma) &= \sum_{s=1}^{j} K_s(t-\tau)K_{j+1-s}(\tau-\sigma) \qquad j = 2,\dots,N-1,
\end{aligned} \tag{3.50}
$$

which will be used below.

Now using (3.47) we can rewrite the optimization problem in the following integral form

$$
\max_{u_1,\dots,u_N} J(u),\; J(u) = \sum_{j=1}^{N} \int_0^{t^*} c_j(\tau)u_j(\tau)d\tau + \gamma \tag{3.51}
$$

subject to the terminal conditions (3.43) and the control constraints (3.45), which can also be rewritten as

$$
\begin{cases}
\int_0^{t^*} g_{11}(\tau)u_1(\tau)d\tau & = h_1,\\
\int_0^{t^*} \Big[g_{21}(\tau)u_1(\tau) + g_{22}(\tau)u_2(\tau)\Big]d\tau & = h_2,\\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & \\
\int_0^{t^*} \Big[g_{N1}(\tau)u_1(\tau) + \dots + g_{NN}(\tau)u_N(\tau)\Big]d\tau & = h_N,
\end{cases} \tag{3.52}
$$

and

$$
|u_k(\tau)| \le 1,\; \tau \in [0,t^*], \quad k = 1,\dots,N,
$$

where

$$
\begin{aligned}
\gamma &= \sum_{k=1}^{N}\sum_{j=1}^{k} p_k^T K_j(t^*)\alpha_{k+1-j} + \sum_{k=1}^{N}\int_0^{t^*} p_k^T K_k(t^*-\tau)Df(\tau)d\tau,\\
c_j(\tau) &= \sum_{k=j}^{N} p_k^T K_{k+1-j}(t^*-\tau)b, \quad j = 1,\dots,N, \quad g_{kj}(\tau) = H_k K_{k+1-j}(t^*-\tau)b, \quad j \le k,\\
h_k &= g_k - \sum_{j=1}^{k} H_k K_j(t^*)\alpha_{k+1-j} - \int_0^{t^*} H_k K_k(t^*-\tau)Df(\tau)d\tau \quad k = 1,...,N.
\end{aligned}
$$

Also we require the following concepts.

Definition 12. *For each fixed* $k,\ 1 \le k \le N$, *the instances* $0 < \tau_{k1} < \tau_{k2} < \cdots < \tau_{km} < t^*$ *are termed supporting and their collection* $\tau_{sup}^k := \{\tau_{k1},\dots,\tau_{km}\}$ *is termed the support of pass* k *for the problem (3.42)–(3.46) if the matrix* $G_{sup}^k := \{g_{kk}(\tau_{k1}),\dots,g_{kk}(\tau_{km})\}$ *is non-singular.*

Note also that from (3.48) we have that $g_{kk}(\tau) = H_k e^{A(t^*-\tau)}b$. Therefore the existence of the support τ^k_{sup} is guaranteed by controllability of the pair $\{A, b\}$.

Definition 13. *A pair $\left\{\tau^k_{sup},\ u_k(t), k = 1, ..., N\right\}$ consisting of a support τ^k_{sup} and admissible control functions $u_k(t), t \in T$ is termed a supporting control function for (3.42)–(3.46).*

Remark 15. *These last two definitions are motivated as follows.*

Often an optimal control problem solution has the so-called bang-bang form, i.e. the control function takes only boundary values in the admissible set U. If $U = \{-1 \le u \le +1\}$ then $u^0(t) = \pm 1$ (the 'switch on/switch off' regime). Also the switching times are constructive elements in the design of the optimal controller. Hence, our goal is to apply these key elements directly to the optimality conditions and hence we use the supporting time instances and control.

Let $\left\{\tau^k_{sup}, u_k(t), k = 1, \dots, N\right\}$ be a support control function and construct a sequence of $m \times 1$ vectors $\left\{\nu^{(k)},\ k = 1, \dots, N\right\}$ by solving following set of linear algebraic equations

$$\begin{cases} (\nu^{(N)})^T G^N_{sup} - c^{(N)}_{sup} & = 0, \\ (\nu^{(N-1)})^T G^{N-1}_{sup} + (\nu^{(N)})^T F^N_{(N-1)sup} - c^{(N-1)}_{sup} & = 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & \\ (\nu^{(1)})^T G^1_{sup} + (\nu^{(2)})^T F^2_{1sup} + \dots + (\nu^{(N)})^T F^N_{1sup} - c^{(1)}_{sup} & = 0 \end{cases} \tag{3.53}$$

where the vectors $c^{(k)}_{sup}$ and matrices F^k_{jsup} are given by

$$\begin{aligned} c^{(k)}_{sup} &:= \big(c_k(\tau_{k1}), \dots, c_k(\tau_{km})\big)^T, \qquad k = 1, \dots, N, \\ F^k_{jsup} &:= \Big(g_{kj}(\tau_{j1}), \dots, g_{kj}(\tau_{jm})\Big), \qquad k > j, \quad j = 1, \dots, N-1. \end{aligned}$$

Now define the so-called co-control function as

$$\Delta(t) = (\Delta_1(t), \dots, \Delta_N(t))^T := \hat{\nu}' \hat{G}(t) - c(t),$$

where $\hat{\nu} = (\nu^{(1)}, \dots, \nu^{(N)})^T, \quad c(t) = (c_1(t), \dots, c_N(t))^T$, and $\hat{G}(t)$ is an $mN \times mN$ matrix function of triangular form, whose rows are the $m \times 1$ vector functions $g_{ij}(t)$ of (3.52). Note also that the special form of the matrix $\hat{G}(t)$ yields that the $mN \times mN$ matrix

$$\tilde{G}_{sup} := \begin{pmatrix} g_{kj}(t), & t \in \tau^k_{sup} \\ j \ge k, & k = 1, ..., N \end{pmatrix} \tag{3.54}$$

is nonsingular. Hence the $mN \times 1$ vector $\nu = (\nu^{(1)}, \dots, \nu^{(N)})^T$ required in (3.53) is given by $\nu = c^T_{sup} \tilde{G}^{-1}_{sup}$.

Definition 14. *We say that the support control function $\{\tau^k_{sup}, u_k(t), k = 1, \dots, N\}$ is non-degenerate for the problem (3.42)–(3.46) if*

$$\frac{d\Delta_k(\tau_j)}{dt} \ne 0 \quad \forall\ \tau_j \in \tau^k_{sup}, \quad k = 1, \dots, N.$$

Remark 16. *Here non-degeneracy means that in a small neighborhood of the supporting points the admissible control can be replaced by constant functions whose values are less than those on the control constraint boundary and satisfy (3.52), i.e. the support control function is non-singular if there exist numbers* $\lambda_0 > 0, \quad \mu_0 > 0, \quad u_j^k(\lambda), \quad j = 1, \ldots, m, \quad k = 1, \ldots, N$ *such that the following equalities*

$$\sum_{j=1}^{k}\sum_{i=1}^{m} u_j^i(\lambda) \int_{\tau_{ij}-\lambda}^{\tau_{ij}+\lambda} g_{kj}(t)dt = \sum_{j=1}^{k}\sum_{i=1}^{m} \int_{\tau_{ij}-\lambda}^{\tau_{ij}+\lambda} g_{kj}(t)u_j(t)dt,$$

$$|u_j^k| \leq 1 - \mu_0, \qquad j = 1, \ldots, m, \quad k = 1, \ldots, N. \tag{3.55}$$

hold for all $\lambda, \quad 0 < \lambda < \lambda_0$ *and* $k, \quad 1 \leq k \leq N$*. Also (3.55) will be used below in the proof of the optimality conditions.*

Associate with each supporting time instance τ_{kj} a small sub-interval T_{kj} from T such that the matrix $G_{gen}^k := \{ \int_{T_{kj}} g_{kk}(\tau)d\tau, j = 1, \ldots, m\}$ is non-singular, and, without loss of generality, we can also assume that τ_{kj} is one or other of the end points of T_{kj} and the supporting control function $u_k(t) = u_j^k$ for $t \in T_{kj}, \quad j = 1, \ldots, N$ are constant over the segments T_{kj}.

Now we have the following result.

Theorem 14. *A supporting control function* $\{\tau_{sup}^k, \quad u_k^0(t), \quad k = 1, \ldots, N\}$ *is an optimal solution of the problem (3.42)–(3.46) if the following conditions are fulfilled*

$$\begin{aligned} &\Delta_k(t) \geq 0 \quad at \quad u_k(t) = -1; \;\; \Delta_k(t) \leq 0 \quad at \quad u_k(y) = +1 \\ &\Delta_k(t) = 0 \quad at \quad -1 < u_k(t) < +1, \;\; k = 1, 2, \ldots, N, \;\; t \in T. \end{aligned} \tag{3.56}$$

Moreover, if this supporting control function is non-degenerate then the above condition is necessary and sufficient, too.

Proof. *Sufficiency.* Let $u_k(t) \neq u_k^0(t), \quad k = 1, \ldots, N$, be an admissible control and $x_k(t)$ the corresponding trajectory of the system (3.42)–(3.43). Then standard transformations yield that the increment, $\Delta J(u) := J(u^0) - J(u,)$ of the cost function can be expressed in the form

$$\Delta J(u) = \int_0^{t^*} \sum_{j=1}^{N} c_j(t)\big[u_j^0(t) - u_j(t)\big]dt = -\sum_{j=1}^{N} \int_0^{t*} \Delta_j(t)\big[u_j^0(t) - u_j(t)\big]dt.$$

Hence, (3.56) yields that $\Delta J(u) \geq 0$ for any admissible control u, i.e. $\{\tau_{sup}^k, u_k^0\}$ is an optimal supporting control function.

Necessity. Let $\{\tau_{sup}^k, \; u_k^0(t), \; k = 1, \ldots, N\}$ be an optimal non-degenerate control but $\exists \; k_*, 1 \leq k_* \leq N$ and $\exists \; t_* \in T$, such that the theorem is not valid. If we suppose that $t_* \in [\tau_{k_*j} - \lambda, \tau_{k_*j} + \lambda]$ where $\lambda > 0$ is a small number, i.e. the instance t_* lies in an neighborhood of some supporting time instance τ_{k_*j}, then using the fact that the supporting control is non-degenerate yields that there exists a control variation $\Delta u_{k_*}^0(t)$, defined on the intervals $[\tau_{k_*j} - \lambda, \tau_{k_*j} + \lambda]$, such that $J(u^0) > 0$, which contradicts the optimality of $u_k^0(t)$. Therefore, we next suppose that $t_* \notin [\tau_{k_*j} - \lambda, \tau_{k_*j} + \lambda] \quad \forall \, j = 1, \ldots, m$ for some small $\lambda > 0$.

Next, without loss of generality, assume that $\Delta^0_{k_*}(t_*) > 0$ and $u_{k_*}(t_*) > 0$. Then by continuity of $\Delta_{k_*}(t)$ and piecewise-continuity of $u_{k_*}(t)$ there exists an neighborhood $T_{k_*}(t_*)$ of t_*, such that $\Delta k_*(t) > 0, \quad u_{k_*}(t) > -1$ for $t \in T_{k_*}(t_*)$. Now, we have to construct the admissible control variation such that the corresponding increment of the cost function satisfies $\Delta J(u) > 0$, which is impossible for the optimal controls $u^0_k(t)$.

Consider now the case of a small real number $\lambda_0 > 0$ (we see below that the existence of such number λ_0 is guaranteed by the fact that the supporting control is non-degenerate) and for all $\lambda,\ 0 < \lambda < \lambda_0$ define the control variation $\Delta u(t) = (\Delta u_1(t), ..., \Delta u_N(t)), \quad t \in T$ as

$$
\begin{aligned}
\Delta u_k(t) &= 0, \quad k < k_*,\ t \in T; \\
\Delta u_{k_*}(t) &= \begin{cases} \theta(-1 - u_{k_*}(t)), & \theta > 0, \quad t \in T_{k_*}(t), \\ 0, & t \in T \setminus \left(\bigcup\limits_{j=1}^{m} [\tau_{k_*j} - \lambda, \tau_{k_*j} + \lambda] \cup T_{k_*}(t) \right). \end{cases}
\end{aligned}
$$

Hence the control variations on the intervals $[\tau_{k_*j} - \lambda, \tau_{k_*j} - \lambda],\ j = 1, \ldots, m$ can be chosen as constant functions $\Delta u_{k_*}(t) \equiv \Delta\vartheta^k_j(\lambda)$. The control variations for the remaining passes $k > k_*$ are defined as

$$
\Delta u_k(t) \equiv 0, \quad k = k_* + 1, \ldots, N, \quad t \in T \setminus \bigcup_{j=1}^{m} [\tau_{kj} - \lambda, \tau_{kj} + \lambda];
$$

$$
\Delta u_k(t) \equiv \Delta\vartheta^k_j(\lambda),\ t \in [\tau_{kj} - \lambda, \tau_{kj} + \lambda], \quad j = 1, \ldots, m, \quad k > k_*
$$

where $\Delta\vartheta^k_j(\lambda)$ are unknown constants which are determined below.

Using (3.52), it follows that the conditions

$$
\int_0^{t^*} \sum_{s=1}^{k} g_{ks}(\tau) \Delta u_s(\tau) d\tau = 0,\ k = 1, \ldots, N, \tag{3.57}
$$

hold for any admissible variation $\Delta u(t)$ and can be re-written in the form

$$\phi_{k_*}(\lambda) := \sum_{j=1}^{m} \int_{\tau_{k_*j}-\lambda}^{\tau_{k_*j}+\lambda} g_{k_*k_*}(\tau)\vartheta_j^{k_*}(\lambda)d\tau = -\theta \int_{T_{k_*}(t_*)} g_{k_*k_*}(\tau)(-1-u_{k_*}(\tau))d\tau,$$

$$\phi_{k_*+1}(\lambda) := \sum_{j=1}^{m} \int_{\tau_{k_*+1j}-\lambda}^{\tau_{k_*+1j}+\lambda} g_{k_*+1k_*+1}(\tau)\vartheta_j^{k_*+1}(\lambda)d\tau$$

$$= -\sum_{j=1}^{m} \int_{\tau_{k_*j}-\lambda}^{\tau_{k_*j}+\lambda} g_{k_*+1k_*}(\tau)\vartheta_j^{k_*}(\lambda)d\tau - \theta \int_{T_{k_*}(t_*)} g_{k_*+1k_*}(\tau)(-1-u_{k_*}(\tau))d\tau, \tag{3.58}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\phi_N(\lambda) := \sum_{j=1}^{m} \int_{\tau_{Nj}-\lambda}^{\tau_{Nj}+\lambda} g_{NN}(\tau)\vartheta_j^{N}(\lambda)d\tau = -\sum_{j=1}^{m} \int_{\tau_{k_*j}-\lambda}^{\tau_{k_*j}+\lambda} g_{Nk_*}(\tau)\vartheta_j^{k_*}(\lambda)d\tau$$

$$-\theta \int_{T_{k_*}(t_*)} g_{Nk_*}(\tau)(-1-u_{k_*}(\tau))d\tau - \cdots - \sum_{j=1}^{m} \int_{\tau_{N-1j}-\lambda}^{\tau_{N-1j}+\lambda} g_{NN-1}(\tau)\vartheta_j^{N-1}(\lambda)d\tau$$

Expanding the function $\phi_{k_*}(\lambda)$ of (3.58) in a Taylor series truncated at the second order and setting $\Delta\vartheta_\lambda^{k_*} = \Delta\vartheta_1^{k_*}(\lambda), \ldots, \Delta\vartheta_m^{k_*}(\lambda)$ leads to

$$2\lambda G_{sup}^{k_*}\Delta\vartheta_\lambda^{k_*} + \frac{\lambda^3}{3}\{\frac{d^2 g_{k_*k_*}(\tau_{k_*j})}{d\tau}, \mathrm{J} = \overline{1,m}\}\Delta\vartheta_\lambda^{k_*} + o_{k_*}(\lambda^3) =$$
$$= -\theta \int_{T_{k_*}(t_*)} g_{k_*k_*}(\tau)(-1-u_{k_*}(\tau))d\tau.$$

where $o_{k_*}(\lambda^3)$ denotes terms of degree 3 and above which are neglected here. Hence the required vector $\Delta\vartheta_\lambda^{k_*}$ can be represented as

$$\Delta\vartheta_\lambda^{k_*} = \frac{1}{\lambda}\theta\hat{u}_{k_*} + \theta O_{k_*}(\lambda), \quad \text{where} \quad \hat{u}_{k_*} = -\frac{1}{2}G_{sup}^{k_*^{-1}} \int_{T_{k_*}(t_*)} g_{k_*k_*}(\tau)(-1-u_{k_*}(\tau))d\tau. \tag{3.59}$$

and $O_{k_*}(\lambda)$ denotes a residual first order term. Using (3.55) and (3.59), it follows that for $\lambda \in (0, \lambda_0)$ there exists the real number $\theta = \theta(\lambda)$, such that $\theta(\lambda) = \mu_{k_*}\lambda \le 1$, where $\mu_{k_*} > 0$ does not dependent on λ, and the following inequalities

$$|u_j^{k_*}(\lambda) + \Delta\vartheta_j^{k_*}(\lambda)| \le 1, \quad j = 1, \ldots, m$$

hold. Here we have exploited the fact that the admissible controls are constants $u_j^k(\lambda)$ over the intervals T_j^k, containing the supporting points τ_{kj}. Hence, the function

$$\bar{u}_{k_*}(t) = \begin{cases} u_j^{k_*}(\lambda) + \Delta\vartheta_j^{k_*}(\lambda), & t \in [\tau_{k_*j}-\lambda, \tau_{k_*j}+\lambda] \\ u_{k_*}(t) + \theta(\lambda)(-1-u_{k_*}(t)), & t \in T_{k_*}(t_*) \end{cases}$$

is an admissible control function for $\theta(\lambda) = \mu_{k_*}\lambda \leq 1$ and a sufficiently small μ_{k_*}.

In order to find $\Delta\vartheta_\lambda^{k_*+1}$ and $\theta(\lambda)$, expand $\phi_{k_*+1}(\lambda)$ as a Taylor series to yield

$$\sum_{j=1}^{m} \int_{\tau_{k_*j}-\lambda}^{\tau_{k_*j}+\lambda} g_{k_*+1k_*}(\tau)\Delta\vartheta_j^{k_*}(\lambda)d\tau = 2\lambda \sum g_{k_*+1k_*}(\xi_j)\Delta\vartheta_j^{k_*}(\lambda) = 2\lambda\tilde{G}_\xi^{k_*+1}\Delta\vartheta_\lambda^{k_*+1}$$

$$= 2\lambda\tilde{G}_\xi^{k_*+1}\left(\frac{1}{\lambda}\mu_{k_*}\lambda\hat{u}_{k_*} + \mu_{k_*}\lambda O_{k_*}(\lambda)\right) = 2\tilde{G}_\xi^{k_*+1}\mu_{k_*}\lambda\hat{u}_{k_*} + \mu_{k_*}\breve{o}_{k_*}(\lambda^3) \quad (3.60)$$

Here the matrix $\tilde{G}_\xi^{k_*+1}$ is constructed from the rows $\{g_{k_*+1k_*}(\xi_j),\ j = 1,\dots,m\}$, where ξ_j are points from the intervals $[\tau_{k_*j} - \lambda, \tau_{k_*j} + \lambda]$.

Next, set $\Delta\vartheta_\lambda^{k_*+1} = (\Delta\vartheta_1^{k_*+1}(\lambda),\dots,\Delta\vartheta_m^{k_*+1}(\lambda))$ to obtain

$$2\lambda G_{sup}^{k_*+1}\Delta\vartheta_\lambda^{k_*+1} + \frac{\lambda^3}{3}\left\{\frac{d^2 g_{k_*+1k_*+1}(\tau_{k_*j})}{d\tau}, j = 1,\dots,m\right\}\Delta\vartheta_\lambda^{k_*+1} + o_{k^*+1}(\lambda^3)$$

$$= -\mu_{k_*}\lambda\left\{\tilde{G}_\xi^{k_*+1}\hat{u}_{k_*} + \int_{T_{k_*}(t_*)} g_{k_*+1k_*+1}(\tau)(-1 - u_{k_*}(\tau))d\tau\right\} + \mu_{k_*}\breve{o}_{k_*}(\lambda^3). \quad (3.61)$$

which means that the required vector $\Delta\vartheta_\lambda^{k_*+1}$ can be expressed as

$$\begin{aligned}\Delta\vartheta_\lambda^{k_*+1} &= \frac{1}{\lambda}\mu_{k_*}\lambda\hat{u}_{k_*+1} + \mu_{k_*}\lambda O_{k^*+1}(\lambda),\\ \hat{u}_{k_*+1} &= -\frac{1}{2}\left(G_{sup}^{k_*+1}\right)^{-1}\left\{\tilde{G}_\xi^{k_*+1}\hat{u}_{k_*} - \int_{T_{k_*}(t_*)} g_{k_*+1k_*+1}(\tau)(-1 - u_{k_*}(\tau))d\tau\right\}. \quad (3.62)\end{aligned}$$

Now choose $\Delta\vartheta_\lambda^{k_*+1}$ such that the following inequalities hold

$$|u_j^{k_*+1}(\lambda) + \Delta\vartheta_j^{k_*+1}(\lambda)| \leq 1, \quad j = 1,\dots,m$$

and hence the values of μ_{k_*} and λ_0 can be decreased as required. Continuing this expansion procedure for the remaining equations in (3.58), we obtain the desired admissible control function in the form

$$\bar{u}(t) = u^0(t) + \Delta u(t) = \left\{u_1^0(t) + \Delta u_1(t),\dots,u_N^0(t) + \Delta u_N(t)\right\}, \quad t \in T.$$

and note here that $\Delta u_k(t) = 0 \quad \forall\ k < k_*$.

Now calculate the increment of the cost function generated by the designed control function $\bar{u}(t)$ as

$$\begin{aligned}\Delta J(u) &= J(\bar{u}) - J(u^0) = \sum_{k=1}^{N}\int_0^{t^*}\Delta_k(t)\Delta u_k(t)dt = \sum_{k=k_*}^{N}\int_0^{t^*}\Delta_k(t)\Delta u_k(t)dt \\ &= -\ \theta\int_{T_{k_*}(t_*)}\Delta_{k_*}(t)\big(-1-u_{k_*}(t)\big)dt - \\ &- \sum_{j=1}^{m}\int_{\tau_{k_*j}-\lambda}^{\tau_{k_*j}+\lambda}\Delta_{k_*}(t)\big[u_j^{k_*}(\lambda)+\Delta\vartheta_j^{k_*}(\lambda)-u_j^{k_*}(t)\big]dt - \\ &- \sum_{s=k_*+1}^{N}\sum_{j=1}^{m}\int_{\tau_{sj}-\lambda}^{\tau_{sj}+\lambda}\Delta_s(t)\big[u_j^s(\lambda)+\Delta\vartheta_j^s(\lambda)-u_j^s(t)\big]dt. \end{aligned} \tag{3.63}$$

Since $\Delta_k(\tau_{kj}) = 0,\ k = k_*,\ \dots, N,\ j = 1,\dots,m$, then again from the Taylor series expansion in λ, we have the following estimate for the integral components

$$\begin{aligned}\int_{\tau_{sj}-\lambda}^{\tau_{sj}+\lambda}\Delta_s(t)\big[u_j^s(\lambda)+\Delta\vartheta_j^s(\lambda)-u_j^s(t)\big]dt &= \int_{\tau_{sj}}^{\tau_{sj}}\Delta_s(t)\big[u_j^s(\lambda)+\Delta\vartheta_j^s(\lambda)-u_j^s(t)\big]dt \\ +2\lambda\Delta_s(\tau_{sj})\big[u_j^s(\lambda)+\Delta\vartheta_j^s(\lambda)-u_j^s(\tau_{sj})\big] \\ +\lambda^2\frac{d\Delta_s(\tau_{sj})}{dt}\big[u_j^s(\lambda) &+ \Delta\vartheta_j^s(\lambda)-u_j^s(\tau_{sj})\big]+o_1(\lambda^2)\cong o(\lambda^2).\end{aligned} \tag{3.64}$$

Hence (3.63) and (3.64) yield

$$\Delta J(u) = -\mu_{k_*}\lambda\int_{T_{k_*}(t_*)}\Delta_{k_*}(t)\big(-1-u_{k_*}(t)\big)dt + o(\lambda) > 0 \tag{3.65}$$

for a sufficiently small $\lambda > 0$, which contradicts the optimality of control functions $u_k^0(t),\ k = 1,..,N$.

