

NONLINEARLY CONSTRAINED OPTIMAL CONTROL PROBLEMS

K. L. TEO¹ and K. H. WONG²

(Received 23 February 1990; revised 2 May 1991)

Abstract

In a paper by Teo and Jennings, a constraint transcription is used together with the concept of control parametrisation to devise a computational algorithm for solving a class of optimal control problems involving terminal and continuous state constraints of inequality type. The aim of this paper is to extend the results to a more general class of constrained optimal control problems, where the problem is also subject to terminal equality state constraints. For illustration, a numerical example is included.

1. Introduction

In [19], a general class of optimal control problems involving inequality and terminal inequality state constraints is considered. This class of optimal control problems can be solved using the gradient-restoration algorithms due to Miele *et al.* (see [5] and [9–11]). It can also be solved by the algorithms of [4] and [18], where a simple constraint transcription is used to convert the continuous inequality constraints into equivalent constraints in canonical form. However, that constraint transcription has the disadvantage that equivalent equality constraints so obtained fail to satisfy the usual constraint qualification. Thus, convergence is not guaranteed and some oscillation may exist in numerical computation. This difficulty is overcome in [19] via a new constraint transcription. However, some practical problems may also be subject to additional terminal equality state constraints. In this paper, the computational algorithm of [19] is extended to this more general class of con-

¹Dept. of Mathematics, University of Western Australia, Nedlands, W.A., Australia .

²Dept. of Applied Mathematics, University of the Witwatersrand, Johannesburg, South Africa.

© Copyright Australian Mathematical Society 1992, Serial-fee code 0334-2700/92

strained optimal control problems, and the convergence properties similar to those reported in Theorem 4.2 and Theorem 4.3 of [19] are also shown to be valid under additional conditions. For illustration, a numerical example is considered.

Note that many interesting theoretical results for the class of problems considered in this paper may be found in [2–3]. For earlier work on control parametrisation, we refer the reader to [6–7], [14–18] and [20–21]. Finally, we wish to note that a general purpose software package known as MISER3 (cf. [8]) is now available for solving constrained optimal control problems including the one considered in this paper.

2. Problem Statement

Consider a process described by the following system of differential equations on the fixed time interval $(0, T]$.

$$\dot{\mathbf{x}}(t) = \mathbf{f}(t, \mathbf{x}(t), \mathbf{u}(t)) \quad (1a)$$

where $\mathbf{x} = [x_1, \dots, x_n]^T \in \mathbb{R}^n$, and $\mathbf{u} = [u_1, \dots, u_r]^T \in \mathbb{R}^r$ are, respectively, the state and control vectors; $\mathbf{f} = [f_1, \dots, f_n]^T \in \mathbb{R}^n$; and the subscript T denotes the transpose.

The initial condition for the differential equation (1a) is:

$$\mathbf{x}(0) = \mathbf{x}^0 \in \mathbb{R}^n. \quad (1b)$$

Define

$$U \equiv \{\mathbf{v} = [v_1, \dots, v_r]^T \in \mathbb{R}^r : \alpha_i \leq v_i \leq \beta_i, \quad i = 1, \dots, r\}, \quad (2)$$

where α_i and β_i , $i = 1, \dots, r$, are real numbers. Clearly, U is a compact and convex subset of \mathbb{R}^r .

Any measurable function defined on $[0, T]$ with values in U is called an admissible control. Let \mathcal{U} be the class of all such admissible controls. Let L_∞^r denote the Banach space $L_\infty([0, T], \mathbb{R}^r)$ of all essentially bounded measurable functions from $[0, T]$ into \mathbb{R}^r . Its norm is defined by

$$\|\mathbf{u}\|_\infty = \text{ess sup}_{t \in [0, T]} \left(\sum_{i=1}^r (u_i(t))^2 \right)^{1/2}.$$

For each $\mathbf{u} \in L_\infty^r$, let $\mathbf{x}(\cdot | \mathbf{u})$ be an absolutely continuous function defined on $[0, T]$ which satisfies the differential equation (1a) almost everywhere on $[0, T]$ and the initial condition (1b). This function is called the solution of system (1) corresponding to the control $\mathbf{u} \in L_\infty^r$.

The terminal state inequality and equality constraints, and continuous state inequality constraints are specified as follows:

$$\phi_i(\mathbf{x}(T|\mathbf{u})) \leq 0, \quad i = 1, \dots, N_I, \tag{3a}$$

$$\psi_i(\mathbf{x}(T|\mathbf{u})) = 0, \quad i = 1, \dots, N_E, \tag{3b}$$

$$g_i(t, \mathbf{x}(t|\mathbf{u})) \leq 0, \quad t \in [0, T], \quad i = 1, \dots, N \tag{3c}$$

where $\phi_i, i = 1, \dots, N_I$, and $\psi_i, i = 1, \dots, N_E$, are real-valued functions defined on \mathbb{R}^n ; and $g_i, i = 1, \dots, N$, are real-valued functions defined on $[0, T] \times \mathbb{R}^n$.

