A METHOD OF CENTERS BASED ON BARRIER FUNCTIONS FOR SOLVING OPTIMAL CONTROL PROBLEMS WITH CONTINUUM STATE AND CONTROL CONSTRAINTS

by

E. Polak, T. H. Yang, and D. Q. Mayne

Memorandum No. UCB/ERL M90/31

2 May 1990

COVER PAGE

A METHOD OF CENTERS BASED ON BARRIER FUNCTIONS FOR SOLVING OPTIMAL CONTROL PROBLEMS WITH CONTINUUM STATE AND CONTROL CONSTRAINTS

by

E. Polak, T. H. Yang, and D. Q. Mayne

Memorandum No. UCB/ERL M90/31

2 May 1990

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

TITLE PAGE

A METHOD OF CENTERS BASED ON BARRIER FUNCTIONS FOR SOLVING OPTIMAL CONTROL PROBLEMS WITH CONTINUUM STATE AND CONTROL CONSTRAINTS

by

E. Polak, T. H. Yang, and D. Q. Mayne

Memorandum No. UCB/ERL M90/31

2 May 1990

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

A METHOD OF CENTERS BASED ON BARRIER FUNCTIONS FOR SOLVING OPTIMAL CONTROL PROBLEMS WITH CONTINUUM STATE AND CONTROL CONSTRAINTS†

by

E. Polak*, T. H. Yang and D. Q. Mayne**

ABSTRACT

This paper describes a method of centers based on barrier functions for solving optimal control problems with continuum inequality constraints on the state and control. The method decomposes the original problem into a sequence of easily solved optimal control problems with control constraints only. The method requires only approximate solution of these problems.

† The research reported herein was sponsored in part by the National Science Foundation grant ECS-8713334, the Air Force Office of Scientific Research contract AFOSR-86-0116 and the State of California MICRO Program grant 532410-19900.

*Department of Electrical Engineering and Computer Sciences and the Electronics Research Laboratory, University of California, Berkeley, CA 94720, U.S.A.

**Department of Electrical Engineering and Computer Sciences, University of California, Davis, CA 95616, U.S.A. and Department of Electrical Engineering, Imperial College of Science and Technology, London, SW7 2BT, Great Britain.

1. Introduction

The difficulty of an optimal control problem is very much a function of the constraints. In the realm of optimal control problems, optimal control problems with control constraints and inequality state space constraints rank close to the top in terms of difficulty, or, alternatively, close to the bottom in terms of tractability. In this paper we make use of ideas contained in recent work on phase I - phase II methods of centers [Pol.2], methods of centers based on barrier functions [Hua.1, Mif.1], and barrier function methods for semi-infinite minimax problems [Pol.3] to construct a reasonably promising optimal control algorithm for solving optimal control problems with both control and inequality state space constraints. An important feature of this algorithm is that it decomposes the original problem into an infinite sequence of highly tractable optimal control problems with integral cost and control constraints only, each of which needs to be solved only approximately.

To help establish the extent to which this paper advances the state of the art, we will now discuss some of the earlier results in this area. We recall that free time problems can always be transcribed into fixed time problems by means of an augmentation of the dynamics (see [War.1] and hence we need to discuss fixed time problems only. First (see [Bak.1]), unconstrained optimal control problems and optimal control problems with inequality end point constraints (both without control constaints), with smooth dynamics, can be formulated as

$$\min_{u \in \mathbf{C}} f^0(u) , \tag{1.1}$$

and

$$\min_{u \in \mathbf{C}} \{ f^0(u) \mid f^j(u) \le 0 , \ j = 1, 2, \ldots, q \} , \tag{1.1b}$$

respectively, with all the functions $f^j(\cdot)$ continuously Frechet differentiable on the prehilbert space $\mathbf{C} \triangleq \{ L_\infty^m[0,1] , \|\cdot\|_2 \}$, where $\|\cdot\|_2$ denotes the norm in $L_2^\infty[0,1]$. These problems can be solved by exact analogs of the Armijo gradient method [Arm.1] (see e.g., [Kle.1]) and of the method of centers (see [Pir.1]), respectively.

Simple optimal control problems with control constraints only, assume the form

$$\min_{u \in \mathbf{U}} f^0(u) , \tag{1.2a}$$

where $\mathbf{U} \triangleq \{ u \in \mathbf{C} \mid u(t) \in U , \ t \in [0,1] \}$, with U is a compact subset of $\mathbb{R}^m$, and $f^0(\cdot)$ is as above. For simple sets U, these problems can be solved by analogs of the slow-to-converge Frank-Wolfe algorithm [Fra.1] and the faster Goldstein-Levitin-Polyak gradient projection method (see

[Ber.1]), as well as by two algorithms which are optimal control specific: the strong variations algorithm in [mayne&polak], and the relaxed control steepest descent algorithm in [War.1]. The addition of control constraints to (1.1b) results in a problem of the form

$$\min_{u \in \mathbf{U}} \{ f^0(u) \mid f^j(u) \le 0 \,, \; j = 1, 2, \ldots, q \} \,, \tag{1.2b}$$

which can be solved by an extension (see [May.2]) of the method of centers [Pir.1]. However, the control constraints result in a considerable increase in difficulty in the search direction finding problem. The addition of equality constraints to (1.2b) can be handled by means of exact penalty functions (see e.g., [May.4,5].

The most difficult optimal control problems have both control and state space constraints, and can assume the following abstract form:

$$\min_{u \in \mathbf{U}} \{ f^0(u) \mid f^j(u) \le 0 \,, \; j = 1, 2, \ldots, q \,, \; \max_{t \in [0,1]} \phi^k(u\,, t) \le 0 \,, \; k = 1, 2, \ldots, r \} \,, \tag{1.3}$$

where the functions $f^j(\cdot)$, and $\phi^k(\cdot\,,\cdot)$ are continuous all continuously differentiable. We recognize these problems as generalizations of finite dimensional semi-infinite programming problems (see [Pol.1]). The presence of the control constraints generates a major obstacle because it precludes the efficient use of minimax theorems in the solution of extremely difficult search direction finding problems (see [Pol.1] for their use in semi-infinite optimization).

Not counting heuristic algorithms, there appear to be only two algorithms in the literature for their solution of problems of the form (1.3). They are both extensions of the method in [Pir.1]; the one in [War.1] is based on the use of relaxed controls while the one in [May.1] is not. Both of these algorithms postulate extremely difficult search direction computations.

The algorithm which we will present in this paper is much simpler in structure than either of the algorithms [War.1] or [May.1]; furthermore, it is easily implemented using existing methods (such as in [Ber.1, Kle.1]). In Section 2, we will present our algorithm in a simplified (conceptual form). In Section 3 and 4 we give full details of two alternative versions of our new algorithm and prove their convergence to feasible stationary points. Computational results are reported in Section 5 and show that the algorithm performs satisfactorily.

2. A Conceptual Phase I - Phase II Method of Centers

We will consider optimal control problems defined in the prehilbert space $L_{\infty,2} \triangleq \{L_\infty^m[0,1], \|\cdot\|_2\}$, consisting of elements in $L_\infty^m[0,1]$, but endowed with the $L_2[0,1]$ scalar product $\langle \cdot, \cdot \rangle_2$ and corresponding norm $\|\cdot\|_2$. The problems are normalized, fixed-time problems with control and state space constraints, of the form

$$\mathbf{P}: \quad \min \{ f^0(u) \mid f^j(u) \le 0 \,,\ j = 1, 2, \dots, q_1 \,,$$

$$\max_{t \in [0,1]} \phi^k(u, t) \le 0 \,,\ k = 1, 2, \dots, q_2 \,,\ u \in G \} \,, \tag{2.1a}$$

where

$$G \triangleq \{ u \in L_{\infty,2} \mid u(t) \in U \,,\ \forall t \in [0,1] \} \,, \tag{2.1b}$$

with $U \subset \mathbb{R}^m$. The *cost function* $f^0 : L_{\infty,2} \to \mathbb{R}$ and the *end-point constraint functions* $f^j : L_{\infty,2} \to \mathbb{R}$, $j = 1, 2, \dots, q_1$ are defined by

$$f^j(u) \triangleq g^j(x^u(1)) \,; \tag{2.2a}$$

while the *state space* constraint functions $\phi^k : L_{\infty,2} \times [0,1] \to \mathbb{R}$, $k = 1, 2, \dots, q_2$, are defined by

$$\phi^k(u, t) \triangleq g^k(x^u(t)) \,, \tag{2.2b}$$

where the functions g^j, $g^k : \mathbb{R}^n \to \mathbb{R}$ and $x^u(\cdot)$ is the solution of the differential equation

$$\dot{x}(t) = h(x(t), u(t)) \,,\ t \in [0,1] \,, \tag{2.3a}$$

$$x(0) = x_0 \,, \tag{2.3b}$$

where $h : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ and $x_0 \in \mathbb{R}^n$ is given.

We assume that the problem $\mathbf{P}$ has a solution. In addition, we will require the following hypotheses which ensure (see [Bak.1]) that (a) the solutions $x^{(\cdot)}$ exist and are locally Lipschitz continuously differentiable, (b) the functions $f^j(\cdot)$ are locally Lipschitz continuously differentiable, and (c) the functions $\phi^j(\cdot, \cdot)$ are continuously differentiable, locally Lipschitz in u:

Assumption 2.1.

(i) The set $U \subset \mathbb{R}^m$ is compact and convex.

(ii) The functions g^j, $g^k : \mathbb{R}^n \to \mathbb{R}$ are locally Lipschitz continuously differentiable.

(iii) The function $h : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ is locally Lipschitz continuously differentiable.

(iv) There is a constant $M < \infty$ such that $\| h(x, u) \| \le M(1 + \| x \|)$ for all $(x, u) \in X \times U$, where X is a sufficiently large, but bounded subset of $\mathbb{R}^n$. □

For the purpose of convergence analysis, it is useful to introduce the relaxed controls closure of the set G.

We recall that a *Radon probability measure* μ on the Borel sets of U (in (2.1b)) is a *positive* measure such that $\mu(U) = 1$. The set of Radon probability measures will be denoted by $rpm(U)$. A *relaxed* control, σ, is a measurable function $\sigma : [0, 1] \to rpm(U)$. We define relaxed controls closure of the set G (in (2.1b) by

$$\overline{G} \triangleq \{ \sigma : [0, 1] \to rpm(U) \mid \sigma \text{ is measurable} \}. \tag{2.4a}$$

We use the *weak**-topology on $L^1([0,1], C(U))^*$ to topologize $\overline{G}$. Consequently, $\{ \sigma_i \} \subset \overline{G}$ converges to $\sigma \in \overline{G}$ if and only if

$$\lim_{\sigma_i \to \sigma} \int_0^1 \int_U \phi(t, u) \sigma_i(t)(du) dt = \int_0^1 \int_U \phi(t, u) \sigma(t)(du) dt, \quad \forall \phi \in L^1([0, 1], C(U)), \tag{2.4b}$$

where $L^1([0, 1], C(U))$ denotes the space of absolutely integrable functions which map the interval $[0, 1]$ into $C(U)$, the space of real value continuous functions defined on U.

In this topology, $\overline{G}$ is compact. We recall that there is an injection of the ordinary controls into the relaxed controls: with each ordinary control, $u \in G$, we associate a relaxed control $\sigma \in \overline{G}$ such that $\sigma(t)(S) = \delta_{u(t)}(S)$ for all measurable sets $S \subset U$, where $\delta_u(S) = 1$ if $u \in S$ and $\delta_u(S) = 0$ otherwise.

Relaxed controls give rise to relaxed dynamics:

$$\dot{x}(t) = \int_U h(x(t), u) d\sigma(t)(u), \quad t \in [0, 1], \tag{2.4b}$$

$$x(0) = x_0, \tag{2.4c}$$

whose solutions we will denote by $\overline{x}^{\sigma}(t)$. We extend this notation also to the functions in (2.1), thus $\overline{f}^j(\sigma) = g^j(\overline{x}^{\sigma}(1))$, and $\overline{\phi}^k(\sigma, t) = g^k(\overline{x}^{\sigma}(t))$.

Our exposition will be simpler if we assume a single form for both the state space and end-point constraints. This requires that the functions $f^j(\cdot)$ be replaced by functions of the form

$\max_{t \in [0,1]} \phi^j(u, t)$ with $\phi^j(u, t) \triangleq g^j(x^u(1))$ for all $t \in [0, 1]$. Then, letting $q = q_1 + q_2$, and replacing the indices k in (2.2b) by $j = q_1 + k$, problem (2.1a) becomes:

$$\mathbf{P}: \quad \min\{ f^0(u) \mid \max_{t \in [0,1]} \phi^j(u, t) \le 0,\ j = 1, 2, \dots, q,\ u \in G \}, \tag{2.5a}$$

or in the even more compact form,

$$\mathbf{P}: \quad \min\{ f^0(u) \mid \psi^j(u) \le 0,\ j = 1, 2, \dots, q,\ u \in G \}, \tag{2.5b}$$

where $\psi^j(u) \triangleq \max_{t \in [0,1]} \phi^j(u, t)$.

We can also state the relaxed control verstion of (2.5b):

$$\overline{\mathbf{P}}: \quad \min\{ \bar{f}^0(\sigma) \mid \bar{\psi}^j(\sigma) \le 0,\ j = 1, 2, \dots, q,\ \sigma \in \bar{G} \}. \tag{2.5c}$$

Since we have assumed that **P** has a solution, the *minimum value* for **P** and $\overline{\mathbf{P}}$ are the same. However, it is conceivable that $\overline{\mathbf{P}}$ has solutions which do not have conterparts in G.

For any $u \in L_{\infty,2}$, let $\psi(u) \triangleq \max_{j \in \underline{q}} \psi^j(u)$ and let $\psi(u)_+ \triangleq \max\{0, \psi(u)\}$. The phase I - phase II methods of centers that we will present in this paper are based on the use of the *unifying* function $F : L_{\infty,2} \times L_{\infty,2} \to \mathbb{R}$, defined by[1]

$$F(u \mid u') = \max_{j \in \underline{q}} \{ f^0(u) - f^0(u') - 2\psi(u')_+,\ \psi^j(u) - \psi(u')_+ \}. \tag{2.6}$$

The following result is obvious.