■

The optimality conditions for the supporting control functions can also be expressed in the maximum principle form. Let $\psi_N(t)$ be the solution of the following differential equations

$$\frac{d\psi_N(t)}{dt} = -A^T\psi_N(t),\quad \psi_N(t^*) = p_N - H_N^T\nu^N,\quad t\in T. \tag{3.66}$$

which can be represented as

$$\psi_N(t) = K_1^T(t^*-t)\psi(t^*),\quad t\in T. \tag{3.67}$$

Hence, the following equalities

$$\begin{aligned}\psi_N^T(t)b &= \big(p_N^T-(\nu^N)^TH_N\big)K_1(t^*-t)b = p_N^TK_1(t^*-t)b \\ &- (\nu^N)^TH_NK_1(t^*-t)b = c_N(t)-(\nu^N)^Tg_{NN}(t) = -\Delta_N(t),\end{aligned} \tag{3.68}$$

hold. In order to verify the validity of the corresponding conditions for subsequent passes we apply (3.50) for the differential equations (3.48). Let $\psi_{N-1}(t),\ t \in T$ be a solution of the differential equation

$$\frac{d\psi_{N-1}(t)}{dt} = -A^T\psi_{N-1}(t) - D^T\psi_N(t), \quad \psi_{N-1}(t^*) = p_{N-1} - H_{N-1}^T\nu^{N-1},\ t \in T. \quad (3.69)$$

Then $\psi_k^T(t)b =$

$$= (p_{N-1}^T - (\nu^{N-1})^T H_{N-1})K_1(t^* - t)b - (p_N^T - (\nu^N)^T H_N)\int_0^t K_1^T(t-\tau)D^T K_1^T(t^*-\tau)bd\tau$$

$$= p_{N-1}^T K_1(t^*-t)b - (\nu^{N-1})^T H_{N-1}K_1(t^*-t)b - (p_N^T - (\nu^N)^T H_N)K_2(t^*-t)b =$$

$$= c_{N-1}(t) - (\nu^{N-1})^T g_{N-1N-1}(t) - (\nu^N)^T g_{NN-1}(t) = -\Delta_{N-1}(t). \quad (3.70)$$

By analogy with the case considered in (3.69)–(3.70), we have

$$\psi_k^T(t)b = -\Delta_k(t), \quad k = 2, \ldots, N, \quad (3.71)$$

where $\psi_k(t), \quad t \in T$ are the solutions of the following differential equations

$$\frac{d\psi_k(t)}{dt} = -A^T\psi_k(t) - D^T\psi_{k+1}(t), \quad \psi_k(t^*) = p_k - H_k^T\nu^k, \quad t \in T. \quad (3.72)$$

For each $k = 1, \ldots, N$ introduce the associated Hamilton function as

$$H_k(x_{k-1}, x_k, \psi_k, u_k) = \psi_k^T\left(Ax_k + Dx_{k-1} + bu_k\right), \quad t \in T. \quad (3.73)$$

Then use (3.71) yields that the optimality conditions (3.56) can be re-formulated in the maximum principle form as follows

Corollary 1. *The admissible supporting control* $\{\tau_{sup}^k,\ u_k^0(t),\ k = 1, \ldots, N\}$ *is optimal if along the corresponding trajectories* $x_k^0(t),\ \psi_k(t)$ *of (3.42)–(3.43) and (3.72) the Hamiltonian function takes the maximum value, i. e.*

$$H_k(x_{k-1}^0(t), x_k^0(t), \psi_k, u_k^0(t)) = \max_{|v| \le 1} H_k(x_{k-1}^0(t), x_k^0(t), \psi_k, v), \quad t \in T \quad (3.74)$$

for $k = 1, ..., N$*. If the admissible supporting control is non-degenerate then this condition is necessary and sufficient.*

In the next section, the maximum principle for arbitrary admissible control functions of the form of (3.42)–(3.46) is established using the sub-optimality conditions.

3.3.2 ϵ- optimality conditions.

Usually, in the design of numerical implementation of optimal control algorithms we exploit approximate solutions with corresponding error estimation. Hence it is necessary to introduce the 'sub-optimality' concept, as it is often sufficient to stop the numerical computations when a satisfactory accuracy level has been achieved.

Assume that $\{u_k^0(t),\ k \in K\}$ is the optimal control for (3.42)–(3.46), and let $J(u^0)$ denote the corresponding optimal cost function value.

Definition 15. *We say that the admissible control function* $\{u_k^\epsilon(t),\ k \in K\}$ *is* ϵ*- optimal, if the corresponding solution* $\{x_k^\epsilon(t),\ t \in T,\ k \in K\}$ *of (3.42)–(3.44) satisfies* $J(u^0) - J(u^\epsilon) \leq \epsilon$.

Now we proceed to calculate an estimate of a supporting control function $\{u_k^0,\ \tau_{sup}^k,\ k \in K,\ t \in T\}$, i.e. a measure of non-optimality of the control. Note also that this estimate can be partitioned into two principal parts: one of which evaluates the degree of non-optimality of the chosen admissible control functions $u_k(t)$, and the second the error produced by non-optimality of the support τ_{sup}^k. This partition is a major advantage in the design of numerically applicable solution algorithms.

Introduce an estimate of optimality $\beta = \beta(\tau_{sup}, u)$ as the value of the maximum increment for the cost function of (3.42)–(3.46) calculated in the absence of the the principal constraints (3.44), i.e. this estimate is given by the solution of the following relaxed optimization problem

$$\Delta J(u) \to \max_{\Delta u_k}, \quad |u_k(t) + \Delta_k u(t)| \leq 1,\ t \in T,\ k = 1, \dots, N. \tag{3.75}$$

It is easy to see that

$$\beta = \beta(\tau_{sup}, u) = \sum_{k=1}^{N} \int_{T_k^+} \Delta_k(t)(u_k(t) + 1)dt + \sum_{k=1}^{N} \int_{T_k^-} \Delta_k(t)(u_k(t) - 1)dt, \tag{3.76}$$

where

$$T_k^+ = \big\{t \in T:\ \Delta_k(t) > 0\big\}, \quad T_k^+ = \big\{t \in T:\ \Delta_k(t) > 0\big\}.$$

and we have the following result.

Theorem 15. *(ϵ-maximum principle) Given any* $\epsilon \geq 0$, *the admissible control* $\{u_k(t),\ t \in T,\ k \in K\}$ *is* ϵ*-optimal for (3.42)–(3.46) if, and only if,* $\exists$ *the support* $\{\tau_{sup}^k,\ k \in K\}$ *such that along the solutions* $x_k(t), \psi_k(t), t \in T,\ k \in K$ *of (3.42)–(3.44) and (3.72) the Hamiltonian attains its* ϵ*- maximum value, i.e.*

$$H_k(x_{k-1}^0(t), x_k^0(t), \psi_k, u_k^0(t)) = \max_{|v| \leq 1} H_k(x_{k-1}^0(t), x_k^0(t), \psi_k, v) - \epsilon_k(t), \quad t \in T, \tag{3.77}$$

where the functions $\epsilon_k(t),\ k \in K$ *satisfy the following inequality*

$$\sum_{k \in K} \int_T \epsilon_k(t)dt \leq \epsilon. \tag{3.78}$$

Proof. *Sufficiency.* Assume that (3.77)–(3.78) hold for an admissible control

$$\{u_k(t), t \in T, k \in K\}.$$

Then by (3.71) the suboptimality estimate can be calculated as $\beta = \beta(\tau_{sup}, u)$:

$$\beta(\tau_{sup}, u) = \sum_{k=1}^{N} \int_{T_k^+} \psi_k^T(t) b\big(-u_k(t) - 1\big) dt + \sum_{k=1}^{N} \int_{T_k^-} \psi_k^T(t) b\big(1 - u_k(t)\big) dt$$

$$= \sum_{k=1}^{N} \int_{T_k^+} \psi_k^T(t)\big(Ax_k(t) + Dx_{k-1}(t) - b\big) dt - \sum_{k=1}^{N} \int_{T_k^+} \psi_k^T(t)\big(Ax_k(t) + Dx_{k-1}(t) + bu_k(t)\big) dt$$

$$= \sum_{k=1}^{N} \int_{T_k^-} \psi_k^T(t)\big(Ax_k(t) + Dx_{k-1}(t) + b\big) dt - \sum_{k=1}^{N} \int_{T_k^-} \psi_k^T(t)\big(Ax_k(t) + Dx_{k-1}(t) - bu_k(t)\big) dt$$

$$+ \sum_{k=1}^{N} \int_{T} \left[\max_{|v| \leq 1} H_k\big(x_{k-1}(t), x_k(t), \psi_k(t), v\big) - H_k\big(x_{k-1}(t), x_k(t), \psi_k(t), u_k(t)\big) \right] dt =$$

$$= \sum_{k=1}^{N} \int_{T} \epsilon_k(t) dt \leq \epsilon.$$

Since the sub-optimal estimate (3.75) has been calculated in the absence of constraints (3.44), then it is obvious that the following inequalities hold

$$J(u^0) - J(u) \leq \beta(\tau_{sup}, u) \leq \epsilon.$$

This proves the ϵ- optimality property of the admissible control $\{u_k(t),\ t \in T,\ k \in K\}$.

Necessity. Let $\{u_k(t),\ t \in T,\ k \in K\}$ be an ϵ-optimal admissible control and let $\{\tau_{sup}^k,\ k \in K\}$ be an arbitrary support. Then the sub-optimal estimate of the control corresponding to the chosen support can be calculated as

$$\beta(\tau_{sup}, u) = \sum_{k=1}^{N} \int_{T} \Delta_k(t) u_k(t) dt + \sum_{k=1}^{N} \int_{T_k^+} \Delta_k(t) dt - \sum_{k=1}^{N} \int_{T_k^-} \Delta_k(t) dt. \tag{3.79}$$

Now introduce the following dual optimization problem

$$I(y, v, w) = \sum_{k \in K} \left[h_k^T y_k + \int_{T} v_k(t) dt + \int_{T} w_k(t) dt \right] \longrightarrow \min_{y,v,w} \tag{3.80}$$

subject to

$$\sum_{s=k}^{N} y_s^T g_{sk}(t) - v_k(t) + w_k(t) = c_k(t), \quad v_k(t) \geq 0,\ w_k(t) \geq 0,\ t \in T,\ k \in K. \tag{3.81}$$

Then it can be shown that (3.80)–(3.81) has an optimal solution if there exists an optimal control for (3.42)–(3.46). Denote the chosen support by $\{\tau_{sup}^k,\ k \in K\}$ and then use (3.53) to construct the vectors $z_k = \{y_k, v_k, w_k,\ k \in K\}$ as

$$y_k = \nu_k;\ v_k(t) = \Delta_k(t),\ w_k(t) = 0 \qquad if \quad \Delta_k(t) \geq 0,$$

$$v_k(t) = 0,\ w_k(t) = \Delta_k(t) \qquad if \quad \Delta_k(t) < 0, \tag{3.82}$$

where, by (3.56), these satisfy the constraint (3.81) of the dual problem.

Let $\{y_k^0, v_k^0(t), w_k^0(t), t \in T, k \in K\}$ denote an optimal solution of (3.80)–(3.81). Then (3.80) and (3.56) yield

$$\begin{aligned}
\beta(\tau_{sup}, u) &= \sum_{k=1}^{N}\sum_{s=k}^{N}\int_T \nu_s^T(t) g_{sk}(t) u_k(t)dt - \sum_{k=1}^{N}\int_T c_k^T(t)u_k(t)dt \\
&+ \sum_{k=1}^{N}\int_T v_k(t)dt - \sum_{k=1}^{N}\int_T w_k(t)dt \\
&= \left[\sum_{k=1}^{N}(\nu^k)^T \sum_{s=1}^{k}\int_T g_{ks}(t)u_s(t)dt + \sum_{k=1}^{N}\int_T v_k(t)dt - \sum_{k=1}^{N}\int_T w_k(t)dt\right] \\
&- \left[\sum_{k=1}^{N}\sum_{s=k}^{N}\int_T (y_s^0)^T g_{sk}(t)u_k^0(t)dt + \sum_{k=1}^{N}\int_T v_k^0(t)dt - \sum_{k=1}^{N}\int_T w_k^0(t)dt\right] \\
&+ \sum_{k=1}^{N}\int_T c_k(t)u_k^0(t)dt - \sum_{k=1}^{N}\int_T c_k(t)u_k(t)dt \\
&= \left[\sum_{k=1}^{N}(\nu^k)^T h_k + \sum_{k=1}^{N}\int_T (v_k(t) - w_k(t))dt\right] - \\
&- \left[\sum_{k=1}^{N}(y_k^0)^T h_k + \sum_{k=1}^{N}\int_T (v_k^0(t) - w_k^0(t))dt\right] \\
&+ \sum_{k=1}^{N}\int_T c_k(t)u_k^0(t)dt - \sum_{k=1}^{N}\int_T c_k(t)u_k(t)dt.
\end{aligned}$$

Finally, the sub-optimal estimate can be written in the form

$$\beta = \beta(\tau_{sup}, u) = \beta_{sup} + \beta_u, \tag{3.83}$$

where

$$\beta_{sup} = \sum_{k=1}^{N} h_k^T(\nu_k - y^{0k}) + \sum_{k=1}^{N}\int_T \Big[(v_k(t) - v_k^0(t)) - (w_k(t) - w_k^0(t))\Big]dt \tag{3.84}$$

denotes the non-optimality measure of the chosen support $\{\tau_{sup}^k,\ k \in K\}$, and

$$\beta_u = \sum_{k=1}^{N}\int_T c_k(t)\big(u_k(t) - u_k^0(t)\big)dt \tag{3.85}$$

denotes the non-optimality measure of the given control function $\{u_k(t),\ t \in T,\ k \in K\}$.

Now choose the support $\tau^0_{sup} = \{\tilde{\tau}^k_{sup},\ k \in K\}$ such that the corresponding collection $z^0_k = \{y^0_k, v^0_k, w^0_k, k \in K\}$ of dual variables is an optimal solution of (3.80)–(3.81). First, we show that the chosen support $\tau^0_{sup} = \{\tilde{\tau}^k_{sup}(\epsilon),\ k \in K\}$ is the one required for the given ϵ- optimal control functions $\{u_k(t),\ k \in K\}$. In particular, since $\beta_{sup} = 0$ then $\beta = \beta(u, \tau^0_{sup}) = \beta_u \leq \epsilon$. Next set

$$\begin{aligned}
\epsilon_k(t) &= \Delta_k(t)(u_k(t)+1), \quad t \in T^+_k, \\
\epsilon_k(t) &= \Delta_k(t)(u_k(t)-1), \quad t \in T^-_k, \\
\epsilon_k(t) &= 0 \qquad if \quad \Delta_k(t) = 0,\ t \in T.
\end{aligned}$$

and note from the definition of $\Delta_k(t)$ that we have

$$\begin{aligned}
\epsilon_k(t) &= -\psi^T_k(t)b(u_k(t)+1) = \psi^T_k(t)(Ax_k(t) + Dx_{k-1}(t) + b(-1)) \\
&- \psi^T_k(t)(Ax_k(t) + Dx_{k-1}(t) + bu_k(t)) \quad \text{if } \psi_k(t)b < 0; \\
\epsilon_k(t) &= \psi^T_k(t)(Ax_k(t) + Dx_{k-1}(t) + b(+1)) \\
&- \psi^T_k(t)(Ax_k(t) + Dx_{k-1}(t) + bu_k(t)) \quad \text{if } \psi_k(t)b > 0; \\
\epsilon_k(t) &= 0 \qquad \text{if} \qquad \psi_k(t)b = 0, \quad t \in T, \quad k \in K.
\end{aligned}$$

Use of the Hamiltonian (3.73) now enables these last expressions to be written in the form

$$\epsilon_k(t) = \max_{|v|\leq 1} H_k\big(x^0_{k-1}(t), x^0_k(t), \psi_k, v\big) - H_k\big(x^0_{k-1}(t), x^0_k(t), \psi_k, u^0_k(t)\big), \quad t \in T,\ k \in K.$$

Adding these last expressions and noting that $\{u_k(t)\}$ is an suboptimal control, yields

$$\begin{aligned}
\sum_{k=1}^{N} \int_T \epsilon_k(t)dt &= \sum_{k=1}^{N} \int_{T^+_k} \Delta_k(t)(u_k(t)+1)dt \\
&+ \sum_{k=1}^{N} \int_{T^-_k} \Delta_k(t)(u_k(t)-1)dt = \beta(u, \tau^0_{sup}) = \beta_u \leq \epsilon.
\end{aligned}$$

which completes the proof.

■

Note now that that maximum principle follows from the theorem above on setting $\epsilon = 0$.

Corollary 2. *The admissible control* $\{u^0_k(t), k \in K,\ t \in T\}$ *is optimal if, and only if, there exists a support* $\{\tau^{0k}_{sup},\ k \in K\}$ *such that the supporting control* $\{u^0_k(t), \tau^{0k}_{sup},\ t \in T,\ k \in K\}$ *satisfies the maximum conditions*

$$\max_{|v|\leq 1} H_k\big(x^0_{k-1}(t), x^0_k(t), \psi_k, v\big) = H_k\big(x^0_{k-1}(t), x^0_k(t), \psi_k, u^0_k(t)\big)$$

for all $k \in K,\ t \in T,$ *where* $\psi_k(t)$ *are the corresponding solutions of (3.72).*

3.3.3 Differential properties of optimal solutions

An important aspect of the optimization theory is sensitivity analysis of optimal control problems. Since, in practice, control problems are often subject to disturbances or perturbations of the system data. In mathematical terms, perturbations can be described by some parameters in the initial data, boundary conditions, control and state constraints and it is clearly important to know how a problem solution depends on these parameters. The aim in this sub-section is to characterize the changes in the solutions developed here due to 'small' perturbations in the parameters, which should enable us to design a fast and reliable real-time algorithms to correct the solutions for these effects. The major advantage of the proposed constructive approach is that the sensitivity analysis and some differential properties of the optimal controls under disturbances can be studied. There are very critical when they are to be applied to control synthesis problems.

Suppose that disturbances influence the initial data for (3.42)–(3.43). In particular, consider the parametric system (3.42)–(3.44) on the interval $T_s = [s, t^*]$ with the initial data $x_k(s) = z_k, k \in K$ where any state parameter z_k belongs to some neighborhood $G_k \subset \mathbb{R}^n$ of the point $x_k = \alpha_k$ and the initial time parameter s belongs to the neighborhood G_0 of the instant $t = 0$. We also assume that the following regularity condition holds: for the given disturbance domain $G_k, k \in K \cup \{0\}$, the structure of the optimal control functions for the non-disturbed data is preserved, i. e. the number of switching instances together with their order are constant.

Using Theorem 14, the optimal controls $\{u_k^0(t, s, z), k \in K\}$ are determined by the supporting time instances $\tau_{kj} = \tau_{kj}(s, z),\ k \in K,\ j = 1, \ldots, m$ which are dependent on the disturbances $(s, z_k),\ s \in G_0,\ z_k \in G_k,\ k \in K$. The aim of this section is to study the differential properties of the functions $\tau_{kj} = \tau_{kj}(s, z),\ k \in K,\ j = 1, \ldots, m$. For ease of notation we set $\tau \equiv \tau(s, z) = \big\{\tau_{kj}(s, z),\quad k \in K,\ j = 1, \ldots, m\big\}, z = \{z_k, k \in K\}$ in what follows.

Theorem 16. *If (3.42)–(3.44) is regular then for any $k \in K$ and $j = 1, ..., m$, the functions $\tau_{kj} = \tau_{kj}(s, z)$ are differentiable in the domain $G_0 \times G_k \subset \mathbb{R} \times \mathbb{R}^n$.*

Proof. For each fixed parameters $(s, z_k),\ s \in G_0,\ z_k \in G_k,\ k \in K$ we have the optimization problem of the type (3.42)–(3.46) where the initial data in (3.43) are

$$x_k(0) = z_k, \quad k \in K, \quad x_0(t) = f(t), \quad t \in T_s \triangleq [s, t*].$$

Using (3.51)–(3.52) and Theorem 14 it follows immediately that the switching instances $\tau_{kj} = \tau_{kj}(s, z),\ k \in K,\ j = 1, \ldots, m$ of the optimal bang-bang control $\{u_k^0(t, s, z),\ k \in K\}$ for the disturbed problem (3.42)–(3.46), are the solutions of the following optimization problem

$$\max_{\tau_{kj}} \sum_{k \in K} R_k(s, z) \sum_{j=1}^{m+1} (-1)^j \int\limits_{\tau_{kj-1}}^{\tau_{kj}} c_k(t) dt \tag{3.86}$$

subject to

$$\sum_{l \in K} R_l(s, z) \sum_{j=1}^{m+1} (-1)^j \int\limits_{\tau_{lj-1}}^{\tau_{lj}} g_{kl}(t) dt = h_k(s, z), \quad k \in K \tag{3.87}$$

Here the constant $R_k(s,z) = \pm 1$ denotes the value ($u = +1$ or $u = -1$) of the optimal control on pass k over the first control interval $t \in [s, \tau_{k1}]$, and

$$h_k(s,z) = g_k - \sum_{j=1}^{k} H_k K_j(t^*) z_{k+1-j} - \int_{s}^{t^*} H_k K_k(t^* - t) D f(t) dt. \tag{3.88}$$

It is obvious that the switching instances $\tau_{kj} = \tau_{kj}(s,z)$ satisfy the following inequalities

$$\tau_{k0} < \tau_{k1} < \tau_{k2} < \cdots < \tau_{km} < \tau_{km+1}, \quad \tau_{k0} = s, \quad \tau_{km+1} = t^*,$$

Since $\{u_k^0, \tau_{sup}^0, k \in K\}$ is optimal supporting control for the non-disturbed problem (3.42)–(3.44) then the optimization problem (3.86)–(3.87) has the optimal solution $\tau_{kj}^0, k \in K, j = 1, \ldots, m\}$ at $s = 0, z_k = \alpha_k, k \in k, j = 1, \ldots, m$. Hence there exist Lagrange multipliers $\lambda_k^0 \in \mathbb{R}^m, k \in K$ which are not simultaneously equal to zero and such that the collection $\{\lambda_k^0, \tau_{kj}^0\}$ is a stationary point for the following Lagrange function associated with the optimization problem (3.86)–(3.87)

$$\begin{aligned} L(\lambda, \tau_{sup}) &= \sum_{k \in K} R_k(s,z) \sum_{j=1}^{m+1} (-1)^j \int_{\tau_{kj-1}}^{\tau_{kj}} c_k(t) dt \\ &+ \sum_{k \in K} \lambda_k \Big[\sum_{l \in K} R_l(s,z) \sum_{j=1}^{m+1} (-1)^j \int_{\tau_{lj-1}}^{\tau_{lj}} g_{kl}(t) dt - h_k(s,z) \Big]. \end{aligned} \tag{3.89}$$

The well known stationary conditions for the Lagrange function L lead to the following equalities

$$2R_k(s,z) \Big[c_k(\tau_{kj}) + \sum_{l=k}^{N} \lambda_l g_{lk}(\tau_{kj}) \Big] = 0, \quad j = 1, \ldots, m, \quad k \in K \tag{3.90}$$

$$\sum_{l=1}^{k} R_l(s,z) \sum_{j=1}^{m+1} (-1)^j \int_{\tau_{lj-1}}^{\tau_{lj}} g_{kl}(t) dt - h_k(s,z) = 0, \quad k \in K \tag{3.91}$$

with respect to the unknown λ_k and $\tau_k(s,z),\ k \in K,\ j = 1, \ldots, m$. The Jacobian matrix D of the mapping (3.90) with respect to variables (λ, τ_{sup}) calculated at $s = 0$ and $z_k = \alpha_k$ can be written in the form

$$D = \prod_{k \in K} 2R_k(0, \alpha) \begin{pmatrix} \hat{G}_{sup} & F \\ 0 & \hat{G}_{sup} \end{pmatrix} \tag{3.92}$$

where the matrix $\hat{G}_{sup}$ is defined as follows

$$\hat{G}_{sup} = \begin{pmatrix} g_{kj}(t), & t \in \tau_{sup}^k \\ j \geq k, & k = 1, \ldots, N \end{pmatrix} \tag{3.93}$$

and the matrix F is formed from the derivatives of the functions $c_k(t), g_{kl}(t)$ taken at the corresponding points. By the definition of the supporting time instances we have that

$\det D \neq 0$ and by implicit function theorem there exists a neighborhood of the point $(0, \alpha_k,\ k \in K)$ where (3.90) has a unique solution $\lambda = \lambda(s,z)$, $\tau_{kj} = \tau_{kj}(s,z)$ where these functions are also differentiable. This completes the proof.

■

The above differential properties of the optimal controls can be used for sensitivity analysis and the solution of the synthesis problem for the repetitive processes considered here. In particular, the supporting control approach can be applied [39] to produce the differential equations for the switching time functions $\tau(s,z)$ necessary to design the optimal controllers. It can be shown that they satisfy the following differential equations

$$G\frac{\partial \tau}{\partial s} + Q = \frac{\partial h}{\partial s}, \quad P\frac{\partial \tau}{\partial z} = \frac{\partial h}{\partial z} \tag{3.94}$$

where $h(s,z) = (h_1(s,z), \ldots, h_m(s,z))$ is an $mN \times 1$-vector given by (3.88) and the matrices G, Q, P are defined by those defining the process dynamics and information associated with the non-disturbed optimal solution. For example

$$G = -\Big(g_{11}(s)sgn\dot{\Delta}_1(\tau_{11}), g_{21}(s)sgn\dot{\Delta}_1(\tau_{11}) + g_{22}(s)sgn\dot{\Delta}_2(\tau_{21}), \ldots, \sum_{j=1}^{N} g_{Nj}(s)sgn\dot{\Delta}_j(\tau_{j1})\Big)^T$$

where the functions $\Delta_j(t)$, $j = 1, \ldots, N$ are designed using the switching moments of the basic optimal control function. Note, that the analogous differential equations can be established for the optimal values of the cost function treated as the function $J(s,z) \equiv J\big(u^0(\tau(s,z)\big)$.

Remark 17. *The equations (3.94) are (sometimes) termed Pfaff differential equations and model an essentially distinct class of continuous nD systems. The main characteristic feature of this model is that it is overdetermined (in the sense that the number of equations exceeds the unknown functions). It can also be shown that the non-degenerate assumption on the supporting control functions leads to the validity of the so-called Frobenious conditions that guarantee the existence and uniqueness of solutions of the Pfaff differential equations [35].*

3.3.4 Example

In order to demonstrate the advantages of the supporting control function approach, we now give the following example.

Consider the following optimal control problem for the repetitive process with $N = 2$ passes, where the superscript $(\cdot)$ is used to denote a particular entry in the state vector $x_k(t) = \big(x_k^{(1)}, x_k^{(2)}(t)\big)$ on the pass k:

$$\max_{u_1,u_2} J(u),\ J(u) := x_1^{(2)}(1) + x_2^{(2)}(1) \tag{3.95}$$

for the system

$$\begin{aligned} \frac{dx_1^{(1)}(t)}{dt} &= x_1^{(2)}(t), \qquad \frac{dx_2^{(1)}(t)}{dt} = x_2^{(2)}(t),\ t \in [s,1] \\ \frac{dx_1^{(2)}(t)}{dt} &= u_1(t), \qquad \frac{dx_2^{(2)}(t)}{dt} = x_1^{(1)}(t) + u_2(t), \end{aligned} \tag{3.96}$$

with boundary conditions of the form

$$x_1^{(1)}(s) = \xi_1^{(1)},\ x_1^{(2)}(s) = \xi_1^{(2)},\ x_2^{(1)}(s) = \xi_2^{(1)},\ x_2^{(2)}(s) = \xi_2^{(2)} \tag{3.97}$$

subject to

$$x_1^{(1)}(1) = 1/8, \quad x_2^{(1)}(1) = 1/384,\ |u_1(t)| \le 1,\ |u_2(t)| \le 1, \tag{3.98}$$

We assume that the parameters $s, \xi_1^{(1)},\ \xi_1^{(2)},\ \xi_2^{(1)},\ \xi_2^{(2)}$ satisfy the regularity conditions formulated in the section above.

