Journal of Nonlinear and Convex Analysis
Volume 5, Number 1, 2004, 1–18

Yokohama Publishers
ISSN 1880-5221 ONLINE JOURNAL

A NUMERICALLY STABLE METHOD FOR CONVEX OPTIMAL CONTROL PROBLEMS

VADIM AZHMYAKOV

Abstract. This paper is concerned with linear and finite-difference approximations for a class of convex optimal control problems with state constraints. We consider control systems governed by ordinary differential equations. For constructive solving the convex optimal control problems we propose a numerical method derived from the proximal-point algorithm. We study convergence properties of the obtained method and show that it can be used to compute approximate optimal controls.

1. Introduction

Linearization and discrete approximation techniques have been long time recognized as a powerful tool for solving optimal control problems (see e.g., [10, 11, 12, 22, 31, 25]. Discrete approximations can be applied directly to the problem at hand or to auxiliary problems used in the solution procedure. The numerical methods for optimal control problems with constraints (with the exception of the works [14] and [25]) are either the methods based on the full discretizations (parametrization of state and control variables), or they are function space algorithms. The first group of methods assumed *a priori* discretization of system equations. The second group of methods is, in fact, theoretical work on the convergence of algorithms which have never been implemented. The major drawback of some numerical schemes from the first group is the lack of the corresponding convergence analysis. This is especially true in regard to the *multiple shooting* and *collocation* methods (see e.g., [8, 29]).

There is a number of results scattered in the literature on discrete approximations that are very often closely related, although apparently independent. Note that the complete convergence analysis for a class of numerical methods for optimal control problems with state constraints is presented in [25].

The *gradient* algorithms [23] can also be applied to optimal control problems with constraints if the problem is discretized a priori and the discretization for states coincides with that for controls. If the second condition does not hold then the optimal control problem can have redundant constraints and as a result the rate of convergence of numerical methods can deteriorate [25]. There are many variants of gradient algorithms depending on whether the problem is a priori discretized in time, and on the optimization solver used. A gradient-based method evaluates gradients of the objective functional. The calculation of second order derivatives of the objective functional can be avoided by applying an *SQP* (Sequential Quadratic Programming)

2000 *Mathematics Subject Classification.* 49J15, 49M25, 90C25, 93B40.
Key words and phrases. optimal control, convex optimization, proximal point method, numerical methods, convergence analysis.

type optimization algorithm in which these derivatives are approximated by quasi-Newton formulas. The application of SQP type methods to optimal control is comprehensively discussed in [20, 9].

Computational methods based on the *Bellman Optimality Principle* were among the first proposed for optimal control problems [4, 7]. These methods are especially attractive when an optimal control problem is discretized a priori and the discrete version of the Bellman equation is used to solve it. Application of necessary conditions of optimal control theory, specifically of the *Pontryagin Maximum Principle*, yields a boundary-value problem with ordinary differential equations. Clearly, the necessary optimality conditions and the corresponding boundary-value problems play an important role in optimal control computations (see e.g., [28, 6]). An optimal control problem with state constraints can also be solved by using some modern numerical algorithms of nonlinear programming. For example, the implementation of the *interior point* method is presented in [34]. The application of the *trust-region* method to optimal control is discussed in [16, 21].

In this paper we are going to analyze the convergent (numerically stable) a priori discretizations of optimal control problems with inequalities state constraints. We consider the optimal control problems under some convexity assumptions. The discretization of optimization problems is an approximation procedure whose accuracy, as it is typical in numerical analysis, depends on the regularity properties of the solutions. In this paper we mainly focus our attention on the application of a regularization method, namely, on the application of a proximal-based method to the discrete approximations of optimal control problems. We are particulary interested in studying the convergence of these discrete approximations. We consider the first order Riemann-Euler approximations [10, 11, 12]. Some alternative discretization procedures are described in [33]. The application of the proximal-based regularization method makes it possible to obtain the convergence results also in the case of relatively easy discretization schemes, namely, for Riemann-Euler approximations. Clearly, the same method can be combined with other discretization procedures.

The proximal-based methods are useful tools for solving convex (see e.g., [1, 19]) and nonconvex (see e.g., [15, 18]) optimization problems. The first application of the proximal-point method to optimal control problems is presented in [5]. A great amount of works is devoted to the classical variant of the proximal point method and its various modifications. One can find a fairly complete review of the main results in [18, 30].

The rest of the paper is organized as follows. In Section 2 we formulate the optiomal control problem and consider the corresponding discretizations. Section 3 contains the relevant technical results. In Section 4 we study convergence properties of some approximation schemes for the abstract convex optimization problems in Hilbert spaces. In Section 5 we apply the results for the convex optimization problems to the optimal control problems with constraints. Section 6 is devoted to computational aspects of the proposed proximal-based algorithm.

2. Statement of the Problem

Consider the following optimal control problems with inequalities constraints

$$
\begin{aligned}
&\text{minimize } J(x(\cdot), u(\cdot)) = \int_0^1 f_0(t, x(t), u(t))dt \\
&\text{subject to } \dot{x}(t) = f(t, x(t), u(t)) \text{ a.e. on } t \in [0,1],\ x(0) = x_0, \\
&u(t) \in U \text{ a.e. on } t \in [0,1], \\
&h_j(x(1)) \leq 0\ \forall j \in I, \\
&q(t, x(t)) \leq 0\ \forall t \in [0,1],
\end{aligned} \tag{1}
$$

where $f_0 : [0,1] \times \mathbb{R}^r \times \mathbb{R}^m \to \mathbb{R}$ is a continuously differentiable function,

$$f : [0,1] \times \mathbb{R}^r \times \mathbb{R}^m \to \mathbb{R}^r,\ h_j : \mathbb{R}^r \to \mathbb{R} \text{ for } j \in I,\ q : [0,1] \times \mathbb{R}^r \to \mathbb{R}$$

and $x_0 \in \mathbb{R}^r$ is a fixed initial state. By I we denote a finite set of index values. The control set

$$U := \{u \in \mathbb{R}^m\ :\ b^i_- \leq u_i \leq b^i_+,\ i = 1, ..., m\},$$

where $b^i_-, b^i_+, i = 1, ..., m$ are constants, is a compact and convex subset of $\mathbb{R}^m$. We assume that the functions $h_j(\cdot),\ j \in I$ and $q(t, \cdot),\ t \in [0,1]$ are continuously differentiable. The ensuing analysis is restricted to a proper convex on $\mathbb{R}^r \times \mathbb{R}^m$ function $f_0(t, \cdot, \cdot),\ t \in [0,1]$. The admissible controls $u : [0,1] \to \mathbb{R}^m$ are square integrable functions in time. Let

$$\mathcal{U} := \{v(\cdot) \in \mathbb{L}^2_m([0,1])\ :\ v(t) \in U \text{ a.e. on } [0,1]\}$$

be the set of admissible control functions. We introduce the following hypothesis:

(i) $f(t, \cdot, \cdot)$ is differentiable,

(ii) f, f_x, f_u are continuous and there exists a constant $S < \infty$ such that

$$||f_x(t, x, u)|| \leq S$$

for all $(t, x, u) \in [0,1] \times \mathbb{R}^r \times U$.

Then for each $u(\cdot) \in \mathcal{U}$ the following initial value problem

$$\dot{x}(t) = f(t, x(t), u(t)) \text{ a.e. on } t \in [0,1],\ x(0) = x_0 \tag{2}$$

has a unique solution (see e.g., [25]). Given an admissible control function the solution to the initial value problem (2) is an absolutely continuous function $x : [0,1] \to \mathbb{R}^r$. It is denoted by $x^u(\cdot)$. We assume that the problem (1) has an optimal solution. The class of optimal control problems of the type (1) is broadly representative [12, 17, 25].