Let \mathcal{F} be the set which consists of all those elements of \mathcal{U} such that the constraints (3) are satisfied. Elements from \mathcal{F} are called feasible controls, and \mathcal{F} is called the class of feasible controls.

We may now state the optimal problem as follows.

PROBLEM (P). Given the system (1), find a control $u \in \mathcal{F}$ such that the cost functional

$$J(\mathbf{u}) = \Phi_0(\mathbf{x}(T|\mathbf{u})) + \int_0^T \mathcal{L}_0(t, \mathbf{x}(t|\mathbf{u}), \mathbf{u}(t)) dt \tag{4}$$

is minimised over \mathcal{F} , where Φ_0 and \mathcal{L}_0 are given real-valued functions. We assume that the following conditions are satisfied throughout.

- (A1) $\mathbf{f}: [0, T] \times \mathbb{R}^n \times \mathbb{R}^r \rightarrow \mathbb{R}^n$ is piecewise continuous on $[0, T]$ for each $(\mathbf{x}, \mathbf{u}) \in \mathbb{R}^n \times \mathbb{R}^r$ and is continuously differentiable with respect to each of the components of \mathbf{x} and \mathbf{u} for each $t \in [0, T]$; furthermore, there exists a constant $K > 0$ such that

$$|\mathbf{f}(t, \mathbf{x}, \mathbf{u})| \leq K(1 + |\mathbf{x}|),$$

for all $(\mathbf{x}, \mathbf{u}) \in \mathbb{R}^n \times U$, where $|\cdot|$ denotes the usual Euclidean norm;

- (A2) for each $i = 1, \dots, N_I, \phi_i: \mathbb{R}^n \rightarrow \mathbb{R}$ is continuously differentiable;
- (A3) for each $i = 1, \dots, N_E, \psi_i: \mathbb{R}^n \rightarrow \mathbb{R}$ is continuously differentiable;
- (A4) for each $i = 1, \dots, N, g_i: [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$ is continuously differentiable;
- (A5) $\Phi_0: \mathbb{R}^n \rightarrow \mathbb{R}$ is continuously differentiable;
- (A6) $\mathcal{L}_0: [0, T] \times \mathbb{R}^n \times \mathbb{R}^r \rightarrow \mathbb{R}$ is piecewise continuous on $[0, T]$ for each $(\mathbf{x}, \mathbf{u}) \in \mathbb{R}^n \times \mathbb{R}^r$ and is continuously differentiable with respect to each of the components of \mathbf{x} and \mathbf{u} for each $t \in [0, T]$.

REMARK 2.1. From the theory of differential equations, we recall that the system (1) admits a unique solution $\mathbf{x}(\cdot|\mathbf{u})$ corresponding to each $\mathbf{u} \in L^\infty_r$, and hence for each $\mathbf{u} \in \mathcal{U}$.

3. Control Parametrisation

As in [4] and [18], the control parametrisation method will be used to solve the problem (P). To be more precise, let $\{\mathcal{S}_p\}_{p=1}^{\infty}$ be a sequence of partitions of the interval $[0, T]$ such that \mathcal{S}_p has $n_p + 1$ elements, \mathcal{S}_{p+1} is a refinement of \mathcal{S}_p and $\|\mathcal{S}_p\| \rightarrow 0$ as $p \rightarrow \infty$, where $\|\mathcal{S}_p\|$ is the length of the largest interval in the partition \mathcal{S}_p . In this paper, we assume that

$$\mathcal{S}_p = \{I_j^p\}_{j=1}^{n_p}$$

where

$$\mathcal{S}_j^p = [t_{j-1}^p, t_j^p], \quad 0 = t_0^p < t_1^p < \dots < t_{N_p}^p = T.$$

Then

$$\|\mathcal{S}_p\| = \max_{j \in \{1, \dots, N_p\}} l(I_j^p),$$

where $l(I_j^p) = t_j^p - t_{j-1}^p$.

Let \mathcal{U}^p be the subset of admissible controls which are piecewise constant and consistent with the partition \mathcal{S}_p . It is clear that $\mathbf{u}^p \in \mathcal{U}^p$ can be written as:

$$\mathbf{u}^p(t) = \sum_{k=1}^{n_p} \sigma_k^p \chi_k^p(t), \quad t \in [0, T],$$

where $\sigma_k^p \in \mathbb{R}^r$, and χ_k^p denotes the characteristic function of I_k^p . This means that each control $\mathbf{u}^p \in \mathcal{U}^p$ can be identified uniquely with a control parameter vector σ^p and vice versa, where

$$\sigma^p = [(\sigma_1^p)^\top, \dots, (\sigma_{N_p}^p)^\top]^\top.$$

Thus, when no confusion can arise, we interchangeably refer to $\mathbf{u}^p \in \mathcal{U}^p$ and $\sigma^p \in \mathcal{U}^p$.

Given a particular control parameter vector $\sigma^p \in \mathcal{U}^p$, let $\mathbf{x}(\cdot | \sigma^p)$ be the solution of the differential equation:

$$\dot{\mathbf{x}} = \hat{\mathbf{f}}(t, \mathbf{x}(t), \sigma^p)$$

with initial condition:

$$\mathbf{x}(0) = \mathbf{x}^0,$$

where $\hat{\mathbf{f}}$ is obtained from \mathbf{f} in an obvious way.