The dynamic here can be written as a stationary differential linear repetitive process of the form

$$\begin{bmatrix} \dot{x}_{k+1}^{(1)}(t) \\ \dot{x}_{k+1}^{(2)}(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_{k+1}^{(1)}(t) \\ x_{k+1}^{(2)}(t) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_k^{(1)}(t) \\ x_k^{(2)}(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_{k+1}(t),\ k = 0, 1. \tag{3.99}$$

Without loss of generality we set $x_0(t) = 0,\ t \in [s, 1]$.

To apply the results developed here to this example we first rewrite (3.95)–(3.98) in the following integral form:

$$\max_{u_1,u_2} \left\{ \xi_2^{(1)} + \xi_2^{(2)} + (1-s)\xi_1^{(1)} + \frac{(1-s)^2}{2}\xi_2^{(2)} + \int_s^1 \frac{(1-t)^2+2}{2} u_1(t)dt + \int_s^1 u_2(t)dt \right\} \tag{3.100}$$

subject to

$$\begin{aligned} &\int_s^1 (1-t)u_1(t)dt = \frac{1}{8} - \xi_1^{(1)} + (1-s)\xi_1^{(2)}, \\ &\int_s^1 \left[\frac{(1-t)^3}{6} u_1(t) + (1-t)u_2(t) \right] dt = \\ &= \frac{1}{384} - \xi_2^{(1)} - (1-s)\xi_2^{(2)} - \frac{(1-s)^2}{2}\xi_1^{(1)} - \frac{(1-s)^3}{6}\xi_1^{(2)}. \end{aligned} \tag{3.101}$$

Hence

$$g_{11}(t) \;=\; 1-t, \quad g_{21}(t) = \frac{(1-t)^3}{6}, \quad g_{22}(t) = 1-t, \tag{3.102}$$

$$c_1(t) \;=\; \frac{(1-t)^2+2}{2}, \quad c_2(t) = 1 \tag{3.103}$$

and the multipliers required to design the co-control function $\Delta_i(t),\ i = 1, 2$ can, noting (3.53), be written as

$$\begin{aligned} \nu^{(2)} g_{22}(\tau_{2sup}) - c_2(\tau_{2sup}) = 0, \\ \nu^{(1)} g_{11}(\tau_{1sup}) + \nu^{(2)} g_{21}(\tau_{1sup}) - c_1(\tau_{1sup}) = 0 \end{aligned} \tag{3.104}$$

Then

$$\Delta_1(t) =(1-t)\left[\frac{1}{1-\tau_{1sup}}+\frac{1-\tau_{1sup}}{2}-\frac{(1-\tau_{1sup})^2}{6(1-\tau_{2sup})}\right]+\frac{(1-t)^3}{6(1-\tau_{2sup})}-\frac{(1-t)^2}{2}-1,$$

$$\Delta_2(t) =\frac{1-t}{1-\tau_{2sup}}-1 \tag{3.105}$$

Now the problem is to find the basic optimal trajectory when all variables in (3.97) are zero, i.e.

$$s=0,\ x_1^{(1)}(0)=0,\ x_1^{(2)}(0)=0,\ x_2^{(1)}(0)=0,\ x_2^{(2)}(0)=0. \tag{3.106}$$

Take the supporting instances as

$$\tau_{1sup}=1-\sqrt{\frac{5}{8}},\quad \tau_{2sup}=1-\sqrt{\frac{131}{256}}. \tag{3.107}$$

Then it follows immediately from Theorem 14 that the optimal control functions for (3.95)–(3.98) with the initial data (3.106) are given by

$$u_1^0(t)=\begin{cases}-1, & 0\le t<1-\sqrt{\frac{5}{8}},\\ +1, & 1-\sqrt{\frac{5}{8}}\le t\le 1\end{cases},\quad u_2^0(t)=\begin{cases}-1, & 0\le t<1-\sqrt{\frac{131}{256}},\\ +1, & 1-\sqrt{\frac{131}{256}}\le t\le 1\end{cases} \tag{3.108}$$

and (3.94) gives the switching functions $\tau_1\equiv\tau_1(\xi_1^{(1)},\xi_1^{(2)},s)$, $\tau_2\equiv\tau_2(\xi_1^{(1)},\xi_2^{(1)},\xi_1^{(2)},\xi_2^{(2)},s)$ have to satisfy the following differential equations

$$-2\frac{\partial\tau_2}{\partial s}(1-\tau_2)-\frac{2(1-\tau_1)^3}{6}\frac{\partial\tau_1}{\partial s}=\frac{(1-s)^2}{2}\xi_1^{(2)}+(1-s)\xi_1^{(1)}+\xi_2^{(2)}-\frac{(1-s)^3}{6}-(1-s),$$

$$-2\frac{\partial\tau_2}{\partial\xi_1^{(1)}}(1-\tau_2)-\frac{(1-\tau_1)^3}{3}\frac{\partial\tau_1}{\partial\xi_1^{(1)}}=-\frac{(1-s)^2}{2},$$

$$-2\frac{\partial\tau_2}{\partial\xi_1^{(2)}}(1-\tau_2)-\frac{(1-\tau_1)^3}{3}\frac{\partial\tau_1}{\partial\xi_1^{(2)}}=-\frac{(1-s)^3}{6},$$

$$-2\frac{\partial\tau_2}{\partial\xi_2^{(1)}}(1-\tau_2)=-1,\quad -2\frac{\partial\tau_2}{\partial\xi_2^{(2)}}(1-\tau_2)=-(1-s), \tag{3.109}$$

with initial conditions

$$\tau_1(0,0,0)=1-\sqrt{\frac{5}{8}},\qquad \tau_2(0,0,0,0,0)=1-\sqrt{\frac{131}{16^2}} \tag{3.110}$$

The solutions of this differential system are

$$\tau_1(z_1^{(1)},\xi_1^{(2)},s)=1-\sqrt{SR_1(\xi_1^{(1)},\xi_1^{(2)},s)}$$

$$\tau^{(2)}(\xi_1^{(1)},\xi_2^{(1)},\xi_1^{(2)},\xi_2^{(2)},s)=1-\sqrt{SR_2(\xi_1^{(1)},\xi_2^{(1)},\xi_1^{(2)},\xi_2^{(2)},s)} \tag{3.111}$$

where

$$SR_1(\xi_1^{(1)}, \xi_1^{(2)}, s) = \frac{5}{8} + (s-1)\xi_1^{(2)} - \xi_1^{(1)} - s + s^2/2,$$

$$SR_2(\xi_1^{(1)}, \xi_2^{(1)}, \xi_1^{(2)}, \xi_2^{(2)}, s) = \frac{131}{256} + \frac{2s^4 - 8s^3 + 59s^2 - 102s}{96} + \tag{3.112}$$

$$\frac{-20s^2 + 40s - 19}{48}\xi_1^{(1)} - \frac{1}{12}\xi_1^{(1)2} + \frac{4s^3 - 12s^2 + 11s - 3}{48}\xi_1^{(2)} + \frac{-s^2 + 2s - 1}{12}\xi_1^{(2)2} +$$

$$+\frac{s\xi_1^{(1)}\xi_1^{(2)}}{6} - \frac{\xi_1^{(1)}\xi_1^{(2)}}{6} - \xi_1^{(2)} + (s-1)\xi_2^{(2)}$$

It easy to see that the solution of the differential equations describing the process dynamics with both u_1 and u_2 constant are

$$\begin{aligned} x_1^{(1)}(t) =& u_1\frac{t^2}{2} + tC_1 + C_2, \\ x_1^{(2)}(t) =& u_1 t + C_1 \\ x_2^{(1)}(t) =& u_1\frac{t^4}{24} + C_1\frac{t^3}{6} + C_2\frac{t^2}{2} + u_2\frac{t^2}{2} + tC_3 + C_4, \\ x_2^{(2)}(t) =& u_1\frac{t^3}{6} + C_1\frac{t^2}{2} + tC_2 + tu_2 + C_3. \end{aligned} \tag{3.113}$$

where the coefficients C_i depend on the parameters $\xi_1^{(1)}, \xi_2^{(1)}, \xi_1^{(2)}, \xi_2^{(2)}, s$.

The formulas (3.113) can be used to have for synthesis optimal control schemes. At the simplest case, we consider now such a disturbance set Ω that the optimal control structure is preserved for the case of zero initial conditions, i.e. $u_1 = -1$ for $t \le \tau_1(\xi_1^{(1)}, \xi_1^{(2)}, s)$, $u_2^0 = -1$, for $t \le \tau_2(\xi_1^{(1)}, \xi_2^{(1)}, \xi_1^{(2)}, \xi_2^{(2)}, s)$ and the inequality

$$\tau_1(\xi_1^{(1)}, \xi_1^{(2)}, s) < \tau_2(\xi_1^{(1)}, \xi_2^{(1)}, \xi_1^{(2)}, \xi_2^{(2)}, s)$$

holds. Using (3.111) we have in this case that the domain Ω is described by

$$0 \le \tau_1(\xi_1^{(1)}, \xi_1^{(2)}, s) < \tau_2(\xi_1^{(1)}, \xi_2^{(1)}, \xi_1^{(2)}, \xi_2^{(2)}, s) \le 1$$

$$SR_1(\xi_1^{(1)}, \xi_1^{(2)}, s) \ge 0, \qquad SR_2(\xi_1^{(1)}, \xi_2^{(1)}, \xi_1^{(2)}, \xi_2^{(2)}, s) \ge 0$$

Drawing the graphic imaging of the disturbance domain Ω and the switching manifolds for optimal synthesis in this case would be an extremely important for highlighting results already presented. However, making this for the both passes $k = 1$ and $k = 2$ is very difficult problem and still ia not available. However some important insights can be also obtained when having the appropriate imaging for the single pass dynamics. Hence, in remainder of this chapter we present the result for the single pass case i.e. $N = 1$, where to simplify notations the superscript $(\cdot)$ that is used above to denote a particular element in the state vector is omitted.

Synthesis of the optimal control can be realized using the switching instance function $\tau = \tau(\xi_1, \xi_2, s)$, which in accordance with (3.111) has the following form

$$\tau(\xi_1, \xi_2, s) = 1 - \sqrt{5/8 + (s-1)\xi_2 - \xi_1 - s + s^2/2}$$

Without loss of generality, assume $s = 0$ and then the optimal switching function is

$$\tau(\xi_1, \xi_2, 0) = 1 - \sqrt{5/8 - \xi_1 - \xi_2}.$$

Figures 3.2 and 3.3 illustrate the form of the optimal synthesis solution below.

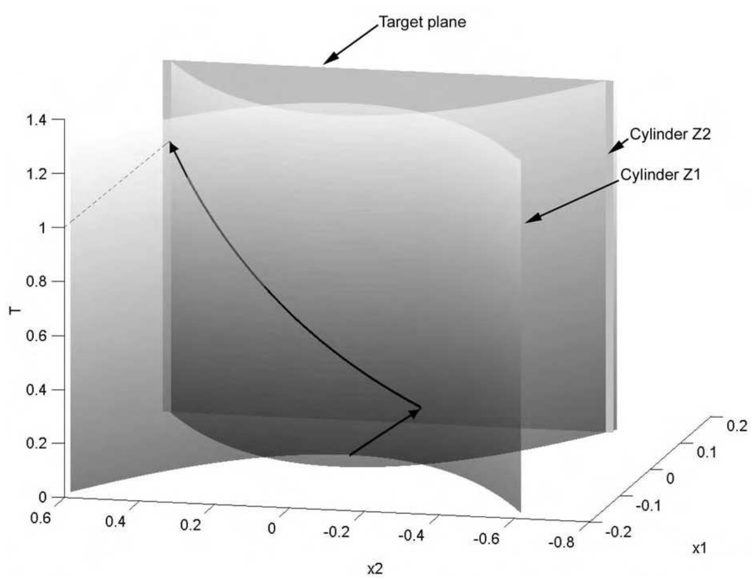

Figure 3.2: Optimal syntethis control

Figure 3.2 shows the state space variables together with additional variable t. The optimal trajectories (3.96) for the pass $k = 1$ corresponding to the bang-bang control law lie on the parabolic cylinders $(Z_1): \ x^{(1)} = -\frac{1}{2}(x^{(2)})^2 + C_1 + C_2$ and $(Z_2): \ x^{(1)} = +\frac{1}{2}(x^{(2)})^2 + \tilde{C}_1 + \tilde{C}_2$ where the constants $C_i, \tilde{C}_i, \ i = 1, 2$ are determined by the initial data $x^{(1)}(0) = \xi_1, \ x^{(2)}(0) = \xi_2$. These cylinders correspond to the solutions of differential equations (3.96) (for $k = 1$) with $u \equiv -1$ or $u \equiv +1$, respectively. It can also be shown that the admissible initial domain for which the problem can be solved is determined by the inequalities: $-\frac{3}{8} \leq \xi_1 + \xi_2 \leq \frac{5}{8}$. The switching manifold Z_h is described in parametric form by

$$\begin{cases} x^{(1)} = -\dfrac{\left(1 - \sqrt{5/8 - \xi_2 - \xi_1}\right)^2}{2} + \xi_2\left(1 - \sqrt{5/8 - \xi_2 - \xi_1}\right) + \xi_1, \\ x^{(2)} = -1 + \sqrt{5/8 - \xi_2 - \xi_1} + \xi_2, \\ T = 1 - \sqrt{5/8 - \xi_2 - \xi_1}, \\ -\frac{3}{8} \leq \xi_1 + \xi_2 \leq \frac{5}{8} \end{cases}$$

Finally, each optimal trajectory consists of two parts — first it evolves along the vertical parabolic cylinder Z_1 until $\tau = 1 - \sqrt{5/8 - \xi_2 - \xi_1}$ when it meets the switching manifold Z_h, and then immediately is switched to continue along the second vertical cylinder Z_2 to meet the target plane $x^{(1)} = 1/8$. Figure 3.2 also shows the optimal trajectory in the space $\mathbb{R}^3$ for zero initial data, and Figure 3.3 shows the projection of this trajectory onto the $x^{(1)}, x^{(2)}$ plane.

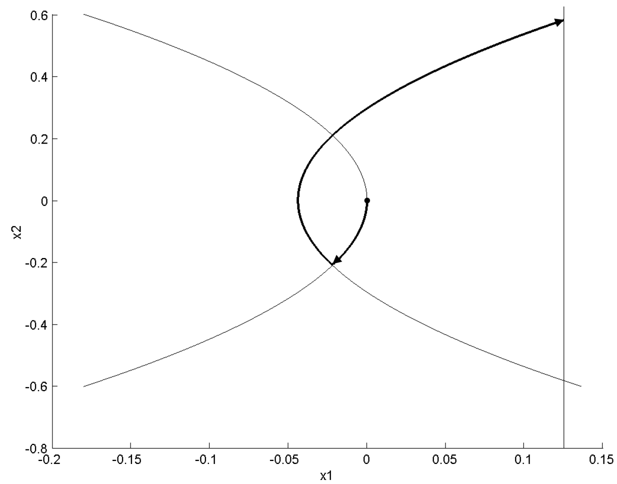

Figure 3.3: Projection on the $0x_1x_2$ plane

Example 2

Now we give an admittedly rather simple example where the advantages of the supporting control function approach is demonstrated for the solution the so-called synthesis problem of optimal control system.

For simplicity we will considered the following optimal control problem: maximize the terminal cost functional

$$\max_{|u|\leq 1} J(u),\;\; J(u) := x_2(1) \tag{3.114}$$

over the control system

$$\begin{array}{c} \frac{dx_1(t)}{dt} = x_2, \frac{dx_2(t)}{dt} = u(t), \\ x_1(s) = z_1,\;\; x_2(s) = z_2,\;\;\; t \in [s, 1],\;\; x_1(t), x_2(t) \in \mathbb{R} \end{array} \tag{3.115}$$

subject to the following constraints on control and state variables

$$|u(t)| \leq 1, \qquad x_1(1) = 1/8, \tag{3.116}$$

Consider the disturbances of the initial state (s, z_1, z_2) in some neighborhood of the point $(s = 0, z_1 = 0, z_2 = 0)$. It is easy to verify that for the case $s = 0$ and $x_1(0) = 0,\; x_2(0) = 0$ the optimal control for (3.114)–(3.116) is given by

$$\begin{aligned} u^0(t) &= -1 \quad \text{for} \quad 0 \leq t \leq 1 - \sqrt{\frac{5}{8}}; \\ u^0(t) &= +1 \quad \text{for} \quad 1 - \sqrt{\frac{5}{8}} < t \leq 1. \end{aligned}$$

Synthesis of the optimal control can be realized using the switching instance function $\tau = \tau(z_1, z_2, s)$, which has to satisfy the following differential equations

$$\frac{\partial \tau}{\partial z_1} = \frac{1}{2(1-\tau)}; \frac{\partial \tau}{\partial z_2} = \frac{1-s}{2(1-\tau)}; \frac{\partial \tau}{\partial s} = \frac{1-s-z_2}{2(1-\tau)} \tag{3.117}$$

with the initial condition

$$\tau(0,0,0) = 1 - \sqrt{\frac{5}{8}},$$

which is a particular case of (3.114). The solution of the Pfaff differential system (3.117) is given by

$$\tau(z_1, z_2, s) = 1 - \sqrt{\frac{5}{8} + (s-1)z_2 - z_1 - s + \frac{s^2}{2}}$$

The obtained formula for the switching instance function $\tau = \tau(z_1, z_2, s)$ allows us to find the optimal control law for any disturbances of the initial data in admissible control domain. The general formula for optimal control function in (3.114)–(3.116) with disturbances is written as

$$u^o(t, s, z_1, z_2) = \begin{cases} -1, & t \in [s, \tau(z_1, z_2, s)]; \\ +1, & t \in [\tau(z_1, z_2, s), 1]. \end{cases}$$

The obtained analytic form of the optimal control law gives a good ability to realize the regular synthesis on the phase plane. Let, for example $s = 0$, then the optimal switching function is

$$\tau(z_1, z_2, 0) = 1 - \sqrt{\frac{5}{8} - z_1 - z_2}.$$

The required synthesis picture is illustrated by the figures below.

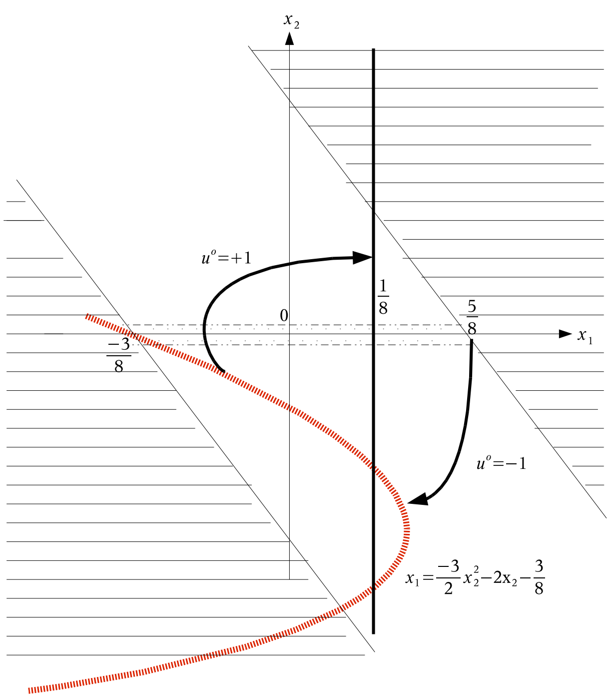

Figure 3.4: Switching curve is parabolic curve

First note that the admissible initial data domain for which the problem can be solved in the case $s = 0$ is determined by the inequalities: $\Omega : -\frac{3}{8} \leq z_1 + z_2 \leq \frac{5}{8}$. In general, the switching curve Z_c for regular synthesis is described in parametric form as

$$\begin{cases} x_1 = -\frac{\left(1-\sqrt{\frac{5}{8}-z_2-z_1}\right)^2}{2} + z_2\left(1-\sqrt{\frac{5}{8}-z_2-z_1}\right) \\ +z_1; \quad x_2 = -1+\sqrt{\frac{5}{8}-z_2-z_1}+z_2, \\ \tau = 1-\sqrt{\frac{5}{8}-z_2-z_1}, \quad -\frac{3}{8} \leq z_1+z_2 \leq \frac{5}{8} \end{cases}$$

The given parametric description can be transformed to the typical regular synthesis pictures.

For example, let the disturbances are realized along the line of the form $\Upsilon_1 : z_2 = \beta, -\infty \leq \beta \leq +\infty$ such that they belong to the admissible control region Ω. Then each optimal trajectory consists of two parts: first it evolves with $u = -1$ along the parabola $Z_1 : \ x_1 = -\frac{1}{2}(x_2)^2 + C_1 + C_2$ until $\tau = 1 - \sqrt{\frac{5}{8} - z_2 - z_1}$ when it meets the switching curve $Z_c(\Upsilon_1) : \ x_1 = -\frac{3}{2}(x_2)^2 - 2x_2(1-\beta) + \beta - \frac{3}{8}$, and then immediately it is switched to $u = +1$ to continue along the second parabola $Z_2 : \ x_1 = +\frac{1}{2}(x_2)^2 + \tilde{C}_1 + \tilde{C}_2$ to meet the target line $x_1 = \frac{1}{8}$. Here the constants $C_i, \tilde{C}_i, \ i = 1, 2$ are determined by the initial data $x_1(0) = z_1, \ x_2(0) = z_2$. These parabolas correspond to the solutions of the differential equations (3.115) with $u \equiv -1$ or $u \equiv +1$, respectively. Fig. 3.4 corresponds the case $\beta = 0$.

If the disturbances are active only along the line $\Upsilon_2 : \ z_1 + z_2 = \alpha$, then the switching curve is the line $Z_c(\Upsilon_2) : x_1 = -\sqrt{5/8 - \alpha}x_2 + \frac{\alpha}{2} - \frac{3}{16}$. Then each optimal trajectory consists of two parts: first it evolves with $u = -1$ along the corresponding parabola until $\tau = 1 - \sqrt{\frac{5}{8} - \alpha}$ when it meets the switching line L_1,and then it immediately is switched to the control law $u = +1$ to continue along the second parabola to meet the target line $x_1 = \frac{1}{8}$. Fig. 3.5 illustrates the case $\alpha = 0$.

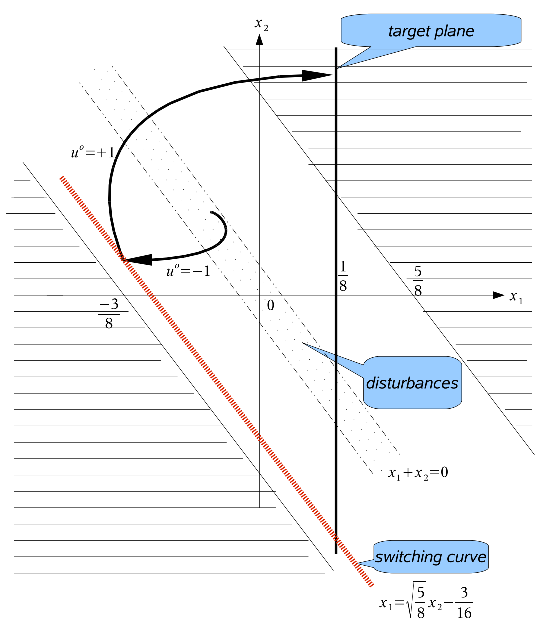

Figure 3.5: Switching curve is line

Remark 18. *In general, we may initially prescribe the possible disturbances curve* Υ*, and then it is easy to determine the required switching curve* $Z_c(\Upsilon)$*. This provides a convenient tool to construct the corresponding automation devices.*

3.4 Conclusions

In this paper the supporting control functions approach has been applied to study the optimal control problems for differential linear repetitive processes. The main contribution is the development of constructive necessary and sufficient optimality conditions in forms which can be effectively used for the design of numerical algorithms. The iterative method developed in this work is based on the principle of decrease of the suboptimality estimate, i. e. the iteration $\{\tau^k_{sup}, \quad u_k(t), \quad k = 1, \dots, N\} \rightarrow \{\hat{\tau}^k_{sup}, \quad \hat{u}_k(t), \quad k = 1, ..., N\}$ is performed in such a way as to achieve $\beta(\hat{\tau}_{sup}, \hat{u}) < \beta(\tau_{sup}, u)$. Also this procedure can be separated into two stages: 1) transformation of the admissible control functions $\{u_k(t), \quad k = 1, ..., N\} \rightarrow \{\hat{u}_k(t), \quad k = 1, ..., N\}$ which decreases the non-optimality measure of the admissible controls $\beta(\hat{u}) < \beta(u)$; and 2) variation of the support $\{\tau^k_{sup}, \quad k = 1, ..., N\} \rightarrow \{\hat{\tau}^k_{sup},, \quad k = 1, ..., N\}$ to again decrease the non-optimality measure of the support, i. e. $\beta(\hat{\tau}_{sup}) < \beta(\tau_{sup})$. These transformations involve, in effect, the duality theory for the problems defined in this work by (3.42)–(3.46) and (3.80)–(3.81) and exploit the ϵ-optimality conditions also developed in this work. These results are first in this general area and work is currently proceeding in a number of follow up areas. One such area is sensitivity analysis of optimal control in the presence of disturbances where in the case of the ordinary linear control systems some work on this topic can be found in, for example, [52].

4

Delay System Approach to Linear Differential Repetitive Processes

It is already known that repetitive processes can be represented in various dynamical system forms, which can, where appropriate, be used to great effect in the control related analysis of these processes. In this chapter, we investigate further the already known links between some classes of linear repetitive processes and delay systems and apply this to analyze control theory problems arising in controllability and optimal control of these repetitive processes. In particular, so-called characteristic mappings introduced in [37] are used to establish controllability properties criteria. Next, time optimal control problems are considered, where it is well known that the separation theorem for convex sets is a useful approach for studying a wide class of extremal problems. Here we adopt this method to establish optimality conditions in the classic form.

It has been conjectured that such a setting is appropriate for development the numerical methods for optimal control problems and related studies on for which very little work has been reported to date. The results developed here provide (part of) the theoretical background for further work aimed at the efficient computation of optimal controllers for these processes. Some areas for further research are also briefly discussed.