Proposition 2.1. (a) For all $u \in L_{\infty,2}$, $F(u \mid u) = 0$. (b) Suppose that $\hat{u} \in G$ is a local optimizer for problem (2.5b). Then $F(u \mid \hat{u}) \ge 0$ for all $u \in G$ near $\hat{u}$, i.e., $\hat{u}$ is a local minimizer for the problem $\min_{u \in G} F(u \mid \hat{u})$. □

Phase I - phase II methods of centers are based on the following geometric notion: given a point $u_i \in G$, its successor u_{i+1} is chosen to be a "center" of the set

$$V(u_i) \triangleq \{ u \in G \mid F(u \mid u_i) \le 0 \}. \tag{2.7}$$

For our methods of centers to work, we must introduce the following, commonly used

[1] The scale factor 2 in the term $2\psi(u')_+$ in (2.6) can be replaced by any other scale factor $\gamma > 1$.

"constraint qualification" type of hypothesis, which is easily interpreted in terms of Proposition 2.1.

Assumption 2.2. For every $\sigma \in \bar{G}$ which is not a solution of $\bar{\mathbf{P}}$, there exists a $u \in G$ such that $F(u \mid \sigma) < 0$. □

The methods differ by the manner in which they define a "center". The simplest, but not practical definition of the "center" u_{i+1} is

$$u_{i+1} = \operatorname*{argmin}_{u \in G} F(u \mid u_i) . \tag{2.8}$$

The solution of (2.8) for u_{i+1} is hardly easier than solving the original problem (2.5b). Hence we will now introduce a much more tractable definition of a "center" based on the parametrized *barrier function* $p_\alpha : L_{\infty,2} \times L_{\infty,2} \to \mathbb{R}$, for the above sets $V(u')$, defined by, for $\alpha > 0$,

$$p_\alpha(u \mid u') \triangleq \frac{1}{\alpha + 2\psi(u')_+ + f^0(u') - f^0(u)} + \sum_{j=1}^{q} \int_0^1 \frac{1}{\alpha + \psi(u')_+ - \phi^j(u, t)} \, dt . \tag{2.9}$$

Because for $\alpha > 0$, $p_\alpha(u' \mid u') < \infty$, the parameter α makes it possible to use the point u' for initializing a descent method in solving $\min_{u \in V(u')} p_\alpha(u \mid u')$.

We begin by establishing that as $\alpha \to 0$, $p_\alpha(\cdot \mid u')$ becomes a barrier function for the set $V(u')$.

Lemma 2.1: There exists a constant $L > 0$, such that for all $u' \in G$, $u \in V(u')$, and $\alpha \geq 0$,

$$p_\alpha(u \mid u') \geq \frac{1}{\alpha + 2\psi(u')_+ + f^0(u') - f^0(u)} + \frac{1}{L} \log \left[1 + \frac{L/2}{(\alpha + \psi(u')_+ - \psi(u))} \right]^2 . \tag{2.10}$$

Proof: It follows from our assumptions that there exists a constant, $L < \infty$, such that each $\phi^j(u, \cdot)$ is uniformly Lipschitz in t on $[0, 1]$ with the same Lipshitz constant L, for all $u \in G$. Without loss of generality, we may assume that $L \geq \psi(u')_+ - \psi(u)$, for all u, $u' \in G$. Now, given $u' \in G$ and $u \in V(u')$, Let $k \in \underline{q}$ and $\hat{t} \in [0, 1]$ be such that $\phi^k(u, \hat{t}) = \psi(u)$, and let $t \in [0, 1]$ be arbitrary. Then we have that

$$\phi^k(u, t) \geq \phi^k(u, \hat{t}) - L \,|t - \hat{t}\,| = \psi(u) - L\,|t - \hat{t}\,| . \tag{2.11}$$

Consequently,

$$p_\alpha(u \mid u') \geq \frac{1}{\alpha + 2\psi(u')_+ + f^0(u') - f^0(u)} + \int_0^1 \frac{1}{(\alpha + \psi(u')_+ - \phi^k(u, t))} \, dt$$

$$\geq \frac{1}{\alpha + 2\psi(u')_+ + f^0(u') - f^0(u)} + \int_0^1 \frac{1}{(\alpha + \psi(u')_+ - \psi(u) + L\,|t - \hat{t}\,|)}\,dt$$

$$\geq \frac{1}{\alpha + 2\psi(u')_+ + f^0(u') - f^0(u)}$$

$$+ \frac{1}{L}\log\left[\frac{(\alpha + \psi(u')_+ - \psi(u) + L\hat{t}\,)(\alpha + \psi(u')_+ - \psi(u) + L(1 - \hat{t}\,))}{(\alpha + \psi(u')_+ - \psi(u))^2}\right]$$

$$\geq \frac{1}{\alpha + 2\psi(u')_+ + f^0(u') - f^0(u)} + \frac{1}{L}\log\left[1 + \frac{L/2}{(\alpha + \psi(u')_+ - \psi(u))}\right]^2, \qquad (2.12)$$

where the last line in (2.12) is obtained by minimizing the preceeding line with respect to $\hat{t}$. □

It follows by inspection of (2.10) that when $u \to u'$, with $u \in V(u')$, then $p_0(u \mid u') \to \infty$, i.e., that $p_0(\cdot\,, u')$ is indeed a barrier function for $V(u')$.

Now consider the following conceptual algorithm for solving the problem **P** (2.5a).

Algorithm 2.1.

Data: $u_0 \in G$ and a sequence $\{\alpha_k\}_{k=0}^{\infty}$ such that $\alpha_k > 0$ for all $k \in \mathbb{N}$ and $\alpha_k \downarrow 0$ as $k \to \infty$.

Step 0: Set $i = 0$ and $k = 0$.

Step 1: Compute

$$u_{i+1} \in A(u_i) \triangleq \underset{u \in V(u_i)}{\operatorname{argmin}}\, p_{\alpha_k}(u \mid u_i)\,. \qquad (2.13)$$

Step 2: If $F(u_{i+1} \mid u_i) = 0$, replace k by $k + 1$ and go to Step 1.

Step 3: Replace i by $i + 1$, k by $k + 1$, and go to Step 1. □

Note that (2.13) defines u_{i+1} as a solution of the simple optimal control problem:

$$\min_{u \in G}\left\{\frac{1}{\alpha_{k_i} + 2\psi(u_i)_+ + g^0(x^{u_i}(1)) - g^0(x^u(1))} + \sum_{j=1}^{q}\int_0^1 \frac{1}{\alpha_{k_i} + \psi(u_i)_+ - g^j(x^u(t))}\,dt\right\}, \qquad (2.14)$$

where, as before $x^u(t)$ is the solution of (2.4a,b). This problem has only control constraints; its cost is of the end point - plus - integral form. Barring possible ill-conditioning, such problems are easily solved by algorithms such as the Goldstein-Levitin-Polyak gradient projection method (see [Ber.1]), or the algorithm described in [Bak.1].

Theorem 2.1. (i) Suppose that Algorithm 2.1 jams up in the loop beween Step 1 and Step 2 at the control u_{i_0}. Then u_{i_0} is a solution of **P**.

(ii) Suppose that $\{ u_i \}_{i=0}^{\infty}$ is a sequence of controls constructed by Algorithm 2.1. If this sequence has an accumulation point $\hat{u} \in G$, then $\hat{u}$ is a solution of **P**.

Proof. (i) For the sake of contradiction, suppose that u_{i_0} is not a solution of **P**. Since the algorithm is cycling the loop beween Step 1 and Step 2, $k \to \infty$. Let $\hat{u}_k \triangleq \text{argmin}_{u \in V(u_{i_0})} p_{\alpha_k}(u \mid u_{i_0})$. Then, since $\alpha_k \to 0$ as $k \to \infty$ and $F(\hat{u}_k \mid u_{i_0}) = 0$ for all k, we conclude that $p_{\alpha_k}(\hat{u}_k \mid u_{i_0}) \to \infty$ as $k \to \infty$. But by Assumption 2.2, since u_{i_0} is not a solution of **P**, there exists a $\hat{u} \in V(u_{i_0})$ such that $F(\hat{u} \mid u_{i_0}) < 0$, which implies that $p_0(\hat{u} \mid u_{i_0}) < \infty$. Since $p_\alpha(\hat{u} \mid u_{i_0})$ is continuous in α, there exists a k_0 such that $p_{\alpha_k}(\hat{u} \mid u_{i_0}) \le 2p_0(\hat{u} \mid u_{i_0})$ for all $k \ge k_0$. But because $p_{\alpha_k}(\hat{u}_k \mid u_{i_0}) \to \infty$, there exists a k_1 such that $p_{\alpha_k}(\hat{u}_k \mid u_{i_0}) > p_{\alpha_k}(\hat{u} \mid u_{i_0})$ for all $k \ge k_1 \ge k_0$, which contradicts the fact that $\hat{u}_k$ is a minimizer.

ii) Suppose that there exists an infinite subset $K \subset \mathbf{N}$ and a $\hat{u} \in G$ such that the subsequence $\{ u_i \}_{i \in K}$ converges to $\hat{u}$; we write this as $u_i \xrightarrow{K} \hat{u} \in G$ as $i \to \infty$. For the sake of contradiction, suppose that $\hat{u}$ is not a solution of **P**.

Case I: There exists an i_0 such that $\psi(u_{i_0}) \le 0$, so that $\psi(u_{i_0})_+ = 0$. Since $F(u_{i_0+1} \mid u_{i_0}) < 0$, we have that

$$f^0(u_{i_0+1}) - f^0(u_{i_0}) < 0\,, \tag{2.15a}$$

$$\psi(u_{i_0+1}) < 0\,. \tag{2.15b}$$

It now follows by induction that for all $i \ge i_0$,

$$f^0(u_{i+1}) - f^0(u_i) < 0\,, \tag{2.16a}$$

$$\psi(u_i) < 0 \,. \tag{2.16b}$$

Hence, since the sequence $\{ f^0(u_i) \}_{i=i_0}^{\infty}$ is monotone decreasing, and since $f^0(u_i) \overset{K}{\rightarrow} f^0(\hat{u})$ as $i \rightarrow \infty$, by continuity, it follows that $f^0(u_i) \rightarrow f^0(\hat{u})$ as $i \rightarrow \infty$ and therefore, since

$$p_{\alpha_{k_i}}(u_{i+1} \mid u_i) > \frac{1}{\alpha_{k_i} + f^0(u_i) - f^0(u_{i+1})} \,, \tag{2.17a}$$

and $\alpha_{k_i} \rightarrow 0$ as $i \rightarrow \infty$, we have that

$$p_{\alpha_{k_i}}(u_{i+1} \mid u_i) \rightarrow \infty \text{ as } i \rightarrow \infty \,. \tag{2.17b}$$

However, since $\hat{u}$ is not a solution to **P**, by Assumption 2.2, there exists a $u^* \in V(\hat{u})$ such that $F(u^* \mid \hat{u}) < 0$ and therefore that $p_0(u^* \mid \hat{u}) < \infty$. Since, by construction, $V(\hat{u}) \subset V(u_i)$ for all $i \geq i_0$, it follows that

$$p_{\alpha_{k_i}}(u_{i+1} \mid u_i) \leq p_{\alpha_{k_i}}(u^* \mid u_i) \,, \;\; \forall i \geq i_0 \,. \tag{2.18}$$

Furthermore, by continuity of $p_\alpha(u^* \mid u)$ in (α, u), $p_{\alpha_{k_i}}(u^* \mid u_i) \overset{K}{\rightarrow} p_0(u^* \mid \hat{u}) < \infty$, and hence there exists an $i_1 \geq i_0$ such that for all $i \in K, i \geq i_1$,

$$p_{\alpha_{k_i}}(u_{i+1} \mid u_i) \leq 2 p_0(u^* \mid \hat{u}) \,, \tag{2.19}$$

which is a contradiction of (2.17b).

Case II: Suppose that $\psi(u_i) > 0$ for all $i \in \mathbb{N}$. Then, since $p_{\alpha_{k_i}}(u_{i+1} \mid u_i) < \infty$ for all $i \in \mathbb{N}$, it follows from (2.10) that $\psi(u_{i+1}) < \psi(u_i)$ for all $i \in \mathbb{N}$, and hence, since by continuity, $\psi(u_i) \overset{K}{\rightarrow} \psi(\hat{u})$ as $i \rightarrow \infty$, that $\psi(u_i) \rightarrow \psi(\hat{u})$ as $i \rightarrow \infty$. Hence, since by (2.10)

$$p_{\alpha_{k_i}}(u_{i+1} \mid u_i) > \frac{1}{L} \log \left[1 + \frac{L/2}{(\alpha_{k_i} + \psi(u_i)_+ - \psi(u_{i+1}))} \right]^2 \,, \tag{2.20a}$$

and $\alpha_{k_i} \rightarrow 0$ as $i \rightarrow \infty$, we have that

$$p_{\alpha_{k_i}}(u_{i+1} \mid u_i) \rightarrow \infty \,, \;\; \text{as } i \rightarrow \infty \,. \tag{2.20b}$$

However, by Assumption 2.2, since $\hat{u}$ is not a solution of **P**, there exists a $u^* \in V(\hat{u})$ such that $F(u^* \mid \hat{u}) < 0$ and hence $p_0(u^* \mid \hat{u}) < \infty$. By continuity of $F(u^* \mid \cdot)$, it now follows that there exists an i_2 such that $F(u^* \mid u_i) < 0$ for all $i \in K, i \geq i_2$, and hence we see that $u^* \in V(u_i)$ for all $i \in K, i \geq i_2$. Hence, again by continuity, there exists an $i_3 \geq i_2$ such that

$$p_{\alpha_*}(u_{i+1} \mid u_i) \le p_{\alpha_*}(u^* \mid u_i) \le 2p_0(u^* \mid \hat{u}) < \infty , \tag{2.21}$$

which contradicts (2.20b). $\square$

Since the set G is not compact, it is entirely possible that a sequence $\{ u_i \}_{i=0}^{\infty}$, constructed by Algorithm 2.1, has no accumulation points in G. In that case, Theorem 2.1 is vacuous. However, the sequence $\{ u_i \}_{i=0}^{\infty}$ must have accumulation points, in the sense of control measures (i.s.c.m.) in the compact set $\overline{G}$. The following result follows directly from the arguments used to prove Theorem 2.1.

