A DUAL METHOD FOR OPTIMAL CONTROL PROBLEMS WITH INITIAL AND FINAL BOUNDARY CONSTRAINTS

O. PIRONNEAU AND E. POLAK

Abstract. This paper presents two new algorithms belonging to the family of dual methods of centers. The first can be used for solving fixed time optimal control problems with inequality constraints on the initial and terminal states. The second one can be used for solving fixed time optimal control problems with inequality constraints on the initial and terminal states and with affine instantaneous inequality constraints on the control. Convergence is established for both algorithms. Qualitative reasoning indicates that the rate of convergence is linear.

1. Introduction. The construction of optimal control algorithms is often hampered by two difficulties. The first is due to the fact that the cost function usually has a gradient only in $L_\infty$, while the convergence of the algorithm must be studied in $L_2$, since control sequences constructed by an optimization algorithm are not likely to converge in $L_\infty$. The second difficulty stems from the fact that "primal-type" subproblems, such as those resulting from a direct application of methods of centers or feasible directions, cannot be solved directly and usually require some sort of "dualization." Both of these sources of difficulty are taken into account in the dual method presented in this paper.

The algorithm in this paper may be classified as a dual method of centers. It has the very nice feature that it is implementable essentially without recourse to heuristics, since both its direction finding and step size finding procedures are finite, in the sense that they require only a finite number of function evaluations per iteration. Since a closely related algorithm presented in [6] converges linearly on finite-dimensional problems, it is reasonably certain that the algorithm presented in this paper also converges linearly on problems in $\mathbb{R}^n$. However, since certain sets, used in the proofs in [6], lose their compactness in general Banach spaces, the proof of rate of convergence given in [6] cannot be extended to general Banach spaces. In spite of this, there are heuristic reasons which lead us to believe that the algorithms presented in this paper do converge linearly at least on a class of optimal control problems with linear dynamics and convex costs.

Basically, the algorithm in this paper is designed for problems with inequality end constraints and without constraints on the control. However, as we shall show, simple constraints on the controls can be handled by means of a minor modification of the algorithm.

The only other algorithms which solve the problems for which our algorithm was designed are based on the use of penalty functions. Since penalty function methods suffer considerably from "ridge paralysis" when the penalty becomes high, we expect our method to be superior in moderate to high precision situations.
2. Optimality and convergence. Because of the peculiar nature of optimal control problems, which necessitates the simultaneous use of both the $L_2$- and the $L_\infty$-norms on a space of regulated functions [3], we need the following abstract structure and accompanying theorems.

Let $V$ be a linear space, let $\| \cdot \|_1$ be a norm on $V$ and let $\langle \cdot , \cdot \rangle_2$ be a scalar product on $V$, such that $\mathcal{B}_1 = \{ V, \| \cdot \|_1 \}$ is a Banach space and $\mathcal{B}_2 = \{ V, \langle \cdot , \cdot \rangle_2 \}$ is a subspace of a Hilbert space. Let $\| \cdot \|_2$ be the norm induced by $\langle \cdot , \cdot \rangle_2$ on $\mathcal{B}_2$ (i.e., $\| z \|_2 = \langle z, z \rangle^{1/2}_2$).

**Assumption 2.1.** There exists a $C > 0$ such that $\| z \|_2 \leq C \| z \|_1$ for all $z \in V$.

Now consider the problem

$$\min \{ f^0(z) | f^j(z) \leq 0, j = 1, 2, \ldots, m \},$$

where $f^j : V \rightarrow \mathbb{R}^1$ for $j = 0, 1, 2, \ldots, m$.

**Assumptions 2.3.**

(i) The functions $f^j(\cdot)$, $j = 0, 1, 2, \ldots, m$, are Fréchet differentiable on $\mathcal{B}_1$, with the Fréchet derivative at $\tilde{z}$ being denoted by $f^j(\tilde{z})(\cdot)$, $j = 0, 1, 2, \ldots, m$.

(ii) The restrictions to $\{ z \in \mathcal{B}_2 | \| z \|_1 < M \}$ of the functions $f^j(\cdot)$, $j = 0, 1, 2, \ldots, m$, are continuous for any $M \in (0, \infty)$ (i.e., they are continuous in $\| \cdot \|_2$ on $\{ z \in V | \| z \|_1 < M \}$).

(iii) There exist functions $Vf^j : V \rightarrow V$, $j = 0, 1, 2, \ldots, m$, with the following properties: (a) the $Vf^j$ are continuous on $\mathcal{B}_1$, (b) the $Vf^j$ have continuous restrictions on $\{ z \in \mathcal{B}_2 | \| z \|_1 < M \}$ for any $M \in (0, \infty)$, (c) the $Vf^j$ satisfy

$$f^j(\tilde{z})(h) = \langle Vf^j(\tilde{z}), h \rangle_2, \quad j = 0, 1, 2, \ldots, m,$$

for any $\tilde{z}, h$ in $\mathcal{B}_1$.

**Theorem 2.5.** Suppose that $\tilde{z} \in \mathcal{B}_1$ is a solution of (2.1) (i.e., $f^j(\tilde{z}) \leq 0$ for $j = 1, 2, \ldots, m$, and $f^0(\tilde{z}) = \min \{ f^0(z) | f^j(z) \leq 0, j = 1, 2, \ldots, m \}$). Then there exist multipliers $\mu^0(\tilde{z}) \geq 0$, $\mu^1(\tilde{z}) \geq 0$, $\ldots$, $\mu^m(\tilde{z}) \geq 0$ such that

$$\sum_{j=0}^m \mu^j(\tilde{z}) Vf^j(\tilde{z}) = 0,$$

$$\mu^j(\tilde{z}) f^j(\tilde{z}) = 0 \quad \text{for} \quad j = 1, 2, \ldots, m,$$

and

$$\sum_{j=0}^m \mu^j(\tilde{z}) = 1.$$

Theorem 2.5 is a straightforward generalization of the well-known F. John condition of optimality [4]. It can be proved in essentially the same manner as the F. John condition (see the proof of Theorem 3.5.11 in [2]).