4.1 Background and Problem statement

The differential linear repetitive processes [66] are defined over $0 \leq t \leq \hat{\alpha}$, $k \geq 0$, by the state space model

$$\begin{aligned} \dot{x}_{k+1}(t) =& \hat{A}x_{k+1}(t) + \hat{B}u_{k+1}(t) + \hat{B}_0 y_k(t) \\ y_{k+1}(t) =& \hat{C}x_{k+1}(t) + \hat{D}u_{k+1}(t) + \hat{D}_0 y_k(t) \end{aligned} \tag{4.1}$$

Here on pass k, $x_k(t)$ is the $n \times 1$ state vector, $y_k(t)$ is the $m \times 1$ pass profile vector, and $u_k(t)$ is the $r \times 1$ vector of control inputs. To complete the process description, it is necessary to specify the boundary conditions, i. e. the state initial vector on each pass and the initial pass profile. Here no loss of generality arises from assuming $x_{k+1}(0) = d_{k+1}$, $k \geq 0$, and $y_0(t) = \hat{g}(t)$, where d_{k+1} is an $n \times 1$ vector of known constant entries and $\hat{g}(t)$ is an $m \times 1$ vector whose entries are known functions of t over $0 \leq t \leq \hat{\alpha}$.

As mentioned before, the repetitive processes posses many other equivalent representations which can be better suitable to the analysis of particular problems as, for example, $1D$ equivalent models enables much simple characterization of the so-called pass controllability or observability [37, 41]. Revisit now a few such examples.

1) Singularly perturbed model with slow and fast modes

$$\begin{aligned} \dot{x}_{k+1}(t) =& \hat{A}x_{k+1}(t) + \hat{B}u_{k+1}(t) + \hat{B}_0 y_k(t) \\ \mu\dot{y}_{k+1}(t) =& \hat{C}_0 y_{k+1}(t) + \hat{C}x_{k+1}(t) + \hat{D}u_{k+1}(t) + \hat{D}_0 y_k(t),\ 0 \le t \le \hat{\alpha},\ k \ge 0. \end{aligned} \tag{4.2}$$

Hence, the standard repetitive process is a limit case of that of (4.2) for $\mu = 0,\ det\hat{C}_0 \neq 0$. This approach is subject of ongoing work and the results will be reported in due course.

2) the Volterra type equation (with respect to the variable k)

$$\dot{x}_{k+1}(t) = \sum_{i-0}^{k} \Big[A_i x_{k+1-i}(t) + B_i u_{k+1-i}(t)\Big] + D_k g(t), x_{k+1}(0) = d_{k+1},\ k \ge 0 \tag{4.3}$$

where

$$A_0 = \hat{A},\ A_i = \hat{B}_0\hat{D}_0^{i-1}\hat{C},\ B_0 = \hat{B},\ B_i = \hat{B}_0\hat{D}_0^{i-1}\hat{D},\ D_i = \hat{B}_0\hat{D}_0^{i-1},\ i \ge 1$$

Discrete Volterra equations and their applications to the discrete repetitive models are given in [14]. The Volterra approach can be also effectively used for the differential case that is no subject of this paper.

To obtain another representation of processes described by (4.1) which is the subject of this paper (for the case $1 \le k \le N$ where N is a fixed positive integer), introduce the new variables $x : [0, \hat{\alpha}N] \to \mathbb{R}^n,\ y : [0, \hat{\alpha}N] \to \mathbb{R}^m,\ u[0, \hat{\alpha}N] \to \mathbb{R}^r$, where

$$x(t) = \begin{cases} x_1(t), & 0 < t < \hat{\alpha} \\ x_2(t-\hat{\alpha}), & \hat{\alpha} < t < 2\hat{\alpha}, \\ \ldots\ldots & \ldots\ldots\ldots \\ x_N(t-\hat{\alpha}(N-1)), & \hat{\alpha}(N-1) < t < \hat{\alpha}N \end{cases}$$

$$y(t) = \begin{cases} y_1(t), & 0 < t < \hat{\alpha} \\ y_2(t-\hat{\alpha}), & \hat{\alpha} < t < 2\hat{\alpha}, \\ \ldots\ldots & \ldots\ldots\ldots \\ y_N(t-\hat{\alpha}(N-1)), & \hat{\alpha}(N-1) < t < \hat{\alpha}N \end{cases},$$

$$u(t) = \begin{cases} u_1(t), & 0 < t < \hat{\alpha} \\ u_2(t-\hat{\alpha}), & \hat{\alpha} < t < 2\hat{\alpha}, \\ \ldots\ldots & \ldots\ldots\ldots \\ u_N(t-\hat{\alpha}(N-1)), & \hat{\alpha}(N-1) < t < \hat{\alpha}N \end{cases}$$

Then, (4.1) can be rewritten in the form of the following delay system

$$\begin{bmatrix} \frac{d}{dt} & 0 \\ 0 & I_m \end{bmatrix} \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} \hat{A} & 0 \\ \hat{C} & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} + \begin{bmatrix} 0 & \hat{B}_0 \\ 0 & \hat{D}_0 \end{bmatrix} \begin{bmatrix} x(t-\hat{\alpha}) \\ y(t-\hat{\alpha}) \end{bmatrix} + \begin{bmatrix} \hat{B} & 0 \\ 0 & \hat{D} \end{bmatrix} \begin{bmatrix} u(t) \\ u(t) \end{bmatrix} \tag{4.4}$$

with initial condition

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} 0 \\ \hat{g}(t) \end{bmatrix}, \quad t \in [-\hat{\alpha}, 0]. \tag{4.5}$$

Here, I_m denotes the identity matrix in $\mathbb{R}^m$. In order to complete the correspondence between the delay system (4.4) and the repetitive process (4.1) we require additional constraints at $t = \hat{\alpha}k,\ k = 1, \dots N-1$, which demand that the solution $x(t)$ is discontinuous and has "jumps/pushes". This leads to the so-called nonlocal conditions of the form

$$x(k\hat{\alpha} + 0) = d_k, \quad k \in K, \tag{4.6}$$

where $x(k\hat{\alpha} + 0)$ denotes $x(t)$ as $t \to k\hat{\alpha}$ from the right. We also assume that the control functions $u(t)$ and pass profile vectors $y(t)$ are continuous from the right hand side at $t = \hat{\alpha}k,\ k = 1, \dots, N-1$.

It is straightforward to see that this last representation is a special singular case of

$$\begin{bmatrix} \frac{d}{dt} & 0 \\ 0 & I_m \end{bmatrix} \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} + \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix} \begin{bmatrix} x(t-\hat{\alpha}) \\ y(t-\hat{\alpha}) \end{bmatrix} + \\ + \begin{bmatrix} B & 0 \\ 0 & D \end{bmatrix} \begin{bmatrix} u(t) \\ u(t) \end{bmatrix} \tag{4.7}$$

which is equivalent to

$$\begin{aligned} \dot{x}(t) =& A_{11}x(t) + A_{12}y(t) + D_{11}x(t-\hat{\alpha}) + D_{12}y(t-\hat{\alpha}) + Bu(t) \\ y(t) =& A_{21}x(t) + A_{22}y(t) + D_{21}x(t-\hat{\alpha}) + D_{22}y(t-\hat{\alpha}) + Du(t) \end{aligned} \tag{4.8}$$

Finally, if the matrix $(I_m - A_{22})$ is nonsingular, then the second equation can be re-arranged to the form

$$\begin{aligned} \dot{x}(t) =& A_{11}x(t) + A_{12}y(t) + D_{11}x(t-\hat{\alpha}) + D_{12}y(t-\hat{\alpha}) + Bu(t) \\ y(t) =& \tilde{A}_{21}x(t) + \tilde{D}_{21}x(t-\tilde{\alpha}) + \tilde{D}_{22}y(t-\tilde{\alpha}) + \tilde{D}u(t) \end{aligned} \tag{4.9}$$

where $\tilde{H} = (I_m - A_{22})^{-1}H$ for H belonging to the set $H \triangleq \{A_{21},\ D_{21},\ D_{22},\ D\}$.

If a linear repetitive process of the form of (4.1) contains time delays such that the resulting process model has the following form over $0 \le t \le \hat{\alpha}$, $1 \le k \le N$,

$$\begin{aligned} \dot{x}_{k+1}(t) =& \hat{A}x_{k+1}(t) + \hat{A}_{-1}x_{k+1}(t-\hat{h}) + \hat{B}u_{k+1}(t) + \hat{B}_0 y_k(t) + \hat{B}_{-1}y_k(t-\hat{h}) \\ y_{k+1}(t) =& \hat{C}x_{k+1}(t) + \hat{C}_{-1}x_{k+1}(t-\hat{h}) + \hat{D}u_{k+1}(t) + \hat{D}_0 y_k(t) + \hat{D}_{-1}y_k(t-\hat{h}) \end{aligned} \tag{4.10}$$

where $\hat{h}$ is a real number such that $0 < \hat{h} \le \hat{\alpha}$. Then such linear repetitive processes can be presented in the multiple delay differential system form of

$$\begin{bmatrix} \frac{d}{dt} & 0 \\ 0 & I_m \end{bmatrix} \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} \hat{A} & 0 \\ \hat{C} & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} + \begin{bmatrix} 0 & \hat{B}_0 \\ 0 & \hat{D}_0 \end{bmatrix} \begin{bmatrix} x(t-\hat{\alpha}) \\ y(t-\hat{\alpha} \end{bmatrix} + \begin{bmatrix} \hat{B} & 0 \\ 0 & \hat{D} \end{bmatrix} \begin{bmatrix} u(t) \\ u(t) \end{bmatrix} \\ + \begin{bmatrix} \hat{A}_{-1} & 0 \\ \hat{C}_{-1} & 0 \end{bmatrix} \begin{bmatrix} x(t-h) \\ y(t-h) \end{bmatrix} + \begin{bmatrix} 0 & \hat{B}_{-1} \\ 0 & \hat{D}_{-1} \end{bmatrix} \begin{bmatrix} x(t-\hat{h}-\hat{\alpha}) \\ y(t-\hat{h}-\hat{\alpha}) \end{bmatrix} \tag{4.11}$$

with initial conditions

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} 0 \\ \hat{g}(t) \end{bmatrix}, \quad t \in [-\hat{\alpha}, 0], \quad \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} \phi(t) \\ \psi(t) \end{bmatrix}, \quad t \in [-\hat{\alpha} - \hat{h}, -\hat{\alpha}) \tag{4.12}$$

and t nonlocal conditions

$$x(k\hat{\alpha} + 0) = d_k, \quad k \in K. \tag{4.13}$$

where $\psi(t), \phi(t), \hat{g}(t)$ are the corresponding initial conditions in (4.11).

4.2 Hybrid delay model for differential repetitive processes

As the basis for further study consider first the case when the nonlocal conditions of (4.6) are absent. This can be realized under the assumption that, for example, the initial condition in (4.1) for the current pass coincides with the end point state of the previous pass, i. e. $x_{k+1}(0) = x_k(\alpha)$, that occur often in machining operations. Such assumption is needed to avoid at the primary stage the presence of a nonlocal impulse initial conditions, which can be the source of significant difficulties.

The system under the consideration is now given by the following pair of differential and difference equations

$$\begin{aligned} \dot{x}(t) =& Ax(t) + A_{-1}x(t-h) + B_0 y(t) + B_{-1}y(t-h) + Bu(t) \\ y(t) =& Cx(t) + C_{-1}x(t-h) + D_{-1}y(t-h) + Du(t),\ t \in T \doteq [0, \alpha] \end{aligned} \tag{4.14}$$

with initial conditions

$$x(t) = f(t),\ t \in [-h, 0),\ x(0) = x_0,\ y(t) = g(t),\ t \in [-h, 0] \tag{4.15}$$

where $x \in \mathbb{R}^n,\ y \in \mathbb{R}^m,\ u \in \mathbb{R}^r$, and α and h are given real numbers such that $h < \alpha$. We also assume that the control function $u(t)$ is piecewise continuous on the interval $[0, \alpha]$. The differential linear repetitive process (4.1) now follows immediately as a special case of this last model structure on choosing the matrices in (4.14) as

$$A = \hat{A},\ A_{-1} = 0,\ B_0 = 0,\ B_{-1} = \hat{B}_0,\ B = \hat{B},\ C = \hat{C},\ C_{-1} = 0,\ D_{-1} = \hat{D}_0,\ D = \hat{D}$$

and $\alpha = \hat{\alpha}N,\ h = \alpha$.

It is well known [43, 62] that a solution of the time delay differential equation can be found by the step method. In other words, by application the standard integration step-to step method on each subinterval $[kh, (k+1)h)$ (with nonnegative integer k) we can construct the solution as the solution of an appropriate ODE. Let us focus on the smoothness property of the solutions as it follows from this procedure. Consider hence on the first delay-interval, and more at the moment $t = 0$. Due to the form of the differential equation (4.14), and since the initial condition (4.15) is chosen arbitrarily, one can say that

$$\dot{x}(t)|_{t=0+} \neq \dot{f}(t)|_{t=0-} = \dot{x}(t)|_{t=0-} \tag{4.16}$$

i. e. there is a discontinuity in the first derivative of the solution $x(t)$ at the moment $t = 0$. Due to this fact we consider the differential equation (4.14) for $t > 0$ and use the separate function value $x(0) = x_0$ in the initial data (4.15). This remark can be extended to the next delay-intervals $[kh, (k+1)h),\ k > 0$, but, note, that the solution is getting smoother from one delay-interval to the next at the moments $t = kh,\ k > 1$. Next, from the difference equation (4.14) it follows that at the moment $t = 0$ we have

$$y(0) = Cx(0) + C_{-1}x(-h) + D_{-1}y(-h) + Du(0), \tag{4.17}$$

i. e. the value $y(0)$ of the pass profile $y(t)$ is determined by the initial data and the value of control function $u(0)$. By this reason we consider the control functions $u(t)$ that are continuous from the right hand side, and put for brevity that the left side limit value $u(0-) = \lim\limits_{t \to 0-} u(t)$ coincides with $g(0)$.

The pair of the functions $\big(x(t), y(t)\big)$ is termed a solution of the system (4.14) — (4.15) for the given control function $u(t)$, if they satisfy the differential equation (4.14) almost everywhere on the interval $[0, \alpha]$ and the difference equation (4.14) for all $t \in [0, h]$. It is known that under the given assumptions the solution $x(t)$ is absolutely continuous and $y(t)$ is piecewise continuous on the interval $[0, \alpha]$.

Systems described by the equations of such form have been discussed in [1, 60], the first results on optimality conditions for the nonlinear version of the system of (4.14) — (4.15) were obtained in [73], and some observability and controllability problems for a particular case of the system can be found in [3]. Here we present in an unified form some results on controllability and optimization that are relevant for the need of a deep theoretical background for control of repetitive processes.

4.2.1 General response formula

The solution of the system (4.14) — (4.15) can be constructed by the step-by-step procedure for each subinterval of the form $[ih, (i+1)h),\ i = 0, 1, \ldots, q_\alpha$, where $q_\alpha = \left[\frac{\alpha}{h}\right]$ denotes the integer part of the fraction $\frac{\alpha}{h}$. First, it is straightforward to show that the recurrent procedure based on the equation (4.14) leads to the following representation of $y(t)$ on the time interval $[0, t],\ t > h,\ t \in [q_t h, (q_t + 1)h)$ where $q_t = \left[\frac{t}{h}\right]$

$$\begin{aligned} y(t) =& Cx(t) + \sum_{j=0}^{q_t-1} M_{j+1} x\big(t - (j+1)h\big) + \sum_{j=0}^{q_t} G_j u(t - jh) \\ &+ K_{q_t} g\big(t - (q_t+1)h\big) + W_{q_t} f\big(t - (q_t+1)h\big) \end{aligned} \tag{4.18}$$

and for $t \in [0, h)$

$$y(t) = Cx(t) + Du(t) + D_{-1} g(t - h\big) + C_{-1} f\big(t - h\big),$$

where

$$M_{j+1} = D_{-1}^j (C_{-1} + D_{-1}C),\ G_j = D_{-1}^j D,\ K_i = D_{-1}^i C_{-1},\ W_i = D_{-1}^{i+1},\ M_0 = C, \quad j = 0, 1, \ldots \tag{4.19}$$

Noting the formula (4.18)—(4.19) and using the recurrent procedure on the intervals $[0, h)$, $[h, 2h), \ldots$ allows us to rewrite (4.14) as

$$\begin{aligned} \dot{x}(t) &= \sum_{j=1}^{q_t+1} H_j x\big(t - (j-1)h\big) + \sum_{j=1}^{q_t+1} V_j u\big(t - (j-1)h\big) \\ &+ Q_{q_t+1} g\big(t - (q_t+1)h\big) + P_{q_t+1} f\big(t - (q_t+1)h\big) \end{aligned} \tag{4.20}$$

where

$$\begin{aligned} H_1 =& A + B_0 C,\ H_2 = A_{-1} + B_0\big(C_{-1} + D_{-1}C\big) + B_{-1}C, \\ H_j =& \big(B_0 D_{-1}^{j-1} + B_{-1} D_{-1}^{j-2}\big)\big(C_{-1} + D_{-1}C\big),\ j = 2, \ldots,\ q_t + 1 \\ V_1 =& B + B_0 D,\ V_j = \big(B_0 D_{-1}^{j-1} + B_{-1} D_{-1}^{j-2}\big) D,\ j = 2, \ldots,\ q_t + 1, \\ P_i =& (B_0 D_{-1} + B_{-1} D_{-1}^{i-2}) C_{-1},\ Q_i = \big(B_0 D_{-1} + B_{-1}\big) D_{-1}^{i-1},\ P_1 = A_{-1} + B_0 C_{-1} \end{aligned} \tag{4.21}$$

The formula (4.20) says, in fact, that the hybrid system of (4.14) can be represented by retarded differential equations with varying number of delays. Amount of delays is increasing with the growth of t.

Next, multiplying both sides of the equation (4.20) by the function $F(t,\tau)$, which is unknown at present, and then integrating yields on the left hand side

$$\int_0^t F(t,\tau)\dot{x}(\tau)d\tau = x(t) - F(t,0)x_0 - \int_0^t \frac{\partial F(t,\tau)}{\partial \tau}x(\tau)d\tau \tag{4.22}$$

where we set that $F(t,\tau) \equiv 0,\ \forall\, \tau > t$, and $F(t,t-0) = I_n$. Next, substituting $s = \tau - ih$ in each of the integrals on the right hand side, and noting that

$$F(t,\tau) \equiv 0,\ \forall\, \tau > t,\ x(t) = f(t),\ t \in [-h,0),\ x(t) \equiv 0,\ \forall t < -h$$

together with (4.22) leads to the following formula

$$x(t) = F(t,0)x_0 + \int_0^t \sum_{j=1}^{q_t+1} F(t,\tau)H_j x\big(\tau-(j-1)h\big)d\tau + \int_0^t \frac{\partial F(t,\tau)}{\partial \tau}x(\tau)d\tau + \tag{4.23}$$

$$+\int_0^t \sum_{j=1}^{q_t+1} F(t,\tau)V_j u\big(\tau-(j-1)h\big)d\tau + \int_0^t Q_{q_t+1}g\big(\tau-(q_t+1)h\big)d\tau + \int_0^t P_{q_t+1}f\big(\tau-(q_t+1)h\big)d\tau$$

$$= F(t,0)x_0 + \sum_{j=1}^{q_t+1}\int_0^t F(t,\tau+(j-1)h)H_j x(\tau)d\tau + \sum_{j=1}^{q_t+1}\int_0^t F(t,\tau+(j-1)h)V_j u(\tau)d\tau +$$

$$+\sum_{j=1}^{q_t+1}\int_{-h}^0 F(t,\tau+(j-1)h)H_j f(\tau)d\tau + \int_{-h}^0 F(t,\tau+(q_t+1)h)\Big[P_{q_t+1}f(\tau) + Q_{q_t+1}g(\tau)\Big]d\tau$$

Now define the required function $F(t,\tau)$ as a solution of the following differential equation

$$\frac{\partial F(t,\tau)}{\partial \tau} = -\sum_{j=1}^{q_t+1} F(t,\tau+(j-1)h)H_j,\ F(t,\tau) \equiv 0,\ \forall\, \tau > t,\ F(t,t-0) = I_n, \tag{4.24}$$

(where $F(t,t-0)$ denotes $F(t,\tau)$ evaluated as $t \to \tau$ from the left) whose properties can be found, for example, in [37]. Finally, noting (4.18), we have the following formula for the solutions of the system (4.14)—(4.15)

$$x(t) = F(t,0)x_0 + \sum_{j=1}^{q_t+1}\int_{-h}^0 F(t,\tau+(j-1)h)H_j f(\tau)d\tau + \sum_{j=1}^{q_t+1}\int_0^t F(t,\tau+(j-1)h)V_j u(\tau)d\tau +$$

$$+\int_{-h}^0 F(t,\tau+(q_t+1)h)\big[P_{q_t+1}f(\tau) + Q_{q_t+1}g(\tau)\big]d\tau,\ t \geq 0; \tag{4.25}$$

$$y(t) = CF(t,0)x_0 + \int_{-h}^{0} CF(t,\tau)H_1 f(\tau)d\tau + \int_{-h}^{0} CF(t,\tau+h)\Big[P_1 f(\tau) + Q_1 g(\tau)\Big]d\tau +$$

$$+ \int_{0}^{t} CF(t,\tau)V_1 u(\tau)d\tau + C_{-1}f(t-h) + D_{-1}g(t-h) + Du(t),\ \ t \in [0,h)$$

$$\begin{aligned} y(t) =& CF(t,0)x_0 + \sum_{j=0}^{q_t-1} M_{j+1}F(t-(j+1)h,0)x_0 + \\ & \sum_{l=0}^{q_t-1}\sum_{j=0}^{q_t-l}\int_{-h}^{0} M_l F(t-lh,\tau+jh)H_{j+1}f(\tau)d\tau + \\ & + \sum_{l=0}^{q_t}\int_{-h}^{0} M_l F(t-lh,\tau+(q_t+1-l)h)\big[P_{q_t+1-l}f(\tau) + Q_{q_t+1-l}g(\tau)\big]d\tau + \\ & + \sum_{l=0}^{q_t-1}\sum_{j=0}^{q_t-l}\int_{0}^{t} M_l F(t-lh,\tau+jh)V_{j+1}u(\tau)d\tau + \\ & + \sum_{j=0}^{q_t} G_j u(t-jh) + K_{q_t-1}g(t-q_t h) + W_{q_t-1}f(t-q_t h), \quad q_t = \left[\frac{t}{h}\right],\ t \geq h \end{aligned} \tag{4.26}$$

which clearly is the general response formula for (4.14).

4.3 Controllability

In this section we consider controllability of hybrid system of (4.14) which clearly must be a fundamental element of a mature systems theory for linear repetitive processes and play a significant role for application area. The formula (4.25)–(4.26) is a required starting point for this study. Here it should also be noted there exist more than one distinct controllability notion, see e. g. [57], and this area is far from being complete for the repetitive processes and delay systems considered here.

4.3.1 Point pass profile controllability

For nD systems as well as for repetitive processes there many possibilities for introducing various controllability notions. In this subsection we introduce and study the following point pass profile controllability, which plays a significant role in further analysis.

Definition 16. *The system (4.14)-(4.15) is said to be pass profile controllable at the given points $\beta_0, \beta_1, \ldots, \beta_\nu$, such that $0 = \beta_0 < \beta_1 < \ldots \beta_\nu \leq \alpha$, if for any $c_i \in \mathbb{R}^m,\ i = 0, \ldots, \nu$ there exists a control vector $u(t),\ t \in [0, \alpha]$ such that the solution $y(t, g, f, x_0, u)$ of the* *system (4.14)—(4.15) corresponding to the zero initial data* $g(t) = 0,\ t \in [-h, 0),\ f(t) = 0,\ t \in [-h, 0),\ x_0 = 0$ *satisfies the following conditions*

$$y(\alpha - \beta_j, 0, 0, 0, u) = c_j,\ j = 0, 1, \dots, \nu. \tag{4.27}$$

We suppose that the admissible control functions $u(t)$ belong to the class of all piecewise continuous functions on the interval $t \in [0, \alpha]$ with values in the space $\mathbb{R}^r$ and is denoted by $U(\cdot)$.

Physical motivation for this form of controllability is the requirement that the pass profile vector take pre-assigned values at particular points along the pass. Note also that some first results concerning observability and controllability problems for particular cases of the system model structure considered here can be found in the earlier paper [3].

From (4.25) we have

$$y(t) = \sum_{j=0}^{i} G_j u(t - jh) + \int_0^t R(t, \tau) u(\tau) d\tau, \qquad t \geq h, \tag{4.28}$$

where

$$R(t, \tau) = \sum_{l=0}^{q_t - 1} \sum_{j=0}^{q_t - l} \int_0^t M_l F(t - lh, \tau + jh) V_{j+1}, \qquad q_t = \left[\frac{t}{h}\right] \tag{4.29}$$

Note that in (4.25) $r(t, g, f, x_0) = 0$ for the zero initial data $g(t) = 0,\ t \in [-h, 0),\ f(t) = 0,\ t \in [-h, 0),\ x_0 = 0$.

Theorem 17. *The system (4.14)—(4.15) is pass profile controllable at the given points* $\beta_0, \beta_1, \dots, \beta_\nu$ *if, and only if, the following equalities*

$$g_i^T G_0 = 0,\ \dots,\ g_i^T G_{q_i} = 0, \qquad g_i^T R(\alpha - \beta_i, \tau) \equiv 0,\ \tau \in [0, \alpha - \beta_i],\ i = 0, 1, \dots, \nu \tag{4.30}$$

hold only when $g_i = 0$*, where* $g_i \in \mathbb{R}^m,\ i = 0, 1, \dots \nu$ *and* $q_i = \left[\frac{\alpha - \beta_i}{h}\right]$.