Let

$$\theta^p = \left\{ \sigma^p \in \mathcal{W}^p : \phi_i(x(T|\sigma^p)) \leq 0, \quad i = 1, \dots, N_I; \right. \\ \left. \max_{t \in [0, T]} g_i(t, \mathbf{x}(t|\sigma^p)) \leq 0, \quad i = 1, \dots, N \right\} \quad (5)$$

and let

$$\mathcal{F}^p = \{ \sigma^p \in \theta^p : \psi_i(\mathbf{x}(t|\sigma^p)) = 0, \quad i = 1, \dots, N_E \}. \quad (6)$$

We define the approximate problem (P(p)) as follows:

PROBLEMS (P(p)). Find a control vector $\sigma^p \in \mathcal{F}^p$ such that the cost functional

$$\hat{J}(\sigma^p) = \Phi_0(\mathbf{x}(T|\sigma^p)) + \int_0^T \hat{\mathcal{L}}_0(t, \mathbf{x}(t|\sigma^p), \sigma^p) dt \quad (7)$$

is minimised over \mathcal{F}^p , where $\hat{\mathcal{L}}_0$ is obtained from \mathcal{L}_0 in an obvious way.

4. Constraint Approximation

For each $i = 1, \dots, N$, the corresponding continuous state inequality constraint (3c), when restricted to the space \mathcal{W}^p , can be written as:

$$G_i(\sigma^p) = \int_0^T \max\{0, g_i(t, \mathbf{x}(t|\sigma^p))\} dt = 0 \quad (8)$$

which is, however, nonsmooth. For convenience, let the approximate problem (P(p)) with (3c) replaced by (8), be again denoted by (P(p)). As in [19], we replace the nonsmooth functions $\max\{0, g_i(t, \mathbf{x}(t|\sigma^p))\}$ by the smooth ones given by $g_{i,\varepsilon}(t, \mathbf{x}(t|\sigma^p))$, where

$$g_{i,\varepsilon}(t, \mathbf{x}(t|\sigma^p)) = \begin{cases} g_i(t, \mathbf{x}(t|\sigma^p)), & \text{if } g_i(t, \mathbf{x}(t|\sigma^p)) > \varepsilon, \\ (g_i(t, \mathbf{x}(t|\sigma^p)) + \varepsilon)^2 / 4\varepsilon, & \text{if } |g_i(t, \mathbf{x}(t|\sigma^p))| \leq \varepsilon, \\ 0, & \text{if } g_i(t, \mathbf{x}(t|\sigma^p)) < -\varepsilon. \end{cases} \quad (9)$$

For each $i = 1, \dots, N$, define

$$G_{i,\varepsilon}(\sigma^p) = \int_0^T g_{i,\varepsilon}(t, \mathbf{x}(t|\sigma^p)) dt. \quad (10)$$

We now define two related approximate problems which will be referred to as Problems $(P_\varepsilon(p))$ and $(P_{\varepsilon,\gamma}(p))$.

PROBLEM $(P_\varepsilon(p))$. Problem (P(p)) with the continuous state inequality constraints (8) replaced by

$$G_{i,\varepsilon}(\sigma^p) = 0, \quad i = 1, \dots, N. \quad (11)$$

Let $\mathcal{F}_\varepsilon^p$ be the feasible region of $(P_\varepsilon(p))$, defined by

$$\begin{aligned} \mathcal{F}_\varepsilon^p = \{ \boldsymbol{\sigma}^p \in \mathcal{U}^p : & G_{i,\varepsilon}(\boldsymbol{\sigma}^p) = 0, \quad i = 1, \dots, N; \\ & \phi_i(\mathbf{x}(T|\boldsymbol{\sigma}^p)) \leq 0, \quad i = 1, \dots, N_I; \\ & \psi_i(\mathbf{x}(T|\boldsymbol{\sigma}^p)) = 0, \quad i = 1, \dots, N_E \}. \end{aligned} \quad (12)$$

Then, it is clear that

$$\mathcal{F}_\varepsilon^p \subset \mathcal{F}^p. \quad (13)$$

Note that the equality constraints (11) fail to satisfy the usual constraint qualification. Thus, we may encounter numerical difficulty. To overcome this difficulty, we consider the second approximation problem as follows:

PROBLEM $(P_{\varepsilon,\gamma}(p))$. Problem $(P(p))$ with (8) replaced by

$$G_{i,\varepsilon}(\boldsymbol{\sigma}^p) \leq \gamma, \quad i = 1, \dots, N. \quad (14)$$

For a given problem (P) , we propose to solve the problem via solving a sequence of problems $(P_{\varepsilon,\gamma}(p))$ as follows:

- (1) Choose a positive integer value for p (say, $p = 10$) and hence obtain the problem $(P(p))$.
- (2) Construct the problem $(P_{\varepsilon,\gamma}(p))$ with $\varepsilon > 0$ and $\gamma > 0$ (say, $\varepsilon = 10^{-3}$ and $\gamma = 10^{-4}$).
- (3) Solve the problem $(P_{\varepsilon,\gamma}(p))$ with decreasing γ . It will be shown that for each $\varepsilon > 0$, there exists a $\gamma_0 > 0$ such that for each γ , $0 < \gamma < \gamma_0$, the optimal control vector of the corresponding problem $(P_{\varepsilon,\gamma}(p))$ is a feasible control vector of the problem $(P(p))$. Consequently, it gives rise to a suboptimal control to the problem (P) .
- (4) Reduce $\varepsilon > 0$ (say, by taking $\varepsilon = 10^{-1}\varepsilon$), and go back to (2). From our numerical experiences, we found that there is no need to reduce ε below 10^{-6} .
- (5) Increase the value of p (say, by taking $p = 2p$), and go back to (1). Again from our numerical experiences, we found that it suffices to take $p = 20$. In some exceptional cases, we may need to increase the value of p to 40.

The approach proposed above works very well in practice. In fact, the general purpose software package MISER3 (cf. [19]) is developed using this idea for solving the problem (P) . The rigorous justification is, however, rather involved. We need to introduce several auxiliary problems. Note that there is no need to solve any of these auxiliary problems; they are introduced solely for the purpose of establishing the relationship between the problem (P) and the problem $(P_{\varepsilon,\gamma}(p))$. In fact, only the approximate problem $(P_{\varepsilon,\gamma}(p))$

is required to be solved in practice. Let us proceed further to define the first two auxiliary problems. They are called the δ -tolerated version of the approximate problems $(P(p))$ and $(P_\varepsilon(p))$. We shall refer to them as the problems $(P^\delta(p))$ and $(P_\varepsilon^\delta(p))$, respectively.

DEFINITION 4.1. A control vector $\sigma^p \in \theta^p$ is said to be δ -tolerated if it satisfies the constraints

$$-\delta \leq \psi_i(t, \mathbf{x}(t|\sigma^p)) \leq \delta, \quad i = 1, \dots, N_E. \tag{15}$$

Let $\mathcal{F}^{p,\delta}$ be the subset of θ^p such that the δ -tolerated constraints (15) are satisfied. Let $\text{int}(\mathcal{F}^{p,\delta})$ be the interior of the set $\mathcal{F}^{p,\delta}$ in the sense that

$$\text{int}(\mathcal{F}^{p,\delta}) = \left\{ \sigma^p \in \mathcal{U}^p : \begin{aligned} &\phi_i(\mathbf{x}(T|\sigma^p)) < 0, \quad i = 1, \dots, N_I; \\ &\max_{t \in [0, T]} g_i(t, \mathbf{x}(t|\sigma^p)) < 0, \quad i = 1, \dots, N; \\ &-\delta < \psi_i(\mathbf{x}(t|\sigma^p)) < \delta, \quad i = 1, \dots, N_E \end{aligned} \right\}. \tag{16}$$

Furthermore, let

$$\mathcal{F}_\varepsilon^{p,\delta} = \left\{ \sigma^p \in \mathcal{U}^p : \begin{aligned} &G_{i,\varepsilon}(\sigma^p) = 0, \quad i = 1, \dots, N; \\ &\phi_i(\mathbf{x}(t|\sigma^p)) \leq 0, \quad i = 1, \dots, N_I; \\ &-\delta \leq \psi_i(\mathbf{x}(T|\sigma^p)) \leq \delta, \quad i = 1, \dots, N_E \end{aligned} \right\}. \tag{17}$$

In view of (13), (12), (16) and (17) it is clear that

$$\mathcal{F}_\varepsilon^p \subset \mathcal{F}^p \subset \mathcal{F}^{p,\delta}, \tag{18}$$

and

$$\mathcal{F}_\varepsilon^p \subset \mathcal{F}_\varepsilon^{p,\delta}. \tag{19}$$

We can now define the δ -tolerated version of the approximate problem $(P(p))$ and $(P_\varepsilon(p))$ as follows:

PROBLEM $(P^\delta(p))$. Find a control vector $\sigma^p \in \mathcal{F}^{p,\delta}$ such that the cost functional (7) is minimised over $\mathcal{F}^{p,\delta}$.

PROBLEM $(P_\varepsilon^\delta(p))$. Find a control vector $\sigma^p \in \mathcal{F}_\varepsilon^{p,\delta}$ such that the cost functional (7) is minimised over $\mathcal{F}_\varepsilon^{p,\delta}$.

Let $\sigma^{p,*}$ and $\sigma_{\varepsilon}^{p,*}$ be optimal controls of the approximate problems $(P(p))$ and $(P_{\varepsilon}(p))$, respectively. Let $\sigma^{p,\delta,*}$ and $\sigma_{\varepsilon}^{p,\delta,*}$ be optimal controls of the corresponding δ -tolerated version of these problems.

We make the following additional assumptions.