In parallel with (2) we examine the corresponding linearized control system

$$\dot{y}(t) = f_x(t, x^u(t), u(t))y(t) + f_u(t, x^u(t), u(t))d(t),\ y(0) = 0, \tag{3}$$

where $u(\cdot) \in \mathcal{U},\ d(\cdot) \in \mathbb{L}^2_m([0,1])$. For each $u(\cdot) \in \mathcal{U}$ and $d(\cdot) \in \mathbb{L}^2_m([0,1])$ the initial value problem (3) has a unique solution [25]. The solution to (3), which depends on $u(\cdot) \in \mathcal{U}$ and $d(\cdot) \in \mathbb{L}^2_m([0,1])$, is written $y^{u,d}(\cdot)$. Evidently,

$$y^{u,d}(t) = \int_0^t \Phi(t, s) f_u(s, x^u(s), u(s))d(s)ds \text{ a.e. } t \in [0,1],$$

where $\Phi(\cdot, s)$ is the fundamental solution matrix prescribed by the initial value problem

$$\frac{\partial}{\partial t}\Phi(t,s) = f_x(t, x^u(t), u(t))\Phi(t,s) \text{ a.e. } t \in [0,1],\ \Phi(s,s) = E,\ s \in [0,1].$$

It is common knowledge that an optimal control problem involving ordinary differential equations can be formulated in various ways as an optimization problem in a suitable function space (see e.g., [12, 17]). The original problem (1) can be expressed as an infinite-dimensional optimization problem

$$\begin{aligned} &\text{minimize } \tilde{J}(u(\cdot)) \\ &\text{subject to } u(\cdot) \in \mathcal{U}, \\ &\tilde{h}_j(u(\cdot)) \leq 0 \ \forall j \in I, \\ &\tilde{q}(u(\cdot))(t) \leq 0 \ \forall t \in [0,1], \end{aligned} \tag{4}$$

with the aid of the functions $\tilde{J} : \mathbb{L}^2_m([0,1]) \to \mathbb{R}$, $\tilde{h}_j : \mathbb{L}^2_m([0,1]) \to \mathbb{R}$ for $j \in I$ and $\tilde{q} : \mathbb{L}^2_m([0,1]) \to \mathbb{C}([0,1])$:

$$\begin{aligned} &\tilde{J}(u(\cdot)) := J(x^u(\cdot), u(\cdot)) = \int_0^1 f_0(t, x^u(t), u(t))dt, \\ &\tilde{h}_j(u(\cdot)) := h_j(x^u(1)) \ \forall j \in I, \\ &\tilde{q}(u(\cdot))(t) := q(t, x^u(t)) \ \forall t \in [0,1]. \end{aligned}$$

Fix $u(\cdot) \in \mathcal{U}$ and consider the linearized optimal control problem as an optimization problem over the set $\mathcal{U} - u(\cdot)$

$$\begin{aligned} &\text{minimize } \hat{J}(d(\cdot)) \\ &\text{subject to } d(\cdot) \in \mathcal{U} - u(\cdot), \\ &\tilde{h}_j(u(\cdot)) + \langle \nabla \tilde{h}_j(u(\cdot)), d(\cdot) \rangle \leq 0 \ \forall j \in I, \\ &\tilde{q}(u(\cdot))(t) + \langle \nabla \tilde{q}(u(\cdot))(t), d(\cdot) \rangle \leq 0 \ \forall t \in [0,1], \end{aligned} \tag{5}$$

where

$$\begin{aligned} &\hat{J}(d(\cdot)) := J(y^{u,d}(\cdot), u(\cdot) + d(\cdot)) = \int_0^1 f_0(t, y^{u,d}(t), u(t) + d(t))dt, \\ &\langle \nabla \tilde{h}_j(u(\cdot)), d(\cdot) \rangle := (h_j)_x(x^u(1))y^{u,d}(1) \text{ for } j \in I, \\ &\langle \nabla \tilde{q}(u(\cdot))(t), d(\cdot) \rangle := q_x(t, x^u(t))y^{u,d}(t) \text{ for } t \in [0,1]. \end{aligned}$$

Note that the objective functional $\hat{J}(d(\cdot))$ in (5) is nonlinear. Since the function $f_0(t, \cdot, \cdot)$, $t \in [0,1]$ is convex and the differential equation (3) is linear, $\hat{J}(d(\cdot))$ is convex. The introduced convex-linear optimization problem (5) provided a basis for numerical solving the original problem (1). This problem can be solved by using some minimization algorithms (e.g., by applying a first order method [23]). For example, the implementation of the method of feasible directions is presented in [25].

Using a discretization of (5) (see e.g., [22]), we obtain a finite-dimensional minimization problem. Let N be a sufficiently large positive integer number and

$$G_N := \{t_0 = 0, t_1, ..., t_N = 1\}$$

be a (possible nonequidistant) partition of $[0, 1]$ with

$$\max_{0 \le k \le N-1} |t_{k+1} - t_k| \le \xi_N.$$

We assume that $\lim_{N\to\infty} \xi_N = 0$. Define $\Delta t_{k+1} := t_{k+1} - t_k,\ k = 0, ..., N-1$ and consider the following finite-dimensional optimization problem

$$\begin{aligned}
&\text{minimize } \hat{J}(d_N(\cdot)) \\
&\text{subject to } d_N(\cdot) \in \mathcal{U}_N - u_N(\cdot), \\
&\tilde{h}_j(u_N(\cdot)) + \langle \nabla \tilde{h}_j(u_N(\cdot)), d_N(\cdot) \rangle \le 0 \ \forall j \in I, \\
&\tilde{q}(u_N(\cdot))(t) + \langle \nabla \tilde{q}(u_N(\cdot))(t), d_N(\cdot) \rangle \le 0 \ \forall t \in [0, 1],
\end{aligned} \tag{6}$$

where

$$\begin{aligned}
&\hat{J}(d_N(\cdot)) := \sum_{k=0}^{N-1} f_0(t_k, y_N(t_k), u^k + d^k)\Delta t_{k+1}, \\
&\mathcal{U}_N := \{v_N(\cdot) \in \mathbb{L}^{2,N}_m(G_N) \ : \ v_N(t) \in U\}, \\
&y_N(t_{k+1}) = y_N(t_k) + \Delta t_{k+1}(f_x(t_k, x^u(t_k), u^k)y_N(t_k) + \\
&\quad + f_u(t_k, x^u(t_k), u^k)d^k),\ y_N(t_0) = 0,
\end{aligned}$$

and

$$u_N(t) := \sum_{k=0}^{N-1} \phi_k(t)u^k,\ u^k = u(t_k),\ d_N(t) := \sum_{k=0}^{N-1} \phi_i(t)d^k,\ d^k = d(t_k),$$

$$t \in [0, 1],\ k = 0, 1, ...N-1\ ,\ \phi_k(t) := \begin{cases} 1 & \text{if } t \in [t_k, t_{k+1}[, \\ 0 & \text{otherwise.} \end{cases}$$

Clearly, $u^k \in U$ and $u^k + d^k \in U$. In effect, we deal with the finite-dimensional Hilbert space $\mathbb{L}^{2,N}_m(G_N)$ of the piecewise constant control functions $u_N(\cdot)$. The scalar product and the norm in the space $\mathbb{L}^{2,N}_m(G_N)$ are defined as follows

$$\langle u_N(\cdot), v_N(\cdot) \rangle_{\mathbb{L}^{2,N}_m(G_N)} := \sum_{k=0}^{N-1} \langle u^k, v^k \rangle_{\mathbb{R}^m},$$

$$||u_N(\cdot)||_{\mathbb{L}^{2,N}_m(G_N)} := (\langle u_N(\cdot), u_N(\cdot) \rangle_{\mathbb{L}^{2,N}_m(G_N)})^{1/2} = (\sum_{k=0}^{N-1} ||u^k||^2_{\mathbb{R}^m})^{1/2}.$$

The space $\mathbb{L}^{2,N}_m(G_N)$ is in one-to-one correspondence with the Euclidean space $\mathbb{R}^{mN}$. The Hilbert space $\mathbb{L}^2_m([0, 1])$ and the set $\mathcal{U}$ are replaced by the finite-dimensional Hilbert space $\mathbb{L}^{2,N}_m(G_N)$ and by $\mathcal{U}_N$, respectively. Evidently, we have a restriction of the function $\hat{J}(\cdot)$ on $\mathbb{L}^{2,N}_m(G_N)$.