Proof. The property to be established here requires that the following set of equations

$$\begin{aligned}
c_0 &= G_0 u(\alpha) + \ldots + G_{q_0} u(\alpha - q_0 h) + \int_0^{\alpha - \beta_0} R(\alpha - \beta_0, \tau) u(\tau) d\tau, \\
c_1 &= G_0 u(\alpha - \beta_1) + \ldots + G_{q_1} u(\alpha - \beta_1 - q_1 h) + \int_0^{\alpha - \beta_1} R(\alpha - \beta_1, \tau) u(\tau) d\tau, \\
&\dots \qquad \qquad \dots\dots\dots\dots \\
c_\nu &= G_0 u(\alpha - \beta_\nu) + \ldots + G_{q_\nu} u(\alpha - \beta_\nu - q_\nu h) + \int_0^{\alpha - \beta_\nu} R(\alpha - \beta_\nu, \tau) u(\tau) d\tau
\end{aligned} \tag{4.31}$$

can be solved with respect to the unknown vector $u(t),\ t \in [0, \alpha]$ with piecewise continuous entries and r-vectors $u(\alpha - \beta_i - q_i h),\ i = 0, \dots, \nu$. Consider therefore the following set

$$\begin{aligned}
Y = \Bigg\{ y = &(y_0, \dots, y_\nu) \in \mathbb{R}^{m(\nu+1)} :\ y_s = \sum_{j=0}^{q_s} G_j v_{js} + \\
&+ \int_0^{\alpha - \beta_s} R(\alpha - \beta_s, \tau) u(\tau) d\tau,\ \forall\ v_{js} \in \mathbb{R}^{m(\nu+1)},\ \forall u(\cdot) \in U(\cdot) \Bigg\}
\end{aligned} \tag{4.32}$$

where $U(\cdot)$ denotes the set of all admissible control vectors. Then it is easy to see that the set $Y \subset \mathbb{R}^{m(\nu+1)}$ is a linear subspace from $\mathbb{R}^{m(\nu+1)}$.

Now suppose that conditions (4.30) hold but the system is not pass profile controllable. Then this means that $Y \neq \mathbb{R}^{m(\nu+1)}$. Since the set Y is linear subspace of $\mathbb{R}^{m(\nu+1)}$, there exists a nontrivial vector $\bar{g} = (\bar{g}_1, \dots, \bar{g}_\nu) \in \mathbb{R}^{m(\nu+1)}$, $\bar{g} \neq 0$, such that $\bar{g} \perp Y$. This, in turn, means that there exists a nontrivial vector $\bar{g} \neq 0$ which satisfies the conditions of (4.30) and a contradiction has been established.

Suppose now the system is controllable but condition (4.30) holds for some nontrivial vector $g^* \in \mathbb{R}^{m(\nu+1)}$. This means that $g^* \perp Y$. Hence $Y \neq \mathbb{R}^{m(\nu+1)}$ which is a contradiction and the proof is complete. ■

Theorem 17, however, is hard to apply for checking controllability. Another approach would be to apply the so-called characteristic equations approach introduced in [37] to obtain the effective criteria to check the controllability properties of the considered model. To obtain the characteristic equations follows apply the Laplace transform to the system (4.14)—(4.15) with zero initial data

$$\begin{aligned} pX(p) &= AX(p) + A_{-1}e^{-ph}X(p) + B_0Y(p) + B_{-1}e^{-ph}Y(p) + BU(p), \\ Y(p) &= CX(p) + C_{-1}e^{-ph}X(p) + D_{-1}e^{-ph}Y(p) + DU(p). \end{aligned} \tag{4.33}$$

In what follows the following substitutions are to be done: replace $X(p), Y(p), U(p)$ by the $(n\times r)$, $(m\times r)$ and $(r\times r)$ - matrices $X_{k-1}(t)$, $Y_{k-1}(t)$, $U_{k-1}(t)$, $k = 1, 2, \dots,$ $t \in [0, \alpha]$; the differention operator p is replaced by the shift operator with respect to discrete variable k, the operator e^{-ph} is replaced by the time delay operator such that the following relations

$$\begin{aligned} X(p) &\longrightarrow X_{k-1}(t),\ e^{-ph}X(p) \longrightarrow X_{k-1}(t-h),\ pX(p) \longrightarrow X_k(t) \\ Y(p) &\longrightarrow Y_{k-1}(t),\ e^{-ph}Y(p) \longrightarrow Y_{k-1}(t-h) \end{aligned} \tag{4.34}$$

hold. This enables rewriting (4.14)—(4.15) in the following form

$$\begin{aligned} X_k(t) =& AX_{k-1}(t) + A_{-1}X_{k-1}(t-h) + B_0Y_{k-1}(t) + B_{-1}Y_{k-1}(t-h) + BU_{k-1}(t) \\ Y_{k-1}(t) =& CX_{k-1}(t) + C_{-1}X_{k-1}(t-h) + D_{-1}Y_{k-1}(t-h) + DU_{k-1}(t),\ \ t \in [0, \alpha] \end{aligned} \tag{4.35}$$

In order to complete this setting it is necessary to determine the initial conditions

$$\begin{aligned} X_0(0) =& 0,\ X_i(t) \equiv 0,\ \forall\ i \leq 0,\ t \leq 0;\ Y_0(0) = 0,\ Y_i(t) \equiv 0,\ \forall\ i \leq 0,\ t \leq 0. \\ U_0(0) =& I_r,\ U_i(t) \equiv 0,\ \forall\ i \neq 0,\ t \neq 0. \end{aligned} \tag{4.36}$$

Now, the following theorem can be stated the proof of which is very strongly motivated by the results of the earlier paper [3].

Theorem 18. *The system (4.14)—(4.15) is pass profile controllable at the given points $\beta_0, \beta_1, \dots, \beta_\nu$ if, and only if, the following rank condition holds*

$$rank \begin{bmatrix} Y_i(t-\beta_0) & \\ Y_i(t-\beta_1) & i = 0, \dots, n(q_\alpha + 1) \\ \dots & \\ Y_i(t-\beta_\nu) & t \in [0, \beta_\nu + (q_\alpha + 1)h] \end{bmatrix} = (\nu + 1)m. \tag{4.37}$$

Proof. The conditions (4.30) of Theorem 17 yield that the process (4.14)—(4.15) is pointwise profile pass controllable if, and only if, the following conditions hold

$$\operatorname{rank}\{\mathcal{G}_0, \mathcal{G}_1, \ldots, \mathcal{G}_{q_\alpha}\} = m(\nu+1) \tag{4.38}$$

and

$$g^T \mathfrak{F}(\alpha,\tau) \not\equiv 0 \quad \tau \in [0,\alpha] \qquad \text{for all} \quad g \in \mathbb{R}^{m(\nu+1)},\ g \neq 0 \tag{4.39}$$

where $\big(m(\nu+1)\times r(\nu+1)\big)-$ matrices $\mathcal{G}_i$ are given by

$$\mathcal{G}_i = diag\{G_i, \ldots, G_i\}, \qquad \text{where} \ \ G_i = D_{-1}^i D,\ i = 0, \ldots, q_\alpha, \tag{4.40}$$

and $[n(\nu+1)\times m(\nu+1)]$ matrix function $\mathfrak{F}(\alpha,\tau)$ is defined as

$$\mathfrak{F}(\alpha,\tau) = \begin{cases} \mathcal{M}_0\mathcal{F}(\alpha,\tau)\mathcal{V}, & \tau \in [\alpha-\beta_1, \alpha) \\ \big(\mathcal{M}_0\mathcal{F}(\alpha,\tau) + \mathcal{M}_1\mathcal{F}(\alpha,\tau+\beta_1)\big)\mathcal{V}, & \tau \in [\alpha-\beta_1, \alpha) \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots\ldots \\ \sum\limits_{i=0}^{\nu} \mathcal{M}_i\mathcal{F}(\alpha,\tau+\beta_i)\mathcal{V}, & \tau \in (\alpha-\beta_\nu, \alpha-\beta_{\nu-1}], \end{cases} \tag{4.41}$$

where

$$\begin{aligned} \mathcal{V} &= [V_1, V_2, \cdots, V_{q_\alpha+1}]^T, \\ \mathcal{M}_k &= \Big[O_{n(q+1)\times mk}, \big[M_0, \ldots, M_q\big]^T, O_{n(q+1)\times m(\nu-k)}\Big], \quad k = 0, \ldots, \nu \end{aligned} \tag{4.42}$$

and

$$\mathcal{F}(\alpha,\tau) = \begin{bmatrix} F(\alpha,\tau) & F(\alpha,\tau+h) & F(\alpha,\tau+2h) & \cdots & F(\alpha,\tau+(\nu-1)h) & F(\alpha,\tau+\nu h) \\ F(\alpha,\tau+h) & F(\alpha,\tau+2h) & F(\alpha,\tau+3h) & \cdots & F(\alpha,\tau+\nu h) & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ F(\alpha,\tau+\nu h) & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

Note that to design the function (4.41) the well-known properties $F(t,\tau) \equiv 0, \ t < \tau, \ \ F(t-s,\tau) = F(t, s+\tau),\ t \le s \le \tau$ of the fundamental matrix $F(t,\tau)$ are used. From the characteristic equation (4.35)—(4.36) it follows immediately that $Y_0(ih) = D_{-1}^i D = G_i$, and, hence,

$$\mathcal{G}_i = diag\{Y_0(ih), \ldots, Y_0(ih)\}$$

which contains a part of the required matrices in (4.37).

To select the remaining matrices in (4.37), consider the $n(q_\alpha+1)$ vector-valued function $\psi(g,\tau) = g^T\mathfrak{F}(\alpha,\tau)$. Then it can be shown that the function $\mathfrak{F}(\alpha,\tau)$ satisfies the following matrix differential equation

$$\frac{\partial \mathfrak{F}(\alpha,\tau)}{\partial \tau} = -\mathfrak{F}(\alpha,\tau)\mathcal{H}, \quad \text{where } \mathcal{H} = \begin{bmatrix} H_1 & 0 & 0 & \cdots & 0 & 0 \\ H_2 & H_1 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ H_{q_\alpha} & H_{q_\alpha-1} & H_{q_\alpha-2} & \cdots & H_2 & H_1 \end{bmatrix} \tag{4.43}$$

Hence the function $\psi(g,\tau)$ satisfies an ordinary homogeneous differential equation of order $n(q_\alpha+1)$, and it is known that such differential equation has the trivial solution if, and only if, its initial conditions are zero. From [37] it now follows that the function considered here is analytic on each sub-interval $(\alpha-kh,\alpha-(k+1)h)$, and dis-continuous at $\tau=\alpha-\beta_i-jh,\ i=0,\ldots,\nu,\ j=1,\ldots,q$ where the jumps are given by

$$\begin{aligned}\Delta\psi^{(s)}(g,\alpha-\beta_i-jh) \doteq&\psi^{(s)}(g,\alpha-\beta_i-jh+0)-\psi^{(s)}(g,\alpha-\beta_i-jh-0)=\\ =&(-1)^{i+j}g^TY_s(\alpha-\beta_i-jh),\ i=0,\ldots,\nu,\ j=1,\ldots,q\end{aligned}\tag{4.44}$$

Thus, $\Delta\psi^{(s)}(g,\alpha-\beta_i-jh)=0$, and, therefore $\psi(g,\tau)\equiv 0$, if, and only if, the matrix (4.37) has the maximal rank, which completes the proof. ■

4.3.2 Pointwise completness and controllability with respect to initial data

In general, for differential systems with retarded arguments and, in particular, for hybrid differential-difference systems the so-called pointwise completness [80, 83] pays a key role. In order to formulate this notion we introduce the following notations. Let $C^n[-h,0],\ h>0$ denotes the vector space of the continuous n-vector function $f:[-h,0]\to\mathbb{R}^n$. The solution of system (4.14)—(4.14) (in the absence of input actions, i. e. with $B=0,\ D=0$) corresponding to the initial data (4.15) where $f\in C^n[-h,0],\ g\in C^m[-h,0],\ x_0\in\mathbb{R}^n$ is denoted by $x(t)=x(t,f,g,x_0),\ y(t)=y(t,f,g,x_0)$. Reachability set for the state variable $x(t)$ of the system (4.14)—(4.15) at the given moment $t^*\in[0,T]$ is defined as follows

$$\mathcal{R}_x(t^*)=\left\{x\in\mathbb{R}^n : x=x(t^*,g,f,x_0),\ \text{for all}\ \ f\in C^n[-h,0],\ g\in C^m[-h,0],\ x_0\in\mathbb{R}^n\right\}\tag{4.45}$$

By analogy, the reachability set for the pass profile $y(t)$ of the system (4.14)—(4.15) at the given moment $t^*\in[0,T]$ is defined as

$$\mathcal{R}_y(t^*)=\left\{y\in\mathbb{R}^m : y=y(t^*,g,f,x_0),\ \text{for all}\ \ f\in C^n[-h,0],\ g\in C^m[-h,0],\ x_0\in\mathbb{R}^n\right\}\tag{4.46}$$

For many cases an essential question is: Can one reach the desired state and/or pass profile position by a proper choice of the initial data? The following definition is a formal description of this problem.

Definition 17. *It is said that the system (4.14)—(4.15) is pointwise complete on the interval* $[0,T]$ *if*

$$\mathcal{R}_x(t)=\mathbb{R}^n\ \ \textit{and}\ \mathcal{R}_y(t)=\mathbb{R}^m\ \ \textit{for all}\ \ t\in[0,T].\tag{4.47}$$

If for some $t^*\in[0,T]$ the conditions (4.47) are not true then the system is called pointwise degenerate at the moment t^*.

The notion of pointwise completness was introduced first in [80] for studying the controllability of linear differential time delay systems. Some details and an overview of existing results can be found, also, in the survey [58]. It is obvious that the ordinary linear differential system of the form $\dot{x}(t)=Ax(t)$ is pointwise complete since for any t^* and $x^*\in\mathbb{R}^n$ there exists $x(0)=x_0\in\mathbb{R}^n$ such that the corresponding solution satisfies the condition $x(t^*,x_0)=x^*$. Also, it is proved that each stationary linear differential system with constant time delay is pointwise complete in the case $n = 2$. The following example shows that the presence "difference" equation in the hybrid system destroys the pointwise completness of differential time delay system with $n = 2$.

Example. Consider the hybrid system of (4.14)—(4.15) on the interval $t \in [0, T]$ where $h \leq T \leq 2h,\ n = 2,\ m = 2, h = \ln 2$ and the following choice of the matrices

$$A = \begin{pmatrix} 0 & 2 \\ 0 & 1 \end{pmatrix},\ A_{-1} = \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix},\ B_{-1} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix},\ B_0 = 0,$$
$$C = \begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix},\ C_{-1} = \begin{pmatrix} 0 & 0 \\ -4 & 0 \end{pmatrix},\ D_{-1} = 0,\ B = 0,\ D = 0 \tag{4.48}$$

Substituting the function $y(t)$ from second equation into the first of the system (4.14)—(4.15) corresponding to the given choice of matrices leads to the following time delay system

$$\dot{x}(t) = \begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix} x(t) + \begin{bmatrix} -2 & 0 \\ -1 & 2 \end{bmatrix} x(t-h) + \begin{bmatrix} 0 & 0 \\ -2 & 0 \end{bmatrix} x(t-2h) \tag{4.49}$$

Thus the state variable of the considered hybrid system (4.48) is described by the retarded differential system (4.49) with multiple delays. For simplicity, denote next the matrices involved in (4.49) by $A,\ A_1,\ A_2$, respectively. It is known (see, [58], for example) that the linear stationary differential system with multiple delays is pointwise complete if, and only if, the following conditions

$$rank M^0 = n + n_1, \qquad \text{where } n_1 = \sum_{i=1}^{N} rank M_i(\lambda_i) \tag{4.50}$$

hold. Here the matrices M^0 and $M_i(\lambda_i)$ are defined by spectral parameters of the operator

$$W(\lambda, e^{-\lambda h}) = (\lambda I - A - e^{-\lambda h} A_1 - e^{-2\lambda h} A_2), \qquad \lambda \in \mathbb{C} \tag{4.51}$$

associated with the system (4.49). In the considered case we have

$$W(\lambda, e^{-\lambda h}) = \begin{bmatrix} \lambda + 2e^{-\lambda h} & -2 \\ e^{-\lambda h} + 2e^{-2\lambda h} & -\lambda - 1 - 2e^{-\lambda h} \end{bmatrix}, \quad \det W(\lambda, e^{-\lambda h}) = \lambda^2 - \lambda. \tag{4.52}$$

Hence, the eigenvalues are $\lambda_1 = 0$ and $\lambda = 1$. Further, noting $h = \ln 2$, we have

$$M_1(\lambda_1) = W(\lambda, e^{-\lambda h})|_{\lambda=0} = \begin{bmatrix} 2 & -2 \\ 3 & -3 \end{bmatrix}, \quad M_2(\lambda_2) = W(\lambda, e^{-\lambda h})|_{\lambda=1} = \begin{bmatrix} 2 & -2 \\ 1 & 1 \end{bmatrix} \tag{4.53}$$

and the constant $(n+1)n \times n^2$ (in this case 6×4) matrix M^0 is defined as

$$M^0 = \begin{bmatrix} M_1(\lambda_1) & O \\ O & M_2(\lambda_2) \\ I & I \end{bmatrix}, \quad \text{where } I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{4.54}$$

It is easy to verify that

$$rank M_1 = rank \begin{bmatrix} 2 & -2 \\ 3 & -3 \end{bmatrix} = 1, \qquad rank M_2 = rank \begin{bmatrix} 2 & -2 \\ 1 & -1 \end{bmatrix} = 1$$

and

$$rank M^0 = rank \begin{bmatrix} 2 & -2 & 0 & 0 \\ 3 & -3 & 0 & 0 \\ 0 & 0 & 2 & -2 \\ 0 & 0 & 1 & -1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} = 3.$$

Hence

$$3 = rank\ M^0 < n + n_1 = 4,$$

which immediately shows that the considered system is not pointwise complete for the delay value $h = \ln 2$.

Note that the eigenfuncions corresponding to the given eigenvalues $\lambda_1 = 0,\ \lambda_2 = 1$ are $\phi_1(t) = (1,1)^T,\ \phi_2(t) = (e^t, e^t)^T$. It is obvious that the rank of the fundamental matrix, the entries of which are the given eigenfunctions, is equal 1. Hence the linear space formed by these basic functions is isomorphic to the space $\mathbb{R}$. This means [83] again that the system under consideration is degenerate.

For the hybrid differential-difference systems there exists the links between the pointwise completness and controllability notions. We start here with particular case of the state controllability with respect initial conditions.

Definition 18. *The system (4.14)—(4.15) (with $B = 0,\ D = 0$) is called to be state controllable with respect to initial data at the given moment $t = T$ if for any n-vector $c_T \in \mathbb{R}^n$ there exist the initial functions $g(t),\ f(t),\ t \in [-h, 0]$ such that the corresponding solution $x(t, g, f, x_0)$ of the system (4.14)—(4.15) satisfies the following condition*

$$x(T, g, f, x_0) = c_T \tag{4.55}$$

Now, the following theorem can be stated.

Theorem 19. *[58] The system (4.14)—(4.15) (with $B = 0,\ D = 0$) is state controllable with respect to initial data at the given moment $t = T$ if, and only if,*

i) system (4.14)—(4.15) is pointwise complete;

ii)

$$rank\{H_0^i[H_1, H_2, \ldots, H_{q_T}, G_{q_T}, P_{q_T}],\ i = 0, \ldots, n\} = n \tag{4.56}$$

where the matrices $H_i, i = 0, \ldots, q_T,\ G_{q_T},\ P_{q_T}$ are defined in (4.21).

The proof of the theorem and other results can be found [58] and, hence, the details are omitted here.

By analogy to Definition 18 profile controllability with respect to initial data can be introduced and studied.

4.4 Optimization

In this section the following time optimal control problem for the process (4.14)—(4.15) is considered.

Optimization Problem. *For the given initial data* $x(t) = f(t),\ y(t) = g(t),\ t \in [-h, 0),\ x(0) = x_0$ *find the minimal time* T *and the control function* $u(t),\ t \in [0, T]$ *such that the corresponding trajectory*

$$x(t) \equiv 0,\ t \in [T-h, T] \tag{4.57}$$

is in the equilibrium state.

In effect, the solution of this problem will drive the system dynamics to the zero equilibrium state as fast as possible.

Note that in this case, subject to some additional assumptions, the control function on the last interval $[T-h, T]$ can be represented in the feedback form. In particular, suppose that the matrix B from (4.14) is invertible. Then from (4.14) we have

$$u(t) = -B^{-1}\big[A_{-1}x(t-h) + B_0 y(t) + B_{-1}y(t-h)\big] \tag{4.58}$$

Substituting this into (4.14), and assuming that there exists $\big[E+DB^{-1}B_0\big]^{-1}$ yields finally

$$u(t) = Nx(t-h) + My(t-h),\ t \in [T-h, T] \tag{4.59}$$

where

$$N = \big[E+DB^{-1}B_0\big]^{-1}\big[C_{-1} - DB^{-1}A_{-1}\big],\ M = \big[E+DB^{-1}B_0\big]^{-1}\big[D_{-1} - DB^2_{-1}\big] \tag{4.60}$$

The representation (4.59) shows also that the complete controllability can be solved for the given particular case on the base of relative (pointwise) controllability formulated above. Indeed, if there exists the control function $u(t),\ t \in [0, T-h]$ such that $x(T-h) = 0$ (for the considered case it is sufficient to choose the single point for $\nu = 0,\ \beta_\nu = h$ and $c_\nu = 0$) then the following setting

$$\bar{u}(t) = \begin{cases} u(t),\ t \in [0, T-h) \\ Nx(t-h) + My(t-h),\ t \in [T-h, T] \end{cases}$$

solves the problem of the complete controllability. Note that this approach is however of a limited significance as in the majority cases the state dimension exceeds considerably the input dimension. Then, the semi inverse approach can be applied, which is the subject of ongoing work.

4.4.1 General optimality conditions

Now let T be a fixed time moment. The class of the admissible controls $u(t),\ t \in [0, T]$ is the set of all piecewisse continuous functions such that $u(t) \in U,\ t \in [0, T]$, where U is a compact convex set from $\mathbb{R}^r$. By analogy to (4.28,) the solution of the process (4.14)—(4.15) can be rewritten in the following form

$$\begin{aligned} x(t) =& s(t, f, g, x_0) + \int_0^t S(t, \tau)u(\tau)d\tau, \\ y(t) =& r(t, f, g, x_0) + \sum_{j=0}^{i} G_j u(t-jh) + \int_0^t R(t, \tau)u(\tau)d\tau, \end{aligned} \tag{4.61}$$

where $r(t,f,g,x_0)$ and $R(t,\tau)$ were defined by (4.29) and

$$s(t,f,g,x_0) = F(t,0)x_0 + \sum_{j=1}^{i+1}\int_{-h}^{0} F(t,\tau+(j-1)h)H_j f(\tau)d\tau +$$

$$+\int_{-h}^{0} F(t,\tau+(i+1)h)\big[P_{i+1}f(\tau)+Q_{i+1}g(\tau)\big]d\tau, \tag{4.62}$$

$$S(t,\tau) = \sum_{j=1}^{i+1} F(t,\tau+(j-1)h)V_j,\ i=\left[\frac{T}{h}\right]$$

Definition 19. *We say that the control function $u(t),\ t\in[0,T]$ is T-admissible for the system (4.14)—(4.15), if the corresponding trajectory satisfies the following condition*

$$x(t)\equiv 0,\ t\in[T-h,T] \tag{4.63}$$

Introduce

$$Z=\left\{x\in\mathbb{R}^n \mid x=s(T-h,f,g,x_0) \text{ for all } (f,g,x_0)\in C_{[-h,0]}\times C_{[-h,0]}\times\mathbb{R}^n\right\}$$

and

$$\mathcal{R}=\left\{s\in Z \mid \text{ such that for } x=s\ \exists \text{ T-admissible control } u(\cdot)\right\}. \tag{4.64}$$

In fact, the set $\mathcal{R}$ is the reachability set for the system (4.14)—(4.15) with (4.63) in place. We assume that $\mathcal{R}\neq\emptyset$, which is true if the system is controllable. Equivalently, we suppose that there exists at least one collection of the initial data

$$x(t)=f(t),\ t\in[-h,0),\ x(0)=x_0,\ y(t)=g(t),\ t\in[-h,0]$$

for which there exists the $T-$ addmisible control functions. Denote next by $U_T(\cdot)$ the set of the all T-admissible control vectors for the system (4.14)—(4.15) corresponding to the set $\mathcal{R}$. Then it is easy to show that $\mathcal{R}$ is closed and convex.

Theorem 20. *For the given initial data $f(t),\ g(t),\ t\in[-h,0), x(0)=x_0$ there exists T-admissible control if, and only if, the following inequality*

$$\max_{\|g\|=1}\left\{g^T s(T-h,f,g,x_0)+\inf_{u\in U_T(\cdot)}\int_0^{T-h} g^T S(T-h,\tau)u(\tau)d\tau\right\}\le 0 \tag{4.65}$$

holds.

Proof. *Necessity.* Let Optimization Problem is solvable for the moment T and $u(t),\ t\in[0,T]$ is a T-admissible control function. This means here that

$$0=x(T-h)=s(T-h,f,g,x_0)+\int_0^{T-h} S(T-h,\tau)u(\tau)d\tau.$$

Multiplying the both sides of the last equality by the vector $g \in \mathbb{R}^n$ yields

$$g^T s(T-h, f, g, x_0) + \int_0^{T-h} g^T S(T-h, \tau) u(\tau) d\tau = 0.$$

Hence

$$g^T s(T-h, f, g, x_0) + \inf_{u \in U_T(\cdot)} \int_0^{T-h} g^T S(T-h, \tau) u(\tau) d\tau \leq 0$$

and (4.65) holds.

Sufficiency. Let the inequality (4.65) holds for the given initial data $(f,\ g, x_0)$. On contrary, assume that for this data there is no any T-admissible control $u(\cdot)$ which solves the problem. This means that the corresponding vector $s^* = s^*(T-h, f,\ g, x_0) \notin \mathcal{R}$ does not belong to the set $\mathcal{R}$, i. e. $s^*(T-h, f,\ g, x_0) \notin \mathcal{R}$. Since $\mathcal{R}$ is a convex set then there exists a supporting hyperplane with the nontrivial normal vector $g* \in \mathbb{R}^n,\ \|g^*\| = 1$ such that the following inequality

$$g^{*T} s^*(t, f, g, x_0) > g^{*T} s, \qquad \forall \ \ s \in \mathcal{R} \tag{4.66}$$

holds. Since $s \in \mathcal{R}$ then there exists a $T-$ admissible control function $u(t),\ t \in [0, T-h]$ such that

$$s + \int_0^{T-h} S(T-h, \tau) u(\tau) d\tau = 0.$$

Hence, (4.66) yields that

$$g^{*T} s^*(T-h, f, g, x_0) + \int_0^{T-h} g^{*T} S(T-h, \tau) u(\tau) d\tau > 0$$

and since s is an arbitrary vector from the set $\mathcal{R}$ then the last inequality is true for all $u(\cdot) \in U_T(\cdot)$. Therefore

$$g^{*T} s^*(T-h, f, g, x_0) + \inf_{u \in U_T(\cdot)} \int_0^{T-h} g^{*T} S(T-h, \tau) u(\tau) d\tau > 0$$

which contradicts (4.65). ■

Next, denote

$$\Lambda(T) = \max_{\|g\|=1} \left\{ g^T s(T-h, f, g, x_0) + \inf_{u \in U_T(\cdot)} \int_0^{T-h} g^T S(T-h, \tau) u(\tau) d\tau \right\}. \tag{4.67}$$

It can be shown that $\Lambda(T)$ is a non decreasing, continuous function, and hence we have the result bellow for which we also require the following definition.