- (A6) There exists $\varepsilon_1 > 0$ such that for each $\sigma^p \in \mathcal{F}^{p,\delta}$ ($\delta > 0$), there exists a control $\bar{\sigma}^p \in \text{int}(\mathcal{F}^{p,\delta}) \cap \mathcal{F}_{\varepsilon_1}^p$ such that

$$\alpha \bar{\sigma}^p + (1 - \alpha) \sigma^p \in \text{int}(\mathcal{F}^{p,\delta}) \cap \mathcal{F}_{\varepsilon_1}^p, \quad \alpha \in (0, 1].$$

- (A7) For any $\varepsilon_1 > 0$, there exists $p_1 > 0$ such that

$$\lim_{\delta \rightarrow 0} \hat{J}(\sigma_{\varepsilon_1}^{p,\delta,*}) = \hat{J}(\sigma_{\varepsilon_1}^{p,*}) \quad (20)$$

uniformly with respect to $p \geq p_1$.

- (A8) For each $p_1 > 0$ and $\varepsilon_1 > 0$

$$\lim_{\delta \rightarrow 0} \hat{J}(\sigma_{\varepsilon_1}^{p_1,\delta,*}) = \hat{J}(\sigma_{\varepsilon_1}^{p_1,*}). \quad (21)$$

REMARK 4.1. The assumption (A6) roughly implies that the feasible set $\mathcal{F}^{p,\delta}$ contains a non-empty interior and is locally convex. It is to ensure that the problem to be solved is “well-posed”. This assumption is clearly not easily verifiable. However, in the absence of this assumption, the feasible region may contain a subregion with empty interior. Such problems are obviously extremely difficult, if not impossible, to solve numerically.

REMARK 4.2. The assumptions (A7) and (A8) roughly mean that small violations in the terminal equality constraints will not produce significant changes in the cost value. This type of assumption is, in fact, rather reasonable, because exact satisfaction of the equality constraints is simply impossible in a computer.

Further to the comments made in Remark 4.1 and Remark 4.2, we note that it is possible to construct examples which fail to satisfy these assumptions. However, the practical value of such examples might not be too significant.

LEMMA 4.1. For each $p > 0$, let $\{\sigma_{\varepsilon}^{p,*}\}$ be a sequence in ε of the optimal control parameters of $(P_{\varepsilon}(p))$. Then

$$\lim_{\varepsilon \rightarrow 0} \hat{J}(\sigma_{\varepsilon}^{p,*}) = \hat{J}(\sigma^{p,*}) \quad (22)$$

where $\sigma^{p,*}$ is the optimal control parameter vector of the problem $(P(p))$.

PROOF. Let $\hat{\delta} > 0$ be given. Let $\sigma^{p, \hat{\delta}, *}$ be the optimal control of the problem $(P^{\hat{\delta}}(p))$. Then, by (A6), there exists a control $\bar{\sigma}^p \in \text{int}(\mathcal{F}^{p, \hat{\delta}}) \cap \mathcal{F}_{\varepsilon_1}^p$ such that

$$\sigma_{\alpha}^p = \alpha \bar{\sigma}^p + (1 - \alpha) \sigma^{p, \hat{\delta}, *}, \quad \alpha \in (0, 1].$$

For any $c > 0$, there exists $\alpha_1 \in (0, 1)$ such that

$$\hat{J}(\sigma^{p, \hat{\delta}, *}) \leq \hat{J}(\sigma_{\alpha_1}^p) < \hat{J}(\sigma^{p, \hat{\delta}, *}) + c. \tag{23}$$

Since $\sigma_{\alpha_1}^p \in \text{int}(\mathcal{F}^{p, \hat{\delta}}) \cap \mathcal{F}_{\varepsilon_1}^p \subset \mathcal{F}_{\varepsilon_1}^{p, \hat{\delta}}$, we have

$$\hat{J}(\sigma_{\varepsilon_1}^{p, \hat{\delta}, *}) < \hat{J}(\sigma_{\alpha_1}^p), \tag{24}$$

where $\sigma_{\varepsilon_1}^{p, \hat{\delta}, *}$ is an optimal control of the problem $(P_{\varepsilon_1}^{\hat{\delta}}(p))$. Thus, from (23) and (24) and the definitions of $\sigma^{p, \hat{\delta}, *}$ and $\sigma_{\varepsilon_1}^{p, \hat{\delta}, *}$, we obtain

$$\hat{J}(\sigma^{p, \hat{\delta}, *}) \leq \hat{J}(\sigma_{\varepsilon_1}^{p, \hat{\delta}, *}) \leq \hat{J}(\sigma^{p, \hat{\delta}, *}) + c. \tag{25}$$

Taking limit as $\hat{\delta} \rightarrow 0$ in (25), it follows from (A7) and (A8) that

$$\hat{J}(\sigma^{p, *}) < \hat{J}(\sigma_{\varepsilon_1}^p) < \hat{J}(\sigma^{p, *}) + c. \tag{26}$$

On the other hand, for any ε , $0 < \varepsilon < \varepsilon_1$, we have

$$\hat{J}(\sigma^{p, *}) \leq \hat{J}(\sigma_{\varepsilon}^p) \leq \hat{J}(\sigma_{\varepsilon_1}^p). \tag{27}$$

Thus, from (26) and (27), we obtain

$$\hat{J}(\sigma^{p, *}) \leq \hat{J}(\sigma_{\varepsilon}^p) \leq \hat{J}(\sigma^{p, *}) + c \tag{28}$$

for all ε , $0 < \varepsilon < \varepsilon_1$. Hence,

$$\hat{J}(\sigma^{p, *}) \leq \lim_{\varepsilon \rightarrow 0} \hat{J}(\sigma_{\varepsilon}^p) \leq \hat{J}(\sigma^{p, *}) + c. \tag{29}$$

Since $c > 0$ is arbitrary, the conclusion of this lemma follows easily from (29).