**Definition 2.9.** Let the set of feasible points $\Omega \subset \mathcal{B}_1$ be defined by

$$\Omega = \{ z \in \mathcal{B}_1 | f^j(z) \leq 0, j = 1, 2, \ldots, m \},$$

and let the set of desirable points $\Delta \subset \Omega$ be the set of points $\tilde{z} \in \Omega$ for which there exist multipliers $\mu^j(\tilde{z})$, $j = 0, 1, \ldots, m$, which satisfy (2.6)–(2.8).

Thus, $\Delta$ is the set of feasible points which satisfy the optimality condition (2.5). Since in general it is not possible to identify points in $\Omega$ which are optimal for (2.2), the best we can hope to achieve is to compute a desirable point.
The algorithm which we shall describe in the next section uses a map \( A : \Omega \to 2^\Omega \) and is of the following form.

**Algorithm Model 2.11.**

*Step 0.* Compute a \( z_0 \in \Omega \), and set \( i = 0 \).

*Step 1.* Compute a \( y \in A(z_i) \).

*Step 2.* If \( f_0(y) < f_0(z_i) \), set \( z_{i+1} = y \), set \( i = i + 1 \), and go to Step 1; else, set \( z = z_i \), and stop.

The convergence properties of our algorithm are summarized by the following result.

**Theorem 2.12.** Suppose that Assumption 2.3(ii) is satisfied, that for every \( M > 0 \), \( \Omega_M \supseteq \{ z \in \Omega \mid \| z \|_1 < M \} \), and that for every \( z \in \Omega \), \( z \notin \Delta \), there exist an \( \xi(z) > 0 \) and a \( \delta(z) < 0 \) such that for every \( M > \| z \|_1 \),

\[
(2.13) \quad f_0(z^\prime) - f_0(z) \leq \delta(z)
\]

for all \( z' \in \{ z' \in \Omega_M \mid \| z' - z \|_2 \leq \xi(z) \} \), for all \( z'' \in A(z') \).

Suppose that \( \{ z_i \} \) is a sequence generated by Algorithm model 2.11. If \( \{ z_i \} \) is finite, then its last element \( z \) is in \( \Delta \). If \( K \subset \{ 0, 1, 2, \ldots \} \) is an infinite subset and \( z^* \in \Omega \) is such that either (i) \( \lim_{i \in K} \| z_i - z^* \|_1 = 0 \) or (ii) \( \lim_{i \in K} \| z_i - z^* \|_2 = 0 \) and \( \| z_i \|_1 < M \) for some \( M > 0 \) and all \( i \in K \), then \( z^* \in \Delta \).

We omit a proof of this theorem since it follows directly from Theorem 1.3.10 in [7] and the Assumption 2.1.

With the preliminaries out of the way, we can now get down to the task of establishing a specific algorithm for finding points in the set \( \Delta \).

### 3. A dual method of centers.

For the algorithm below to make sense, we need the following additional hypothesis, as is usual in conjunction with methods of centers and methods of feasible directions (see §§ 4.2 and 4.3 in [7]).

**Assumption 3.1.** The set \( \check{\Omega} = \{ z \in \mathcal{B}_1 \mid f_j(z) < 0, j = 1, 2, \ldots, m \} \) is not empty.  \(^1\)

**Algorithm 3.2.** (\( \beta \in (0, 1) \) is a step size parameter.)

*Step 0.* Compute a \( z_0 \in \Omega \), and set \( i = 0 \).

*Step 1.* Compute \( \mu(z_i) = (\mu_1(z_i), \mu_2(z_i), \ldots, \mu_m(z_i))^{T} \in \mathbb{R}^{m+1} \) to be a solution of the quadratic programming problem

\[
(3.3) \quad \phi(z_i) = \max \left\{ \sum_{j=1}^{m} \mu_j f_j(z_i) - \frac{1}{2} \left\| \sum_{j=0}^{m} \mu_j \nabla f_j(z_i) \right\|_2^2 : \sum_{j=0}^{m} \mu_j = 1, \mu \geq 0 \right\}.
\]

*Step 2.* If \( \phi(z_i) = 0 \), set \( \bar{z} = z_i \), and stop; else, set

\[
(3.4) \quad h(z_i) = -\sum_{j=0}^{m} \mu_j(z_i) \nabla f_j(z_i)
\]

and go to Step 3.

*Step 3.* Compute the smallest nonnegative integer \( k(z_i) \) such that

\[
(3.5) \quad \theta(\beta^{k(z_i)}, h(z_i), z_i) - \frac{1}{2} \beta^{k(z_i)} \phi(z_i) \leq 0,
\]

\(^1\) When \( \check{\Omega} \) is empty, the algorithm below stops at \( z_0 \) and hence is useless.
where \( \theta: \mathbb{R}^1 \times \mathcal{B}_1 \times \mathcal{B}_1 \to \mathbb{R}^1 \) is defined by
\[
\theta(\lambda, h, z) = \max \{ f^0(z + \lambda h) - f^0(z); f^j(z + \lambda h), j = 1, 2, \ldots, m \}.
\]

Step 4. Set \( z_{i+1} = z_i + \beta^{k(z_i)}h(z_i) \), set \( i = i + 1 \), and go to Step 1.

The following result is obvious.

**Proposition 3.7.** Let \( \phi: \mathcal{B}_1 \to \mathbb{R}^1 \) be defined as in (3.3) and let \( z \in \bar{\Omega} \) be arbitrary. Then \( \phi(z) \leq 0 \), and \( \phi(z) = 0 \) if and only if \( z \in \Delta \).