Definition 20. *Let $p : \mathbb{R} \to \mathbb{R}$ be a function. Then we say that $Z^0 \in \mathbb{R}$ is the minimal root of equation $p(z) = 0$ if $p(z^0) = 0$ and there is no $z^* \in \mathbb{R}$ such that $z^* < z^0$ and $p(z^*) = 0$.*

Theorem 21. *Given initial data* $f(t), g(t),\ t \in [-h, 0), x(0) = x_0$, *the moment* T^0 *is optimal if, and only if,* T^0 *is a minimal root of the equation*

$$\Lambda(T) = 0. \tag{4.68}$$

Proof. *Necessity.* Let $u^0(\cdot)$ be the optimal control for optimization Problem. Then Theorem 20 gives $\Lambda(T^0) \le 0$. At first, suppose that $\Lambda(T^0) < 0$. Since $\Lambda(T)$ is a non decreasing and continuous function than $\exists \bar{T},\ \bar{T} < T^0$ such that $\Lambda(T^0) \le \Lambda(\bar{T}) \le 0$. In accordance with Theorem 20, optimization Problem is solvable with $\bar{T} < T^0$ which is impossible. Thus, T^0 is a root for the equation (4.68). The minimality of T^0 can be shown analogously.

Sufficiency. Let for the control function $u^0(t),\ t \in [0, T^0 - h]$ T^0 is the minimal root of $\Lambda(T) = 0$. Suppose now that this control function is not optimal for the given initial data. Hence, there is the $\bar{T}$-admissible control function $\bar{u}(t),\ t \in [0, \bar{T} - h]$ where $\bar{T} < T^0$. Then Theorem 20 yields $\Lambda(\bar{T}) \le 0$. On the other hand, noting non decreasing the function $\Lambda(T)$, we have $\Lambda(\bar{T}) \ge \Lambda(T^0) = 0$, which contradicts the minimality of the root T^0, which completes the proof. ■

Finally, the optimal time T^0 is given by the equality (4.68) and the optimal control function $u^0(t)$ is determined as

$$\min_{u \in U_T(\cdot)} \int_0^{T-h} g^{0T} S(T-h, \tau) u(\tau) d\tau = \int_0^{T-h} g^{0T} S(T-h, \tau) u^0(\tau) d\tau \tag{4.69}$$

where g^0 is the vector which maximizes (4.67).

These optimality conditions can be presented in a more practically usable form for some particular sets of admissible controls $U(\cdot)$. In the next section, for example, the time optimal control problem subject to integral control constraints is solved.

4.4.2 Time optimal problem subject to integral control constraints

Consider the following optimization problem: Minimize

$$T \longrightarrow \min_{u \in U(\cdot)} \tag{4.70}$$

over the solutions of the process

$$\begin{aligned} \dot{x}(t) =& Ax(t) + A_{-1}x(t-h) + B_0 y(t) + B_{-1} y(t-h) + Bu(t) \\ y(t) =& Cx(t) + C_{-1}x(t-h) + D_{-1}y(t-h) + Du(t),\ t \in [0, T] \end{aligned} \tag{4.71}$$

with initial conditions

$$x(t) = f(t),\ t \in [-h, 0),\ x(0) = x_0,\ y(t) = g(t),\ t \in [-h, 0] \tag{4.72}$$

subject to state constraints

$$x(t) \equiv 0,\ t \in [T-h, T] \tag{4.73}$$

and integral control constraints

$$U(\cdot) \triangleq \left\{ u(\cdot) : \int_0^T u^T(\tau) u(\tau) d\tau \le 1 \right\} \tag{4.74}$$

In accordance with Theorem 21 optimal time control function has to satisfy the equation $\Lambda(T) = 0$ where the function $\Lambda(T)$ is defined in (4.67). Noting (4.62) reduces our task to calculate minimum in (4.67) as

$$\mathcal{M}(g) \triangleq \int_0^{T-h} g^T \sum_{j=1}^{i+1} F(T-h, \tau + (j-1)h) V_j u(\tau) d\tau \longrightarrow \inf_{u \in U_T(\cdot)} \tag{4.75}$$

subject to

$$\int_0^T u^T(\tau) u(\tau) dt \leq 1 \tag{4.76}$$

Using (4.59) and (4.61) allows rewriting (4.76) as

$$\int_0^{T-h} u^T(\tau) u(\tau) d\tau + \int_{T-h}^{T} \big(Nx(\tau-h) + My(\tau-h)\big)^T \big(Nx(\tau-h) + My(\tau-h)\big) d\tau =$$

$$= \Upsilon + \int_0^{T-h} u^T(\tau) \big[I_m + \mathcal{G}(\tau)\big]^T u(\tau) d\tau + \int_0^{T-h} \big[\psi(\tau) + \varphi(\tau)\big]^T u(\tau) d\tau + \tag{4.77}$$

$$+ \int_0^{T-h} \int_0^{T-h} u^T(\tau) \big[\Psi(\tau,\theta) + \Phi(\tau,\theta)\big] u(\theta) d\tau \leq 1$$

where I_m is the identity $(m \times m)$-matrix, and

$$\Upsilon = \int_{T-h}^{T} \big[Ns(\tau-h) + Mr(\tau-h)\big]^T \big[Ns(\tau-h) + Mr(\tau-h)\big] dt$$

$$\psi(\tau) = \int_{T-2h}^{T-h} \Big\{ S^T(\theta,\tau) N^T \big[Ns(\theta) + Mr(\theta)\big] + \big[s^T(\theta) N^T + r^T(\theta) M^T\big] MR(\theta,\tau) \Big\} d\theta$$

$$\Phi(\theta,\tau) = \int_{T-2h}^{T-h} \Big\{ S^T(t,\tau) N^T \big[NS(t,\theta) + MR((t,\theta)\big] + R^T(t,\tau) M^T MR((t,\theta) \Big] d\tau,$$

$$\Psi(\tau,\theta) = \begin{cases} \big[S^T(\theta,\tau)N^T + R^T(\theta,\tau)M^T\big] MG_0, & \tau \in (T-2h, T-h] \\ \big[S^T(\theta+h,\tau)N^T + R^T(\theta+h,\tau)M^T\big] MG_1, & \tau \in (T-3h, T-2h] \\ \dots\dots\dots\dots\dots & \dots \\ \big[S^T(\theta+(q_T-1)h,\tau)N^T + R^T(\theta+(q_T-1)h,\tau)M^T\big] MG_{q_T-1}, & \tau \in [0,h], \end{cases}$$

$$\varphi(\tau)=\begin{cases}\big(s^T(\tau)N^T+r^T(\tau)M^T\big)MG_0, & \tau\in(T-2h,T-h]\\ \big(s^T(\tau+h)N^T+r^T(\tau+h)M^T\big)MG_1, & \tau\in(T-3h,T-2h]\\ \dots\dots\dots\dots\dots & \dots\\ \Big(s^T(\tau+(q_T-1)h)N^T+r^T(\tau+(q_T-1)h)M^T\Big)MG_{q_T-1}, & \tau\in[0,h]\end{cases} \tag{4.78}$$

and

$$\mathcal{G}(\tau)=\begin{cases}G_0^TM^TMG_0+G_0M^TMG_1e^{-ph}+\dots & \\ +G_0^TM^TMG_{g_T-1}e^{-(q_T-1)ph}, & \tau\in(T-2h,T-h]\\ G_1^TM^TMG_1+G_1M^TMG_2e^{-ph}+\dots & \\ +G_1^TM^TMG_{g_T-2}e^{-(q_T-2)ph}, & \tau\in(T-3h,T-2h]\\ \dots\dots\dots\dots\dots & \dots\\ G_{g_T-2}^TM^TMG_{g_T-2}+G_{g_T-2}^TM^TMG_{g_T-1}e^{-ph}, & \tau\in(h,2h]\\ G_{g_T-1}^TM^TMG_{g_T-1} & \tau\in[0,h]\end{cases}$$

Here e^{-kph} denotes the shift operator such that $(e^{-kph}u)(\tau)=u(\tau-kh)$. Using the Lagrange multiplier method leads to the functional

$$\Pi(u)=\int_0^{T-h}g^TS(T-h,\tau)u(\tau)d\tau+\lambda\Bigg\{\Upsilon+\int_0^{T-h}u^T(\tau)\big[I_m+\mathcal{G}(\tau)\big]^Tu(\tau)d\tau+$$
$$+\int_0^{T-h}\big[\psi(\tau)+\varphi(\tau)\big]^Tu(\tau)d\tau+\int_0^{T-h}\int_0^{T-h}u^T(\tau)\big[\Psi(\tau,\theta)+\Phi(\tau,\theta)\big]u(\theta)d\tau\Bigg\} \tag{4.79}$$

which is subject of minimization with respect to unknown λ and $u(t)$. Now, it is to find the first variation $\delta\Pi$ for $\Pi(u)$, which can be represented as

$$\delta\Pi(u)=\frac{\partial\Pi(u+\alpha v)}{\partial\alpha}|_{\alpha=0}=\int_0^{T-h}v^T(\tau)S^T(T-h,\tau)gd\tau+$$
$$+\int_0^{T-h}\lambda\Bigg\{u^T(\tau)\big[I_m+\mathcal{G}(\tau)\big]^Td\tau+\big[\psi(\tau)+\varphi(\tau)\big]^T+ \tag{4.80}$$
$$+\int_0^{T-h}(\tau)K(\theta,\tau)u(\theta)d\theta\Bigg\}^Tv(\tau)d\tau$$

where

$$K(\theta,\tau)=(\Psi(\tau,\theta)+\Phi(\tau,\theta))+(\Psi^T(\tau,\theta)+\Phi^T(\tau,\theta))$$

Since $\delta\Pi(u) = 0 \quad \forall\, v(\tau)$ for the optimal solution then (4.80) yields

$$S^T(T-h,\tau)g + \lambda\left\{ 2\big[I_m + \mathcal{G}(\tau)\big]u(\tau) + \psi(\tau) + \varphi(\tau) + \int\limits_0^{T-h} K(\theta,\tau)u(\theta)d\theta \right\} = 0 \quad (4.81)$$

The solution of (4.81) can be represented as the following sum

$$u_g(t) = u_1(t) + u_2(t), \qquad \text{where} \quad u_2(t) = \frac{1}{\lambda}L(t)g. \quad (4.82)$$

Here the vector $u_1(t)$ and the $(n \times r)$- matrix $L(t)$ satisfy the following integral equations

$$2u_1(t)(I + \mathcal{G}(t)) + \psi(t) + \varphi(t) + \int\limits_0^{T-h} K(\theta,t)u_1(\theta)d\theta = 0 \quad (4.83)$$

and

$$2L(t) + S(T-h,t) + \int\limits_0^{T-h} K(\theta,t)L(\theta)d\theta = 0 \quad (4.84)$$

To show this it is sufficient to substitute (4.100) into (4.81), which gives

$$S^T(T-h,\tau)g + \lambda\left\{ 2(u_1(t) + \frac{1}{\lambda}L(t)g))\big[I_m + \mathcal{G}(\tau)\big] + \psi(\tau) + \varphi(\tau) + \right.$$

$$\left. + \int\limits_0^{T-h} u^T(\tau)K(\theta,\tau)(u_1(\theta) + \frac{1}{\lambda}L(\theta)g)d\theta \right\} = \Big(S^T(T-h,\tau) + L(\tau) +$$

$$+ \int\limits_0^{T-h} K(\theta,\tau)L(\theta)d\theta \Big) g + \lambda\left[2u_1(\tau)(I_m + \mathcal{G}(\tau)) + \psi(\tau) + \varphi(\tau) + \int\limits_0^{T-h} u^T(\tau)K(\theta,\tau)u_1(\theta) \right] = 0$$

The unknown multiplier λ can be determined by that the required control function belongs to the admissible set $U(\cdot)$, i. e. $\int\limits_0^T u^T(\tau)u(\tau)d\tau = 1$. Hence

$$\Upsilon + \int\limits_0^{T-h} \big[u_1(\tau) + \frac{1}{\lambda}L(\tau)g)\big]^T \big(I_m + \mathcal{G}(\tau)\big)\big[u_1(\tau) + \frac{1}{\lambda}L(\tau)g)\big]d\tau =$$

$$+ \int\limits_0^{T-h} \big(\psi(\tau) + \varphi(\tau)\big)^T \big[u_1(\tau) + \frac{1}{\lambda}L(\tau)g)\big]d\tau + \quad (4.85)$$

$$+ \int\limits_0^{T-h}\int\limits_0^{T-h} \big[u_1(\tau) + \frac{1}{\lambda}L(\tau)g)\big]^T \big(\Psi(\theta,\tau) + \Phi(\theta,\tau)\big)\big[u_1(\tau) + \frac{1}{\lambda}L(\tau)g)\big]d\theta d\tau = 1$$

This leads to the following equation for λ

$$a\frac{1}{\lambda^2} + 2b\frac{1}{\lambda} + c = 0 \quad (4.86)$$

where the required coefficients are

$$a = \int_0^{T-h} g^t L(\tau)^T L(\tau) g d\tau + \int_0^{T-h}\int_0^{T-h} g^T L^T(\tau) K(\theta,\tau) L(\theta) d\theta d\tau = -\tfrac{1}{2}\int_0^{T-h} S^T(T-h,\tau) L(\tau) d\tau,$$

$$b = \int_0^{T-h} u_1^T(\tau) L(\tau) g d\tau + \int_0^{T-h} [\psi(\tau)+\varphi(\tau)] L(\tau) g d\tau + \int_0^{T-h}\int_0^{T-h} g^T L^T(\tau) K(\theta,\tau) u_1(\theta) d\tau d\theta,$$

$$c = \Upsilon - 1 + \int_0^{T-h} u_1^T(\tau)\big(I+\mathcal{G}(\tau)\big) u_1(\tau) d\tau + 2\int_0^{T-h} (\psi(\tau)+\varphi(\tau)) u_1(\tau) d\tau +$$

$$+ \int_0^{T-h}\int_0^{T-h} u_1^T(\tau)\big[\Phi(\theta,\tau)+\Psi(\theta,\tau)\big] u_1(\theta) d\tau d\theta = \Upsilon - 1 + \int_0^{T-h} \big(\psi(\tau)+\varphi(\tau)\big) u_1(\tau) d\tau.$$

Thus the required λ is the positive root of the equation (4.86), and the optimal control for the given T then is defined by (4.100). Substituting the obtained control function $u_g(t)$ of (4.100) into the basic condition (4.67) and noting Theorem 21 reduce the time optimisation problem to the following:

find the minimal root T^0 for the equation

$$\max_{\|g\|=1} \mathcal{L}(g,T) = 0 \tag{4.87}$$

where

$$\mathcal{L}(g,T) = g^T s(T-h,f,g,x_0) + \int_0^{T-h} g^T S(T-h,\tau) u_g(\tau) d\tau$$

and the function $u_g(t)$ is given by (4.100).

Hence, the following theorem has been proved

Theorem 22. *Optimal time T^0 in optimisation problem (4.70)—(4.74) is the minimal root of the equation (4.104) and the corresponding optimal control is*

$$u^0(t) = \begin{cases} u_{g^0}(t), \ t \in [0, T^0-h) \\ Nx^0(t-h) + My^0(t-h), \ t \in [T^0-h, T^0] \end{cases} \tag{4.88}$$

where the vector g^0 realizes maximum in (4.104), the function $u_g(t)$ is given by (4.100) and the matrices M, N are defined by (4.4.3).

4.4.3 Illustrative Examples

In order to demonstrate our approach consider the following test example.

Example 1. Consider the time delayed differential equation with control input

$$\dot{x}(t) = -x(t-\frac{\pi}{2}) + u(t), \quad t \in [0, \frac{3\pi}{2}] \tag{4.89}$$

with the initial data

$$x(t) \equiv 0, \quad t \in [-\frac{\pi}{2}, 0], \quad x(0) = 1 \tag{4.90}$$

The considered control system is a particular case of the introduced differential-algebraic system (4.71), where

$$A_{-1} = 1, \quad B = 1, \quad h = \frac{\pi}{2}, \quad T = \frac{3\pi}{2}$$

and others coefficients in (4.71) are zero.

Consider the following optimization problem: minimize the cost functional

$$J(u) \to \min_u, \quad J(u) = \int_0^{\frac{3\pi}{2}} u^2(t)dt \tag{4.91}$$

over the solution of (4.89)—(4.90) and subject to

$$x(t) \equiv 0, \quad t \in [\pi, \frac{3\pi}{2}] \tag{4.92}$$

Remark 19. *Let $M^0 = J(u^0)$ is the optimal cost value for the problem (4.89)—(4.92).*

Consider the following time optimal problem: minimize

$$T \to \min \tag{4.93}$$

over the solutions of the control system

$$\dot{x}(t) = -x(t - \frac{\pi}{2}) + u(t), \ t \in [0, T] \tag{4.94}$$

with the initial conditions (4.90) and the constraints of the form

$$\int_0^T u^2(t)dt \leq M^0 \tag{4.95}$$

subject to

$$x(t) \equiv 0, \quad t \in [T - \frac{\pi}{2}, T] \tag{4.96}$$

It is easy to show that the optimal solution for the problem (4.93)—(4.96) is $T^0 = \frac{3}{2}\pi$. Hence, in some sense these optimization problems are equivalent.

Next, find the fundamental solution $F(t, \tau)$ for the differential equation (4.24) that in this case is

$$\frac{\delta F(t, \tau)}{\delta \tau} = F(t, \tau + \frac{\pi}{2}), \quad F(t, \tau) \equiv 0, \quad \tau > t, \quad F(t, t) = 1. \tag{4.97}$$

It is easy to check that the function

$$F(t, \tau) = \begin{cases} e^{-i(t-\tau)}, & if \quad \tau \leq t, \\ 0, & if \quad \tau > 0, \end{cases} \tag{4.98}$$

satisfies (4.97) where i means imaginary unit ($i^2 = -1$). Thus, the solution of the system (4.89) with the initial data (4.90) is given for $t \in [0, \frac{3\pi}{2}]$ as

$$x(t) = F(t, 0)x(0) + \int_0^t F(t, \tau)u(\tau)d\tau = e^{-it} + \int_0^t e^{-i(t-\tau)}u(\tau)d\tau = e^{-it}(1 + \int_0^t e^{i(\tau)}u(\tau)d\tau) \tag{4.99}$$

Problem statement says that we consider the real valued functions. Since in accordance with Euler formula $e^{i\varphi} = \cos\varphi + i\sin\varphi$, then from (4.99) we can extract the real part of $x(t)$ when it is necessary. In order to determine the optimal solution of the problem under consideration we need to calculate the function $\Phi(\tau,\theta)$ and $K(\tau,\theta)$, which are required in the formulas (4.84) — (4.85) and the constants γ, a, b, c also. Note that in our case

$$H_1 = 0,\ H_2 = 1,\ H_j = 0, V_1 = 1,\ V_j = 0,\ P_j =,\ M_j = 0$$

and hence the function $\varphi(\tau) = 0$, and

$$\psi(t) = \int_{\frac{\pi}{2}}^{\pi} e^{-i(t-\tau)}x(0)e^{-it}dt = e^{i\tau}\int_{\frac{\pi}{2}}^{\pi} e^{-2it}dt = e^{i\tau}(\frac{1}{-2i})e^{-2it}\Big|_{\frac{\pi}{2}}^{\pi} =$$

$$\frac{e^{i\tau}}{2i}[e^{-2\pi i} - e^{-\pi i}] = \frac{e^{i\tau}}{2i}[\cos 2\pi - i\sin 2\pi - (\cos\pi - i\sin\pi)] = \frac{e^{i\tau}}{i}[1+1] = -ie^{i\tau}.$$

Further

$$\frac{1}{2}K(\tau,\theta) = \int_{\frac{\pi}{2}}^{\pi} F(t,\tau)F(t,\theta)dt = \int_{\frac{\pi}{2}}^{\pi} e^{i(\tau-t)}e^{i(\theta-t)}dt =$$

$$= \begin{cases} 0, & if \quad 0 \le \tau,\ \ \theta \le \frac{\pi}{2} \\ \int_{\tau}^{\pi} e^{i(\tau-t)}e^{i(\theta-t)}dt, & if \quad \tau \ge \theta,\ \ \frac{\pi}{2} \le \tau, \theta \le \frac{3\pi}{2} \\ \int_{\theta}^{\pi} e^{i(\tau-t)}e^{i(\theta-t)}dt, & if \quad \tau > \theta,\ \ \frac{\pi}{2} \le \tau, \theta \le \frac{3\pi}{2}, \end{cases} =$$

$$= \begin{cases} 0, & if \quad 0 \le \tau,\ \ \theta \le \frac{\pi}{2} \\ e^{i(\tau+\theta)}(\frac{1}{-2i})e^{2it}\Big|_{\tau}^{\pi} = e^{i(\tau+\theta)}[e^{2\pi i} - e^{2i\tau}], & if \quad \tau \ge \theta,\ \ \frac{\pi}{2} \le \tau, \theta \le \frac{3\pi}{2} \\ e^{i(\tau+\theta)}(-\frac{1}{2i})e^{2it}\Big|_{\theta}^{\pi} = e^{i(\tau+\theta)}[e^{2\pi i} - e^{2i\theta}], & if \quad \tau > \theta,\ \ \frac{\pi}{2} \le \tau, \theta \le \frac{3\pi}{2}, \end{cases} =$$

$$= \begin{cases} 0, \\ e^{i(\tau+\theta)}[\cos(-2\pi) - i\sin(2\pi) - (\cos 2\tau - i\sin 2\tau)] = e^{i(\tau+\theta)} - e^{i(\theta-\tau)}, \\ e^{i(\tau+\theta)}[1 - e^{-2i\theta}] = e^{i(\tau+\theta)} - e^{i(\tau-\theta)}, \end{cases} =$$

$$= \begin{cases} 0, & if \quad 0 \le \tau,\ \ \theta \le \frac{\pi}{2} \\ e^{i\theta}(e^{i\tau} - e^{-i\tau}), & if \quad \tau \ge \theta,\ \ \frac{\pi}{2} \le \tau, \theta \le \frac{3\pi}{2} \\ e^{i\tau}(e^{i\theta} - e^{-i\theta}), & if \quad \tau > \theta,\ \ \frac{\pi}{2} \le \tau, \theta \le \frac{3\pi}{2}, \end{cases} = \begin{cases} 0, & if \quad 0 \le \tau,\ \ \theta \le \frac{\pi}{2} \\ e^{i\theta}2i\sin\tau, & if \quad \tau \ge \theta,\ \ \frac{\pi}{2} \le \tau, \theta \le \frac{3\pi}{2} \\ e^{i\tau}2i\sin\theta, & if \quad \tau > \theta,\ \ \frac{\pi}{2} \le \tau, \theta \le \frac{3\pi}{2}. \end{cases}$$

It is easy to see that $K(\tau,\theta) = K(\theta,\tau)$.