The discrete minimization problem (6) approximates the continuous-time problem (5). We mainly focus our attention on the application of a proximal-based method to the convex-linear optimization problem (5) and on the convergence of the corresponding discrete approximation schemes (6). For this purpose we examine approximations for an abstract convex optimization problem in Hilbert space. At first, we establish the convergence properties of these approximations. We next apply the results for the abstract optimization problem to problem (5).

3. Proximal point method for convex optimization

This section contains some preliminary results. Let Z be a real Hilbert space. Consider the following problem of convex minimization

$$\begin{aligned} &\text{minimize } f(z) \\ &\text{subject to } z \in Q, \end{aligned} \tag{7}$$

where $Q \subset Z$ is a bounded, convex, closed set, $f : Z \rightarrow \bar{\mathbb{R}}$ is a proper convex functional and $\bar{\mathbb{R}} := \mathbb{R}\bigcup\{\infty\}$. In addition to these assumptions we suppose that $f(\cdot)$ is bounded on $Q + \epsilon B$. Here B is the open unit ball of Z and $\epsilon > 0$. Clearly, $Q + \epsilon B \subset \text{int dom} f := \{z \in Z \mid f(z) < \infty\}$.

The existence of an optimal solution of the convex minimizing problem (7) is guaranteed. Since $f(\cdot)$ is bounded on $Q+\epsilon B$, it follows that $f(\cdot)$ is a continuous on $Q+\epsilon B$ functional (see e.g., [17]). That is, $f(\cdot)$ is lower semicontinuous. The proper convex, lower semicontinuous functional $f(\cdot)$ attains its minimum on the bounded, convex, closed set Q [13]. Thus (7) has an optimal solution $z^{opt} \in Q$.

Note that a convex function on a convex subset of an infinite dimensional topological vector space does not need to be continuous on the interior of its domain. For instance, any discontinuous linear functional on an infinite dimensional topological vector space provides such an example.

Let F be the set of optimal solutions of (7). The next result is an immediate consequence of the boundedness hypothesis.

Lemma 1. *The functional $f(\cdot)$ is Lipschitz on Q.*

Proof. See [26], Theorem 10.4. □

We introduce the proximal mapping

$$\mathcal{P}_{f,Q,\chi} : \alpha \rightarrow \text{Argmin}_{z\in Q}[f(z) + \frac{\chi}{2}||z-\alpha||^2], \;\; \chi > 0, \;\; \alpha \in Z$$

and define the classical proximal point method [27, 18]

$$\begin{aligned} &z^{i+1}_{cl} \approx \mathcal{P}_{f,Q,\chi_i}(z^i_{cl}), \\ &z^0_{cl} \in Q, \;\; i = 0, 1, \dots \,, \end{aligned}$$

where $\{\chi_i\}$ is a given sequence with $0 < \chi_i \leq C < \infty$. Thus the original problem of convex minimization is replaced by a sequence of the auxiliary problems

$$\begin{aligned} &f(z) + \frac{\chi_i}{2}||z - z^i_{cl}||^2 \rightarrow \min, \\ &z \in Q, \;\; i = 0, 1, \dots \end{aligned}$$

with strong convex objective functionals. Evidently, for constructive solving the given problem of convex minimization the proximal point method must be combined with some numerical procedures for the auxiliary problems.

Suppose that the approximation $\{z_{cl}^i\},\ i = 0, 1, ...$ satisfies the estimate

$$||z_{cl}^{i+1} - \mathcal{P}_{f,Q,\chi_i}(z_{cl}^i)|| \leq \epsilon_i,\ \sum_{i=0}^{\infty} \frac{\epsilon_i}{\chi_i} < \infty,$$

for all $i = 0, 1, ...$. Under these conditions the sequence $\{z_{cl}^i\}$ converges in the weak topology to a point $z^{opt} \in F$ [18]. Besides, $\{z_{cl}^i\},\ i = 0, 1, ...$ is a minimizing sequence. In specific cases we have the strong convergence of the corresponding minimizing sequences (see e.g., [27, 3, 2]).

4. Convex approximations

This section is devoted to constructing some approximation schemes for the convex problems of the type (7). Let $\{Z_N\},\ N \in \mathbb{N}$ be a sequence of subspaces of Z such that

$$Z \supseteq ... \supseteq Z_{N+1} \supseteq Z_N, ..., \supseteq Z_1\ ,\ N \in \mathbb{N}.$$

The norm $||\cdot||_{Z_N}$ of the subspace Z_N is induced by the norm $||\cdot||_Z$ of the Hilbert space Z. In parallel with (7) we consider the sequence of the minimization problems

$$\begin{aligned} &\text{minimize } f(z) \\ &\text{subject to } z \in Q_N, \end{aligned} \tag{8}$$

where $\{Q_N\},\ Q_N \subset Z_N$ is a sequence of bounded, convex, closed subsets of Z_N. We assume that

$$Q \supseteq ... \supseteq Q_{N+1} \supseteq Q_N, ..., \supseteq Q_1\ ,\ N \in \mathbb{N}$$

and $\text{int}Q_N \neq \emptyset$. In fact we deal with some restrictions of the function $f(\cdot)$ on Z_N, however, we use the same notation $f(\cdot)$. Denote

$$f^{opt} := \inf_{z \in Q} f(z),\ f^{opt,N} := \inf_{z \in Q_N} f(z).$$

We shall use the familiar concept.

Definition 1. The sequence of problems (8) is called an approximating sequence for (7) if

$$\lim_{N \to \infty} f^{N,opt} = f^{opt}.$$

Let $z_N^{opt} \in Q_N$ be an optimal solution of (8). The following theorem is an extension of the convergence results of Vasil'ev [32].

Theorem 1. *Assume that for all $z^{opt} \in F$ there exists a mapping*

$$P_N : Z \to Z_N,\ N \in \mathbb{N}$$

such that $f(P_N(z^{opt})) - f(z^{opt}) \leq \gamma_N$, where $\lim_{N\to\infty} \gamma_N = 0$. Then (8) is an approximating sequence for problem (7).