**Lemma 3.8.** Suppose that \( z_i \in \bar{\Omega} \) is such that \( \phi(z_i) \neq 0 \), and let \( h(z_i) \) be defined as in (3.4). Then
\[
\max \{ \langle \nabla f^0(z_i), h(z_i) \rangle_2, f^j(z_i), j = 1, 2, \ldots, m \} \leq \phi(z_i) - (1/2)\|h(z_i)\|^2 < 0.
\]

This lemma follows directly from the fact that the dual of (3.3) is
\[
\phi(z) = \min \{ (1/2)\|h\|^2 + \max \{ \langle \nabla f^0(z_i), h(z_i) \rangle_2, f^j(z_i), j = 1, 2, \ldots, m \} \}.
\]

**Corollary 3.11.** Suppose that \( z_i \in \bar{\Omega} \) is such that \( \phi(z_i) < 0 \). Then there exists an integer \( k(z_i) \geq 0 \) such that (3.5) holds.

**Proof.** This corollary follows directly from the definition of a Fréchet differential, (2.4) and the fact that by (3.9), \( \langle \nabla f^0(z_i), h(z_i) \rangle_2 \leq \phi(z_i) \), and \( \langle \nabla f^j(z_i), h(z_i) \rangle \leq \phi(z_i) \) for all \( j = 1, 2, \ldots, m \) such that \( f^j(z_i) = 0 \).

**Theorem 3.12.** Let \( \{z_i\} \) be a sequence generated by Algorithm 3.2 in the process of searching the set \( \bar{\Omega} \) for a point in \( \Delta \) (see Definition 2.9), and suppose that Assumptions 2.1, 2.3 and 3.1 are satisfied. Then, either \( \{z_i\} \) is finite and its last point \( z \in \Delta \), or \( \{z_i\} \) is infinite, in which case any \( z^* \in \bar{\Omega} \) satisfying either \( \lim_{i \to \infty} \|z_i - z^*\|_2 = 0 \) or \( \lim_{i \to \infty} \|z_i - z^*\|_2 = 0 \), where \( K \) is an infinite subset of \( \{0, 1, 2, \ldots\} \), also satisfies \( z^* \in \Delta \), provided there exists an \( M \in (0, \infty) \) such that \( \|z_i\|_2 \leq M \) for all \( i \in K \).

**Proof.** Algorithm 3.2 is of the form of Algorithm 2.11, with \( A(\cdot) \) defined as follows. Let \( S: \bar{\Omega} \to 2^\mathbb{K} \) be defined by
\[
S(z) = \left\{ -\sum_{j=0}^m \mu^j\nabla f^j(z) | \mu \geq 0, \sum_{j=0}^m \mu^j = 1 ; \right\}
\]
\[
\sum_{j=1}^m \mu^j f^j(z) - \frac{1}{2} \left\| \sum_{j=0}^m \mu^j\nabla f^j(z) \right\|_2^2 = \phi(z).
\]

Then \( A: \bar{\Omega} \to 2^\mathbb{K} \) is given by
\[
A(z) = \{ z' = z + \beta^{k(z, h)}h | h \in S(z) \},
\]
where \( k(z, h) \) is the smallest nonnegative integer which satisfies (3.5) for \( z_i = z, h(z_i) = h \) and \( k(z_i) = k(z, h) \). (Since by (3.6) and (3.5), \( f^j(z') \leq (1/2)\beta^{k(z, h)}h(z') \leq 0 \) for \( j = 1, 2, \ldots, m \), it is clear that \( A(\cdot) \) maps \( \bar{\Omega} \) into \( 2^\mathbb{K} \).) To complete our proof, we only need to show that (2.13) is satisfied by the maps \( f^0(\cdot) \) and \( A(\cdot) \), as defined.
in (3.14). Since this is straightforward, we omit it.

4. An application to an optimal control problem. We shall now show that the algorithm, presented in the preceding section, can be used to solve the following problem:

\[ \min \left\{ \int_{t_0}^{t_f} h^0(x(t), u(t), t) \, dt \left| \begin{array}{c}
  x(t) = h(x(t), u(t), t), \\
  t \in [t_0, t_f], \\
  g_0(x(t_0)) \leq 0, \\
  g_f(x(t_f)) \leq 0; \\
  u \in L^2_{\infty}[t_0, t_f]
\end{array} \right. \} \]

(4.1)

where \( h^0 : \mathbb{R}^n \times \mathbb{R}^s \times [t_0, t_f] \to \mathbb{R}^1 \), \( h : \mathbb{R}^n \times \mathbb{R}^s \times [t_0, t_f] \to \mathbb{R}^n \), \( g_0 : \mathbb{R}^n \to \mathbb{R}^m_0 \), \( g_f : \mathbb{R}^n \to \mathbb{R}^m_f \), and \( L^2_{\infty}[t_0, t_f] \) is the space of equivalence classes of essentially bounded integrable functions from \([t_0, t_f]\) into \( \mathbb{R}^n \).