The statement of optimization problem is given in the real valued functions terms. Hence, we are needed to pick the real parts in the obtained functions. Thus

$$\psi(\tau) = -ie^{i\tau} = \mathrm{Re}\psi(\tau) + i\cdot\mathrm{Im}\psi(\tau), \quad \text{where } \mathrm{Re}\psi(\tau) = \sin\tau,\ \mathrm{Im}\psi(\tau) = -\cos\tau,$$

and

$$K(\tau,\theta)=\begin{cases}0, & if \quad 0\le\tau,\ \ \theta\le\frac{\pi}{2}\\ e^{i\theta}i\sin\tau, & if \quad \tau\ge\theta,\ \ \frac{\pi}{2}\le\tau,\theta\le\frac{3\pi}{2}\\ e^{i\tau}i\sin\theta, & if \quad \tau>\theta,\ \frac{\pi}{2}\le\tau,\theta\le\frac{3\pi}{2}\end{cases} =$$

$$=\begin{cases}0, & if \quad 0\le\tau,\ \ \theta\le\frac{\pi}{2}\\ (\cos\theta+i\cdot\sin\theta)i\sin\tau, & if \quad \tau\ge\theta,\ \ \frac{\pi}{2}\le\tau,\theta\le\frac{3\pi}{2}\\ (\cos\tau+i\cdot\sin\tau)i\sin\theta, & if \quad \tau>\theta,\ \frac{\pi}{2}\le\tau,\theta\le\frac{3\pi}{2}\end{cases} =$$

$$=\begin{cases}0, & if \quad 0\le\tau,\ \ \theta\le\frac{\pi}{2}\\ -\sin\theta\sin\tau+i\cdot\cos\theta\sin\tau, & if \quad \tau\ge\theta,\ \ \frac{\pi}{2}\le\tau,\theta\le\frac{3\pi}{2}\\ -\sin\tau\sin\theta+i\cos\tau\sin\theta, & if \quad \tau>\theta,\ \frac{\pi}{2}\le\tau,\theta\le\frac{3\pi}{2}\end{cases}$$

Thus the real part of the function K is

$$K(\tau,\theta)=-\sin\theta\sin\tau \ \text{ for all } \ \frac{\pi}{2}\le\tau,\theta\le\frac{3\pi}{2}$$

According to the formulas (4.100) the optimal control function is represented as follows

$$u_g(t)=v(t)+w(t),\ t\in[0,\frac{3\pi}{2}], \qquad \text{where} \quad w(t)=\frac{1}{\lambda}L(t)g. \tag{4.100}$$

Here the scalar function $v(t)$ satisfies the following integral equation

$$2v(\tau)+2\psi(\tau)+\int_0^{\pi}K(\tau,\theta)v(\theta)d\theta=0,\ \tau\in[0,\frac{3\pi}{2}]. \tag{4.101}$$

1) If $\tau\in[0,\frac{\pi}{2}]$ then $K(\tau,\theta)\equiv 0$ and hence from (4.100) follows that

$$v(\tau)=\varphi(\tau), \quad \tau\in[0,\frac{\pi}{2}].$$

It should be noted that, here we are considered only real valued functions, and hence

$$v(\tau)=\text{Re}\psi(\tau)=\sin\tau, \quad for \ \ \tau\in[0,\frac{\pi}{2}].$$

2) If $\tau\in[\frac{\pi}{2},\frac{3\pi}{2}]$ then for the unknown function $v(\tau)$ we have the following integral equation

$$v(\tau)+\sin\tau-\int_0^{\pi}\sin\tau\sin\theta\cdot v(\theta)d\theta=0$$

or

$$\sin\tau\int_0^{\pi}\sin\theta v(\theta)d\theta-v(\tau)=\sin\tau. \tag{4.102}$$

Denote $\int\limits_0^\pi \sin\theta v(\theta)d\theta \doteq A$. Multiplying (4.102) by $\sin\tau$ and integrating then the obtained relation with respect τ over the interval $[0,\pi]$, we have

$$\int\limits_0^\pi \sin^2\tau d\tau \cdot \int\limits_0^\pi \sin\theta v(\theta)d\theta - \int\limits_0^\pi \sin\tau v(\tau)d\tau = \int\limits_0^\pi \sin^2\tau d\tau.$$

Noting that $\int\limits_0^\pi \sin\tau v(\tau)d\tau = A$ also, we obtain the following algebraic equation with respect to the unknown value A:

$$\int\limits_0^\pi \sin^2\tau d\tau \cdot A - A = \int\limits_0^\pi \sin^2\tau d\tau.$$

Since

$$\int\limits_0^\pi \sin^2\tau d\tau = \int\limits_0^\pi \frac{1-\cos 2\tau}{2} d\tau = \frac{1}{2}\tau\Big|_0^\pi - \frac{1}{4}\sin 2\tau\Big|_0^\pi = \frac{\pi}{2} - 0 = \frac{\pi}{2}$$

then

$$A(\frac{\pi}{2}-1) = \frac{\pi}{2} \quad \Rightarrow A = \frac{\frac{\pi}{2}}{(\frac{\pi}{2}-1)} = \frac{2\pi}{2(\pi-2)} = \frac{\pi}{\pi-2}.$$

Hence $A = \int\limits_0^\pi \sin\theta v(\theta)d\theta = \frac{\pi}{\pi-2}$. Then from (4.102) it follows

$$v(\tau) = \sin\tau A - \sin\tau = (A-1)\cdot\sin\tau = [\frac{\pi}{\pi-2} - 1]\sin\tau = \frac{2}{\pi-2}\sin\tau.$$

Thus, the asked function $v(\tau)$ is

$$v(\tau) = \begin{cases} sin\tau, & for \quad \tau \in [0, \frac{\pi}{2}), \\ \frac{2}{\pi-2}\sin\tau, & for \quad \tau \in [\frac{\pi}{2}, \frac{3\pi}{2}] \end{cases}$$

By analogue with $v(\tau)$ we can find the function $L(\tau)$ that satisfies to the following integral equation:

$$2L(\tau) + \int\limits_0^\pi K(t,\theta)L(\theta)d\theta + S(\frac{3\pi}{2} - \frac{\pi}{2}, \tau) = 0$$

In our case

$$S(\frac{3\pi}{2} - \frac{\pi}{2}, \tau) = F(\pi,\tau)\cdot 1 = \begin{cases} e^{-i(\pi-\tau)}, & if \quad \tau \le \pi, \\ 0, & if \quad \tau > \pi \end{cases} = \begin{cases} e^{-i\pi}e^{i\tau}, & if \quad \tau \le \pi, \\ 0, & if \quad \tau > \pi \end{cases} =$$

$$= \begin{cases} (\cos\tau + i\sin\tau)(\cos(-\pi) + i\sin\pi), & if \quad \tau \le \pi \\ 0, & if \quad \tau > \pi \end{cases} = \begin{cases} -\cos\tau - i\sin\tau, & if \quad \tau \le \pi, \\ 0, & if \quad \tau > \pi \end{cases}$$

Hence, the real part is

$$S(\pi,\tau) = \begin{cases} -\cos\tau, \tau \le \pi \\ 0, \quad \tau > \pi. \end{cases}$$

Thus the integral equation with respect to $L(\tau)$ is

$$L(\tau) - \int_0^\pi \sin\tau \sin\theta L(\theta)d\theta - \cos\tau = 0, \quad \text{for } 0 \le \tau \le \pi$$

and

$$L(\tau) - \int_0^\pi \sin\tau \sin\theta L(\theta)d\theta + 0 = 0, \quad \text{for} \frac{3\pi}{2} > \tau > \pi.$$

For $\frac{3\pi}{2} > \tau > \pi$ we have the following equation

$$L(\tau) - \sin\tau \int_0^\pi \sin\theta L(\theta)d\theta = 0$$

Multiplying by $\sin\tau$ and then integrating from 0 to π yields

$$\int_0^\pi \sin\tau L(\tau)d\tau - \int_0^\pi \sin^2\tau d\tau \cdot \int_0^\pi \sin\theta L(\theta)d\theta = 0.$$

Denoting $\int_0^\pi \sin\tau L(\tau)d\tau = B$ and since $\int_0^\pi \sin^2\tau d\tau = \frac{\pi}{2}$, then we have the algebraic equation with respect to B

$$B(1 - \frac{\pi}{2}) = 0 \Rightarrow \;B = 0.$$

Then from the equation (4.103) we obtain

$$L(\tau) = 0, \quad \tau \in [\pi, \frac{\pi}{\frac{3\pi}{2}}].$$

Let now $\tau \in [0, \pi]$:

$$L(\tau) - sin\tau \int_0^\pi \sin\theta L(\theta)d\theta = \cos\tau, \quad \tau \in [0, \pi] \tag{4.103}$$

Multiplying by $\sin\tau$ and then integrating from 0 to π:

$$\int_0^\pi \sin\tau L(\tau)d\tau - \int_0^\pi \sin^2\tau d\tau \int_0^\pi \sin\theta L(\theta)d\theta = \int_0^\pi \cos\tau \sin\tau d\tau$$

or

$$B - \frac{\pi}{2}B = \frac{1}{2}\int_0^\pi \sin 2\tau d\tau,$$

Since

$$\int_0^\pi \sin 2\tau d\tau = -\frac{1}{2}\cos 2\tau\Big|_0^\pi = -\frac{1}{2}[1-1] = 0,$$

then $B(1-\frac{\pi}{2})=0 \Rightarrow B=0$. Then from the equation (4.103) we have $L(\tau)=\cos\tau,\ \tau\in[0,\pi]$. Thus

$$L(\tau)=\begin{cases}\cos\tau, \tau\in[0,\pi]\\ 0, \quad \tau\in[\pi,\frac{3\pi}{2}].\end{cases}$$

Hence the optimal control is given as

$$u_g(t)=v(t)+\frac{1}{\lambda}L(t)g$$

where the parameter λ is determined as a positive solution of the following algebraic equation

$$a\frac{1}{\lambda^2}+2b\frac{1}{\lambda}+c=0$$

where

$$a=-\frac{1}{2}\int_0^{T-h}S(T-h,\tau)L(\tau)d\tau=-\frac{1}{2}\int_0^{\pi}\cos\tau\cos\tau d\tau=$$

$$=-\frac{1}{2}\int_0^{\pi}\cos^2\tau d\tau=-\frac{1}{2}\int_0^{\pi}\frac{1+\cos 2\tau}{2}d\tau=-\frac{1}{4}[\tau+\sin 2\tau]\Big|_0^{\pi}=-\frac{1}{4}\pi,$$

$$b=0,$$

$$c=\gamma-1+\int_0^{T-h}(\psi(\tau)+\varphi(\tau))v(\tau)d\tau=\gamma-1+\int_0^{\pi}\sin\tau v(\tau)d\tau=$$

$$=\gamma-1+\int_0^{\frac{\pi}{2}}\sin\tau 2\sin\tau d\tau+\int_{\frac{\pi}{2}}^{\pi}\sin\tau\frac{2}{\pi-2}\sin\tau d\tau=\gamma-1+2\int_0^{\frac{\pi}{2}}\sin^2\tau d\tau+\frac{2}{\pi-2}\int_{\frac{\pi}{2}}^{\pi}\sin^2\tau d\tau=$$

$$=\gamma-1+\frac{2}{2}[\tau-\frac{1}{2}\cos 2\tau]_0^{\frac{\pi}{2}}+\frac{2}{\pi-2}\frac{1}{2}[\tau-\frac{1}{2}\cos 2\tau]_{\frac{\pi}{2}}^{\pi}=$$

$$=\gamma-1+[\frac{\pi}{2}+\frac{1}{2}+\frac{1}{2}]+\frac{1}{\pi-2}[\pi-\frac{1}{2}-\frac{\pi}{2}-\frac{1}{2}]=$$

$$=\gamma-1+\pi/2+1+\frac{\pi}{2(\pi-2)}-\frac{1}{\pi-2}=\gamma+\frac{\pi-1}{2}.$$

Here

$$\gamma=\int_{T-h}^{T}\left[Ns(\tau-h,x_0)+Mr(\tau-h)\right]^T\left[Ns(\tau-h,x_0)+Mr(\tau-h)\right]dt$$

and the required coefficients are

$$N=\left[E+DB^{-1}B_0\right]^{-1}\left[C_{-1}-DB^{-1}A_{-1}\right],\ M=\left[E+DB^{-1}B_0\right]^{-1}\left[D_{-1}-DB^2_{-1}\right]$$

In our case

$$r(\tau)=0,\ s(t,x_0)=e^{-it}\cdot 1=\int_{\frac{\pi}{2}}^{\pi}e^{-it}e^{-it}dt=\int_{\frac{\pi}{2}}^{\pi}e^{-2it}dt=-\frac{1}{-2i}e^{-2it}\Big|_{\frac{\pi}{2}}^{\pi}=$$

$$=\frac{i}{2}[e^{-2i\pi}-e^{-i\pi}]=\frac{i}{2}[\cos 2\pi+i\sin 2\pi-\cos\pi+i\sin\pi]=i$$

Then $\gamma = i$ such that $\mathrm{Re}\gamma = 0$, $\mathrm{Im}\gamma = 1$. Since we are consider only the real valued parameters, then we have $c = 0 + \frac{\pi-1}{2} = \frac{\pi-1}{2}$, $b = 0$, $a = -\frac{\pi}{4}$ and the required algebraic equation is

$$-\frac{\pi}{4\lambda^2} + \frac{\pi - 1}{2} = 0 \Rightarrow \lambda^2 = \frac{\pi}{2(\pi - 1)}$$

The positive solution of the last equation is $\lambda^* = \sqrt{\frac{\pi}{2\pi-2}}$.

Thus the optimal control function is given as

$$u_g(t) = v(t) + \frac{1}{\lambda^*}L(t)g = \begin{cases} \sin t + \frac{g}{\lambda^*}\cos t, & if \quad t \in [0, \frac{\pi}{2}] \\ \frac{\pi}{\pi-2}\sin t + \frac{g}{\lambda^*}\cos t, & if \quad t \in [\frac{\pi}{2}, \pi] \\ \frac{\pi}{\pi-2}\sin t + 0, & if \quad t \in [\pi, \frac{3\pi}{2}]. \end{cases}$$

According to Theorem 20 the unknown scalar g is determined by the inequality

$$\max_{\|g\|=1} \mathcal{L}(g, T) \le 0$$

where

$$\mathcal{L}(g, T) = g^T s(T - h, x_0) + \int_0^{T-h} g^T F(T - h, \tau)u_g(\tau)d\tau$$

In our case we have

$$\max_{g\neq 0}\{g\cos\pi + \int_0^{\pi} gF(\pi, \tau)u_g(\tau)d\tau\} \le 0.$$

Simplifying

$$\max_{g\neq 0}\{-g + \int_0^{\frac{\pi}{2}} g(-\cos\tau)(\sin\tau + \frac{g}{\lambda^*}\cos\tau)d\tau + \int_{\frac{\pi}{2}}^{\pi} g(-\cos\tau)(\frac{\pi}{\pi-2}\sin\tau + \frac{g}{\lambda^*}\cos\tau)d\tau\} =$$

$$= \max_{g\neq 0}\{-g - \frac{g}{2}\int_0^{\frac{\pi}{2}} \sin 2\tau d\tau - \frac{g^2}{\lambda^*}\int_0^{\frac{\pi}{2}} \cos^2\tau d\tau - \frac{1}{2}g\frac{\pi}{\pi-2}\int_{\frac{\pi}{2}}^{\pi} \sin 2\tau d\tau - \frac{g^2}{\lambda^*}\int_{\frac{\pi}{2}}^{\pi} \cos^2\tau d\tau\} =$$

$$= \max_{g\neq 0}\{-g + \frac{g}{4}\cos 2\tau\Big|_0^{\frac{\pi}{2}} - \frac{g^2}{2\lambda^*}(\tau + \frac{\sin 2\tau}{2})\Big|_0^{\frac{\pi}{2}} + \frac{1}{2}\frac{\pi}{\pi-2}g\frac{1}{2}\cos 2\tau\Big|_{\frac{\pi}{2}}^{\pi} - \frac{g^2}{2\lambda^*}(\tau + \frac{\sin 2\tau}{2})\Big|_{\frac{\pi}{2}}^{\pi}\} =$$

$$= \max_{g\neq 0}\{-g - \frac{g}{2} - \frac{g^2}{2\lambda^*}(\frac{\pi}{2} + 0) + g\frac{\pi}{2(\pi-2)} - \frac{g^2}{2\lambda^*}(\frac{\pi}{2} + 0)\} =$$

$$= \max_{g\neq 0}\{g(-frac32 + \frac{\pi}{2(\pi-2)}) - g^2(\frac{\pi}{2\lambda^*})\} =$$

$$= \max_{g\neq 0}\{-g^2\sqrt{\frac{\pi+2}{\pi}} - g(\frac{3\pi - 6 - \pi}{2(\pi-2)})\} = \max_{\|g\|=1}\{-g^2\sqrt{1 + \frac{2}{\pi}} - g\frac{2\pi - 6}{2\pi - 4}\} \le 0$$

This inequality gives the following solution $g^0 = \frac{6-2\pi}{2\pi-4}\sqrt{\frac{\pi}{\pi+2}}$

Finally, the optimal control is given as

$$u_g^0(t) = v(t) + \frac{1}{\lambda^*}L(t)g^0 = \begin{cases} \sin t + \frac{6-2\pi}{2\pi-4}\sqrt{\frac{2\pi-2}{\pi+2}}\cos t, & if \quad t \in [0, \frac{\pi}{2}] \\ \frac{\pi}{\pi-2}\sin t + \frac{6-2\pi}{2\pi-4}\sqrt{\frac{2\pi-2}{\pi+2}}\cos t, & if \quad t \in [\frac{\pi}{2}, \pi] \\ \frac{\pi}{\pi-2}\sin t, & if \quad t \in [\pi, \frac{3\pi}{2}] \end{cases}$$

Example 2

$$\begin{gathered} \dot{x}(t) = -x(t-1) + u(t), \\ f(t) \equiv 0, \quad t \in [-1, 0), \quad x(0) = 10, \quad T = 3, \\ \int_0^3 u^2(t)dt \le M \to \min_{u \in U_T(\cdot)} \end{gathered} \tag{4.104}$$

The problem is to find the function $u^0(t)$ such that the trajectory of the control system satisfies $x(t) \equiv 0$ on the last time interval $t \in [2, 3]$ and minimizing the functional

$$\min_{u \in U_T(\cdot)} \int_0^3 u^2(t)dt$$

Using the method illustrated above we can establish that the optimal control have the following form:

$$u^0(\tau) = \begin{cases} e^{\frac{\sqrt{3}}{2}\tau}(2.6521 \sin\frac{\tau}{2} - 2.5115 \cos\frac{\tau}{2}) - e^{-\frac{\sqrt{3}}{2}\tau}(2.3118 \sin\frac{\tau}{2} + 4.6325 \cos\frac{\tau}{2}), \ \tau \in [0, 1] \\ e^{\frac{\sqrt{3}}{2}\tau}(1.1407 \sin\frac{\tau}{2} - 1.0301 \cos\frac{\tau}{2}) + e^{-\frac{\sqrt{3}}{2}\tau}(12.2625 \sin\frac{\tau}{2} + 1.0359 \cos\frac{\tau}{2}), \ \tau \in [1, 2] \\ e^{\frac{\sqrt{3}}{2}\tau}(0.4901 \sin\frac{\tau}{2} - 0.4219 \cos\frac{\tau}{2}) - e^{-\frac{\sqrt{3}}{2}\tau}(17.2647 \sin\frac{\tau}{2} - 23.6179 \cos\frac{\tau}{2}) - \\ \qquad -8.3334\tau + 19.9980, \ \tau \in [2, 3]. \end{cases} \tag{4.105}$$

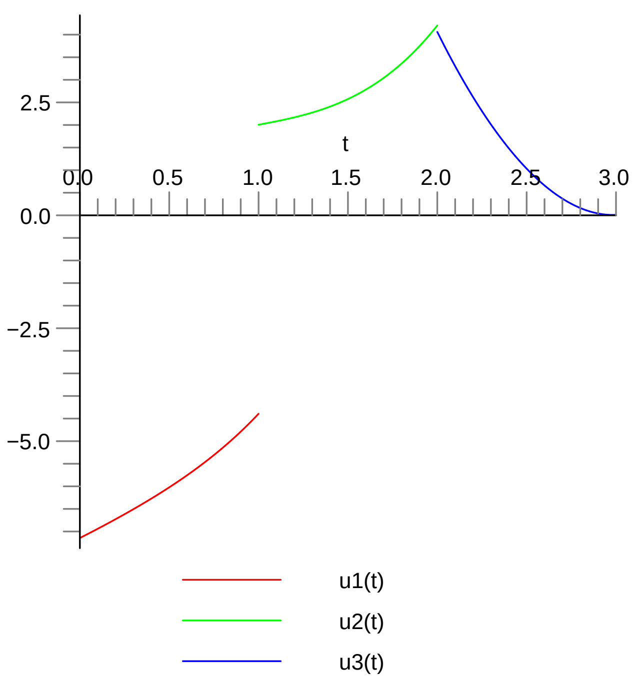

Figure 4.1: Optimal control

4.5 Conclusions

In this chapter differential repetitive processes are studied from the perspective of differential delayed systems. The new mathematical models for this class of systems have been introduced and primary analysis is provided. First of all, the controllability and time optimal control have been outlined. It is necessary to add that this note covers only first attempts to investigate the differential repetitive processes from that point of view, and hence a rich material remains to be the subject for further work. For example, new controllability and observability notions are of a significant interest for further investigations. In particular, the controllability notion which includes so-called functional controllability (see, for example, [56]) when is required to drive the state variables at the final interval $[\alpha, \alpha+h]$ to the pre-assigned functions $x(t) = \varphi(t), y(t) = \psi(t), t \in [\alpha, \alpha+h]$. This notion can be given as follows

Definition 21. *The process (4.14)—(4.15) is said to be complete controllable if for any initial data* $g(t),\ t \in [-h, 0],\ f(t),\ t \in [-h, 0),\ x_0 = 0$ *of (4.15) there exist the moment* $t_1 < +\infty$ *and the control function* $u(t),\ t \geq 0,\ u(t) \equiv 0,\ t \geq t_1$ *such that the corresponding solutions* $x(t, g, f, x_0, u),\ y(t, g, f, x_0, u)$ *of the system (4.14)—(4.15) satisfy the following conditions*

$$x(t) \equiv 0,\ y(t) \equiv 0,\ t \geq t_1 \tag{4.106}$$

In fact, it is required to drive the system at the interval $[t_1, t_1 + h]$ to the zero position and to keep it during the time $t > t_1 + h$. Related analysis for ordinary time delay system can be found, for example, in [57, 67] and some results on controllability of the multiconnected system have been also given in [37, 53].

The developed in this chapter results for the linear process (4.14)—(4.15) can be also extended to obtain the necessary conditions for optimal control of nonlinear models. As known, the cost functional increment method [37] is based on the estimate for trajectory variation generated by the corresponding control function variation, and in fact uses the linear part of the model. For this purpose, we can consider the following nonlinear optimisation problem

$$\dot{x}(t) = F\big(x(t), x(t-h), y(t), y(t-h), u(t)\big) \tag{4.107}$$

$$y(t) = G\big(x(t), x(t-h), y(t-h), u(t)\big),\ t \in T \doteq [0, \alpha] \tag{4.108}$$

with the initial conditions

$$x(t) = f(t),\ t \in [-h, 0),\ x(0) = x_0,\ y(t) = g(t),\ t \in [-h, 0] \tag{4.109}$$

and $x \in \mathbb{R}^n,\ y \in \mathbb{R}^m,\ u \in \mathbb{R}^r$. Here it is necessary to assume that the control functions are piecewise continuous on the interval $[0, \alpha]$ and $u(t) \in U$ for all $t \in T$, where $U \subset \mathbb{R}^r$ is some prescribed set. The couple of the functions $\big(x(t), y(t)\big)$ is a solution of the system (4.107)—(4.109) for the given control function $u(t)$, if they satisfy the differential equation (4.107) almost everywhere on the interval $[0, \alpha]$ and the difference equation (4.108) for all $t \in [0, h]$.

Let $\beta_0, \beta_1, \ldots, \beta_\nu$, be given time moments such that $0 = \beta_0 < \beta_1 < \ldots \beta_\nu \leq \alpha$, and, $M_i \subset \mathbb{R}^m,\ i = 0, \ldots, \nu$ be given convex closed sets from $\mathbb{R}^m$. The optimal control problem, hence, is to minimize the cost functional of the form

$$J(u) = \varphi(y(\alpha - \beta_0), y(\alpha - \beta_1), \ldots, y(\alpha - \beta_\nu)) \tag{4.110}$$

subject to the constraints

$$y(\alpha - \beta_i) \in M_i, \; i = 0, \ldots, \nu \tag{4.111}$$

over the solutions of the system (4.107)—(4.109)). Here $\varphi(x_1, \ldots, x_{\nu+1})$ is a continuously differentiable function. The introduction of such kind optimisation problem corresponds to the notion of the pass controllability for the given points when it is necessary to optimize the final pass profile running through the pre-assigned value set at the specified time moments. However, the major task is to solve the general problem with nonlocal initial conditions omitted here. These problems are subject of ongoing work and will be reported in due course.

Summary

This work has presented a panorama of modern dynamical system and optimization theory and its application. This theory has been enhanced by new developments, many of which appear here for the first time in relation to engineering problems. The presentation is not encyclopedic in nature and does not cover all areas of system theory. The text is intended to concisely present a class of networks (graph models) and multidimensional models, with its own character, theory, techniques and tools. Particular attention has been given to the principal important subclass of 2-D systems such as the repetitive processes and their links with differential-algebraic systems. Although this subclass of the models is restrictive, it is sufficient to solve a broad spectrum of applications as has been demonstrated in this work.

This contribution has developed some new results on the simulation of gas transportation networks and industrial phenomena with repetitive operation governed by linear discrete and continuous dynamical systems, in general. The aim of the work is to develop a new approach and enhance the existing ones to the mathematical description of multidimensional (2-D) and repetitive models of physical processes on the base of a 'mature' optimization theory and to solve related control problems.

In particular, the new optimization methods for the control systems under consideration are given on a strong basis by exploiting the constructive approach in the modern optimization theory. The obtained results indicates the natural way to study the large scale distributed transportation networks by a multistage modelling when the simple graph models are used at the first stage and the 2-D system setting can be realized then for the second stage for the more detailed description. For both graph and 2-D models the new optimization methods are proposed. It is the main advantage of this mixed representation that a rather general class of the complicated problems of technical area can be represented in a conceptually simple fashion. It is the starting point for the design of suitable discrete models and their efficient computer implementation.

In the work it is proposed a general operator setting to study the control theoretic properties which allows for a very significant generalization and extension of previous results for this class of systems. This leads to major new results on controllability and optimization, in particular, by feedback action in this general setting.

The contribution of this work is a beautiful, unified and powerful theory. According to which the following principal results are:

- In Chapter 1, the new algorithm for optimization of graph networks used the iterative procedure based on the principle of the decreasing suboptimality estimate

and exploiting the duality theory. Also, 2-D system setting approach for studying optimal control programm for gas pressure and gas flow at the pipeline unit;

- In Chapter 2, the strong mathematical background for the $2-D$ control optimization problem where the existence and uniqueness of optimal solution (Theorem 5) is proved, the various representations form for optimal control (Theorem 7) including feedback control law (Theorem 10) are established. Also, a specific optimization problem via boundary data (Theorem 8) is solved;
- In Chapter 3, the optimality conditions in classic principle maximum form (Theorem 13) for a nonstationary repetitive process with general intermediate constraints. Also, the new constructive optimality (Theorem 14) and sub-optimality (Theorem 15) conditions for the particular stationary repetitive models;
- In Chapter 4, extension of the delay time system approach for studying differential repetitive processes and investigation of system theoretic properties such as controllability (Theorem 18) and optimality (Theorem 21) for the hybrid (time delayed differential - algebraic) with integral constraints (Theorem22).

The application of the proposed methods to real-world problems has been demonstrated by the illustrative examples.

Zusammenfassung

In der vorliegenden Arbeit wurde ein Überblick über moderne dynamische Systeme sowie der Optimierungstheorie und deren Anwendung vorgestellt. Diese Theorie konnte durch einige neue Verbesserungen weiterentwickelt werden, wobei ein Großteil davon im Rahmen dieser Arbeit das erste Mal in Verbindung mit technischen Problemstellungen angewendet wird. Die vorgestellte wissenschaftliche Arbeit stellt keine allumfassende Ausarbeitung dar und ist somit nicht auf alle Bereiche der System Theory übertragbar. Das grundlegende Ziel meiner Dissertation besteht vorrangig darin, eine kurze und klare Darstellung von Netzwerkklassen (Graphenmodelle) und mehrdimensionalen Modellen, mit ihrem jeweils eigenen Charakter, zu zeigen. Besondere Aufmerksamkeit wurde dabei auf den allgemein wichtigen Teilbereich der 2-D Systeme, wie beispielsweise den repetitiven Prozessen sowie deren Verbindungen zu differentiellen algebraischen Systemen, gerichtet. Obwohl die Anwendbarkeit der Modelle in diesem Bereich teilweise eingeschränkt ist, kann mit ihnen ein weiträumiges Spektrum an entsprechenden technischen Aufgabenstellungen gelöst werden.