Proof. The sequence $\{f(z_N^{opt}) - f(z^{opt})\}$ is a monotonically decreasing and bounded below sequence. This means that $\{f(z_N^{opt}) - f(z^{opt})\}$ is a convergent sequence. Moreover, $f(z_N^{opt}) - f(z^{opt}) \geq 0$ and

$$\gamma_N \geq f(P_N(z^{opt})) - f(z^{opt}) \geq f(z_N^{opt}) - f(z^{opt}) \geq 0.$$

We conclude that

$$0 = \lim_{N\to\infty} \gamma_N \geq \lim_{N\to\infty} f(z_N^{opt}) - f(z^{opt}) \geq 0.$$

In other words, (8) is an approximating sequence for (7). □

Let $\mathcal{P}^n_{f,Q_N,K_N}(z_N) := (\underbrace{\mathcal{P}_{f,Q_N,K_N} \circ ... \circ \mathcal{P}_{f,Q_N,K_N}}_{n})(z_N),\ n \in \mathbb{N}$. Using the classical proximal point method, we define the sequence of approximations

$$\begin{aligned} & z_1 \in Q_1, \\ & z_{N+1} = \mathcal{P}_{f,Q_{N+1},K_{N+1}}(\mathcal{P}^n_{f,Q_N,K_N}(z_N)), \\ & \infty > \tilde{K} > K_N \downarrow K > 0,\ n \in \mathbb{N},\ N = 1,2,..., \end{aligned} \tag{9}$$

where $\tilde{K},\ K$ are constants and

$$\mathcal{P}_{f,Q_{N+1},K_{N+1}}(\mathcal{P}^n_{f,Q_N,K_N}(z_N)) := \text{Argmin}_{z\in Q_{N+1}}[f(z)+ \\ + \frac{K_{N+1}}{2}||z - \mathcal{P}^n_{f,Q_N,K_N}(z_N)||^2_{Z_{N+1}}],\ z_N \in Q_N.$$

We now examine some properties of the introduced sequence of approximations (9).

Lemma 2. *Let $\{z_N\}$ be the sequence generated by the method (9). Then the following inequality*

$$||z_{N+1} - \mathcal{P}^n_{f,Q_N,K_N}(z_N)||_{Z_{N+1}} \leq \sqrt{\frac{2}{K_{N+1}}(f(z_N) - f(z_{N+1}))}, \tag{10}$$

holds.

Proof. By definition of $\mathcal{P}_{f,Q_N,K_N}(\cdot)$,

$$\mathcal{P}_{f,Q_N,K_N}(z_N) := \text{Argmin}_{z\in Q_N}[f(z) + \frac{K_N}{2}||z - z_N||^2_{Z_N}].$$

Hence

$$\begin{aligned} & f(z_N) = f(z_N) + \frac{K_N}{2}||z_N - z_N||^2_{Z_N} \geq f(\mathcal{P}_{f,Q_N,K_N}(z_N))+ \\ & + \frac{K_N}{2}||\mathcal{P}_{f,Q_N,K_N}(z_N) - z_N||^2_{Z_N} \geq f(\mathcal{P}_{f,Q_N,K_N}(z_N)) = f(\mathcal{P}_{f,Q_N,K_N}(z_N))+ \\ & + \frac{K_N}{2}||\mathcal{P}_{f,Q_N,K_N}(z_N) - \mathcal{P}_{f,Q_N,K_N}(z_N)||^2_{Z_N} \geq f(\mathcal{P}^2_{f,Q_N,K_N}(z_N))+ \\ & + \frac{K_N}{2}||\mathcal{P}^2_{f,Q_N,K_N}(z_N) - \mathcal{P}_{f,Q_N,K_N}(z_N)||^2_{Z_N} \geq ... \geq f(\mathcal{P}^n_{f,Q_N,K_N}(z_N)) \geq \\ & \geq f(\mathcal{P}^{n+1}_{f,Q_N,K_N}(z_N)) + \frac{K_N}{2}||\mathcal{P}^{n+1}_{f,Q_N,K_N}(z_N) - \mathcal{P}^n_{f,Q_N,K_N}(z_N)||^2_{Z_N}. \end{aligned}$$

The sequence $\{K_N\}$ is a monotonically decreasing sequence $K_N \downarrow K$. Since

$$\inf_{z\in Q_N}[f(z) + \frac{K_N}{2}||z-\alpha||^2_{Z_{N+1}}] \geq \inf_{z\in Q_{N+1}}[f(z) + \frac{K_{N+1}}{2}||z-\alpha||^2_{Z_{N+1}}],\ \alpha \in Q_N,$$

we conclude

$$f(\mathcal{P}^{n+1}_{f,Q_N,K_N}(z_N)) + \frac{K_N}{2}||\mathcal{P}^{n+1}_{f,Q_N,K_N}(z_N) - \mathcal{P}^{n}_{f,Q_N,K_N}(z_N)||^2_{Z_{N+1}} \geq$$
$$\geq f(z_{N+1}) + \frac{K_{N+1}}{2}||z_{N+1} - \mathcal{P}^{n}_{f,Q_N,K_N}(z_N)||^2_{Z_{N+1}} \geq f(z_{N+1}).$$

We have

$$f(z_N) \geq f(z_{N+1}) + \frac{K_{N+1}}{2}||z_{N+1} - \mathcal{P}^{n}_{f,Q_N,K_N}(z_N)||^2_{Z_{N+1}}.$$

It follows that

$$f(z_N) - f(z_{N+1}) \geq \frac{K_{N+1}}{2}||z_{N+1} - \mathcal{P}^{n}_{f,Q_N,K_N}(z_N)||^2_{Z_{N+1}} \geq 0$$

and

$$\sqrt{\frac{2}{K_{N+1}}(f(z_N) - f(z_{N+1}))} \geq ||z_{N+1} - \mathcal{P}^{n}_{f,Q_N,K_N}(z_N)||_{Z_{N+1}}.$$

Thus the inequality (10) holds. □

Using the uniform convexity of the function

$$f(\cdot) + \frac{K_N}{2}||\cdot - \alpha||_{Z_{N+1}},\ \alpha \in Q_N,$$

one can obtain the following estimate [2]

$$||z_{N+1} - \mathcal{P}^{n}_{f,Q_N,K_N}(z_N)||_{Z_{N+1}} \leq \frac{2}{\sqrt{3K_{N+1}}}\sqrt{(f(z_N) - f(z_{N+1}))}.$$

By Lemma 1, the function $f(\cdot)$ is Lipschitz on Q. Denote by L the corresponding Lipschitz constant. We see that

$$||z_{N+1} - \mathcal{P}^{n}_{f,Q_N,K_N}(z_N)||_{Z_{N+1}} \leq \sqrt{\frac{2L}{K_{N+1}}||z_N - z_{N+1}||_{Z_{N+1}}}.$$

Our next result presents a convergence property of the introduced sequence of approximations (9).

Theorem 2. *Let $\{z_N\}$ be the sequence generated by the method (9). Assume that for all $z^{opt} \in F$ there exists a mapping $P_N : Z \to Z_N,\ N \in \mathbb{N}$ such that*

$$f(P_N(z^{opt})) - f(z^{opt}) \leq \gamma_N,\ \lim_{N\to\infty} \gamma_N = 0.$$

Then $\lim_{N\to\infty} \lim_{n\to\infty} f(z_N) = f^{opt}$ and

$$\lim_{N\to\infty} \lim_{n\to\infty} ||z_N - \mathcal{P}^{n}_{f,Q_N,K_N}(z_N)||_Z = 0.$$

Proof. From the minimizing properties of $\mathcal{P}_{f,Q_N,K_N}(\cdot)$ and $\mathcal{P}_{f,Q_{N+1},K_{N+1}}(\cdot)$ we deduce

$$\begin{aligned}
&f(\mathcal{P}^n_{f,Q_N,K_N}(z_N)) = f(\mathcal{P}^n_{f,Q_N,K_N}(z_N))+\\
&+\frac{K_N}{2}||\mathcal{P}^n_{f,Q_N,K_N}(z_N)-\mathcal{P}^n_{f,Q_N,K_N}(z_N)||^2_{Z_N} \geq f(\mathcal{P}^{n+1}_{f,Q_N,K_N}(z_N))+\\
&+\frac{K_N}{2}||\mathcal{P}^{n+1}_{f,Q_N,K_N}(z_N)-\mathcal{P}^n_{f,Q_N,K_N}(z_N)||^2_{Z_N} \geq\\
&\geq f(z_{N+1})+\frac{K_{N+1}}{2}||z_{N+1}-\mathcal{P}^n_{f,Q_N,K_N}(z_N)||^2_{Z_{N+1}} \geq f(z_{N+1}).
\end{aligned}$$