Corollary 2.1. Suppose that $\{ u_i \}_{i=0}^{\infty}$ is a sequence of controls constructed by Algorithm 2.1. If this sequence has an accumulation point $\hat{\sigma} \in \overline{G}$, then $\hat{\sigma}$ is a solution of $\overline{\mathbf{P}}$. $\square$

3. An Implementable Phase I - Phase II Method of Centers

The main objection to Algorithm 2.1 is that the update operation in (2.13) is not implementable. We will now develop an implementable algorithm which replaces (2.13) by an approximate stationarity condition and which incorporates a feature which enables us to use the point u_i as a starting point in computing u_{i+1} by an algorithm such as Algorithm 5.14 in [Bak.1].

First, referring to [Bak.1] we see that the Frechet differentials (of the functions $f^0(\cdot)$, $\phi^k(\cdot\, , t)$, in (2.5a)) $df^0(u\, ; u' - u)$, $d\phi^k(u\, , t\, ; u' - u)$ exist and can be expressed in terms of scalar products with gradients $\nabla f^0(u)$ and $\nabla_u \phi^j(u\, , t)$ which are in $L_{\infty, 2}$, i.e., $df^0(u\, ; u' - u) = \langle \nabla f^0(u)\, , u' - u \rangle_2$, and $d\phi^j(u\, , t\, ; u' - u) = \langle \nabla_u \phi^j(u\, , t)\, , u' - u \rangle_2$. Neither our proofs nor our algorithm require formulae for these gradients.

Next we define an *optimality function* $\theta : G \to \mathbb{R}$ for the problem **P** (2.5a), which is a first order convex approximation to the function $F(u' \mid u)$, by

$$\theta(u) \triangleq \min_{u' \in G} \max \Big\{ \tfrac{1}{2}\|u' - u\|_2^2 + \{ -2\psi(u)_+ + df^0(u\, ; u' - u)\, ,$$

$$\phi^j(u\, , t_j) - \psi(u)_+ + \langle \nabla_u \phi^j(u\, , t_j)\, , u' - u \rangle_2\, ,\ t_j \in [0\, , 1]\, ,\ j \in \underline{q}\ \} \Big\} . \tag{3.1}$$

At one point in our convergence proof we will need to bring in the relaxed controls topology. Hence we need the relaxed controls extension of $\theta(u)$. For this purpose, it is useful to recall that alternative formulae for $\langle \nabla f^0(u)\, , u' - u \rangle_2$ and $\langle \nabla \phi^j(u\, , t)\, , u' - u \rangle_2$ are given by

$$\langle \nabla f^0(u), u' - u \rangle_2 = \langle \nabla g^0(x^u(1)), \delta x^{(u'-u)}(1) \rangle , \tag{3.2}$$

$$\langle \nabla \phi^j(u, t), u' - u \rangle_2 = \langle \nabla g^0(x^u(t)), \delta x^{u',u}(t) \rangle , \tag{3.3}$$

where $\delta x^{u',u}(t)$ is the solution of the first variational equation:

$$\dot{\delta x}(t) = \left[\frac{\partial h(x^u(t), u(t))}{\partial x} \right] \delta x(t) + \left[\frac{\partial h(x^u(t), u(t))}{\partial u} \right] [u'(t) - u(t)] , \quad \delta x(0) = 0 . \tag{3.4}$$

Refering to [Wil.1, Bak.1], given a relaxed control $\sigma \in \overline{G}$, with corresponding solution $\overline{x}^{\sigma}(\cdot)$ of (2.4b), and any continuous function $s : U \to C^m[0, 1]$, we define $\delta\overline{x}^{\sigma,s}(t)$ to be the solution of

$$\dot{\delta x}(t) = \int_U \left[\frac{\partial h(\overline{x}^{\sigma}(t), u)}{\partial x} \right] \sigma(t)(du)\, \delta x(t) + \int_U \left[\frac{\partial h(x^{\sigma}(t), u)}{\partial u} \right] w(u)(t)\sigma(t)(du) , \tag{3.5a}$$

$$\delta x(0) = 0 . \tag{3.5b}$$

Next, we say that a *search direction function* $s : U \to C^m[0, 1]$ is *admissible* if for all $u(t) \in U$, $u(t) + s(u)(t) \in U$ for almost all $t \in [0, 1]$. We define S to be the set of all admissible search direction functions. With these definitions, the relaxed controls extension $\overline{\theta} : \overline{G} \to \mathbb{R}$, of $\theta(\cdot)$, is defined by

$$\overline{\theta}(\sigma) \triangleq \min_{s \in S} \max \Big\{ ½\|s\|_2^2 + \{ -2\overline{\psi}(\sigma)_+ + \langle \nabla g^0(\overline{x}^{\sigma}(1)), \delta\overline{x}^{\sigma,s}(1) \rangle ,$$

$$\overline{\phi}^j(\sigma, t_j) - \overline{\psi}(\sigma)_+ + \langle \nabla g^j(\overline{x}^{\sigma}(t_j)), \delta\overline{x}^{\sigma,s}(t_j) \rangle , \ t_j \in [0, 1], \ j \in \underline{q} \} \Big\} . \tag{3.6}$$

The following result can be found in [Bak.1]:

Theorem 3.1.

(a) The optimality functions $\theta(\cdot)$ amd $\overline{\theta}(\cdot)$ are well defined and continuous.

(b) If $\sigma \in \overline{G}$ corresponds to the ordinary control $u \in G$, then $\overline{\theta}(\sigma) = \theta(u)$.

(c) If $\hat{u} \in G$ is an optimal solution to the problem **P**, then $\theta(\hat{u}) = 0$.

(d) If $\hat{\sigma} \in \overline{G}$ is an optimal solution to the problem $\overline{\mathbf{P}}$, then $\overline{\theta}(\hat{\sigma}) = 0$. □

In our proofs we will find it convenient to use an alternative formula for $\theta(u)$. Let the set of Radon probability measures on the interval $[0, 1]$ be denoted by $rpm([0, 1])$, let V denote the set of mesurable functions $v : [0, 1] \to rpm([0, 1])$, and let

$$W \triangleq \{ w = (w^0, w^1, \dots, w^q) \in \mathbf{R}^{m+1} \mid \textstyle\sum_{i=0}^{q} w^i = 1 \,,\; w^i \geq 0 \,, i = 0, 1, \dots, q \} . \tag{3.7a}$$

Finally, with $rfm([0,1])$ denoting the space of Radon finite measures, let Σ be the set of measurable functions $\mu(\cdot) : [0,1] \to [rfm([0,1])]^{q+1}$ defined by

$$\Sigma \triangleq \{ \mu \in [rfm([0,1])]^{q+1} \mid \mu^j = w^j v^j \,,\; j = 0, 1, \dots, q \,,\; w \in W \,,\; v \in V \} . \tag{3.7b}$$

Then it is obvious that

$$\theta(u) \triangleq \min_{u' \in G} \max_{\mu \in \Sigma} \Bigg\{ \tfrac{1}{2} \| u' - u \|_2^2 + \{ \int_{[0,1]} [-2\psi(u)_+ + df^0(u \,; u' - u)] \, \mu^0(t)(dt)$$

$$+ \sum_{j=1}^{q} \int_{[0,1]} [\phi^j(u, t) - \psi(u)_+] \, \mu^j(t)(dt)$$

$$+ \sum_{j=1}^{q} \langle \int_{[0,1]} \nabla_u \phi^j(u, t) \, \mu^j(t)(dt) \,, u' - u \rangle_2 \} \Bigg\} . \tag{3.7c}$$

We will require the following assumption which is usually required for methods of centers and feasible directions:

Assumption 3.1.

(a) The $\bar{G}$-closure of the set $\{ u \in G \mid \psi(u) \leq 0 \}$ is equal to the $\bar{G}$-closure of its interior.

(b) For all $u \in G$ such that $\psi(u) > 0$, $\theta(u) < 0$. □

Next, referring to the definition (2.9), we conclude that

$$\nabla_u p_\alpha(u \mid u') = \frac{\nabla f^0(u)}{[\alpha + 2\psi(u')_+ + f^0(u') - f^0(u)]^2} + \sum_{j=1}^{q} \int_0^1 \frac{\nabla_u \phi^j(u, t)}{[\alpha + \psi(u')_+ - \phi^j(u, t)]^2} \, dt \,. \tag{3.8a}$$

To evaluate $\nabla_u p_\alpha(u \mid u')$ one does not use the cumbersome formula (3.8a), rather, one uses the following computationally efficient formula:

$$\nabla_u p_\alpha(u \mid u')(t) = \left[\frac{\partial h(x^u(t), u(t))}{\partial u} \right]^T \lambda^{u'u}(t) \,, \tag{3.8b}$$

where $\lambda^{u'u}(t)$ is the solution of the adjoint system:

$$\dot{\lambda}(t) = -\left[\frac{\partial h(x^u(t), u(t))}{\partial x} \right]^T \lambda(t) - \sum_{j=1}^{q} \frac{\nabla_x g^j(x^u(t))}{[\alpha + \psi(u')_+ - g^j(x^u(t))]^2}, \quad t \in [0, 1], \tag{3.8c}$$

$$\lambda(1) = \frac{1}{[\alpha + 2\psi(u')_+ + g^0(x^{u'}(1)) - g^0(x^u(1))]^2} \nabla g^0(x^u(1)). \tag{3.8d}$$

Algorithm 2.1 now gives rise to the following implementable algorithm:

Algorithm 3.1.

Data: $u_0 \in G$, $\varepsilon > 0$, and $\{ \alpha_k \}_{k=0}^{\infty}$ such that $\alpha_k > 0$ for all $k \in \mathbb{N}$ and $\alpha_k \downarrow 0$ as $k \to \infty$.

Step 0: Set $i = 0$ and $k = 0$.

Step 1: Use any descent algorithm to generate a $u_{i+1} \in V(u_i)$ such that

$$0 \geq \min_{u \in G} \langle \nabla_u p_{\alpha_k}(u_{i+1} \mid u_i), u - u_{i+1} \rangle_2 + \tfrac{1}{2} \| u - u_{i+1} \|^2 \geq -\varepsilon. \tag{3.9}$$

Step 2: If $F(u_{i+1} \mid u_i) = 0$, replace k by $k+1$ and go to Step 1.

Step 3: Replace i by $i + 1$, k by $k+1$, and go to Step 1. □

The proof of convergence of Algorithm 3.1 depends on the following two lemmas:

Lemma 3.1. Suppose that Assumption 2.1 holds, that $\gamma > 0$, that $u' \in G$, and that $u \in V(u')$. For $j \in \underline{q}$, let $T_\gamma^j(u') \subset [0, 1]$ be defined by

$$T_\gamma^j(u) = \{ t \in [0, 1] \mid \phi^j(u, t) \geq \psi(u) - \gamma \}. \tag{3.10a}$$

Then for each $t \notin T_\gamma^j(u)$, with $t \in [0, 1]$ and $j \in \underline{q}$,

$$\frac{1}{(\psi(u')_+ - \phi^j(u, t))} \leq \frac{1}{(\psi(u') - \phi^j(u, t))} \leq \frac{1}{\gamma}. \tag{3.10b}$$

Proof. Because $t \notin T_\gamma^j(u)$, $\phi^j(u, t) < \psi(u) - \gamma$. Hence,

$$\psi(u') - \phi^j(u, t) > \psi(u') - \psi(u) + \gamma > \gamma, \tag{3.11}$$

and the desired inequality follows. □

Lemma 3.2. Suppose that Assumption 2.1 holds. Then for all $\alpha > 0$, there exist a constant $\delta_\alpha > 0$ such that for all $u' \in G$, for all $u \in V(u')$,

$$(\alpha + \psi(u') - \psi(u)) \sum_{j \in \underline{q}} \int_{[0, 1]} \frac{1}{(\alpha + \psi(u') - \phi^j(u, t))^2} dt \geq \delta_\alpha > 0. \tag{3.12}$$

Proof. Since G is bounded, it follows from Assumption 2.1 that there exists a Lipschitz constant, $L < \infty$, such that each $\phi^j(u, \cdot)$ is uniformly Lipschitz in t on $[0, 1]$, for all $u \in G$. Without loss of generality, we may assume that $L \geq \psi(u') - \psi(u)$, for all $u, u' \in G$. Let $j \in \underline{q}$ be such that $T_0^j(u)$ is nonempty, let $t_u \in T_0^j(u)$ be given and let $t \in [0, 1]$. Then we have that,

$$\phi^j(u, t) \geq \phi^j(u, t_u) - L \mid t - t_u \mid = \psi(u) - L \mid t - t_u \mid . \tag{3.13a}$$

Now suppose that $0 < \gamma \leq L + \alpha$. Then $\{ t \in [0, 1] \,/\, \mid t - t_u \mid \leq \gamma/(L + \alpha) \} \subset T_\gamma^j(u)$, and hence $m(T_\gamma^j(u)) \geq \gamma/(L + \alpha)$, where $m(\cdot)$ denotes the Lebesgue measure on $\mathbb{R}$. Hence we conclude that

$$\sum_{j \in \underline{q}} \int_{[0,1]} \frac{\alpha + \psi(u') - \psi(u)}{(\alpha + \psi(u') - \phi^j(u, t))^2} dt \geq \int_{T_\gamma^j(u)} \frac{\alpha + \psi(u') - \psi(u)}{(\alpha + \psi(u') - \phi^j(u, t))^2} dt$$

$$\geq \frac{\gamma}{(L + \alpha)} \frac{\alpha + \psi(u') - \psi(u)}{(\alpha + \psi(u') - \psi(u) + \gamma)^2} . \tag{3.13b}$$

Setting $\gamma = \alpha + \psi(u') - \psi(u)$, and $\delta = 1/4(L + \alpha)$, we obtain the desired result. □

Theorem 3.2. (i) Suppose that Algorithm 3.1 jams up in the loop beween Step 1 and Step 2 at the control u_{i_0}. Then u_{i_0} is a solution of **P**.

(ii) Suppose that $\{ u_i \}_{i=0}^{\infty}$ is a sequence of controls constructed by Algorithm 3.1. If this sequence has an accumulation point $\hat{u} \in G$, then $\psi(\hat{u}) \leq 0$ and $\theta(\hat{u}) = 0$.

Proof. (i) The proof of this part is essentially the same as for (i) of Theorem 2.1 and hence is omitted.

(ii) Since u_{i+1} is constructed from u_i by a descent method, it follows from $F(u_{i+1} \mid u_i) < 0$ that if $\psi(u_i) \leq 0$, then $\psi(u_{i+1}) \leq 0$ also. Hence our proof breaks down into the examination of two cases.