We must begin by transcribing problem (4.1) into the form of problem (2.2). Therefore, let \( V = \{(\xi, u) | \xi \in \mathbb{R}^n, u \in L^2_{\infty}[t_0, t_f]\} \), let the norm \( \| \cdot \|_1 : V \to \mathbb{R}^1 \) be defined by

\[ \| (\xi, u) \|_1^2 = |\xi|^2 + \text{ess sup}_{t \in [t_0, t_f]} |u(t)|^2, \]

(4.2)

where \( | \cdot | \) denotes the Euclidean norm, and finally, let the scalar product \( \langle \cdot, \cdot \rangle_2 \) on \( V \) be defined by

\[ \langle (\xi, u), (\xi', u') \rangle_2 = \langle \xi, \xi' \rangle + \int_{t_0}^{t_f} \langle u(t), u'(t) \rangle \, dt, \]

(4.3)

where \( \langle \cdot, \cdot \rangle \) denotes the Euclidean scalar product. Then we see that the space \( \mathcal{B}_1 = \{ V, \| \cdot \|_1 \} \) is a Banach space and the space \( \mathcal{B}_2 = \{ V, \langle \cdot, \cdot \rangle_2 \} \) is a subspace of a Hilbert space. Furthermore, setting \( \| \cdot \|_2 = \sqrt{\langle \cdot, \cdot \rangle_2} \), it is not difficult to show that there exists a \( C \in (0, \infty) \) such that \( \| \cdot \|_2 \leq C \| \cdot \|_1 \). Next, let \( x(t, \xi, u), t \in [t_0, t_f] \), denote the solution of the differential equation

\[ \frac{d}{dt} x = h(x, u, t), \quad x(t_0) = \xi, \quad t \in [t_0, t_f], \]

(4.4)

corresponding to a given \((\xi, u) \in V\). Then we define the functions \( f^0 : V \to \mathbb{R}^1 \), \( f_1 : V \to \mathbb{R}^{m_1} \) and \( f_2 : \mathbb{R}^{m_2} \) as follows:

\[ f^0(\xi, u) = \int_{t_0}^{t_f} h^0(x(t, \xi, u), u(t), t) \, dt, \]

(4.5)

\[ f_1(\xi, u) = g_0(\xi), \]

(4.6)

\[ f_2(\xi, u) = g_f(x(t_f, \xi, u)). \]

(4.7)

With the above definitions, problem (4.1) can be written as follows, setting \( z = (\xi, u) \):

\[ \min \{ f^0(z), f_1(z) \leq 0, f_2(z) \leq 0 \}, \]

(4.8)

i.e., it can be written in the form (2.2).
Assumptions 4.9.

(i) For every $(\xi, u) \in V$, the solution $x(\cdot, \xi, u)$ of (4.4) exists and is unique.

(ii) The functions $h^0$ and $h$ are continuously differentiable in $x$ and in $u$, and $h^0, h, \partial h^0/\partial x, \partial h/\partial u, \partial h/\partial x, \partial h/\partial u$ are piecewise continuous in $t$.

(iii) The functions $g_0$ and $g_1$ are continuously differentiable.

(iv) The set $\{z = (\xi, u) \in V | f_i(z) < 0, f_j(z) < 0\}$ is not empty.

Lemma 4.10. Suppose that Assumptions 4.9(i)–4.9(iii) are satisfied. Then the functions $f^0, f_1$ and $f_2$, defined in (4.5)–(4.7), are Fréchet differentiable on $\mathcal{B}$, with their differentials $f^0, f_1, f_2$, defined as follows:

\begin{align}
 f^0_i(z')(h) &= \langle \nabla f^0(z'), h \rangle_2, \\
 f^i_k(z')(h) &= \langle \nabla f^i(z'), h \rangle_2, \quad i = 1, 2, \ldots, m_k, \quad k = 1, 2, \\
 \text{where, for } z' = (\xi', u'), \\
 \nabla f^0(z') &= (-q_0(t_0, \xi', u'), -\frac{\partial h^0}{\partial u}(x(\cdot, \xi', u'), u'(-)), \cdot)^T \\
 \cdot q_0(\cdot, \xi', u') + \frac{\partial h^0}{\partial u}(x(\cdot, \xi', u'), u'(-))^T, \\
 \nabla f^i(z') &= \begin{pmatrix} \frac{\partial g_i^0}{\partial x}(\xi') \cdot 0 \end{pmatrix}, \quad i = 1, 2, \ldots, m_1, \\
 \text{and} \\
 \nabla f^i(z') &= (-q_i(t_0, \xi', u'), -\frac{\partial h}{\partial u}(x(\cdot, \xi', u'), u'(-))^T q_i(\cdot, \xi', u')), \\
 i = 1, 2, \ldots, m_2, \\
 \text{with } q_i(t, \xi', u'), i = 0, 1, 2, \ldots, m_2 \text{ defined (with } \delta_{i0} = 0 \text{ if } i \neq 0 \text{) by} \\
 \frac{d}{dt} q_i(t, \xi', u') &= -\left[ \frac{\partial h}{\partial x}(x(t, \xi', u'), u'(t), t) \right]^T q_i(t, \xi', u') \\
 + \delta_{i0} \left[ \frac{\partial h^0}{\partial x}(x(t, \xi', u'), u'(t), t) \right]^T, \\
 t \in [t_0, t_f], \quad q_i(t_f, \xi', u') = 0, \\
 q_i(t_f, \xi', u') &= -\left[ \frac{\partial g_i^0}{\partial x}(x(t, \xi', u')) \right]^T.
\end{align}

Corollary 4.17. For any $M > 0$, the functions $f^0, f_1, f_2, \nabla f^1_i, i = 1, 2, \ldots, m_1,$ and $\nabla f^2_i, i = 1, 2, \ldots, m_2,$ have continuous restrictions on $\{z \in \mathcal{B} \land \|z\|_1 < M\}$.

This lemma and the corollary follow directly from Theorem 10.7.1 in [3] and from Theorem A1 in [5]. We therefore omit their proofs.

Thus we see that the functions $f^0, f_1$ and $f_2$ satisfy the Assumptions 2.3. We now show that the set $\Delta$ defined in Definition 2.9, with $f^j = f^1_j$, for $j = 1, 2, \ldots, m_1,$
and \( f^{i+m_1} = f^i_{\frac{1}{2}} \), for \( j = 1, 2, \ldots, m_2 \), is the set of initial states and controls for which the Pontryagin-maximum-principle-in-differential-form\(^2\) is satisfied.