Using the convergence properties of the classical proximal point method, we obtain $\lim_{n\to\infty}|f(\mathcal{P}^n_{f,Q_N,K_N}(z_N))-f(z_N^{opt})|=0$. Hence

$$\lim_{n\to\infty}|f(z_{N+1})-f(z_N^{opt})|=0.$$

By Theorem 1, we have $\lim_{N\to\infty} f(z_N^{opt}) = f^{opt}$. Since $f(z_{N+1}) \geq f^{opt}$, we conclude that $\lim_{N\to\infty}\lim_{n\to\infty} f(z_{N+1}) = f^{opt}$. By Lemma 2,

$$||z_{N+1}-\mathcal{P}^n_{f,Q_N,K_N}(z_N)||_{Z_{N+1}} \leq \sqrt{\frac{2}{K_{N+1}}(f(z_N)-f(z_{N+1}))}.$$

It follows that $\lim_{N\to\infty}\lim_{n\to\infty}||z_{N+1}-\mathcal{P}^n_{f,Q_N,K_N}(z_N)||_Z = 0$, as claimed. Thus the assertions of the theorem are proved. □

The sequence of solutions of the auxiliary problems

$$\begin{aligned}
&f(z)+\frac{K_{N+1}}{2}||z-\mathcal{P}^n_{f,Q_N,K_N}(z_N)||^2_{Z_{N+1}} \to \min,\\
&\text{subject to } z \in Q_{N+1},
\end{aligned}$$

is a minimizing sequence for the initial problem (7) subject to the condition that $n \to \infty$.

Corollary 1. *Let $\{z_N\}$ be the sequence generated by the method (9). Assume that for all $z^{\mathrm{opt}} \in F$ there exists a mapping $P_N : Z \to Z_N,\ N \in \mathbb{N}$ such that*

$$f(P_N(z^{\mathrm{opt}}))-f(z^{\mathrm{opt}}) \leq \gamma_N,\ \lim_{N\to\infty}\gamma_N = 0.$$

Then there exists a sequence of numbers $\{n_i\},\ n_i \in \mathbb{N}$ such that

$$\lim_{N\to\infty}\lim_{n_i\to\infty}||z_{N+1}-z_N^{opt}||_Z = 0.$$

Proof. Using the convergence properties of the classical proximal point method, we establish the convergence in the weak topology of the sequence

$$\{\xi_n\},\ \xi_n := \mathcal{P}^n_{f,Q_N,K_N}(z_N)$$

to an optimal solution z_N^{opt} of (8). Evidently, there exists a subsequence

$$\{\xi_{n_i}\},\ \xi_{n_i} := \mathcal{P}^{n_i}_{f,Q_N,K_N}(z_N),\ n_i \in \mathbb{N}$$

such that $\lim_{n_i\to\infty}||\mathcal{P}^{n_i}_{f,Q_N,K_N}(z_N) - z^{opt}_N||_Z = 0$. By Lemma 2,

$$||z_{N+1} - \mathcal{P}^{n_i}_{f,Q_N,K_N}(z_N)||_{Z_{N+1}} \leq \sqrt{\frac{2}{K_{N+1}}(f(z_N) - f(z_{N+1}))}.$$

Theorem 2 implies that

$$\lim_{N\to\infty}\lim_{n_i\to\infty}||z_{N+1} - \mathcal{P}^{n_i}_{f,Q_N,K_N}(z_N)||_Z = 0.$$

Then by using the triangle inequality we see that

$$\lim_{N\to\infty}\lim_{n_i\to\infty}||z_{N+1} - z^{opt}_N||_Z = 0.$$

The corollary is proved. □

5. Discretizations of the Optimal Control Problem

Let us apply the convex approximations given above to the optimal control problem (5). We describe an algorithm for numerical solving the optimization problem (5) and a convergence result associated with it. Our algorithm is based on the proximal-like method (9). We will consider the first order method for solving the auxiliary problems but similar results can also be obtained for a second order scheme. Note that the convex-linear minimization problem (5) has a solution (see Section 3).

Let $d^{opt}(\cdot)$ be an optimal solution of (5). We introduce the sequence of spaces

$$\mathbb{L}^2_m([0,1]) \supseteq ... \supseteq \mathbb{L}^{2,N+1}_m(G_N) \supseteq \mathbb{L}^{2,N}_m(G_N) \supseteq ... \supseteq \mathbb{L}^{2,1}_m(G_N),$$

and the sequence of sets $\mathcal{U} \supseteq ... \supseteq \mathcal{U}_{N+1} \supseteq \mathcal{U}_N \supseteq ... \supseteq \mathcal{U}_1,\ N \in \mathbb{N}$ such that $\mathcal{U}_N := \mathcal{U}\bigcap\mathbb{L}^{2,N}_m(G_N)$. Let us consider the mapping

$$P_N : \mathbb{L}^2_m([0,1]) \to \mathbb{L}^{2,N}_m(G_N),\ N \in \mathbb{N}$$

defined as follows

$$P_N(d^{opt}(\cdot)) = (d^0, ..., d^{N-1})^T,\ d^k := \frac{1}{\Delta t_{k+1}}\int_{t_k}^{t_{k+1}} d^{opt}(t)dt,$$
$$k = 0, ..., N-1.$$

We now establish an important fact: the value $P_N(d^{opt}(\cdot))$ belongs to the set $\mathcal{U}_N - u_N(\cdot)$ for all $N \in \mathbb{N}$.

Lemma 3. *For every $N \in \mathbb{N}$ and for every solution $d^{opt}(\cdot)$ of problem (5),*

$$P_N(d^{opt}(\cdot)) \in \mathcal{U}_N - u_N(\cdot).$$

Proof. We have $u(\cdot) \in \mathcal{U}$ (see Introduction). Evidently, $u(t) + d^{opt}(t) \in U$. Since U is convex and closed, it follows [32] that

$$(u(t) + \frac{1}{\Delta t_{k+1}}\int_{t_k}^{t_{k+1}} d^{opt}(t)dt) \in U,\ k = 0, ..., N-1$$

and

$$(u(t_k) + \frac{1}{\Delta t_{k+1}}\int_{t_k}^{t_{k+1}} d^{opt}(t)dt) \in U,\ k = 0, ..., N-1.$$

Moreover, $u_N(\cdot) + P_N(d^{opt}(\cdot)) \in \mathbb{L}^{2,N}_m(G_N)$. Hence

$$(u_N(\cdot) + P_N(d^{opt}(\cdot))) \in \mathbb{L}^{2,N}_m(G_N) \bigcap \mathcal{U} = \mathcal{U}_N.$$

We conclude that $P_N(d^{opt}(\cdot)) \in \mathcal{U}_N - u_N(\cdot)$. The proof is finished. □

Fix $N \in \mathbb{N}$, $u_N(\cdot) \in \mathcal{U}_N$ and consider the set $\mathcal{U}_{N+M} - u_N(\cdot)$ and the mapping

$$P_{N+M}(d^{opt}(\cdot)),\ M = 1, 2, \dots .$$

Let $G_{N+M} := \{t_0 = 0, t_1, \dots, t_{N+M} = 1\}$ be the corresponding partition of $[0, 1]$. We assume that $G_N \subset G_{N+M}$.