Case I: Suppose $u_i \xrightarrow{K} \hat{u} \in G$ as $i \to \infty$ and that there exists an i_0 such that $\psi(u_i) \leq 0$ for all $i \geq i_0$. Then, by construction, the sequence $\{ f^0(u_i) \}_{i=i_0}^{\infty}$ is monotone decreasing, and since it is bounded, it must converge to $f(\hat{u})$. The fact that $\psi(\hat{u}) \leq 0$ follows directly from the continuity of $\psi(\cdot)$.

Now, for $i \geq i_0$, $j \in \underline{q}$ and $t \in [0, 1]$, let

$$\rho_i^j(t) \triangleq \frac{[\alpha_{k_i} + f^0(u_{i-1}) - f^0(u_i)]^2}{(\alpha_{k_i} - \phi^j(u_i, t))^2} . \tag{3.14a}$$

Finally, let

$$\nu_i \triangleq 1 + \sum_{j=1}^{q} \int_0^1 \rho_i^j(t)\, dt \ , \tag{3.14b}$$

It now follows from (3.8) and (3.9) that for all $i \geq i_0$,

$$0 \geq \Theta'(u_i) \triangleq \min_{u \in G} \frac{1}{\nu_i} \left\{ df^0(u_i\ ; u - u_i) + \sum_{j=1}^{q} \int_0^1 \rho_i^j(t) d\phi^j(u_i\ , t\ ; u - u_i) + \tfrac{1}{2}\|u - u_i\|_2^2 \right\}$$

$$\geq -\varepsilon_i \triangleq -\varepsilon \frac{[\alpha_{k_i} + f^0(u_{i-1}) - f^0(u_i)]^2}{\nu_i} . \tag{3.15a}$$

Since $\nu_i \geq 1$ for all $i \geq i_0$ and since $f^0(u_i) \to f^0(\hat{u})$ and $\alpha_{k_i} \to 0$ as $i \to \infty$, it follows that $\varepsilon_i \to 0$ as $i \to \infty$. Next, it follows from (3.7c) that for all $i \geq i_0$, $\theta(u_i) \geq \Theta'(u_i)$, and hence we find that for all $i \geq i_0$,

$$0 \geq \theta(u_i) \geq \Theta'(u_i) + \frac{1}{\nu_i} \sum_{j=1}^{q} \int_0^1 \rho_i^j(t) \phi^j(u_i\ , t) dt \geq -\varepsilon_i + \frac{1}{\nu_i} \sum_{j=1}^{q} \int_0^1 \rho_i^j(t) \phi^j(u_i\ , t) dt\ . \tag{3.15b}$$

It now follows from (3.14a) and the relation, $0 \geq \dfrac{\phi^j(u_i\ , t)}{(\alpha_{k_i} - \phi^j(u_i\ , t))^2} \geq \dfrac{\phi^j(u_i\ , t)}{(-\phi^j(u_i\ , t))^2}$, and Lemma 3.1 that for any $j \in \underline{q}$, and any $\gamma > 0$,

$$0 \geq \frac{1}{\nu_i} \int_0^1 \rho_i^j(t) \phi^j(u_i\ , t) dt = \frac{1}{\nu_i} [\alpha_{k_i} + f^0(u_{i-1}) - f^0(u_i)] \int_0^1 \frac{\phi^j(u_i\ , t)}{(\alpha_{k_i} - \phi^j(u_i\ , t))^2}\, dt$$

$$\geq \frac{\alpha_{k_i} + f^0(u_{i-1}) - f^0(u_i)}{\nu_i} \left\{ \int_{t \in T_\gamma^j(u_i)} \frac{1}{\phi^j(u_i\ , t)}\, dt + \int_{t \in T_\gamma^j(u_i)^c} \frac{1}{\phi^j(u_i\ , t)}\, dt \right\}$$

$$\geq -\left[\frac{\alpha_{k_i} + f^0(u_{i-1}) - f^0(u_i)}{\nu_i} \right] \left[\gamma + \frac{1}{\gamma} \right] . \tag{3.16}$$

Since $[\alpha_{k_i} + f^0(u_{i-1}) - f^0(u_i)]/\nu_i \to 0$ as $i \to \infty$, and since by Theorem 3.1, $\theta(\cdot)$ is continuous, it now follows that $\theta(\hat{u}) = 0$.

Case II: We now suppose that $\psi(u_i) > 0$ for all $i \in \mathbb{N}$. Then we must have that the sequence $\{ \psi(u_i) \}_{i=0}^{\infty}$ is monotonically decreasing, and hence it must converge to $\psi(\hat{u})$. In this case, we define

$$\rho_i^0(t) \triangleq \frac{(\alpha_{k_i} + \psi(u_{i-1}) - \psi(u_i))}{(\alpha_{k_i} + 2\psi(u_{i-1}) + f^0(u_{i-1}) - f^0(u_i))^2}, \tag{3.17a}$$

$$\rho_i^j(t) \triangleq \frac{(\alpha_{k_i} + \psi(u_{i-1}) - \psi(u_i))}{(\alpha_{k_i} + \psi(u_{i-1}) - \phi^j(u_i, t))^2}, \quad j = 1, 2, \ldots, q. \tag{3.17b}$$

It now follows from Lemma 3.2 that

$$\nu_i \triangleq \sum_{j=0}^{q} \int_0^1 \rho_i^j(t) dt \geq \delta_{\alpha_{k}} > 0, \tag{3.18}$$

and from (3.9) that

$$0 \geq \Theta''(u_i) \triangleq \min_{u \in G} \frac{1}{\nu_i} \left\{ \int_0^1 \rho_i^0(t)\, df^0(u_i; u - u_i)\, dt + \sum_{j=1}^{q} \int_0^1 \rho_i^j(t) d\phi^j(u_i, t; u - u_i) dt + \tfrac{1}{2}\| u - u_i \|_2^2 \right\}$$

$$\geq -\varepsilon_i \triangleq -\varepsilon \frac{[\alpha_{k_i} + \psi(u_{i-1}) - \psi(u_i)]}{\nu_i}. \tag{3.19}$$

Clearly, $\varepsilon_i \to 0$ as $i \to \infty$. Next, by the same argument as for in (3.15b), we obtain that

$$0 \geq \theta(u_i) \geq \Theta''(u_i) + \frac{1}{\nu_i} \left\{ \int_0^1 -2\rho_i^0(t)\psi(u_{i-1})\, dt + \sum_{j=1}^{q} \int_0^1 \rho_i^j(t)[\phi^j(u_i, t) - \psi(u_{i-1})]\, dt \right\}$$

$$\geq -\varepsilon_i + \frac{1}{\nu_i} \left\{ \int_0^1 -2\rho_i^0(t)\psi(u_{i-1})\, dt + \sum_{j=1}^{q} \int_0^1 \rho_i^j(t)[\phi^j(u_i, t) - \psi(u_{i-1})]\, dt \right\}. \tag{3.20}$$

First, it follows from the same arguments used in part (a) that

$$\frac{1}{\nu_i} \sum_{j=1}^{q} \int_0^1 \rho_i^j(t)\, [\phi^j(u_i, t) - \psi(u_{i-1})]\, dt \to 0, \text{ as } i \to \infty. \tag{3.21}$$

Hence to complete our proof, we only need to show that $\psi(u_i) \to 0$, as $i \to \infty$.

For the sake of contradiction, suppose that $\psi(\hat{u}) > 0$. We now have two possibilities. The first is that $2\psi(u_i) + f^0(u_{i-1}) - f^0(u_i) \to 0$ as $i \to \infty$, which implies that for all sufficiently large i, $f^0(u_i) \geq f^0(u_{i-1}) + \psi(\hat{u})$, and hence that $f^0(u_i) \to \infty$ as $i \to \infty$. Since the set G is bounded, this is clearly impossible. Hence, we consider the second possibility: there exists an infinite subset $K' \subset \mathbb{N}$ and a $\delta > 0$, such that $2\psi(u_i) + f^0(u_{i-1}) - f^0(u_i) \geq \delta$ for all $i \in K'$. In turn, this implies that $\int_0^1 \rho_i^0(t)\, dt \xrightarrow{K'} 0$ as $i \to \infty$. Furthermore, we may assume that $u_i \xrightarrow{K'} \sigma^* \in \overline{G}$, i.s.c.m., as $i \to \infty$. Clearly, we must have $\overline{\psi}(\sigma^*) = \psi(\hat{u}) > 0$. In view of the above, (3.20), (3.21), and the continuity of $\overline{\theta}(\cdot)$, this implies that $\theta(\sigma^*) = 0$. However, this contradicts Assumption 3.1 and hence our proof is complete. $\square$

4. A Special Case

It is not at all uncommon for the set U in (2.1b) to have description in terms of convex inequalities, as follows:

$$U \triangleq \{ z \in \mathbb{R}^m \mid s^l(z) \leq 0 \,,\; l = 1, 2, \ldots q_3 \} \,, \tag{4.1}$$

where the $s^l : \mathbb{R}^m \to \mathbb{R}$ are all Lipschitz continuously differentiable convex functions.

If, for $j = q_1 + q_2 + 1, \ldots, q_1 + q_2 + q_3$, we define the functions $\phi^j : L_{\infty,2} \times [0, 1] \to \mathbb{R}$ by

$$\phi^j(u, t) \triangleq s^{j-q_1-q_2}(u(t)) \,, \tag{4.2}$$

we find that $\phi^j(\cdot, t)$ is differentiable on $L_{\infty,2}$, with $d_u\phi^j(u, t; u' - u) = \langle \nabla s^{j-q_1-q_2}(u(t)), u'(t) - u(t) \rangle$.

Now, let $q \triangleq q_1 + q_2 + q_3$, then (2.5a) becomes

$$\mathbf{P}': \;\; \min \{ f^0(u) \mid \max_{t \in [0,1]} \phi^j(u, t) \leq 0 \,,\; j = 1, 2, \ldots, q \,,\; u \in L_{\infty,2} \} \,, \tag{4.3a}$$

or in the even more compact form (see (2.5b)),

$$\mathbf{P}': \;\; \min \{ f^0(u) \mid \psi^j(u) \leq 0 \,,\; j = 1, 2, \ldots, q \,,\; u \in L_{\infty,2} \} \,, \tag{4.3b}$$

where $\psi^j(u) \triangleq \max_{t \in [0,1]} \phi^j(u, t)$.

For problem $\mathbf{P}'$, letting $\psi(u) = \max \{ \psi^j(u) \,,\; j = 1, \ldots, q \}$ and $\psi(u)_+ = \max \{ 0, \psi 8u) \}$, we define the optimality function $\theta' : L_{\infty,2} \to \mathbb{R}$ by

$$\theta'(u) \triangleq \min_{u' \in L_{\infty,2}} \{ \tfrac{1}{2}\| u' - u \|_2^2 + \max \{ df(u\,; u' - u) - 2\psi_+(u),$$

$$\phi^j(u\,,t) - \psi(u)_+ + d_u\phi^j(u\,,t\,; u' - u)\,,\ t \in [0\,,1]\,,\ j \in \underline{q} \} \}. \tag{4.4}$$

which is a first order convex approximation to the unifying function $F'(u' \mid u)$, defined by (2.6) with q and the functions $\phi^j(\cdot\,,\cdot)$ redefined as above.

We begin by establishing a relationship between the functions $\theta(\cdot)$ and $\theta'(\cdot)$.

Theorem 4.1: Suppose that

$$\overset{o}{U} \triangleq \{ z \in U \mid s^j(z) < 0\,,\ j = 1\,,2\,,\dots\,,q_3 \}\,, \tag{4.5}$$

where $\overset{o}{U}$ denotes the interior of U, and that $\overset{o}{U}$ is not empty. Then, for any $u \in G$, $\theta(u) = 0$ if and only if $\theta'(u) = 0$.

Proof: $\Longrightarrow$ Suppose that $\theta(\hat{u}) = 0$ but $\theta'(\hat{u}) < 0$. Since $\theta'(\hat{u}) < 0$, there exists a $\bar{u} \in L_{\infty,2}$ such that $F'(\bar{u} \mid \hat{u}) < 0$. If $\psi(\hat{u})_+ = 0$, then because of $F'(\bar{u} \mid \hat{u}) < 0$, we have that

$$f^0(\bar{u}) - f^0(\hat{u}) - 2\psi(\hat{u})_+ = f^0(\bar{u}) - f^0(\hat{u}) < 0\,, \tag{4.6a}$$

$$\phi^j(\bar{u}\,,t) - \psi(\hat{u})_+ = \phi^j(\bar{u}\,,t) < 0\,,\ \forall t \in [0\,,1]\,,\ \text{for } j = 1\,,\dots\,,q_1 + q_2\,, \tag{4.6b}$$

$$\phi^j(\bar{u}\,,t) - \psi(\hat{u})_+ = \phi^j(\bar{u}\,,t) < 0\,,\ \forall t \in [0\,,1]\,,\ \text{for } j = q_1 + q_2 + 1\,,\dots\,,q_1 + q_2 + q_3. \tag{4.6c}$$

By (4.6c), $\bar{u} \in G$. It now follows from (4.6a) and (4.6b) that $\theta(\hat{u}) < 0$, which contradicts our hypothesis.

Next, we need to prove that $\psi(\hat{u})_+ = 0$. Suppose that $\psi(\hat{u})_+ > 0$. Since $\hat{u} \in G$, $\phi^j(\hat{u}\,,t) \le 0$ for all $j = Q_3 \triangleq \{ q_1 + q_2 + 1\,,\dots\,,q_1 + q_2 + q_3 \}$ and for all $t \in [0\,,1]$. Hence $\psi(\hat{u})_+ > 0$ implies that there exist $j_0 \in Q_{1,2} \triangleq \{ 1\,,\dots\,,q_1 + q_2 \}$ and $t_0 \in [0\,,1]$ such that $\psi(\hat{u}) = \phi^{j_0}(\hat{u}\,,t_0) > 0$. However, in this case Assumption 3.1 ensures that $\theta(\hat{u}) < 0$ which is a contradiction and, therefore, $\psi(\hat{u})_+ = 0$.