In der vorliegenden Abhandlung konnten einige neue Ergebnisse im Bezug auf die Simulation von Netzwerken zur Gasbeförderung und zu industriellen Aufgabenstellungen mit repetitiven Operationen im Allgemeinen erzielt werden. Die dabei zugrundeliegenden mathematischen Modelle wurden durch lineare, diskrete und kontinuierliche dynamische Systeme beschrieben. Das Ziel der Arbeit besteht darin, einen neuen Denkansatz zu entwickeln und die bereits existierenden Methoden hinsichtlich der mathematischen Beschreibung von mehrdimensionalen (2-D) und repetitiven Modellen von physikalischen Prozessen, basierend auf der 'mature' Optimierungstheorie, zu verbessern sowie die damit verbundene Regelungsproblematik zu lösen.

Dabei werden die neuen Optimierungsmethoden für die betrachteten Regelungssysteme insbesondere unter starker Verwendung von konstruktiven Denkansätzen der modernen Optimierungstheorie dargestellt. Die daraus hervorgehenden Ergebnisse weisen darauf hin, die großräumig verteilten Transportnetzwerke auf die ursprüngliche Art durch mehrstufige Modelbildung zu untersuchen. Die "simple graph" Modelle werden dabei in der ersten Stufe verwendet und die 2-D Einstellung zur detaillierteren Beschreibung kann dann im zweiten Schritt erfolgen. Für die beiden Graphen und die 2-D Modelle werden neue Optimierungsmethoden vorgeschlagen. Der Hauptvorteil dieser vermischten Darstellungsweise liegt darin begründet, dass eine eher allgemeinere Klasse von komplizierten technischen Problemen in eine konzeptionell einfache Form übertragen werden kann. Daraus ergibt sich die Gestaltung der diskreten Modelle sowie deren effiziente Einbindung in die entsprechende Computersoftware.

In dieser Arbeit wird eine allgemeine Operatoreinstellung vorgeschlagen, um die Steuerung bzw. Regelung der theoretischen Eigenschaften zu untersuchen, die eine wesentliche Verallgemeinerung sowie eine Ergänzung zu bisherigen Ergebnissen von Systemen für diese Kategorie zulassen. Durch die entsprechenden Feed-Back Reaktionen führt diese allgemeine Einstellung insbesondere zu bedeutenden neuen Erkenntnissen im Bereich der Regelbarkeit und Optimierung.

Im Rahmen dieser Dissertation konnte eine hervorragende, vereinheitlichende und leistungsfähige Theory entwickelt werden. Die damit verbundenen, prinzipiellen Ergebnisse sind im folgenden formuliert:

- In Kapitel 1 wird ein neuer Algorithmus für die Optimierung von Graphen Netzwerken entwickelt. In diesem Zusammenhang werden iterative Prozesse, die auf dem Prinzip der Verringerung der "suboptimality estimate" und der Verwendung der dualen Theorie basieren, verwendet. Ebenfalls wird eine 2-D Systemeinstellungsmethode zur Untersuchung eines optimalen Regelungsprogramms für den Gasdruck und die Gasströmung in einer Fernleitungsanlage benutzt.

- In Kapitel 2 wird zunächst der fundierte mathematische Hintergrund für 2-D Regelungsoptimierungsaufgaben, eingeführt wobei bereits die Existenz und die Vereinheitlichung von optimalen Lösungen nachgewiesen ist, (Theorem 5). Des weiteren werden verschiedene Darstellungsformen für optimale Regelungen (Theorem 7), einschließlich der Feed-Back Regelungsgesetze (Theorem 10) behandelt. Ein spezifisches Optimierungsproblem wird mit Hilfe von Grenzwertdaten (Theorem 8) gelöst.

- In Kapitel 3 werden die optimalen Bedingungen in "classic maximum principle form" (Theorem 13) für instationäre repetitive Prozesse mit allgemeinen Zwischenbegrenzungen präsentiert. Die neuen "constructive optimality" (Theorem 14) und "suboptimality" (Theorem 15) Bedingungen für stationäre repetitive Modelle werden dargestellt.

- In Kapitel 4 wird die Erweiterung der Verzögerungszeit-Systemmethode zur Untersuchung von differentiellen repetitiven Prozessen sowie die Ermittlung von systembezogenen theoretischen Eigenschaften, wie bspw. Regelbarkeit (Theorem 18) und die Optimierung (Theorem 21) für die Hubride (zeitverzögerte differenziell algebraisch) mit integralen Begrenzungen (Theorem 22), dargestellt.

Die Anwendung der vorgeschlagenen Methodik auf reale Problemstellungen wurde durch illustrative Beispiele demonstriert.

Bibliography

[1] Abolinia V.E and Myshkis A. D.: Mixed problem for an almost linear hyperbolic system in the plane, 1960, Mathematic sbornik, vol. 12, pp. 42 3–442 (in Russian).

[2] Amann, N. and Owens, D. H. and Rogers, E.: Predictive optimal iterative learning control, 1998, International Journal of Control, **69**(2), pp. 203–226.

[3] Akhundov A.: Observability of a hybrid system, 1975, Izvestia Akademii Nauk Belarusi, Fiz.-Mathem. Seria, No. 5, pp. 53–57 (in Russian)

[4] Azevedo-Perdicoulis T.P., G. Jank. Iterative solution of algebraic matrix Riccati equations in open loop Nash games, Int. J. Robust Nonlin. Cont. 15:55-62, (www.interscience.wiley.com), 2005. DOI: 10.1002/rnc.996

[5] Bose N.K. Applied Multidimensional System Theory, New York: Van Nostrand Reinhold Company, 1982.

[6] Bisiacco M. and Fornasini E. Optimal control of two dimensional systems. SIAM J.Control and Application, v. 28, No 3, pp.582-601, 1990.

[7] Benton, S. E. Analysis and Control of Linear Repetitive Processes, PhD Thesis, University of Southampton, UK, 2000.

[8] Collins W. D. Controllability and canonical forms for multipass systems described by the ordinary dif. equations. - IMA J. Math. Cont. and Informat. v. 1, pp. 1-25, 1984.

[9] Chesari L.: Optimization—theory and applications: problems with ordinary differential equations, 1983, New Yoork, Berlin.: Springer-Verlag.

[10] S. Dymkou, G. Jank, 2-D Control Theory in Gas Transport Network Modeling, Proceedings of Forth International Workshop on Multidimensional Systems (NDS 2005), Wuppertal, Germany, pp. 241 — 246, 2005.

[11] S. Dymkou, E. Rogers, K. Galkowski, D. H. Owens, Optimal control for a class of differential repetitive processes, Proceedings of 15-th MTNS Symposium (MTNS 2002), CD Rom, 2002.

[12] M. Dymkov, I. Gaishun, E. Rogers, K. Galkowski, D. H. Owens, Z-Transform and Volterra Operator Based Approaches to Controllability and Observability Analysis for Discrete Linear Repetitive Processes, 2003, Multidimensional Systems and Signal Processing, 14, pp. 365–395.

[13] Daniel De Wolf, Yves Smeers, The gas transmission problem solved by an extension of the Simplex algorithm, GREMARS, University de Lille 3, France and SMG, University libre de Bruxelles, Belgique , CORE, University Catholique de Louvain, Belgium. March 1, 1999.

[14] Dymkov M. P. and Gaishun, I. and Rogers, E. and Galkowski, K. and Owens, D. H.: Exponential stability of discrete linear repetitive processes, 2002, International Journal of Control, 2002, **75**(12), pp. 861–869.

[15] Dymkov M. P. and Gaishun, I. and Rogers, E. and Galkowski, K. and Owens, D. H.: $z-$ Transform and Volterra based approaches to controllability and observability analysis for discrete linear repetitive processes, 2003, Multidimensional Systems and Signal Processing, Vol. 14, pp. 365 - 395.

[16] Dymkou S. and Rogers, E. and Dymkov M. E. and Galkowski, K. and Owens, D. H.: Controllability and Optimization for Differential Linear Repetitive Processes, 2003, 42-th IEEE Conference on Decision and Control (CDC-2003), Hawai, USA

[17] D'Andrea R., Dullerud G.E.: Distributed control design fir spatially interconnected systems, 2003, IEEE Trans on AC, Vol.49, No.8, 1478-1495.

[18] Demidenko N. D. Modeling and Optimization Heat- and Mass Transfer Processes In Chemistry Industry. Moscow, Nauka, 1991.

[19] Dymkou S., Jank G., Azevedo-Perdicoulis T.P, Graph and 2-D systems approach in gas transport network modeling. Proceedings Int. 13th IFAC Workshop on Control Applications of Optimisation, Paris,CAO 2006.

[20] S. Dymkou, G. Jank, 2-D Control Theory in Gas Transport Network Modeling, Proceedings of Forth International Workshop on Multidimensional Systems (NDS 2005), Wuppertal, Germany, pp. 241 — 246, 2005.

[21] Dymkou S., Jank G. Time optimal control problem for a class of linear differential-algebraic systems with delay (Sixteenth International Symposium on MTNS 2004, Katholieke Universiteit Leuven, Belgium July 5-9, 2004)

[22] S. Dymkou, Rogers E, Galkowski K., Dymkov M., Owens D.H. *Optimal Control for a Class of Differential Linear Repetitive Processes*. CD Rom Proceedings of MTNS-2002 Conference, Indiana, USA, August, 12-16, 2002.

[23] Dymkou S., Rogers E, Galkowski K., Dymkov M., Owens D.H. *An approach to contrallability and optimization problems for repetitive processes*. Proceedings Int.Conference "Stability and control processes" (SCP-2005). Saint-Ptersburg, Russia, Vol.3, pp. 1504–1515, 2005.

[24] Dymkou S., Rogers E, Dymkov M. P., Galkowski K., Owens D. H.: Delay System Approach to Linear Differential Repetitive Processes. Proceedings of the IFAC Workshop on Time Delay Systems (TDS-2003), 8-10 September, 2003, France.

[25] Dymkou S. Approximations of the constrained nonsmooth optimal control problems and their applications for the robot planning motion. Proceedings of the ICIAM 2003 Congress, Sydney, Australia, July 7-11, 2003.

[26] Dymkou S. Robot Optimal Control in the Presence of State Constraints. Proceedings of the Ballarat Conference, November 3, 2002, Victoria, Australia.

[27] Dymkou S. Approximation of the optimal control problem with the nonsmooth state constraints, Proceedings of Industrial Optimization Symposium (WACEIO), Perth, Australia, September 30 - October 3, 2002.

[28] Dymkou S. and Minchenko L. Optimal control problem for multipass systems. Proceedings of Intern. Conf. "Erugins readings-VIII". Brest, Belarus, May 20-23. 2002, 134-135.

[29] Dymkou S. and Minchenko L. Optimization problem for repetitive control systems. Proceedings of Intern. Conf. "Analytic methods of analysis and differential equations" ("AMADE-2001"), February 15-19, Minsk, 2001, pp.113-114.

[30] Dymkou S. A linear optimization problem for repetitive processes, Bulletin of the National Academy of Sciences of Belarus, Physics and Math. Series, 4, (2002), 17-26.

[31] Dymkou S. and Minchenko L. Suboptimality and differential properties of optimal solutions for repetitive control systems, Reports of the National Academy of Sciences of Belarus, 46, 3, (2002), pp.5-8 (Russian).

[32] Edwards, J. B.: Stability problems in the control of multipass processes, Proceedings of The Institution of Electrical Engineers, Vol **121**(11), pp. 1425–1431, 1974.

[33] A.E. Fincham, G.H. Goldwater. Simulation Models for Gas Transmission Networks, Trans. Inst. MC vol. 1, No. 1, 1979.

[34] Fornasini, E. and Marchesini, G.: Doubly-indexed dynamical systems: state space models and structural properties, Mathematical Systems Theory, 12, pp. 59–72, 1978.

[35] Gaishun I. Completely integrable multivariable differential equations, 1983, Minsk, Nauka i Technika (in Russian).

[36] Galkowski K. and Wood J.(Eds), Multidimensional Signals, Circuits and Systems, Taylor and Francis, London, 2001.

[37] R. Gabasov, F.M. Kirillova, O.I. Kostyukova, Constructive methods of optimization, Part 3 (Network problems). University Press BSU, Minsk, 1986 (Russian).

[38] R. Gabasov, Kirillova F. M. and Prischepova S. V. Optimal Feedback Control, Springer-Verlag Lecture Notes in Control and Information Sciences Serie, vol. 207, 1995.

[39] Gabasov R., Kirillova F. M.: Software Optimization, Plenum Press, USA, 1988.

[40] Gabasov, R. and Kirillova, F. M. The Qualitative Theory of Optimal Processes, Marcel Dekker Inc., USA. 1976.

[41] Galkowski, K. and Rogers, E. and Owens, D. H.: Matrix rank based conditions for reachability/controllability of discrete linear repetitive processes, Linear Algebra and its Applications, 275-276, pp. 201–224, 1998.

[42] Galkowski, K. and Rogers, E. and Owens, D. H.: New 2D models and a transition matrix for discrete linear repetitive processes, International Journal of Control, 72(15), pp. 1365–1380, 1999.

[43] Hale, J. K. : Theory of Functinal Differential Equations, 1976, Springer-Verllag, New York.

[44] R. Hartl, S. Sethi and R. Vicson, "A survey of the maximum principles for optimal control problems with state constraints", SIAM Rev. 37(1995), 181-218.

[45] T. Kaczorek Linear Control Systems. Research Studies Press Series and J. Willey, New York, v. 1-2, 1993.

[46] Hartman P.: Ordinary Differential Equations, John Willey, New York, 1964.

[47] G. Jank. Controllability, Observability and Optimal Control of Continuous-Time 2-D Systems. App. Math. Comput. Science, 12, 2, pp.181-195, 2002.

[48] G. Jank. Disturbance attenuation in hyperbolic 2D-systems. University of Aveiro. Preprints of Research Report CM03/I-20, 2003.

[49] J. Kralik et al, Modelling the Dynamics of Flow in Gas Pipelines, IEEE Trans. Syst., Man and Cyb., vol.14, no.4, 1984.

[50] V.V. Kirillov, Research a Model of Gas Transporting Network, http://www.laboratory.ru

[51] Kaczorek T. and Klamka J. Local controllability and minimum energy control of n-D linear systems. 1987, Bull. Pol. Acad.Techn. Sci., v. 35, No. 11-12, pp. 679-685.

[52] Kostyukova O. I.: Parametric optimal control problem with variable index, 2003, Computational Mathematics and Mathematical Physics, **43(1)**, pp. 26–41 (translated from Journal Vychislitelnoi Mathematiki i Mathematicheskoi Fiziki, 2003, v. 43, No. 1, pp. 26–41)

[53] Kirillova, F and Kopeikina, T and Marchenko V and Minyuk S. Controllability of linear multiconnected nonstationary heredirety systems, 1976, Proc. IFAC Simposium, Udine, Italy, pp. 255-263.

[54] Jan Mohring, Jochen Hoffmann, Thomas Halfmann, Automated Model Reduction of Complex Gas Pipeline Networks. PSIG Annual Meetting, Palm Springs, California, 20-22 October, 2004.

[55] Mordukhovich B. S.: Optimal control of difference, differential, and differential-difference inclusions, SIAM J. Contr. Optimiz. **33**(1995) 882-915.

[56] Manitius A. and Triggiani R.: Function space controllability of linear retarted systems: a derivation from abstract operator conditions, 1978, SIAM J. Control and Optimization, vol. 16, No. 4 pp. 599–645. pp. 81–89.

[57] Marchenko, V.: Controllability of the linear retarted systems, 1977, Soviet Mathematics, **236**(5), pp. 1083-1086.

[58] Metelskii A. I.: Pointwise completness problem in control theory of differential-difference systems, 1994, Achivments in Mathematics, Vol. **49(2)** , pp.103–140 (in Russian).

[59] B. S. Mordukhovich, *Approximation methods in problems of optimization and control*(Moscow, Nauka, 1988).

[60] Myshkis A. D.: Linear differential equations with retarted arguments (in Russian), 1972, Nauka, Moscow.

[61] Natanson I.: Constructive Function Theory, 1965, New York.

[62] Niculescu S. I : Delay Effects on Stability. A Robust Control Approach, 2001, Springer-Verlag, Berlin.

[63] Nieplocha J. Discrete-time quadratic optimal control of gas network systems.- Foundations of control engineering, Vol.13, No.4, pp. 175-186, 1988.

[64] Andrzej J. Osiadacz. Simulation and Analysis of Gas Network. Gulf Publishing Company, 1987.

[65] Owens, D. H.: Stability of multipass processes, 1977, Proceedings of The Institution of Electrical Engineers, **124**(11), pp. 1079–1082.

[66] Owens, D. H. and Rogers, E.: Stability analysis for a class of 2D continuous-discrete linear systems with dynamic boundary conditions, 1999, Systems and Control Letters, **37**, pp. 55–60.

[67] Rabah, R. and Malabre, M.: On the structure at infinity of linear time delay systems with application to the disturbance decoupling problem, 1999, Kibernetika, **35**(6), pp. 668-680.

[68] Roberts, P. D.: Numerical investigations of a stability theorem arising from 2-dimensional analysis of an iterative optimal control algorithm, 2000, Multidimensional Systems and Signal Processing, **11**(1/2), pp. 109–124.

[69] P. D. Roberts. Computing the Stability of Iterative Optimal Control Algorithms through the use of Two-Dimensional Systems Theory, Proceedings of UKACC International Conference, Control 96, Vol 2, pp 981-986, 1996.

[70] Roesser, R. P.: A discrete state space model for linear image processing, 1975, IEEE Transactions on Automatic Control, **20**(1), pp. 1–10.

[71] Rogers, E. and Owens, D. H: Stability Analysis for Linear Repetitive Processes, 1992, Springer-Verlag Lecture Notes in Control and Information Sciences Series, Vol 175, Berlin.

[72] E. Rogers, K. Galkowski and D. H. Owens. Control Systems Theory and Applications for Linear Repetitive Processes, Springer-Verlag Lecture Notes in Control and Information Sciences Series, 2005.

[73] Streltzov, S.: Necessary optimality conditions for a hybrid system, 1976, Differentzialnye uravnenya, **12**(5), pp.940—943 (in Russian).

[74] Simone Software:Equations and methods. Personal contacts.

[75] A.K. Sod, G.L Funk, A.C. Delmastro. Dynamic optimization of natural gas pipeline using a gradient search technique, Int. J. of Control. No. 6, 1971.

[76] R. Stelter, H.A. Nour Eldin, Gas network operation optimization using singular pertubation model reduction, Proc. of the IMACS/FAC Inter. Symp. on Modelling and Simulation of Distributed Parameter Systems. Hiroshima. Oct. 6-9,1987.

[77] M.A. Stoner, Analysis and control of insteady flows in natural gas piping systems, J. of Basic Eng., no. 9, 1969.

[78] K. L. Teo, C.J.Goh and K.H. Wong, *A unified computational approach to optimal control problems,*(Pitman monographs and surveys in Pure and Applied mathematics, Bath Press, Avon, 55, 1991).

[79] P.J. Wong, R.E. Larson, Optimization of natural gas pipeline via dynamic programming, IEEE Trans. on AC, No. 5, 1968.

[80] Weiss L.: On controllability of delay differential systems, 1967, SIAM J. Control, Vol. 5, No. 4, pp. 57 –587.

[81] Zerz E. Topics in Multidimensional Linear Systems Theory Springer Lecture Notes in Control and Information Sciences 256 Springer, London (2000).

[82] Zerz E. Multidimensional Behaviors: An Algebraic Approach to Control Theory for PDE International Journal of Control 77, 812-820 (2004)

[83] Zverkin A. M.: On pointwise completness of time delay systems, 1973, Differential Equations, Vol. 9, No. 3, pp. 430–436.

List of Figures

Curriculum vitae

Name Siarhei M. DYMKOU

Sex Male

Date and Place of Birth May 07, 1978
Minsk, Belarus

Family Status Married, one child

Home Address Siarhei Dymkou
Junkersstr. 20, Erlangen
91058, Germany

Office Martensstrasse 3, Raum 12.152
Phone: (+49 9131) 85 29960
Fax: (+49 9131) 85 28126
E-mail: dymkou@am2.am.uni-erlangen.de (dymkov78@mail.ru)

Education

Secondary school(Belarus, Minsk) 1985-93
Average record 5.0 (maximum 5.0)

Lyceum of Byelorussian State University
(physics and mathematics classes)
1993-1995
Average record 4.2 (maximum 5.0)

Byelorussian State University (Belarus, Minsk)
1995-2000
Average record 4.0 (maximum 5.0)
Faculty : Mechanics and Mathematics,
Department : Theoretical Mechanics and Robot Engineering

M. S. Degree 2000
" Optimal feedback control in linear systems with terminal constraints"
Supervisor: Dozent, PhD A.I. Lutov, BSU

Professional experience

Byelorussian State University
of Informatics and Radioelectronics (Belarus, Minsk) 2000-2001
Scientific Assistant at the Department of Informatics,
Supervisors: Prof. Dr. Leonid Minchenko

Centre for Informatics and Applied Optimization (CIAO)
University of Ballarat (Australia, Ballarat)

2001—2003

Scientific Assistant at School of Information Technology and Mathematical Sciences

M.S. Dissertation

" Approximation for some classes of optimal control problems with state constraints and repetitive control systems "

Principal Supervisor: Prof. A. Rubinov

Associate Supervisor: Dr. A. Bagirov

Lehrstuhl II für Mathematik RWTH Aachen
(Germany, Aachen) 2003– 2006

Scientific Assistant at Lehrstuhl II für Mathematik RWTH Aachen

Supervisors: Univ.Prof. Dr. Gerhard Jank,
Prof. Dr. Eva Zerz,

Ph.D Degree September , 2006

Ph.D Dissertation:

" Graph and 2D Optimization Theory and their application for discrete simulation of gas transportation networks and industrial processes with repetitive operations. "

Research Fields

- ✓ 2D system theory and repetitive processes
- ✓ Optimal Control Theory
- ✓ Nonsmooth analysis and nonsmooth optimization
- ✓ Mathematical modeling
- ✓ Numerical Analysis
- ✓ Computer simulation

Additional Notes

Languages: Russian, English, German (Grundstufe)

Programming: C/C++, Fortran, 1C, Matlab, Maple

Sports: Swimming

Interests: Books

List of Publications by Siarhei Dymkou

1. S. Dymkou, E.Rogers, M.Dymkov, K.Galkowski, Constrained Optimal Control Theory for Differential Linear Repetitive Processes. SIAM J. Optimization (submitted).

2. S. Dymkou, E.Rogers, M.Dymkov, K.Galkowski, Optimal Control of Non-stationary Differential Linear Repetitive Processes. Integral Equations and Operator Theory (submitted).

3. S. Dymkou, G. Jank, T.P. Azevedo-Perdicoulis, Graph and 2-D systems approach in gas transport network modeling. (Proceedings of IFAC CAO'06, E.N.S. Cachan, France/, pp 165-171 April 26-28, 2006.

4. Dymkou S., Rogers E, Galkowski K., Dymkov M., Owens D.H. An approach to contrallability and optimization problems for repetitive processes. Proceedings Int.Conference "Stability and control processes" (SCP-2005). Saint-Ptersburg, Russia, Vol.3, pp. 1504–1515, 2005.

5. S. Dymkou, G. Jank, 2-D Control Theory in Gas Transport Network Modeling, Proceedings of Forth International Workshop on Multidimensional Systems (NDS 2005), Wuppertal, Germany, pp. 241 246, 2005.

6. Dymkou S., Jank G., Azevedo-Perdicoulis T.P, Graph and 2-D systems approach in gas transport network modeling. Accepted for International Journal of Tomography and Statistics.

7. S. Dymkou. 2-D theory settings and correct discretization in the gas network simulation. Abstracts of the International Conference "Erugins Readings -X", Mogilev, Belarus, May 24-26, 2005, p.109-110.

8. Dymkou S., Jank G.: Time optimal control problem for a class of linear differential-algebraic systems with delay (Sixteenth International Symposium on MTNS 2004, Katholieke Universiteit Leuven, Belgium July 5-9, 2004)

9. Dymkou S., Rogers E, Dymkov M. P., Galkowski K., Owens D. H.: Controllability and Optimization for Differential Linear Repetitive Processes (42-th IEEE Conference on Decision and Control (CDC-2003), December 912, 2003, Hawai, USA.

10. Dymkou S., Rogers E, Dymkov M. P., Galkowski K., Owens D. H.: Delay System Approach to Linear Differential Repetitive Processes. Proceedings of the IFAC Workshop on Time Delay Systems (TDS-2003), 8-10 September, 2003, France.

11. Dymkou S.,Approximations of the constrained nonsmooth optimal control problems and their applications for the robot planning motion. Proceedings of the ICIAM 2003 Congress, Sydney, Australia, July 7-11, 2003

12. Dymkou S. Robot Optimal Control in the Presence of State Constraints.pp. Proceedings of the Ballarat University Conference, November 3, 2002, Victoria, Australia

13. Dymkou S. Approximation of the optimal control problem with the nonsmooth state constraints, Industrial Optimization Symposium (WACEIO, Perth, Australia, September 30 - October 3, 2002)(Proceedings of the Conference)

14. Dymkou S., Rogers E., Dymkov M. P., Galkowski K., Owens D. H.: Optimal Control for a Class of Differential Linear Repetitive Processes, CD Rom Proceedings of MTNS-2002 Conference, Indiana, USA, August, 12-16, 2002

15. Dymkou S. A linear optimization problem for repetitive processes, Bull. of the National Academy of Sciences of Belarus, Physics and Math. Series, 4, (2002), pp. 35-41(Russian).

16. Dymkou S. and Minchenko L. Suboptimality and differential properties of optimal solutions for repetitive control systems, Reports of the National Academy of Sciences of Belarus, 46, 3, (2002), pp. 5-8 (Russian).

17. Dymkou S. and Minchenko L. Optimal control problem for multipass systems. Proceedings of Intern. Conf. "Erugins readings-VIII". Brest, Belarus, May 20-23. 2002, pp. 134-135.

18. Dymkou S. and Minchenko L. Optimization problem for repetitive control systems. Proceedings of Intern. Conf. "Analytic methods of analysis and differential equations" ("AMADE-2001"), February 15-19, Minsk, 2001, pp.113-114. (Russian).