Lemma 4. *For every $N \in \mathbb{N}$ and for every solution $d^{opt}(\cdot)$ of problem (5),*

$$P_{N+M}(d^{opt}(\cdot)) \in \mathcal{U}_{N+M} - u_N(\cdot),\ M = 1, 2, \dots .$$

Proof. The lemma can be proved in the same way as Lemma 3. □

For fixed $N \in \mathbb{N}$, $u_N(\cdot)$ and chosen $d_N(\cdot) \in \mathcal{U}_N - u_N(\cdot)$ we introduce the discrete approximations of the type (8) for problem (5)

$$\begin{aligned}
&\text{minimize } \hat{J}(d_{N+M}(\cdot))\\
&\text{subject to } d_{N+M}(\cdot) \in \mathcal{U}_{N+M} - u_N(\cdot),\\
&\tilde{h}_j(u_N(\cdot)) + \langle \nabla \tilde{h}_j(u_N(\cdot)), d_{N+M}(\cdot)\rangle \le 0\ \forall j \in I,\\
&\tilde{q}(u_N(\cdot))(t) + \langle \nabla \tilde{q}(u_N(\cdot))(t), d_{N+M}(\cdot)\rangle \le 0\ \forall t \in [0, 1],\\
&M = 1, 2, \dots ,
\end{aligned}$$

where

$$\begin{aligned}
&\hat{J}(d_{N+M}(\cdot)) = \sum_{k=0}^{N+M-1} f_0(t_k, y_{N+M}(t_k), u_N(t_k) + d_{N+M}(t_k))\Delta t_{k+1},\\
&y_{N+M}(t_{k+1}) = y_{N+M}(t_k) + \Delta t_{k+1}(f_x(t_k, x^{u_N}(t_k), u_N(t_k))y_{N+M}(t_k) +\\
&+ f_u(t_k, x^{u_N}(t_k), u_N(t_k))d_{N+M}(t_k)),\ y_{N+M}(t_0) = 0,\ k = 0, 1, \dots, N + M - 1.
\end{aligned}$$

By $x^{u_N}(\cdot)$ we denote here the solution to the initial value problem (2) for $u_N(\cdot)$. The proximal-like algorithm in this case is:

$$\begin{aligned}
&d_{N+M}(\cdot) = \mathcal{P}_{\hat{J}, \mathcal{U}_{N+M} - u_N(\cdot), K_{N+M}}(\mathcal{P}^n_{\hat{J}, \mathcal{U}_{N+M-1} - u_N(\cdot), K_{N+M-1}}(d_{N+M-1}(\cdot))),\\
&n \in \mathbb{N},\quad \infty > \tilde{K} > K_{N+M} \downarrow K > 0,\ M = 1, 2, \dots ,
\end{aligned} \tag{11}$$

where

$$\begin{aligned}
&\mathcal{P}_{\hat{J}, \mathcal{U}_{N+M} - u_N(\cdot), K_{N+M}}(\mathcal{P}^n_{\hat{J}, \mathcal{U}_{N+M-1} - u_N(\cdot), K_{N+M-1}}(d_{N+M-1}(\cdot))) =\\
&= \text{Argmin}_{d(\cdot) \in \mathcal{U}_{N+M} - u_N(\cdot)}[\hat{J}(d(\cdot)) + \frac{K_{N+M}}{2}||d(\cdot) -\\
&- \mathcal{P}^n_{\hat{J}, \mathcal{U}_{N+M-1} - u_N(\cdot), K_{N+M-1}}(d_{N+M-1}(\cdot))||^2_{\mathbb{L}^{2,N+M}_m(G_{N+M})}].
\end{aligned}$$

Using the method (11), we obtain a sequence of discrete approximations

$$\{d_{N+M}(\cdot)\},\ M = 1, 2, \dots .$$

The following theorem establishes the convergence properties of these finite-dimensional approximations.

Theorem 3. *Let $\{d_{N+M}(\cdot)\}$ be the sequence generated by the method (11). Then*

$$\lim_{M\to\infty}\lim_{n\to\infty}\hat{J}(d_{N+M}(\cdot)) = \hat{J}(d^{opt}(\cdot))$$

and

$$\lim_{M\to\infty}\lim_{n\to\infty}||d_{N+M}(\cdot) - \mathcal{P}^n_{\hat{J},\mathcal{U}_{N+M-1}-u_N(\cdot),K_{N+M-1}}(d_{N+M-1}(\cdot))||_{\mathbb{L}^2_m([0,1])} = 0.$$

Proof. The objective functional

$$\hat{J}(d(\cdot)) = \int_0^1 f_0(t, y^{u,d}(t), u(t) + d(t))dt$$

is proper convex (with respect to $d(\cdot)$). We establish that the objective functional $\hat{J}(d(\cdot))$ is bounded on

$$\mathcal{U} - u(\cdot) + \epsilon B_{\mathbb{L}^2_m([0,1])},$$

where $B_{\mathbb{L}^2_m([0,1])}$ is the open unit ball of $\mathbb{L}^2_m([0,1])$ and $\epsilon > 0$.

We have introduced the following hypothesis (see Introduction):

(i) $f(t,\cdot,\cdot)$ is differentiable,

(ii) f, f_x, f_u are continuous and there exists a constant $S < \infty$ such that

$$||f_x(t,x,u)|| \le S \quad \forall\ (t,x,u) \in [0,1]\times\mathbb{R}^r\times U.$$

That is, the unique solution $y^{u,d}(\cdot)$ of (3) is bounded. More precisely, there exists a constant $c > 0$ such that

$$||y^{u,d}(\cdot)||_\infty \le c||d(\cdot)||_\infty,$$

where $d(\cdot) \in \mathcal{U} - u(\cdot)$ (see e.g., [25]). Since the set $\mathcal{U} - u(\cdot)$ is bounded, it follows that $y^{u,d}$ is bounded and $y^{u,d}(t)$ belongs to a compact set $A \subset \mathbb{R}^r$. The continuous function f_0 is bounded on the compact set $[0,1]\times A\times\tilde{U}$. This implies the boundedness of the functional $\hat{J}(d(\cdot))$ on $\mathcal{U} - u(\cdot) + \epsilon B_{\mathbb{L}^2_m([0,1])}$. By Lemma 1, the functional $\hat{J}(d(\cdot))$ is a Lipschitz-continuous on $\mathcal{U} - u(\cdot)$ functional. The sets

$$\mathcal{U} - u(\cdot),\ \mathcal{U}_N - u_N(\cdot),\ N\in\mathbb{N} \text{ and } \mathcal{U}_{N+M} - u_N(\cdot),\ M = 1,2,...$$

are bounded, convex and closed subsets of

$$\mathbb{L}^2_m([0,1]),\ \mathbb{L}^{2,N}_m(G_N) \text{ and } \mathbb{L}^{2,N+M}_m(G_{N+M}),$$

respectively. Thus all assumption of Section 3 and Section 4 are satisfied.