$\Longleftarrow$ Suppose that $\theta'(\hat{u}) = 0$. For the sake of contradiction, suppose that $\theta(\hat{u}) < 0$. Let $\xi(u \mid \hat{u})$ be defined by

$$\xi(u \mid \hat{u}) = \tfrac{1}{2}\| u - \hat{u} \|_2^2 + \max \{ -2\psi(\hat{u})_+ + df^0(\hat{u} ; u - \hat{u}),$$

$$\phi^j(\hat{u}, t) - \psi(\hat{u})_+ + d_u\phi^j(\hat{u}, t ; u - \hat{u}),\ t \in [0, 1],\ j = 1, \ldots, q_1 + q_2 \} \tag{4.7a}$$

Since $\theta(\hat{u}) < 0$, there exists a $\bar{u} \in G$ such that,

$$\theta(\hat{u}) = \min_{u \in G} \xi(u \mid \hat{u}) = \xi(\bar{u} \mid \hat{u}) \triangleq -2\delta < 0. \tag{4.7b}$$

For $\alpha \in (0, 1)$, let $u_\alpha \triangleq \hat{u} + \alpha(\bar{u} - \hat{u})$. Then $u_\alpha \in G$ for all $\alpha \in (0, 1)$ because the set G is convex. Since $-2\psi(\hat{u})_+ \le 0$ and $\phi^j(\hat{u}, t) - \psi(\hat{u})_+ \le 0$ for all j, we have that for all $\alpha \in (0, 1)$,

$$\xi(\bar{u} \mid \hat{u}) \le \xi(u_\alpha \mid \hat{u}) \le -2\alpha\delta < 0. \tag{4.7c}$$

Let $\breve{u}$ be any control such that $\breve{u}(t) \in \overset{o}{U}$ for all $t \in [0, 1]$. Then, by (4.5), there exist $\bar{\delta}_j$ such that $\max_{t \in [0,1]} \phi^j(\breve{u}, t) = -\bar{\delta}_j < 0$ for all $j \in Q_3$. Let $u'_{\alpha,\beta} \triangleq (1-\alpha)\hat{u} + \alpha(\bar{u} + \beta(\breve{u} - \bar{u})) = u_\alpha + \alpha\beta(\breve{u} - \bar{u})$ where $\beta \in (0, 1)$. Obviously, $u'_{\alpha,\beta} \in G$. By the continuity of $\xi(\cdot \mid \hat{u})$, for any $\alpha \in (0, 1)$, there exist a $\beta_\alpha \in (0, 1)$ such that

$$\theta(\hat{u}) \le \xi(u'_{\alpha,\beta} \mid \hat{u}) \le -\alpha\delta < 0,\ \ \forall 0 < \beta \le \beta_\alpha. \tag{4.7d}$$

Since the functions $\phi^j(\cdot, t)$ are convex for all $j \in Q_3$ and for all $t \in [0, 1]$, we conclude that for all $j \in Q_3$, for all $t \in [0, 1]$, for all $\alpha \in (0, 1)$, and for all $\beta \in (0, 1)$,

$$\begin{aligned} \phi^j(u'_{\alpha,\beta}, t) &\le (1-\alpha)\phi^j(\hat{u}, t) + \alpha\phi^j(\bar{u} + \beta(\breve{u} - \bar{u}), t) \\ &\le (1-\alpha)\phi^j(\hat{u}, t) + \alpha\{ (1-\beta)\phi^j(\bar{u}, t) + \beta\phi^j(\breve{u}, t)) \} \\ &\le -\alpha\beta\bar{\delta}_j < 0. \end{aligned} \tag{4.8a}$$

The last inequality is valid because $\phi^j(\hat{u}, t) \le 0$, $\phi^j(\bar{u}, t) \le 0$, and $\phi^j(\breve{u}, t) \le -\bar{\delta}_j < 0$. Also, because all the gradients $\nabla_u\phi^j(\cdot, t)$ are Lipschitz continuous,

$$\begin{aligned} \phi^j(u'_{\alpha,\beta}, t) &= \phi^j(\hat{u}, t) + \langle \nabla_u\phi^j(\hat{u}, t), u'_{\alpha,\beta}(t) - \hat{u}(t) \rangle \\ &\quad + \int_0^1 \langle (\nabla_u\phi^j(\hat{u} + s(u'_{\alpha,\beta} - \hat{u}), t) - \nabla_u\phi^j(\hat{u}, t)), u'_{\alpha,\beta} - \hat{u} \rangle ds \end{aligned}$$

$$\geq \phi^j(\hat{u}, t) + \langle \nabla_u \phi^j(\hat{u}, t), u'_{\alpha,\beta}(t) - \hat{u}(t) \rangle$$

$$-L_j \alpha^2/2 \, (\| \bar{u}(t) - \hat{u}(t) \|^2 + \beta^2 \| \tilde{u}(t) - \bar{u}(t) \|^2) , \tag{4.8b}$$

where $L_j \triangleq \max_{t \in [0,1]} L_j(t)$ and $L_j(t)$ is a Lipschitz constant of $\nabla_u \phi^j(\cdot, t)$. Combining (4.8a) with (4.8b), we can conclude that there exist α_0, $\beta_0 \in (0, 1)$ such that for all $0 < \alpha \leq \alpha_0$, for all $0 < \beta \leq \beta_0$, and for all $t \in [0, 1]$,

$$½\| u'_{\alpha,\beta}(t) - \hat{u}(t) \|^2 + \phi^j(\hat{u}, t) + \langle \nabla \phi^j(\hat{u}, t), u'_{\alpha,\beta}(t) - \hat{u}(t) \rangle \leq -\alpha\beta\bar{\delta}_j/2 < 0 , \ \forall\, j \in Q_3 . \tag{4.8c}$$

Since (4.7d) and (4.8c) imply that, given $0 < \alpha \leq \alpha_0$, for all $0 < \beta \leq \min \{ \beta_\alpha, \beta_0 \}$,

$$\theta'(\hat{u}) \leq \max \{ -\alpha\delta , -\alpha\beta\bar{\delta}_j / 2 , \ j \in Q_3 \} < 0. \tag{4.8d}$$

we obtain a contradiction of our hypothesis, and hence our proof is complete. □

Clearly, Algorithm 3.1 is applicable to **P′** and it may be initialized with a control which is not in G. However, we must ammend Assumption 3.1, as follows. Since relaxed controls must be associated with bounded controls, we introduce an arbitrarily large compact set $U^* \subset \mathbb{R}^n$, and we define G^* by (2.1b) with U replaced by U^*. We noted the corresponding set of relaxed controls by $\bar{G}^*$.

Assumption 4.1.

(a) The $\bar{G}^*$-closure of the set $\{ u \in G^* \mid \psi(u) \leq 0 \}$ is equal to the $\bar{G}^*$-closure of its interior.

(b) For all $u \in G^*$ such that $\psi(u) > 0$, $\theta'(u) < 0$. □

At this point, the following result should be obvious:

Theorem 4.2. (i) Suppose that Algorithm 3.1 is applied to problem **P′**, and jams up in the loop beween Step 1 and Step 2 at the control u_{i_0}. Then u_{i_0} is a local solution of **P′** and hence also of **P**.

(ii) Suppose that $\{ u_i \}_{i=0}^{\infty}$ is a sequence of controls constructed by Algorithm 3.1 in solving **P′**. If this sequence has an accumulation point $\hat{u} \in L_{\infty,2}$, then $\psi(\hat{u}) \leq 0$ (so that $\hat{u} \in G$) and $\theta'(\hat{u}) = \theta(\hat{u}) = 0$. □

There is considerable programming convenience in using the formulation **P′** over **P** when possible. Our computational results, in the next section, show that the use of the formulation **P′** does not result in any penalty in terms of computing times.

5. Numerical Results

We will now present two examples which illustrate the performance of Algorithm 3.1. In our experiments, the computations in Step 1 of Algorithm 3.1 were carried out using Algorithm A in [Pol.3], which is of the Gauss-Newton type. All the computations were performed in double precision on a Sun 3/140 Workstation with a floating point accelerator. The sequence $\{\alpha_k\}_{k=0}^{\infty}$ was defined by $\alpha_{k+1} = \alpha_k/1.1$, with $\alpha_0 = 0.005$. Our experiments suggest that larger values of α_0 result in an increase in the number of iterations needed to solve a problem.

Example 5.1: Our first problem is a minimum time brachistochrone problem with a state variable inequality constraint, described in [Bry.1]. We treat this problem in fixed-time scaled form, where the scale variable T corresponds to the actual final time.

$$\mathbf{P}: \min_{\gamma \in L_{\infty,2}} \{ T^2 \mid \max_{t \in [0,1]} \phi^j(\gamma, t) \le 0 \,, \; j = 1,2 \,, \; \forall t \in [0,1] \} \,. \tag{5.1a}$$

where

$$\phi^1(\gamma, t) = y(t) - x(t)\tan\theta - h \,, \tag{5.1b}$$

$$\phi^2(\gamma, t) = \tfrac{1}{2}(x(1) - l)^2 - \xi \,, \tag{5.1c}$$

where horizontal distance, x, vertical distance, y, are defined by

$$\dot{x}(t) = T\sqrt{2gy(t)}\cos\gamma(t) \,, \tag{5.2a}$$

$$\dot{y}(t) = T\sqrt{2gy(t)}\sin\gamma(t) \,, \tag{5.2b}$$

where g is the acceleration due to gravity, and γ is the path angle to the horizontal, θ and h are constants, and ξ is a small tolerance by which we are willing to relax the requirement $x(1) = l$.

Since the system equations (5.2a) and (5.2b) cannot be integrated explicitely, one must use a numerical integration scheme, which discretizes the time interval $[0, 1]$. The discretization may be either fixed, or variable. We used 40 uniformly spaced points in conjunction with the Runge-Kutta second order method ([Ral.1]). The gradients used were those corresponding the discretized dynamics imposed by the integration scheme. The results that we obtained converge to the expected results, obtained analytically in [Bry.1, pp120]. We stopped our computations when the constraints were satisfied and the difference in the cost value between successive iterations was less than 1×10^{-5}.

We used two sets of initial conditions: (a) $(x(0), y(0), T) = (0.0, 0.3, 2.0)$, $\theta = 0.2$, $h = 0.6$, $l = 4.0$, $\xi = 0.0005$ and (b) $(x(0), y(0), T) = (0.0, 1.0, 2.0)$, $\theta = 0.2$, $h = 2.0$, $l = 10.0$ and

$\xi = 0.0005$.

Figure 1a presents a plot of the values $f^0(u_i)$ versus iteration number i. Figures 1b and 1c show state space trajectories and inputs at various iterations, respectively, obtained using the first set of initial conditions. The minimum time is determined to be 0.99971 after 12 iterations. The results of the computations using the second set of initial conditions are shown in Figures 2a, 2b, and 2c. The final minimum time, 1.63998 seconds, is obtained after 10 iterations.

Example 5.2: Our second problem is a fixed-time minimum final error problem with a state variable inequality constraint and bounds on the control:

$$\mathbf{P}: \min_{u \in G} \{ \tfrac{1}{2}\|x(1)\|^2 \mid x^2(t) - l \le 0,\ \forall t \in [0,1] \} , \tag{5.3a}$$

where the state is determined by the scaled differential equation

$$\dot{x}(t) = \begin{bmatrix} \dot{x}^1(t) \\ \dot{x}^2(t) \end{bmatrix} = T \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x(t) + T \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t) , \tag{5.3b}$$

with $T > 0$ the actual final time. The input $u(\cdot)$ is scalar valued and $u \in G \triangleq \{ u \in L_{\infty,2} \mid \|u\|_\infty \le 1.0 \}$. We used the initial state are given by $x(0) = (-3.0\,, 0.1)^T$. We set $l = 0.15$ and $T = 20$. With $s(\cdot)$ defined by $s(t) = u(t)^2 - 1.0$, the problem **P** (5.3a) becomes

$$\mathbf{P}': \min_{u \in L_{\infty,2}} \{ \tfrac{1}{2}\|x(1)\|^2 \mid \max_{t \in [0,1]} \phi^j(u,t) \le 0 ,\ j = 1,2 \} , \tag{5.4a}$$

where

$$\phi^1(u,t) = x^2(t) - l , \tag{5.4b}$$

$$\phi^2(u,t) = s(t) . \tag{5.4c}$$

We applied Algorithm 3.1 to both problems **P** and **P'**.

Figures 3a, 3b, and 3c present plots of the cost $f^0(u_i)$ versus iteration number i, as well as corresponding state space trajectories and inputs at various iterations using the formulation (2.1a). Similar results using the formulation (4.3b) are shown in Figures 4a, 4b, and 4c. As we can see, the results obtained are almost same.

It is clear from our experimental results that Algorithm 3.1 is quite effective in solving optimal control problems with continuum state and control constraints. Also, we can see that when we have a special description on the set of controls, we can take advantage of it without any penalty.

6. Conclusion

We have presented two versions of a phase I - phase II method of centers type algorithm for the solution of optimal control problems with control and state space constraints. The computational advantages of these algorithms derive from the fact that we used barrier functions for defining an approximate center to be computed at each iteration. Although, at first glance, the algorithms appear to have potential for failure due to illconditioning, preliminary computational results show that this is not so, and in fact, that the algorithms are highly effective. This observation agrees with the numerical results reported in [Pol.3] for a related algorithm which solves semi-infinite minimax problems.

7. References

[Arm.1] Armijo, L., "Minimization of functions having Lipschitz continuous first partial derivatives", *Pacific Journal of Mathematics* Vol 16, pp 1-3, 1966.

[Bak.1] Baker, T. E. and Polak, E., "On the Optimal Control of Systems Described by Evolution Equations", *University of California, Electronics Research Laboratory*, Memo UCB/ERL M89/113, 27 September 1989.

[Ber.1] Bertsekas, D. P., "On the Goldstein-Levitin-Polyak Gradient Projection Method," *IEEE Trans. AC*, Vol. AC-21, pp. 174-184, 1976.

[Bry.1] Bryson, A. E. Jr. and Ho, Yu-Chi, *Applied optimal control; optimization, estimation, and control* Waltham, Mass., Blaisdell Pub. Co. 1969.

[Fra.1] Frank, M. and Wolfe, Ph., "An Algorithm for Quadratic Programming", *Naval Res. Logistics Quart.*, No. 3, pp. 950110, 1956.

[Hua.1] Huard, P., "Programmation Mathematique Convexe," *Rev. Francaise Inf. Rech. Oper.*, Vol. 7, pp. 43-59, 1968.

[Kle.1] Klessig, R. and Polak, E., "An Adaptive Algorithm for Unconstrained Optimization with Applications to Optimal Control", *SIAM J. Control*, Vol. 11, No. 1, pp. 80-94, 1973.