**Proposition 4.18.** Let \( \Omega \) and \( \Delta \) be defined as in Definition 2.9, with \( f^j = f^i_{\frac{1}{2}} \), for \( j = 1, 2, \ldots, m_1 \), and \( f^{i+m_1} = f^i_{\frac{1}{2}} \), for \( j = 1, 2, \ldots, m_2 \), and let \( m = m_1 + m_2 \). If \((\xi, \tilde{u}) \in \Delta\), then there exists a multiplier function \( \lambda : [t_0, t_f] \to \mathbb{R}^n \), and a scalar \( \lambda_0 \leq 0 \) such that

\[
\frac{d}{dt} \lambda(t) = -\frac{\partial h}{\partial x}(x(t, \xi, \tilde{u}), \tilde{u}(t), t)^T \lambda(t)
\]

\[
+ \lambda_0 \frac{\partial h^0}{\partial x}(x(t, \xi, \tilde{u}), \tilde{u}(t), t)^T,
\]

\( t \in [t_0, t_f] \),

\[
\lambda(t_0) = \frac{\partial g_0(\tilde{u})}{\partial x} \tilde{v}_0,
\]

\[
\lambda(t_f) = \frac{\partial g_f(x(t_f, \xi, \tilde{u}))}{\partial x} \tilde{v}_f,
\]

where \( \tilde{v}_0 \geq 0 \), \( \tilde{v}_f \geq 0 \) are such that \((\tilde{v}_0, \tilde{v}_f) \neq 0 \),

\[
\langle \tilde{v}_0, g_0(\tilde{u}) \rangle = \langle \tilde{v}_f, g_f(x(t_f, \xi, \tilde{u})) \rangle = 0,
\]

and

\[
\frac{\partial}{\partial \tilde{u}}\bigg[ \lambda_0 h^0(x(t, \xi, \tilde{u}), \tilde{u}(t), t) + \langle \lambda(t), h(x(t, \xi, \tilde{u}), \tilde{u}(t), t) \rangle \bigg] = 0.
\]

(Since \((\xi, \tilde{u}) \in \Delta \subseteq \Omega\), we must have \( g_0(\xi) \leq 0 \) and \( g_f(x(t_f, \xi, \tilde{u})) \leq 0 \) by definition.)

For the problem (4.1), Algorithm 3.2 assumes the following specific form.

**Algorithm 4.24.** (Solves (4.1); \( \beta \in (0, 1) \) is a step size parameter.)

**Step 0.** Compute \((\xi_0, u_0) \in \mathcal{B}_1\) such that \( g_0(x_0) \leq 0 \) and \( g_f(x(t_f, \xi_0, u_0)) \leq 0 \), and set \( i = 0 \).

**Comment.** The Algorithm 4.24 can be used to compute such an \((\xi_0, u_0(\cdot))\) by solving the problem

\[
\text{min} \int_{t_0}^{t_f} x^0(t) dt \quad \text{subject to} \quad \frac{d}{dt} x = h(x, u, t);
\]

\[
g_j(x(t_o)) - x^0(t_0) \leq 0, \quad j = 1, 2, \ldots, m_1;
\]

\[
g_j(x(t_f)) - x^0(t_f) \leq 0, \quad j = 1, 2, \ldots, m_2,
\]

where \( x = (x^0, x) \) and \( h = (0, h) \), and for which an initial point \((\xi_0, \tilde{u}_0(\cdot))\), \( \xi_0 = (x^0_0, \xi_0) \), can be chosen as follows: let \( \xi_0, \tilde{u}_0(\cdot) \) be arbitrary, and let

\[
\xi_0 = \max \{ g_0(\xi_0), j = 1, 2, \ldots, m_1; g_f(x(t_f, \xi_0, u)), j = 1, 2, \ldots, m_2 \}.
\]

Since the optimal value of (4.25) is strictly negative, a \((\xi_0, u_0(\cdot))\) for Step 0 above can be computed by means of a finite number of iterations.

\(^2\) In the maximum principle in differential form, the condition of maximum on the Hamiltonian is replaced by the condition \((\partial H(x, \dot{u}, \dot{v}, t)/\partial u)\dot{u} \leq 0\) for all admissible \( \dot{u} \) and for almost all \( t \in [t_0, t_f] \).
Step 1. For $z_i = (\xi_i, u_i)$, compute $Vf^0(z_i), \nabla f^0_j(z_i), j = 1, 2, \ldots, m_1$, $\nabla f^1_j(z_i), j = 1, 2, \ldots, m_2$, according to (4.13)-(4.16).

Step 2. Compute $\mu^0(z_i), \mu^j_1(z_i), j = 1, 2, \ldots, m_1$, $\mu^j_2(z_i), j = 1, 2, \ldots, m_2$, as a solution of

$$
\phi(z_i) = \max \left\{ \sum_{j=1}^{m_1} \mu^j_1 g_0^j(\xi_i) + \sum_{j=1}^{m_2} \mu^j_2 g^j(z(t_f, \xi_i, u_i)) \right\}
$$

$$
- \frac{1}{2} \left\| \mu^0 \nabla f^0(z_i) + \sum_{j=1}^{m_1} \mu^j_1 \nabla f^1_j(z_i) + \sum_{j=1}^{m_2} \mu^j_2 \nabla f^1_j(z_i) \right\|_2^2
$$

$$
\mu^0 + \sum_{j=1}^{m_1} \mu^j_1 + \sum_{j=1}^{m_2} \mu^j_2 = 1, \mu^0 \geq 0, \mu^j_1 \geq 0, j = 1, 2, \ldots, m_1,
$$

$$
\mu^j_2 \geq 0, j = 1, 2, \ldots, m_2\right\},
$$

where $\|z\|_2 = \langle z, z \rangle_2$ is defined as in (4.3).

Step 3. If $\phi(z_i) = 0$, set $\xi = \xi_i, \hat{u}(\cdot) = u_i(\cdot)$ and stop; else, go to Step 4 (see (4.13)-(4.16)).