By Lemma 4, the value of the introduced mapping $P_{N+M}(d^{opt}(\cdot))$ belongs to the set $\mathcal{U}_{N+M} - u_N(\cdot)$. The square integrable function $d(\cdot) \in \mathcal{U} - u(\cdot)$ is bounded on $[0,1]$. Moreover, $\lim_{M\to\infty}\Delta t_k = 0$, where $t_k \in G_{N+M}$. Therefore

$$\lim_{M\to\infty}||P_{N+M}(d^{opt}(\cdot)) - d^{opt}(\cdot)||_{\mathbb{L}^2_m([0,1])} =$$

$$= \lim_{M\to\infty}(\int_0^1 ||P_{N+M}(d^{opt}(\cdot))(t) - d^{opt}(t)||^2_{\mathbb{R}^m}dt)^{1/2} =$$

$$= \lim_{M\to\infty}(\int_0^1 ||\sum_{k=0}^{N+M-1}\phi_k(t)\frac{1}{\Delta t_{k+1}}\int_{t_k}^{t_{k+1}} d^{opt}(\tau)d\tau - d^{opt}(t)||^2_{\mathbb{R}^m}dt)^{1/2} = 0.$$

The convex functional $\hat{J}(\cdot)$ is continuous (Lipschitz-continuous). Therefore,

$$\lim_{M\to\infty} |\hat{J}(P_{N+M}(d^{opt}(\cdot))) - \hat{J}(d^{opt}(\cdot))| = 0$$

This means that for each d^{opt} there exists a sequence $\{\gamma_M\}$ such that

$$\hat{J}(P_{N+M}(d^{opt}(\cdot))) - \hat{J}(d^{opt}(\cdot)) \leq \gamma_M, \ \lim_{M\to\infty} \gamma_M = 0.$$

Thus the mapping $P_{N+M}(d^{opt})$ introduced above satisfies the assumptions of Theorem 1 and Theorem 2. By Theorem 2, we have

$$\lim_{M\to\infty} \lim_{n\to\infty} \hat{J}(d_{N+M}(\cdot)) = \hat{J}(d^{opt}(\cdot)).$$

It also follows from Theorem 2 that

$$\lim_{M\to\infty} \lim_{n\to\infty} ||d_{N+M}(\cdot) - \mathcal{P}^n_{\hat{J},\mathcal{U}_{N+M-1}-u_N(\cdot),K_{N+M-1}}(d_{N+M-1}(\cdot))||_{\mathbb{L}^2_m([0,1])} = 0.$$

The proof is finished. □

In other words, the sequence $\{d_{N+M}(\cdot)\}$ generated by the proximal-like algorithm (11) is a minimizing sequence for the optimal control problem (5) subject to the condition that $n \to \infty$. The presented proximal-like method (11) is a numerically stable procedure for solving the linearized problem (5). This method can be used as a tool for numerical treating the initial optimal control problem (1).

6. The Numerical Aspect

For fixed $N \in \mathbb{N}$ and $u_N(\cdot)$ we examine the sequence generated by the introduced method (11). Let $d_{N+M}(\cdot) \in \mathcal{U}_{N+M} - u_N(\cdot)$ be an element of this sequence such that

$$|\hat{J}(d_{N+M}(\cdot)) - \hat{J}(d^{opt}(\cdot))| < \delta_{M,N},$$

where $\delta_{M,N} > 0$ is a sufficiently small real number. We now define the next approximation of the control function for the nonlinear control problem (1) in the following way $u_{N+M}(\cdot) := u_N(\cdot) + d_{N+M}(\cdot)$. It is evident that $u_{N+M}(\cdot) \in \mathcal{U}_{N+M}$. Given this control function the solution to the initial value problem (2) is denoted by $x^{u_{N+M}}(\cdot)$. Using the computed trajectory $x^{u_{N+M}}(\cdot)$, we can consider the next linearization step

$$\dot{y}(t) = f_x(t, x^{u_{N+M}}(t), u_{N+M}(t))y(t) + f_u(t, x^{u_{N+M}}(t), u_{N+M}(t))d(t),$$
$$y(0) = 0$$

and the corresponding linear optimization problem (5). If we replace N in (6) by $N + M$, then we have the possible next discrete approximation for the obtained linear problem with respect to the higher-order partition G_{N+M} of $[0, 1]$. Note that in this paper we don't consider the convergence properties of the linear approximations given above. Convergence of some linearization schemes for optimal control problems is examined in [31, 25].

The proximal-like method (11) must be combined with an effective algorithm for computing the solution of the following auxiliary optimization problems

$$\hat{J}(d(\cdot)) + \frac{K_{N+M}}{2}||d(\cdot) - \alpha||^2_{\mathbb{L}^{2,N+M}_m(G_{N+M})} \to \min_{d(\cdot)\in\mathcal{U}_{N+M}-u_N(\cdot)}, \quad \alpha \in \mathbb{L}^{2,N+M}_m(G_{N+M}) \tag{12}$$

and

$$\hat{J}(d(\cdot)) + \frac{K_{N+M-1}}{2}||d(\cdot) - \alpha||^2_{\mathbb{L}^{2,N+M-1}_m(G_{N+M-1})} \to \min_{d(\cdot)\in\mathcal{U}_{N+M-1}-u_N(\cdot)}, \quad \alpha \in \mathbb{L}^{2,N+M-1}_m(G_{N+M-1}). \tag{13}$$

We compute the solution of (12) in each step M and the solution of (13) in each step n of the algorithm (11). In our paper we use a variant of the reduced gradient method for this purpose (see e.g., [23, 31, 32]). The reduced gradient $\nabla_k\hat{J}(d_{N+M}(\cdot))$ for $t_k \in G_{N+M}$ can be computed as follows

$$\nabla_k\hat{J}(d_{N+M}(\cdot)) = H_d(t_k, y_{N+M}(t_k), u_N(t_k), d_{N+M}(t_k), p(t_{k+1})),$$
$$k = 0, ..., N + M - 1 \ ,$$

where $p(t_k),\ k = 1, ..., N + M$ are the adjoint variables

$$p(t_k) = -H_y(t_k, y_{N+M}(t_k), u_N(t_k), d_{N+M}(t_k), p(t_{k+1})),\ k = 1, ...N + M - 1,$$
$$p(t_{N+M}) = 0$$

and

$$H(t_k, y_{N+M}(t_k), u_N(t_k), d_{N+M}(t_k), p(t_{k+1})) :=$$
$$= f_0(t_k, y_{N+M}(t_k), u_N(t_k) + d_{N+M}(t_k))\Delta t_{k+1}-$$
$$- \langle p(t_{k+1}), \Delta t_{k+1}(f_x(t_k, x^{u_N}(t_k), u_N(t_k))y_{N+M}(t_k)+$$
$$+ f_u(t_k, x^{u_N}(t_k), u_N(t_k))d_{N+M}(t_k) + y_{N+M}(t_k)\rangle_{\mathbb{R}^n},\ k = 0, ..., N + M - 1 \ ,$$

is the Hamiltonian of the discrete convex-linear control problem. Note that the full reduced gradient in the step $N + M$ of the algorithm (11) is

$$\nabla\hat{J}(d_{N+M}(\cdot)) + K_{N+M}(d_{N+M}(\cdot) - \mathcal{P}^n_{\hat{J},\mathcal{U}_{N+M-1}-u_N(\cdot),K_{N+M}}(d_{N+M-1}(\cdot))).$$

As an example we consider the following ill posed optimal control problem

$$\text{minimize } J(u(\cdot)) = \int_0^1 x^2(t)dt$$
$$\text{subject to } \dot{x} = u \text{ a.e. on } t \in (0, 1],\ x(0) = 0,$$
$$u(\cdot) \in \mathbb{L}^2_m([0, 1]),\ |u(t)| \le 1 \text{ a.e.},\ x(1) \le 0.$$

This optimal control problem has a unique optimal solution $u^{opt}(t) = 0$ a.e., however the following minimizing sequence $u_r(t) = \sin(2\pi rt),\ r \in \mathbb{N}$ does not converge. Evidently, $x_r(t) = \frac{1}{2\pi r}(1 - \cos(2\pi rt)),\ x_r(1) = 0$ and

$$\lim_{r\to\infty} J(u_r(\cdot)) = \lim_{r\to\infty} \frac{3}{8\pi^2 r^2} = J(u^{opt}(\cdot)) = 0.$$

The objective functional $J(u(\cdot)) = \int_0^1 (\int_0^t u(\tau)d\tau)^2 dt$ is convex and the set of admissible control functions $\mathcal{U} = \{v(\cdot) \in \mathbb{L}^2_m([0,1]) \ : \ |v(t)| \leq 1 \text{ a.e. on } [0,1]\}$ is bounded, convex and closed. We apply the proximal-like algorithm (11) for $1 \leq N \leq 50$ and $1 \leq M \leq 50$. For $N = 50, \ M = 50$ the computed optimal control $\{u_{N+M}(\cdot)\}$ has the following property

$$\max_{0 \leq k \leq N+M} |u_{N+M}(t_k)| = 0.0016521.$$

The constraints were satisfied with tolerance 10^{-4}. The computed optimal objective value is $8 \cdot 10^{-7}$. The implementation of the algorithm, described above, was carried out, using the "Numerical Recipes in C" package [24] and the author program written in C.