[May.1] E. Polak and D. Q. Mayne, "First Order, Strong Variations Algorithms for Optimal Control Problems with Terminal Inequality Constraints", *JOTA*, Vol. 16, No. 3/4, pp. 303-325, 1975.

[May.2] D. Q. Mayne and E. Polak, "A Feasible Directions Algorithm for Optimal Control Problems with Terminal Inequality Constraints", *IEEE Transactions on Automatic Control*, Vol.

AC-22, No. 5, pp. 741-751, 1977.

[May.3] D. Q. Mayne and E. Polak "An Exact Penalty Function Algorithm for Control Problems with State and Control Constraints", *IEEE Trans. on Automatic Control,* Vol. AC-32 No. 5, pp. 380-387, 1987.

[May.4] D. Q. Mayne and E. Polak "An Exact Penalty Function Algorithm for Optimal Control Problems with Control and Terminal Equality Constraints, Part 1", *JOTA,* Vol. 32 No. 2, pp. 211-246, 1980.

[May.5] D. Q. Mayne and E. Polak "An Exact Penalty Function Algorithm for Optimal Control Problems with Control and Terminal Equality Constraints, Part 2", *JOTA,* Vol. 32 No. 3, pp. 345-363, 1980.

[Mif.1] R. Mifflin, "Rates of Convergence for a Method of Centers Algorithm", *Journal of Optimization Theory and Applications,* Vol. 18, No. 2, pp. 199-228, February 1976.

[Pir.1] Pironneau, O. and Polak, E., "On the Rate of Convergence of Certain Methods of Centers", *Mathematical Programming,* Vol. 2, No. 2, pp. 230-258, 1972.

[Pir.2] Pironneau, O. and Polak, E., "A Dual Method for Optimal Control Problems with Initial and Final Boundary Constraints", *SIAM J. Control,* Vol. 11, No. 3, pp. 534-549, 1973.

[Pol.1] Polak, E., "On the Mathematical Foundations of Nondifferentiable Optimization in Engineering Design", *SIAM Review,* pp. 21-91, March 1987.

[Pol.2] Polak, E. and He, L., "A Unified Phase I Phase II Method of Feasible Directions for Semi-infinite Optimization", *University of California, Electronics Research Laboratory,* Memo UCB/ERL M89/7, 3 February 1989. To appear in *JOTA.*

[Pol.3] E. Polak, J. Higgins and D. Q. Mayne, "A Barrier Function Method for Minimax Problems", *University of California, Electronics Research Laboratory,* Memo UCB/ERL M88/64, 20 October 1988. *Math. Programming,* in press.

[Pol.4] Polak, E., *Computational Methods in Optimization: A unified Approach,* Academic Press, 1972.

[Pol.5] E. Polak, "On the Mathematical Foundations of Nondifferentiable Optimization in Engineering Design", *SIAM Review,* Vol. 29(1), pp. 21-89, March 1987.

[Pol.6] E. Polak and D. Q. Mayne, "An Algorithm for Optimization Problems with Functional Inequality Constraints", *IEEE Transactions on Automatic Control,* Vol. AC-21, pp. 184-193, April 1976.

[Ral.1] Ralston, A., *A First course in Numerical Analysis*,McGraw-Hill, New York. 1965.

[War.1] Warga, J., *Optimal Control of Differential Equations and Functional Equations*, Academic Press, New York, 1972.

[War.2] Warga, J., "Steepest Descent with Relaxed Controls", *SICON*, Vol. 15, pp. 674-682, 1977.

[War.3] Warga, J., "Iterative Procedures for Constrained and Unilateral Optimization Problems", *SICON*, Vol. 20, pp. 360-367, 1982.

[War.4] Warga, J., "Iterative Optimization with Equality Constraints", *Math. of Oper. Res.*, Vol. 9, pp. 592-605, 1984.

[Wil.1] Williamson, L. J., and Polak, E., "Relaxed Controls and the Convergence of Optimal Control Algorithms", *SIAM J. Control*, Vol. 14, No. 4, pp. 737-757, 1976.

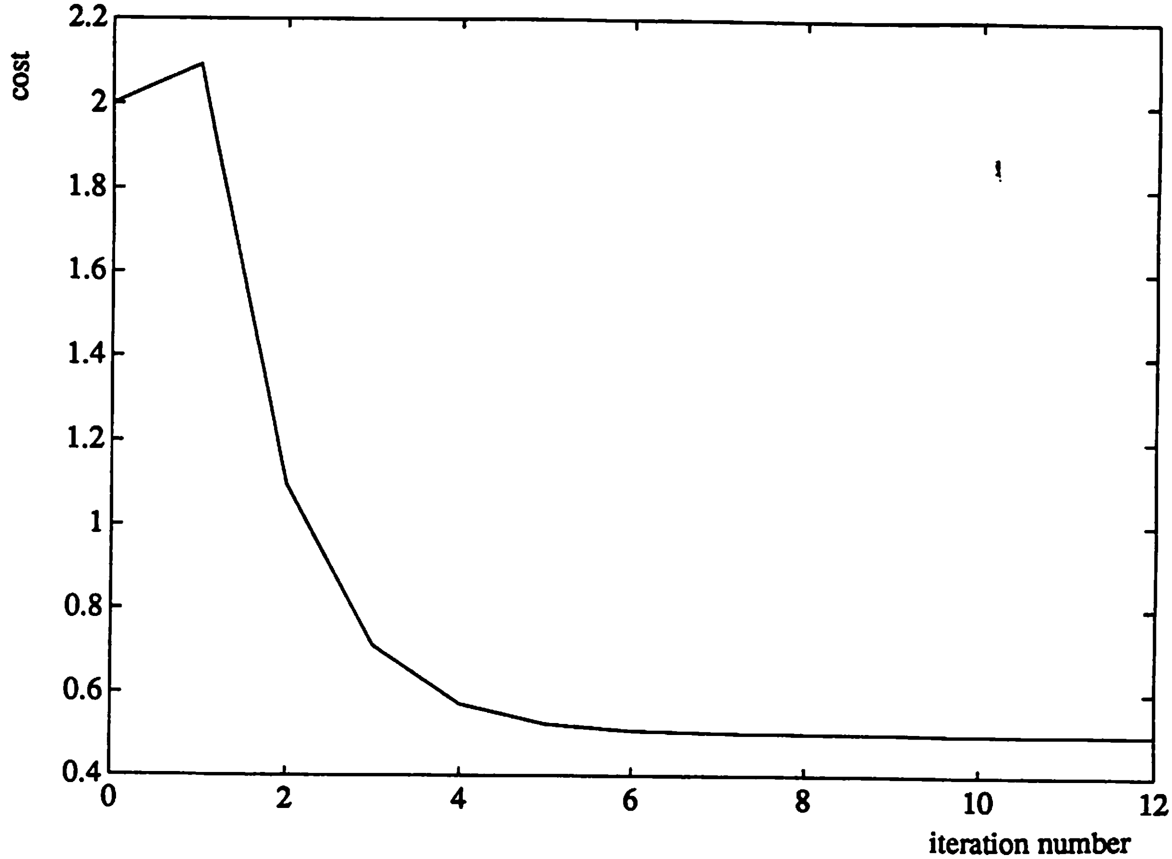

Figure 1a. Cost versus iteration number for Example 5.1(a).

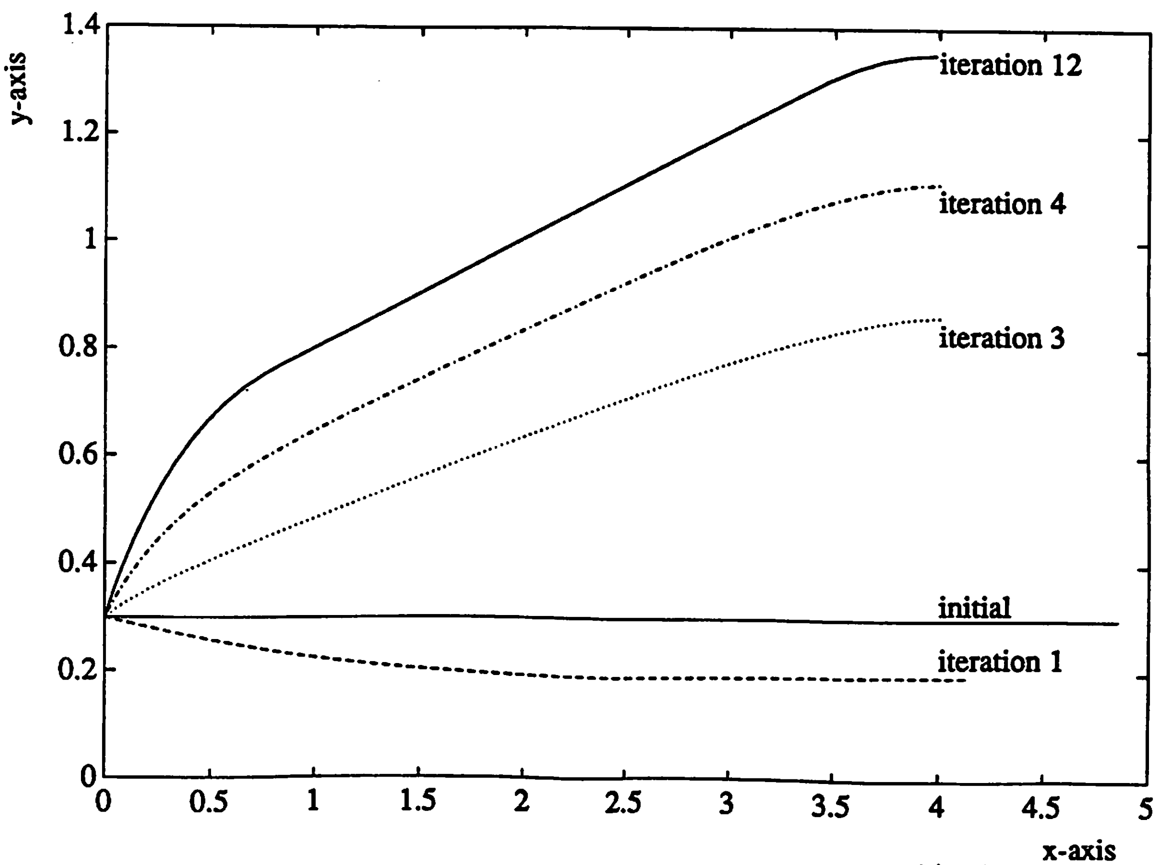

Figure 1b. State space trajectories for Example 5.1(a).

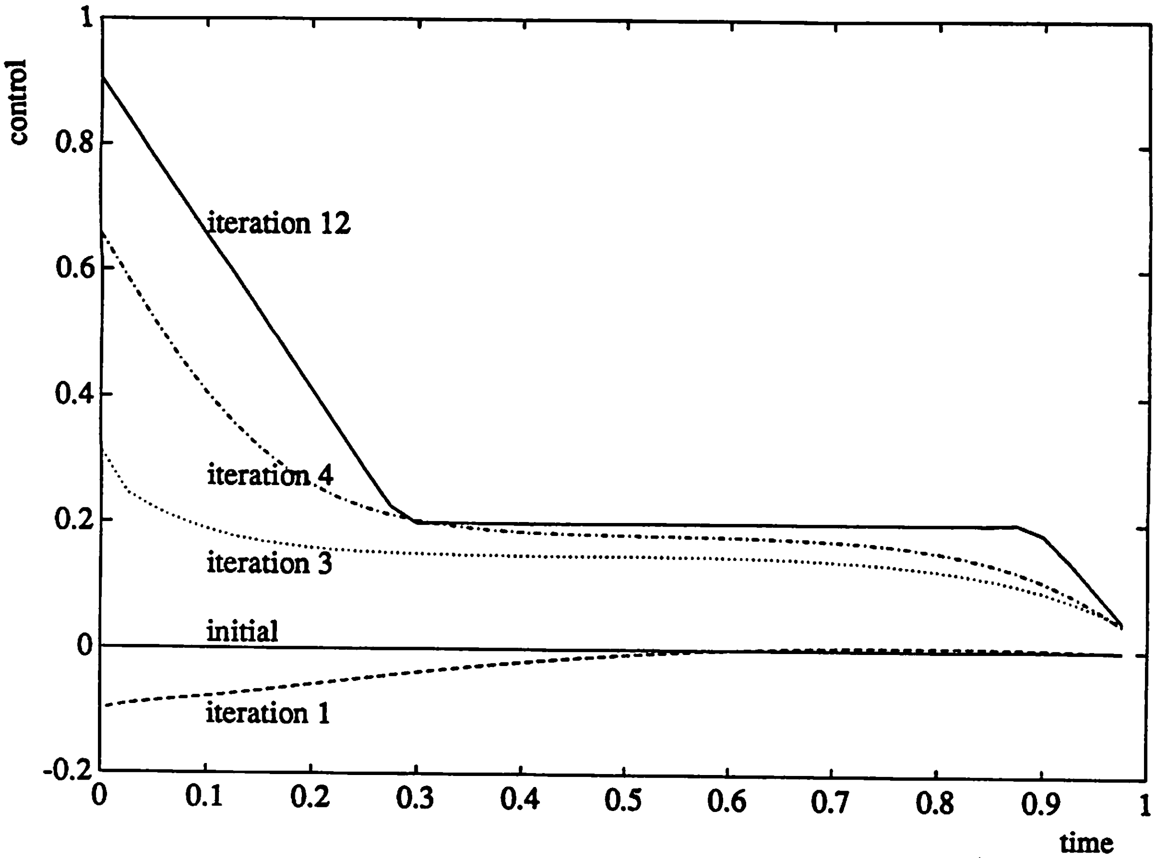

Figure 1c. Controls at various iterationsfor Example 5.1(a).