Step 4. Set

$$
\omega_i = \mu^0(z_i)q_0(t_0, \xi_i, u_i) - \sum_{j=1}^{m_1} \mu^j_1(z_i) \frac{\partial g_0^j(\xi_i)}{\partial x} + \sum_{j=1}^{m_2} \mu^j_2(z_i)q_0(t_0, \xi_i, u_i),
$$

$$
v_1(\cdot) = \mu^0(z_i) \left[ \frac{\partial h^0}{\partial u}(x(\cdot, \xi_i, u_i), u(\cdot, \cdot))^T q_0(\cdot, \xi_i, u_i) - \frac{\partial h^0}{\partial u}(x(\cdot, \xi_i, u_i), u(\cdot, \cdot))^T q_0(\cdot, \xi_i, u_i) \right]
$$

$$
+ \sum_{j=1}^{m_2} \mu^j_2(z_i) \frac{\partial h^0}{\partial u}(x(\cdot, \xi_i, u_i), u(\cdot, \cdot))^T q_0(\cdot, \xi_i, u_i),
$$

and go to Step 5.

Step 5. Compute the smallest integer $k$, such that

$$
\max \left\{ \int_{t_0}^{t_f} \left[ h^0(x(t, \xi_i + \beta^k \omega_i, u_i + \beta^k v_i), u_i + \beta^k v_i) - h^0(x(t, \xi_i, u_i), u_i, t) \right] \, dt \right\}
$$

$$
g_0^j(\xi_i + \beta^k \omega_i), j = 1, 2, \ldots, m_1; g_1^j(x(t_f, \xi_i + \beta^k \omega_i, u_i + \beta^k v_i)),
$$

$$
j = 1, 2, \ldots, m_2 \right\} - (\beta^k/2)\phi(z_i) \leq 0.
$$

Step 6. Set $\xi_{i+1} = \xi_i + \beta^k \omega_i$, set $u_{i+1}(\cdot) = u_i(\cdot) + \beta^k v_i(\cdot)$, set $i = i + 1$, and go to Step 1.

The following result is obvious.

**Proposition 4.30.** Theorem 3.12 holds for Algorithm 4.24, with the set $\Delta$ defined as the set of feasible points $(\xi, \hat{u}) \in \mathcal{B}_1$ satisfying the Pontryagin-maximum-principle-in-differential-form for problem (4.1).
5. An extension to problems with instantaneous constraints on the control.

We shall now show that Valentine-type transformations [9] can be used to adapt Proposition 4.30 for the solution of the following optimal control problem:

\[
\min \left\{ \int_{t_0}^{t_f} h^0(x(t), u(t), t) \, dt \mid \frac{d}{dt} x(t) = h(x(t), u(t), t), \right. \\
t \in [t_0, t_f]; \quad g_0(x(t_0)) \leq 0, \quad g_f(x(t_f)) \leq 0; \quad u \in L^2_{o}(t_0, t_f); \\
\left. b^k \leq \langle a_k, u(t) \rangle \leq c^k, \quad k = 1, 2, \cdots, r, \quad \text{for all } t \in [t_0, t_f] \right\},
\]

where \( h^0, h, g_0, g_f \) are as in (4.1); \( a_k \in \mathbb{R}^s \) for \( k = 1, 2, \cdots, r; \ c_k \in \mathbb{R}^1 \) for \( k = 1, 2, \cdots, r, \ b_k \in \mathbb{R}^1 \) for \( k = 1, 2, \cdots, r', \ r' \leq r, \) and \( b_k = -\infty \) for \( k = r' + 1, \cdots, r. \)

Assumptions 5.2.

(i) We shall assume that (4.9) is satisfied.
(ii) The vectors \( a_k, k = 1, 2, \cdots, r, \) are linearly independent.
(iii) There exists a control \( \bar{u} \in L^2_{o}(t_0, t_f) \) and an initial state \( \xi \in \mathbb{R}^n \) such that \( g_0(\xi) \leq 0, \ g_f(x(t_f), \xi, u) \leq 0, \) and \( b^k < \langle a_k, \bar{u}(t) \rangle < c^k \) for \( k = 1, 2, \cdots, r \) and all \( t \in [t_0, t_f]. \)

To apply the Valentine trick, we must use certain substitutions for the inequalities on the control. Thus, consider the constraints

\[
(5.3a) \quad b^k \leq \langle a_k, u(t) \rangle \leq c^k, \quad k = 1, 2, \cdots, r', \quad t \in [t_0, t_f], \\
(5.3b) \quad \langle a_k, u(t) \rangle \leq c^k, \quad k = r' + 1, \cdots, r, \quad t \in [t_0, t_f].
\]

Suppose that \( u \in L^2_{o}(t_0, t_f) \) satisfies (5.2) and (5.3). Then we can associate with this \( u \) functions \( v^k : [t_0, t_f] \to \mathbb{R}^1, \ k = 1, 2, \cdots, r', \) and \( w^k : [t_0, t_f] \to \mathbb{R}^1, \ k = 1, 2, \cdots, r - r', \) such that

\[
(5.4) \quad \cos v^k(t) = \frac{2}{c^k - b^k} \langle a_k, u(t) \rangle - \frac{c^k + b^k}{c^k - b^k}, \quad k = 1, 2, \cdots, r', \quad t \in [t_0, t_f], \\
(5.5) \quad (w^k(t))^2 = c^{k+r'} - \langle a_{k+r'}, u(t) \rangle, \quad k = 1, 2, \cdots, r - r', \quad t \in [t_0, t_f].
\]