7. Concluding Remarks

In this paper, we have shown that the proximal approach can be used for creating numerically stable discrete approximations for optimal control problems with state constraints. The introduced proximal-like method (11) can be combined not only with the used gradient-type method but also with a second order optimization method (see [23]). We have considered the optimal control problem (1) under convexity assumptions. Using a variant of the proximal point method for nonconvex optimization [18], one can extend the presented proximal-type methods (9) and (11) to some classes of nonconvex optimal control problems of the type (1).

Acknowledgements: The author is grateful to Professor A. Kaplan and to Professor W.H. Schmidt for their helpful comments and suggestions.

References

[1] A. Auslender, J.P. Crouzeix and P. Fedit, *Penalty proximal methods in convex programming*, Journal of Optimization Thery and Applications **55** (1987), 1-21.

[2] V. Azhmyakov and W.H. Schmidt, *Strong convergence of a proximal-based method for convex optimization*, Mathematical Methods of Operations Research **57** (2003), 393-407.

[3] A.B. Bakushinskii, *Solution methods for monotonous variational inequalities founded on the principle of iterative regularization*, U.S.S.R. Computational Mathematics and Mathematical Physics **17** (1977), 12-24.

[4] R. Bellman and S.E. Dreyfus, *Applied Dynamic Programming*, Princeton University Press, Princenton, 1962.

[5] H. Benker, A. Hamel and C. Tammer, *A proximal point algorithm for control approximation problems*, Mathematical Methods of Operations Research **43** (1996), 261-280.

[6] J.F. Bonnans, *On an algorithm for optimal control using Pontryagin's maximum principle*, SIAM Journal of Control and Optimization **24** (1986), 579-588.

[7] A.E. Bryson and Y.C. Ho, *Applied Optimal Control*, Wiley, New York, 1975.

[8] R. Bulirsch, F. Montrone and H.J. Pesch *Abort landing in the presence of windshear as a minimax optimal control problem, Part 2: Multiple chooting and homotopy*, Journal of Optimization Theory and Applications **70** (1991), 223-254.

[9] C. Büskens and H. Maurer, *SQP-Methods for solving optimal control problems with control and state constraints: adjoint variables, sensitivity analysis and real-time control*, Journal of Computational and Applied Mathematics **120** (2000), 85-108.

[10] A.L. Dontchev and F. Lempio, *Difference methods for differential inclusions: a survey*, SIAM Review **34** (1992), 263-294.

[11] A.L. Dontchev, *Discrete approximations in optimal control*, in: *Nonsmooth Analysis and Geometric Methods in Deterministic Optimal Control* (B.S.Mordukhovich and H.J.Sussmann, eds.), Springer-Verlag, New-York, 1996, pp.59-80.

[12] J.C. Dunn, *On state constraint representations and mesh-depent gradient projection convergence rates for optimal control problems*, SIAM Journal of Control and Optimization **39** (2000), 1082-1111.

[13] I. Ekeland and R. Temam, *Convex Analysis and Variational Problems*, North-Holland, Amsterdam, 1976.

[14] R.P. Fedorenko, *Priblizhyonnoye Reshenyie Zadach Optimalnogo Upravlenya* (Russian), Nauka, Moscow, 1978.

[15] M. Fukushima and H. Mline, *A generalized proximal point algrithm for certain non-convex minimization prolems*, International Journal of Systems Science **12** (1981), 989-1000.

[16] M. Fukushima and Y. Yamamoto *A second-order algorithm for continuous-time nonlinear optimal control problems*, IEEE Transaction on Automatic Control **31** (1986), 673-676.

[17] A.D. Ioffe and V.M. Tihomirov, *Theory of Extremal Problems*, North Holland, Amsterdam, 1979.

[18] A. Kaplan and R. Tichatschke, *Proximal point methods and nonconvex optimization*, Journal of Global Optimization **13** (1998), 389-406.

[19] A. Kaplan and R. Tichatschke, *A general view on proximal point methods to variational inequalities in Hilbert spaces-iterative regularization and approximation*, Journal of Nonlinear and Convex Analysis **2** (2001), 305-332

[20] F.S. Kupfer and E. Sachs, *Reduced SQP method for nonlinear heat conduction control problems*, International Series of Numerical Mathematics **111** (1993), 145-160.

[21] F. Leibfritz and E. Sachs, *Optimal static output feedback design using a trust region interior point method*, Presentation at the First Workshop on Nonlinear Optimization "Interior Point and Filter Methods", Coimbra, Portugal.

[22] K. Malanowski, *Finite difference approximations to constrained optimal control problems*, in: *Optimization and Optimal Control, Lecture Notes in Control and Information Sciences*, Springer-Verlag, New-York, 1981, pp.243-254.

[23] E. Polak, *Optimization*, Springer-Verlag, New York, 1997.

[24] W.H. Press, S.A. Teukolsky, W.T. Vetterling and B.P. Flannery, *Numerical Recipes in C*, Cambridge University Press, Cambridge, 1992.

[25] R. Pytlak, *Numerical Methods for Optimal Control Problems with State Constraints*, Springer-Verlag, Berlin, 1999.

[26] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.

[27] R.T. Rockafellar, *Monotone operators and the proximal point algorithm*, SIAM Journal on Control and Optimization **14** (1976), 877-898.

[28] Y. Sakawa, Y. Shindo and Y. Hashimoto, *Optimal control of a rotary crane*, Journal of Optimization Theory and Applications **35** (1981), 535-557.

[29] O.von Stryk, *Numerische Lösung Optimaler Steuerungsproblems: Diskretisierung, Parameteroptimierung und Berechnung der Adjungierten Variablen*, VDI-Verlag, Düsseldorf, 1995.

[30] M. Thera and R. Tichatschke (eds.) *Ill-Posed Variational Problems and Regularization Techniques*, Springer-Verlag, Berlin, 1999.

[31] K.L. Teo, C.J. Goh and K.H. Wong, *A Unifed Computational Approach to Optimal Control Problems*, Wiley, New York, 1991.

[32] F.P. Vasil'ev, *Methods for Solving the Extremal problems* (Russian), Nauka, Moscow, 1981.

[33] V.M. Veliov, *Second order discrete approximations to linear differential inclusions*, SIAM Journal of Numerical Analysis **29** (1992), 439-451.

[34] S.J. Wright, *Interior point method for optimal control of discrete-time systems*, Journal of Optimization Theory and Applications **77** (1993), 161-187.

Manuscript received August 20, 2003

Vadim Azhmyakov
Institute of Mathematics and Computer Sciences, Ernst-Moritz-Arndt University of Greifswald, Jahnstr. 15a, D-17487 Greifswald, Germany
E-mail address: azmjakow@uni-greifswald.de