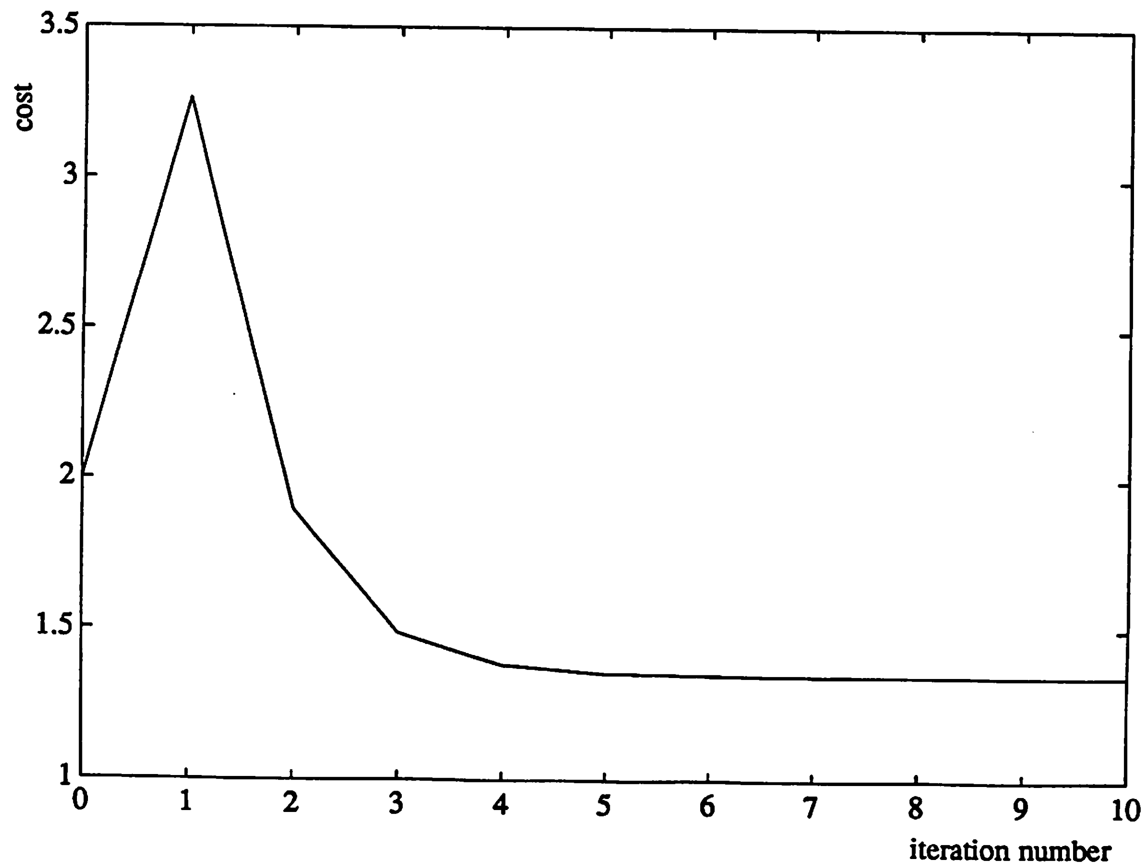

Figure 2a. Cost versus iteration number for Example 5.1(b).

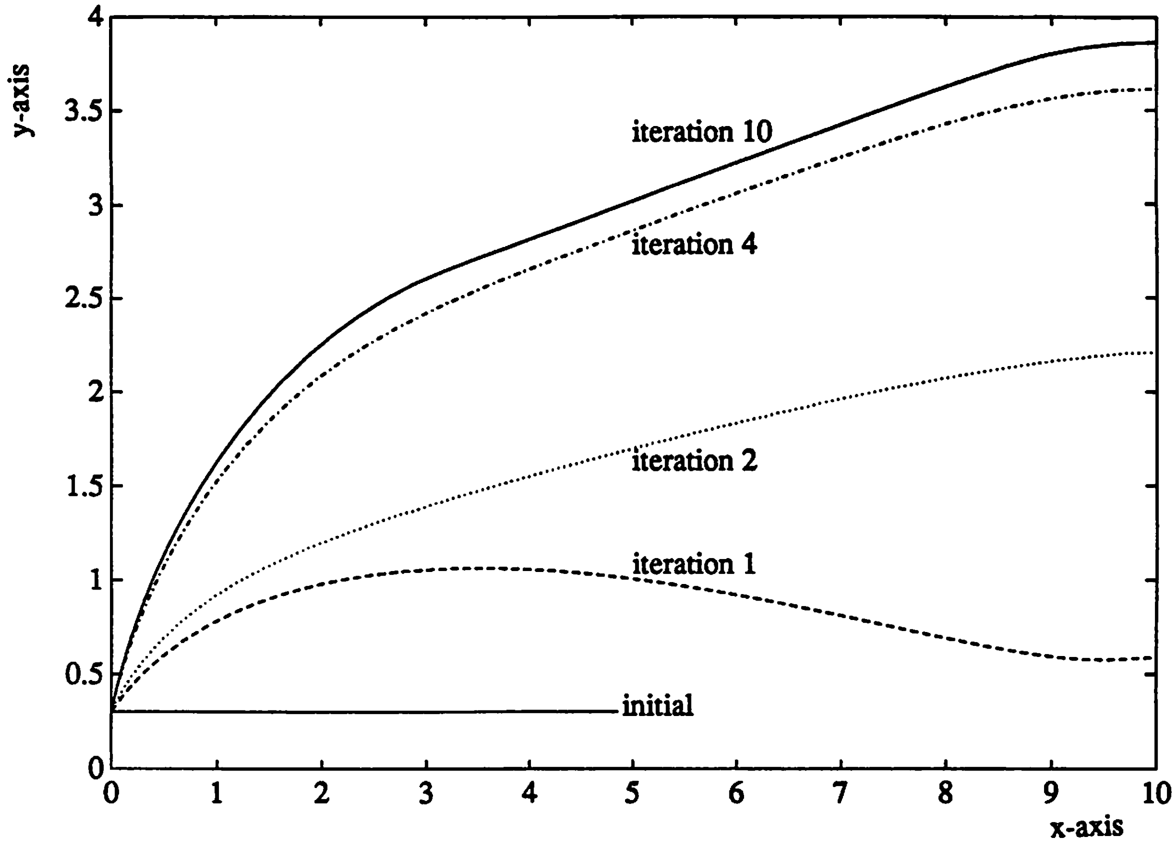

Figure 2b. State space trajectories for Example 5.1(b).

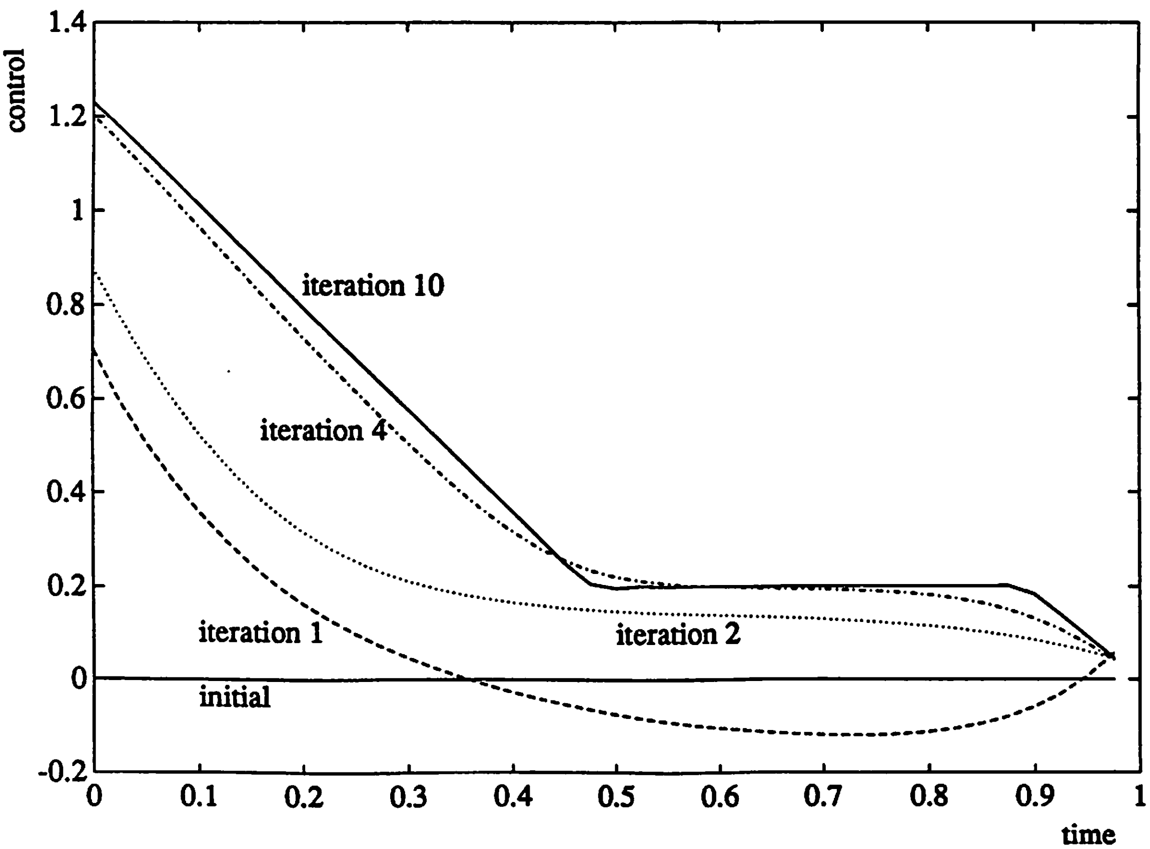

Figure 2c. Controls at various iterationsfor Example 5.1(b).

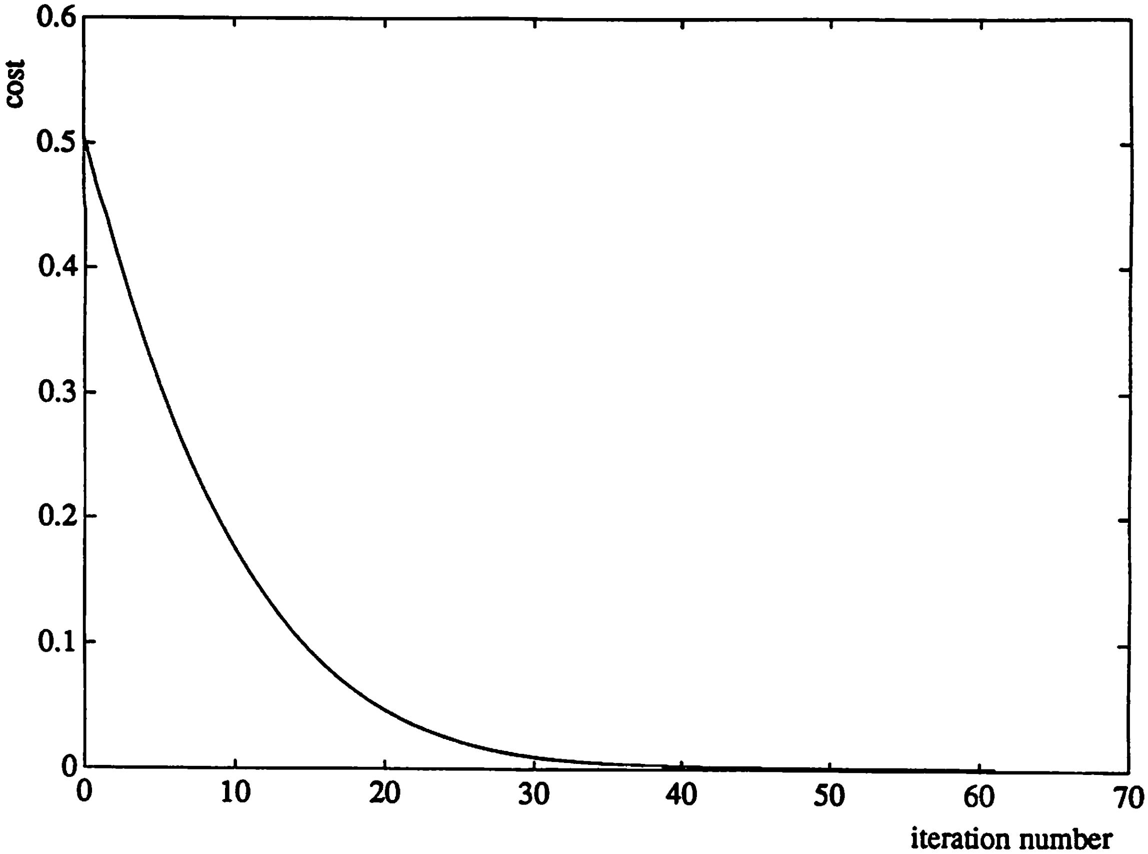

Figure 3a. Cost versus iteration number for Example 5.2 using penalized control.

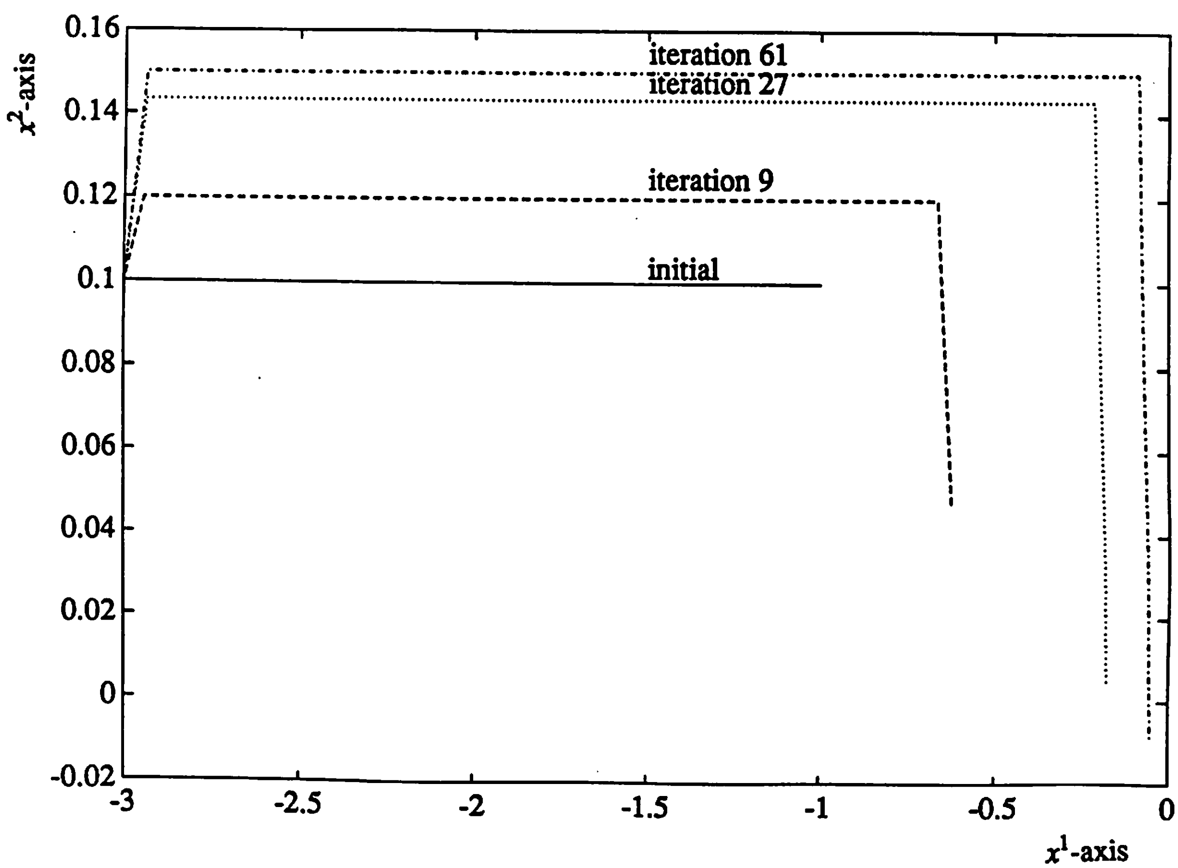

Figure 3b. State space trajectories for Example 5.2 using penalized control.