We shall now use these functions to construct a problem which is equivalent to (5.1). Let \( A^T \) be the \( s \times r \) matrix whose columns are \((2/(c^k - b^k))a_k, k = 1, 2, \cdots, r', \) and \(-a_k, \ k = r' + 1, r' + 2, \cdots, r. \) Then its transpose, \( A, \) is an \( r \times s \) matrix which can be partitioned as follows: \( A = [A^r, A^r] \) (rearranging the components of \( u(\cdot) \), if necessary), where \( A^r \) is an \( r \times r \) nonsingular matrix. We partition \( u \) similarly, i.e., we set \( u = (u', u^r), \) with \( u' \in \mathbb{R}^r \) and \( u^r \in \mathbb{R}^{r'-r}. \) Then, if \( u, \ v^k, w^k \) satisfy (5.4) and (5.5), we obtain

\[
u''(t) = A''^{-1} \left[ (\cos v^1(t), \cdots, \cos v^{r'}(t), \cos^1(t)^2, \cdots, \cos^{r'-r'}(t)^2)^T \\
+ \left( \frac{c^1 + b^1}{c^1 - b^1}, \cdots, \frac{c^{r'} + b^{r'}}{c^{r'} - b^{r'}}, \ - c^r u, \cdots, - c^{r'} \right)^T \right] \\
- A''^{-1} A' u'(t) \triangleq u''(u'(t), v(t), w(t)).
\]
Let \( \tilde{u} = (u', v, w) \), let
\[
\tilde{h}^0(x, \tilde{u}, t) = h^0(x, (u', u'^{2}(u', v, w)), t),
\]
and consider the problem
\[
\min \left\{ \int_{t_0}^{t_f} \tilde{h}^0(x, \tilde{u}, t) \, dt \mid \dot{x} = h(x, \tilde{u}, t), \; t \in [t_0, t_f] ; \right. \\
g_0(x(t_0)) \leq 0; \; g_f(x(t_f)) \leq 0; \; \tilde{u} \in L^\infty_{x} [t_0, t_f] \left. \right\},
\tag{5.7}
\]
where all the quantities are as in (5.1) and (5.6). It is trivial to show that if \( (\xi, u', \phi, \tilde{w}) \) is any optimal solution of (5.7), \( (\xi, (u', \cos \phi, \tilde{w}^2)) \) is also optimal for problem (5.1). However, not all the points \( (\xi, u', \phi, \tilde{w}) \) which satisfy the Pontryagin principle for problem (5.7) result in a pair \( (\xi, (u', \cos \phi, \tilde{w}^2)) \) which satisfy the maximum principle for problem (5.1). Hence, although Algorithm 4.24 is directly applicable to problem (5.7), it is desirable to modify it so as to prevent convergence to points which do not satisfy the optimality conditions for problem (5.1). This can be done by modifying the step length rule in Step 5 of Algorithm 4.24. To explain how this is done, without loss of generality, we assume that there is only one constraint of the form (5.3a) and only one constraint of the form (5.3b), i.e., we assume that (5.3a) and (5.3b) have the following specific form:
\[
-1 \leq u^{r-1}(t) \leq 1, \quad u^0(t) \geq 0, \quad t \in [t_0, t_f],
\tag{5.8}
\]
and that the remaining components of \( u(t) \) are unconstrained. For this specific case, given \( z_i \), Algorithm 4.24 computes the following feasible direction \( (\omega_i, v_i(\cdot)) \):
\[
\omega_i = \mu^0(z_i) \rho_i(t_0) - \sum_{k=1}^{m_1} \mu_k^1(z_i) \frac{\partial g_0}{\partial x}(\xi_i) + \sum_{l=1}^{m_2} \mu_l^2(z_i) q_{i,l}(t_0),
\tag{5.9}
\]
\[
v_i^T = (v_i^T, v_i^{r-1}, v_i^r)
\tag{5.10}
\]
with
\[
v_i^r(\cdot) = \mu^0(z_i) \left[ \frac{\partial h}{\partial u}(x_i(\cdot), u_i(\cdot), \cdot)^T p_i(\cdot) - \frac{\partial h^0}{\partial u}(x_i(\cdot), u_i(\cdot), \cdot)^T \right]
\]
\[
+ \sum_{l=1}^{m_2} \mu_l^2(z_i) \frac{\partial h}{\partial u}(x_i(\cdot), u_i(\cdot), \cdot)^T q_{i,l}(\cdot),
\]
\[
v_i^{r-1}(\cdot) = -\sin \phi_i(\cdot) \left\{ \mu^0(z_i) \left[ \frac{\partial h}{\partial u^{r-1}}(x_i(\cdot), u_i(\cdot), \cdot)^T q_{i,0}(\cdot) \right.ight.
\]
\[
- \frac{\partial h^0}{\partial u^{r-1}}(x_i(\cdot), u_i(\cdot), \cdot)^T \right] + \sum_{l=1}^{m_2} \mu_l^2(z_i) \frac{\partial h}{\partial u^{r-1}}(x_i(\cdot), u_i(\cdot), \cdot)^T q_{i,0}(\cdot),
\]
\[
v_i^s(\cdot) = 2w_i(\cdot) \left\{ \mu^0(z_i) \left[ \frac{\partial h}{\partial u}(x_i(\cdot), u_i(\cdot), \cdot)^T p_i(\cdot) \right.ight.
\]
\[
- \frac{\partial h^0}{\partial u}(x_i(\cdot), u_i(\cdot), \cdot)^T \right] + \sum_{l=1}^{m_2} \mu_l^2(z_i) \frac{\partial h}{\partial u}(x_i(\cdot), u_i(\cdot), \cdot)^T q_{i,l}(\cdot) \right\},
\]
where the $q_{l,t}(\cdot)$, $l = 0, 1, 2, \cdots, m_2$, are computed by solving (4.16) for $\xi' = \xi_i$, $x' = u_i$ and $h, h^0$ as in (5.1). The new step length rule can be stated as a substitute Step 5, which, for the problem in hand, becomes:

**Step 5’**: Compute the smallest integer $k_i$ satisfying

$$
\max \left\{ \int_0^t \left[ h^0(x(t), \xi_i + \beta^k v_i, (u_i + \beta^k v_i, \cos (v_i + \beta^k v_i^{-1}), (w_i + \beta^k v_i^2)) \right] dt, \right.
$$

$$
\left. g_0(\xi_i + \beta^k \omega_i), g_j(x(t), \xi_i + \beta^k \omega_i),
\right.
$$

$$
\left. (u_i + \beta^k v_i, \cos (v_i + \beta^k v_i^{-1}), (w_i + \beta^k v_i^2)) \right\} - (\beta^k/2) \phi(z_i) \leq 0,
$$

(5.12)

**Lemma 5.13.** The conclusions of Theorem 3.12 remain true for the modified Algorithm 4.24 with respect to problem (5.7).