LEMMA 4.2. *There exists a $\gamma_0(\varepsilon) > 0$ such that for all γ , $0 < \gamma < \gamma_0(\varepsilon)$, any feasible control vector $\sigma_{\varepsilon, \gamma}^p$ of problem $(P_{\varepsilon, \gamma}(p))$, i.e.,*

$$G_{i, \varepsilon}(\sigma_{\varepsilon, \gamma}^p) \leq \gamma, \quad i = 1, \dots, N, \tag{30}$$

$$\phi_j(\mathbf{x}(T | \sigma_{\varepsilon, \gamma}^p)) \leq 0, \quad i = 1, \dots, N_I, \tag{31}$$

$$\psi_i(\mathbf{x}(T | \sigma_{\varepsilon, \gamma}^p)) = 0, \quad i = 1, \dots, N_E \tag{32}$$

is also a feasible control parameter vector of the problem $(P(p))$.

PROOF. The proof is similar to that given for Lemma 3.3 of [19].

REMARK 4.3. In view of Lemma 4.2, it is clear that Algorithm A2 of [19] can be easily modified so that we can use it to generate a sequence $\{\sigma_{\varepsilon,\gamma}^{p,*}\}$ of control parameter vectors in ε such that each element is in the feasible region of $(P(p))$.

REMARK 4.4. From Remark 4.2 of [19], the problem $(P_{\varepsilon,\gamma}(p))$ is essentially a nonlinear mathematical programming problem in control parameters, which can be solved by any standard optimisation software package such as NLPQL (see [13]).

The following Theorem is a direct consequence of Theorem 4.1 of [19].

THEOREM 4.1. Let $\{\sigma_{\varepsilon,\gamma}^{p,*}\}$ be a sequence in ε of control parameter vectors produced by the algorithm. Then,

$$\hat{J}(\sigma_{\varepsilon,\gamma}^{p,*}) \rightarrow \hat{J}(\sigma^{p,*}),$$

where $\sigma^{p,*}$ is an optimal control of Problem $(P(p))$. Furthermore, any accumulation point of $\{\sigma_{\varepsilon,\gamma}^{p,*}\}$ is a solution of Problem $(P(p))$.

5. Some Convergence Results

In this section, we shall investigate some convergence properties of the sequence of approximate optimal controls to the true optimal control.

Let $\{\sigma^{p,*}\}_{p=1}^{\infty}$ be a sequence of optimal control vectors to the sequence of finite-dimensional problems $\{(P(p))\}$; and let $\{u^{p,*}\}_{p=1}^{\infty}$ be the corresponding sequence of controls in \mathcal{U} . Clearly,

$$J(u^{p+1,*}) \leq J(u^{p,*}).$$

A control $u \in \theta$ is said to be δ -tolerated feasible if it satisfies (15) with σ^p being replaced by u . Let \mathcal{F}^δ be the subset of θ which consists of all those δ -tolerated feasible controls. Let $\text{int}(\mathcal{F}^\delta)$ be defined from $\text{int}(\mathcal{F}^{\delta,p})$ by replacing \mathcal{U}^p and σ^p in (1b) by \mathcal{U} and u , respectively. Let the δ -tolerated version of the original problem (P) , denoted by (P^δ) , be defined from $(P^\delta(p))$ by replacing σ^p and $\mathcal{F}^{p,\delta}$ in the definition of $(P^\delta(p))$ by u and \mathcal{F}^δ , respectively. Let $u^{\delta,*}$ and u^* be optimal controls of (P^δ) and (P) , respectively.

We need the following additional assumptions:

(A9) For each $\mathbf{u} \in \mathcal{F}^\delta$, there exists a control $\bar{\mathbf{u}} \in \text{int}(\mathcal{F}^\delta)$ such that

$$\alpha \bar{\mathbf{u}} + (1 - \alpha)\mathbf{u} \in \text{int}(\mathcal{F}^\delta), \quad \alpha \in (0, 1].$$

(A10) $\lim_{\delta \rightarrow 0} J(\mathbf{u}^{\delta,*}) = J(\mathbf{u}^*)$.

Note that similar comments made for (A7) and (A8) are applicable to (A9) and (A10).

REMARK 5.1. The problems $(P^\delta(p))$ and (P^δ) involve only continuous state inequality constraints. Hence, they have the same structure as the problem $(P(p))$ and (P) in [19], respectively. Thus the convergence result of Theorem 4.2 of [19] can be applied to the problem $(P^\delta(p))$ as $p \rightarrow \infty$.