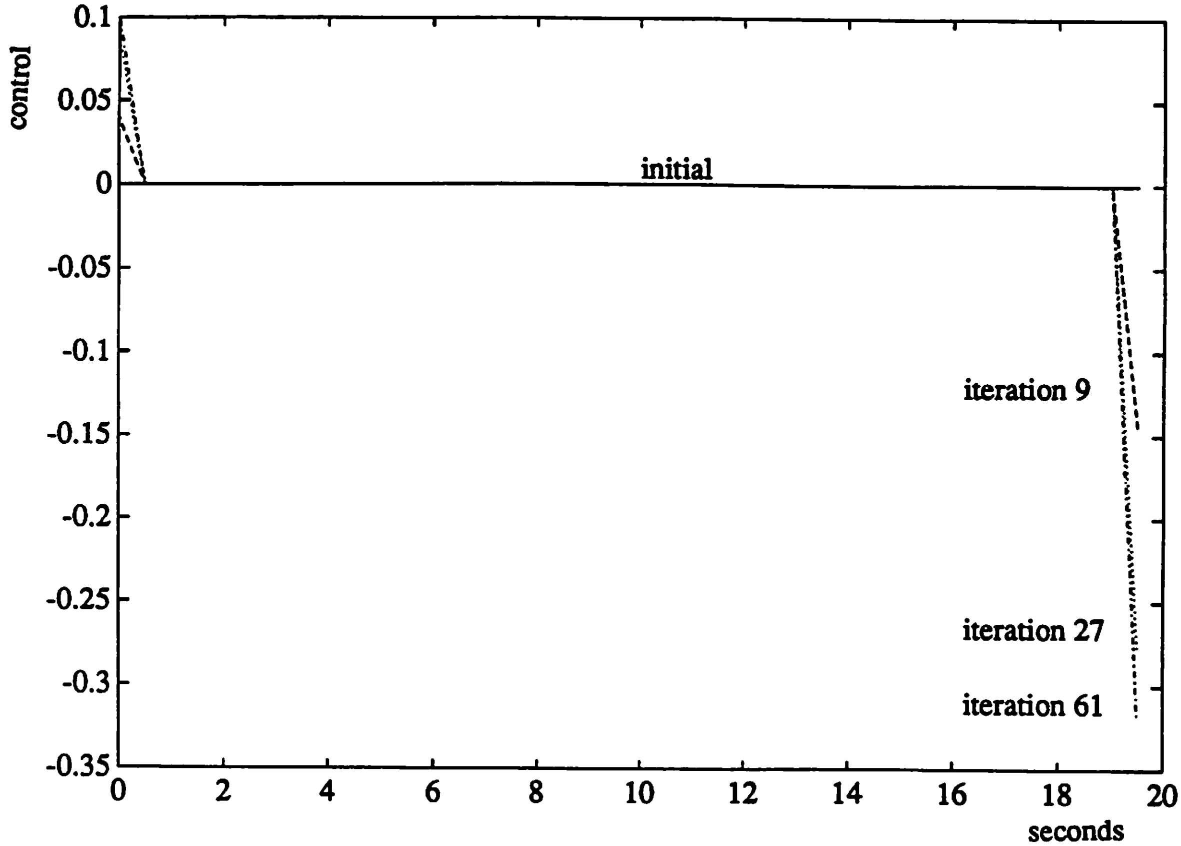

Figure 3c. Controls at various iterationsfor Example 5.2 using penalized control.

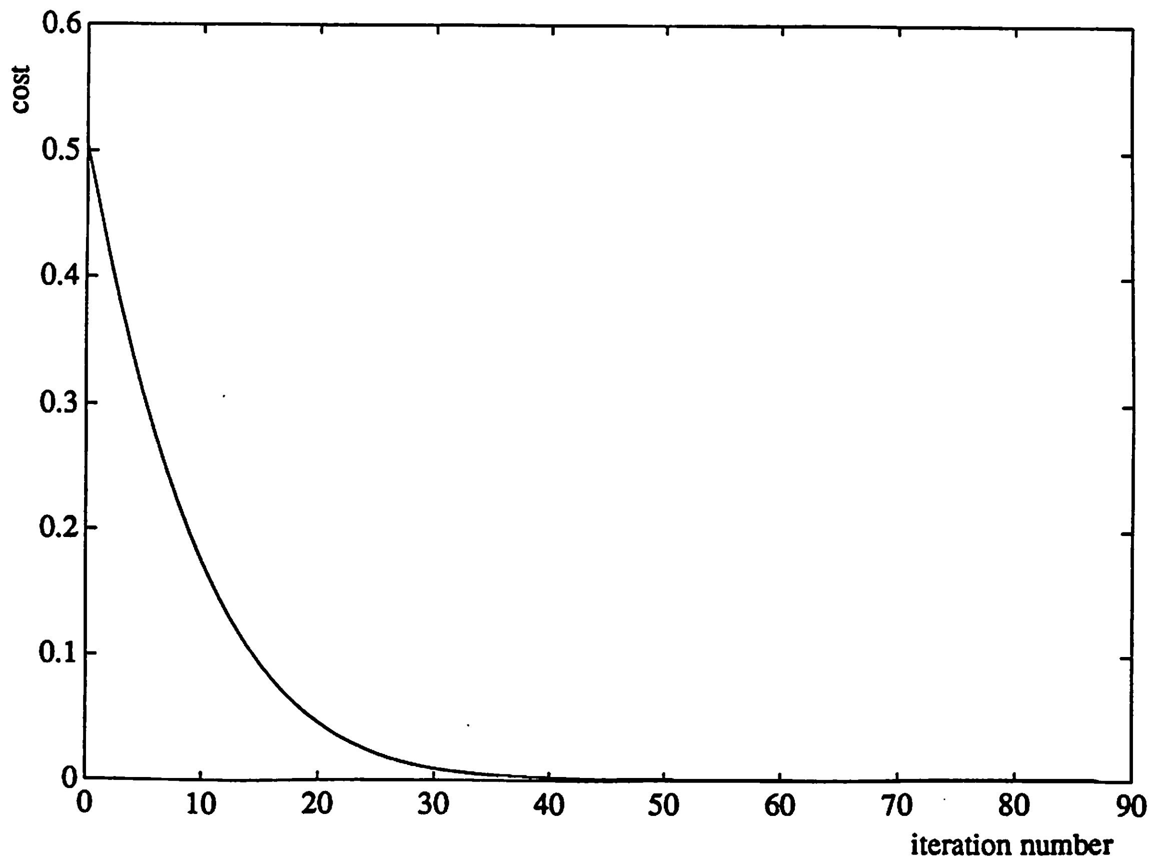

Figure 4a. Cost versus iteration number for Example 5.2 using non-penalized control.

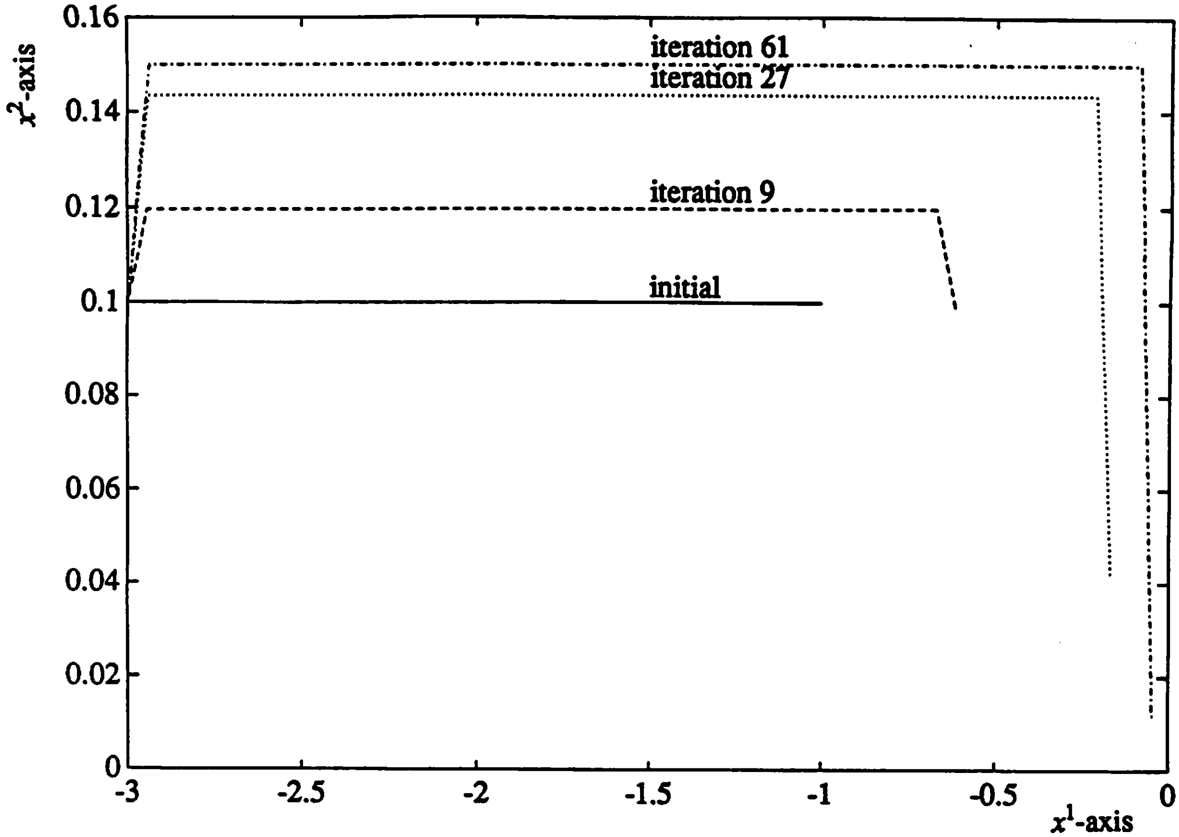

Figure 4b. State space trajectories for Example 5.2 using non-penalized control.

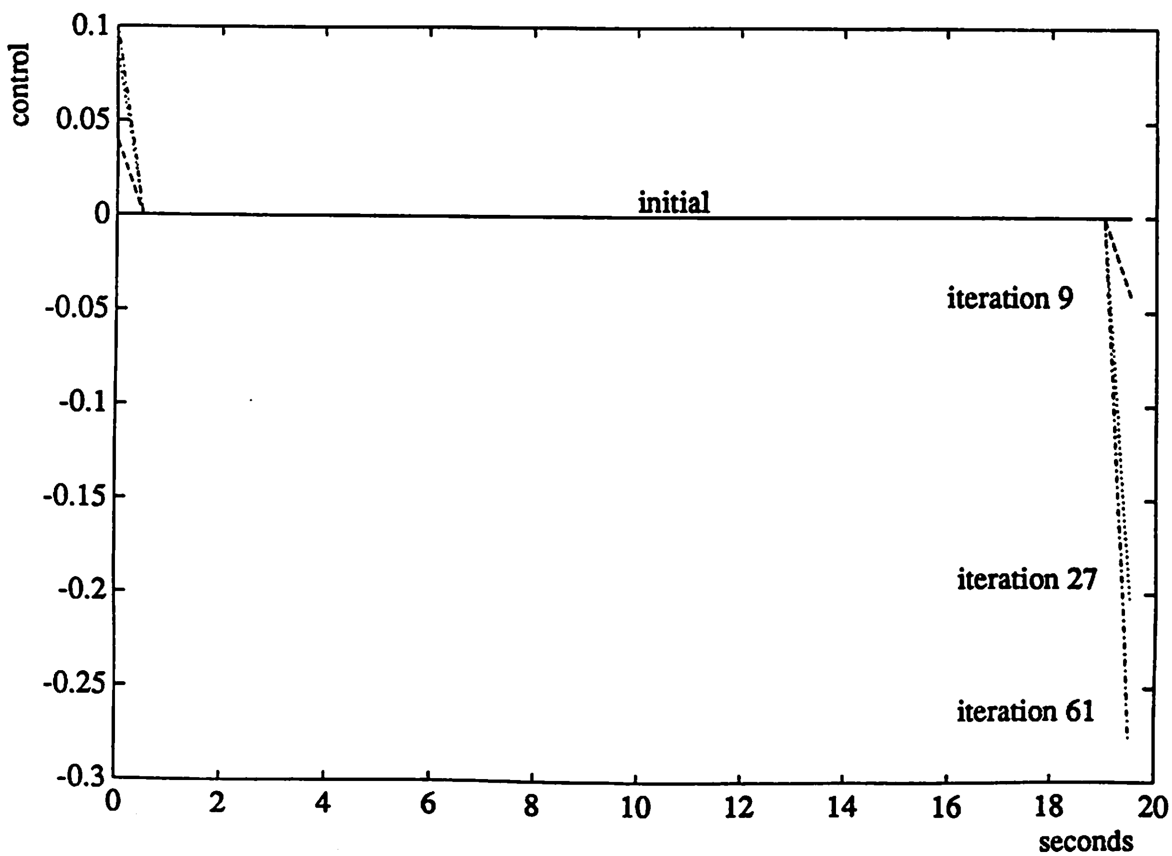

Figure 4c. Controls at various iterationsfor Example 5.2 using non-penalized control.