THEOREM 5.1. Let $\mathbf{u}^{p,*}$ be an optimal control of the approximate problem $(P(p))$. Suppose that \mathbf{u}^* is an optimal control of the problem (P) . Then,

$$\lim_{p \rightarrow \infty} J(\mathbf{u}^{p,*}) = J(\mathbf{u}^*).$$

PROOF. From (A10) and (A7), it is clear that for any $c > 0$, there exists $\hat{\delta} > 0$ such that

$$0 \leq J(\mathbf{u}^*) - J(\mathbf{u}^{\hat{\delta},*}) \leq c \tag{33}$$

and

$$0 \leq J(\mathbf{u}^{p,*}) - J(\mathbf{u}^{p,\hat{\delta},*}) \leq c \tag{34}$$

for all $p > p_1$. On the other hand, in view of Remark 5.1 and Theorem 4.2 of [19], we have

$$\lim_{p \rightarrow \infty} J(\mathbf{u}^{p,\hat{\delta},*}) = J(\mathbf{u}^{\hat{\delta},*}). \tag{35}$$

Thus, from (33), (34) and (35), we get

$$\left| \lim_{p \rightarrow \infty} J(\mathbf{u}^{p,*}) - J(\mathbf{u}^*) \right| \leq c. \tag{36}$$

Since $c > 0$ is an arbitrary constant, we have $\lim_{p \rightarrow \infty} J(\mathbf{u}^{p,*}) = J(\mathbf{u}^*)$. This completes the proof.

THEOREM 5.2. Let $\mathbf{u}^{p,*}$ be an optimal control of Problem $(P(p))$. Let \mathbf{u}^* be an optimal control of the problem (P) . Suppose that

$$\lim_{p \rightarrow \infty} \mathbf{u}^{p,*} = \bar{\mathbf{u}}, \quad \text{a.e. on } [0, T].$$

Then, \bar{u} is an optimal control of the problem (P).

PROOF. Since $u^{p,*} \rightarrow \bar{u}$ a.e. in $[0, T]$, it follows from Lemma 4.4 of [21] that

$$\lim_{p \rightarrow \infty} J(u^{p,*}) = J(\bar{u}). \tag{37}$$

An argument similar to that in the proof of Lemma 3.2 of [19] shows that \bar{u} is a feasible control of problem (P). However, from Theorem 5.1, we get

$$\lim_{p \rightarrow \infty} J(u^{p,*}) = J(u^*). \tag{38}$$

Thus, the proof of this theorem follows easily from (37) and (38).

6. A Numerical Example

To demonstrate the applicability of the proposed approach for solving the constrained optimal control problem (P), we consider a problem of transferring containers from a ship to a cargo truck. The container crane is driven by a hoist motor and a trolley drive motor. The aim is to minimise the swing during and at the end of the transfer. Note that this problem was originally formulated in [12].

Without going into the details of the modelling aspect, we summarise the problem after appropriate normalisation as follows:

$$\text{minimise } \left\{ g_0 = 4.5 \int_0^1 [(x_3(t))^2 + (x_6(t))^2] dt \right\} \tag{39}$$

subject to the dynamical equations

$$\dot{x}_1(t) = 9x_4(t), \tag{40a}$$

$$\dot{x}_2(t) = 9x_5(t), \tag{40b}$$

$$\dot{x}_3(t) = 9x_6(t), \tag{40c}$$

$$\dot{x}_4(t) = 9(u_1(t) + 17.2656x_3(t)), \tag{40d}$$

$$\dot{x}_5(t) = 9u_2(t), \tag{40e}$$

$$\dot{x}_6(t) = -9[u_1(t) + 27.0756x_3(t) + 2x_5(t)x_6(t)]/x_2(t), \tag{40f}$$

where

$$x(0) = [0, 22, 0, 0, -1, 0]^T, \tag{41a}$$

$$x(1) = [10, 14, 0, 2.5, 0, 0]^T \tag{41b}$$

and

$$|u_1(t)| \leq 2.83374, \quad \forall t \in [0, 1], \quad (42a)$$

$$-0.80865 \leq u_2(t) \leq 0.71265, \quad \forall t \in [0, 1] \quad (42b)$$

with continuous state inequality constraints

$$|x_4(t)| \leq 2.5, \quad \forall t \in [0, 1], \quad (43a)$$

$$|x_5(t)| \leq 1.0, \quad \forall t \in [0, 1]. \quad (43b)$$

Define

$$h_1(t, \mathbf{x}(t)) \equiv x_4(t) + 2.5, \quad (44a)$$

$$h_2(t, \mathbf{x}(t)) \equiv 2.5 - x_4(t), \quad (44b)$$

$$h_3(t, \mathbf{x}(t)) \equiv x_5(t) + 1.0, \quad (44c)$$

and

$$h_4(t, \mathbf{x}(t)) \equiv 1.0 - x_5(t). \quad (44d)$$

Furthermore, let

$$\mathcal{L}_\varepsilon(t, \mathbf{x}(t)) \equiv \sum_{i=1}^4 \mathcal{L}_{i,\varepsilon}(t, \mathbf{x}(t)) \quad (45)$$

where for each $i = 1, \dots, 4$, $\mathcal{L}_{i,\varepsilon}(t, \mathbf{x}(t))$ is constructed from $h_i(t, \mathbf{x}(t))$ according to (9), and $h_i(t, \mathbf{x}(t))$, $i = 1, \dots, 4$, are defined by (44a)–(44d), respectively. The results are summarised in the following table.

TABLE 6.1. Numerical Results

ε	γ	$g_0(u)$	$\int_0^1 \mathcal{L}_\varepsilon dt$	$\int_0^1 \min \sigma\{h_i, 0\} dt$	reason for termination
10^{-3}	10^{-4}	0.5385×10^{-2}	0.49×10^{-2}	-0.41×10^{-8}	normal
10^{-4}	4×10^{-5}	0.5452×10^{-2}	-0.50×10^{-3}	0	normal
10^{-5}	4×10^{-6}	0.5381×10^{-2}	-0.50×10^{-4}	0	normal
10^{-6}	4×10^{-7}	0.5361×10^{-2}	-0.50×10^{-5}	0.10×10^{-12}	normal

As can be seen from Table 6.1, we experienced good convergence behaviour of the optimisation routine with the normal stopping condition. Note that the obtained numerical results are superior to those reported in [4] and [12].

Acknowledgement

This research is partially supported by a research grant from the Australian Research Council.

References

- [1] W. E. Bosarge Jr. and O. G. Johnson, "Direct method approximation to the state regulator control problem using at Ritz-Treffitz suboptimal control," *IEEE Trans. Auto. Control* Vol. AC-15 (1970) 627–631.
- [2] A. E. Bryson Jr. and Y. C. Ho, *Applied optimal control*, (Halsted Press, New York, 1969).
- [3] L. Cesari, *Optimization: theory and applications*, (Springer-Verlag, New York, 1983.)
- [4] C. J. Goh and K. L. Teo, "Control parametrization: a unified approach to optimal control problems with general constraints," *Automatica* **24** (1988) 3–18.
- [5] S. Gonzalez and A. Miele, "Sequential gradient-restoration algorithm for optimal control problems with general boundary conditions," *JOTA* **26** (1978) 395–425.
- [6] L. Hasdorff, *Gradient optimization and nonlinear control*, (John Wiley, 1976).
- [7] G. A. Hicks and W. H. Ray, "Approximation methods for optimal control systems," *Can. J. Chem. Eng.* **49** (1971) 522–528.
- [8] L. S. Jennings, M. E. Fisher, K. L. Teo and C. J. Goh, MISER3: Optimal control software, theory and user manual, EMCOSS, Perth, 1990.
- [9] A. Miele, "Recent advances in gradient algorithms for optimal control problems," *JOTA* **17** (1975) 361–430.
- [10] A. Miele and T. Wang, "Primal-dual properties of sequential gradient-restoration algorithms for optimal control problems, Part 1, Basic Problems," in F. R. Payne et al., (ed.), *Integral Methods in Science and Engineering*, (Hemisphere Publishing Corporation, Washington, DC, 1986) 577–607.
- [11] A. Miele and T. Wang, "Primal-dual properties of sequential gradient-restoration algorithms for optimal control problems, Part 2, general problem," *J. Math. Anal. Appl.* **119** (1986) 21–54.
- [12] Y. Sakawa and Y. Shindo, "Optimal control of container cranes," *Automatica* **18** (1982) 257–266.
- [13] K. Schittkowski, "NLPQL: a FORTRAN subroutine solving constrained nonlinear programming problems," *Operations Research Annals* **5** (1985) 485–500.
- [14] H. R. Sirisena, "Computation of optimal controls using a piecewise polynomial parametrization," *IEEE Transl. Auto. Control* AC-18 (1973) 409–411.
- [15] H. R. Sirisena and F. S. Chou, "An efficient algorithm for solving optimal control problems with linear terminal constraints," *IEEE Trans. Auto. Control* Vol. AC-21 (1976) 275–277.
- [16] H. R. Sirisena and F. S. Chou, "Convergence of control parameterization Ritz method for nonlinear optimal control problems," *JOTA* **19** (1979) 369–382.
- [17] H. R. Sirisena and K. S. Tan, "Computation of constrained optimal controls using parametrization techniques," *IEEE Trans. Auto. Control* AC-19 (1974) 431–433.
- [18] K. L. Teo and C. J. Goh, "A computational method for combined optimal parameter selection and optimal control problems with general constraints," *J. Math. Soc. Aust. Ser. B.* **30** (1989) 350–364.
- [19] K. L. Teo and L. S. Jennings, "Nonlinear optimal control problems with continuous state inequality constraints," *JOTA* **63** (1989) 1–22.
- [20] K. L. Teo, W. H. Wong and D. J. Clements, "Optimal control computation for linear time-lag systems with linear terminal constraints," *JOTA* **44** (1984) 509–526.
- [21] K. H. Wong, D. J. Clements and K. L. Teo, "Optimal control computation for nonlinear time-lag systems," *JOTA* **47** (1985) 91–107.