**Proof.** The modified algorithm differs from Algorithm 4.24 only in the additional bound (5.12) on $\beta^k$. Now, since this bound, in turn, has a denominator which can be bounded from above by a continuous function of $(\xi_i, u_i, v_i, w_i)$, i.e., since

$$
\max \left\{ \frac{v_i^{-1}(t)}{\sin v_i(t)}, - \frac{v_i(t)}{w_i(t)} \right\}
$$

$$
\leq \sum_{k = s-1}^s \left\{ \left| q_{0,t}(\cdot)^T \frac{\partial h}{\partial u^k}(x(t), u(t), t) - \frac{\partial h^0}{\partial u^k}(x(t), u(t), t) \right| + \sum_{l=1}^{m_2} q_{l,t}(\cdot)^T \frac{\partial h}{\partial u^k}(x(t), u(t), t) \right\},
$$

(5.14)

it follows from arguments essentially duplicating the proof of Theorem 3.12 that the conditions of Theorem 2.12 are satisfied.

**Lemma 5.15.** Let $\{(\xi_i, u_i, v_i, w_i)\}_{i=0}^\infty$ be a sequence generated by Algorithm 4.24, modified to use Step 5’ instead of Step 5 in the process of solving problem (5.7). Suppose $K$ is an infinite subset of the positive integers such that $\lim_{i \to K} \|(\xi_i, u_i, v_i, w_i) - (\xi, u, \vartheta, \hat{w})\|_2 = 0$ and $\sup_{i \in K} \|(\xi_i, u_i, v_i, w_i)\|_1 < \infty$. Furthermore, let $K'$ be an infinite subset of $K$, such that $\lim_{i \to K'} \{\mu^0(\hat{z}_i), \mu_1(\hat{z}_i), \mu_2(\hat{z}_i)\} = \{\mu^0, \mu_1, \mu_2\}$, where $\hat{z}_i = (\xi_i, u_i, v_i, w_i)$ and $\mu^0(\hat{z}_i), \mu_1(z)$ and $\mu_2(z)$ are defined as in Step 2 of Algorithm 4.24 (for problem (5.7)). Then,

$$
\left\{ \frac{\partial h^0}{\partial u^k}(\xi(t), \hat{u}(t), t) - \frac{\partial h}{\partial u^k}(\xi(t), \hat{u}(t), t)^T \right\}
$$

$$
\cdot (\mu^0 \hat{q}(t) + \sum_{i=1}^{m_2} \mu_2 \hat{q}_i(t)) \right\} (-1)^k \leq 0,
$$

(5.16a)

---

The form of (5.10) is due to the fact that we have used the chain rule to calculate the needed partial derivatives.
for almost all $t \in \{t|\theta(t) = k\pi\}, k = 0, 1, \text{ and}$

$$\frac{\partial h^0}{\partial u^*}(\tilde{x}(t), \tilde{u}(t), t) - \frac{\partial h}{\partial u^*}(\tilde{x}(t), \tilde{u}(t), t)^T \left( \beta_0 \tilde{q}_0(t) + \sum_{l=1}^{m_2} \beta_{l}^2 \tilde{q}_l(t) \right) \geq 0$$

for almost all $t \in \{t|\hat{\theta}(t) = 0\}$, where $\hat{u} = (\hat{u}', \cos \hat{v}, \hat{v}^2)$, \( \hat{x}(t) \equiv x(t, \hat{\xi}, \hat{\xi}), \) and \( \hat{q}_l, l = 0, 1, \ldots, m_2, \) are defined by (4.16) for $u'(t) \equiv \hat{u}(t), x(t, \xi', u') \equiv \hat{x}(t), \xi' = \hat{\xi}.$

Proof. Let $H: \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ be defined by

$$H(x, u, \psi, t) = -h^0(x, u, t) + \langle \psi, h(x, u, t) \rangle.$$

Then, from (5.10) and the instructions in Step 6 of Algorithm 4.24, we find that for $t \in [t_0, t_f]$ and with

$$\psi_i(t) = \mu^0(z_i)q_{0,i}(t) + \sum_{l=1}^{m_2} \mu^1(z_i)q_{i,l}(t), \quad t \in [t_0, t_f],$$

we must have

$$v_{i+1}(t) = v_i(t) - \beta^k_i \sin v_i(t) \frac{\partial H}{\partial u^{-1}}(x_i(t), u_i(t), \psi_i(t), t),$$

$$w_{i+1}(t) = w_i(t) + 2\beta^k_i w_i(t) \frac{\partial H}{\partial u}(x_i(t), u_i(t), \psi_i(t), t).$$

Now, the purpose of the bound (5.12) on $\beta^k_i$ was to ensure that for almost all $t \in [t_0, t_f]$,

$$0 < \frac{1}{2}v_i(t) \leq v_{i+1}(t) \leq \frac{1}{2}(\pi + v_i(t)) < \pi,$$

$$\frac{1}{2}w_i(t) \leq w_{i+1}(t).$$

Since we must have $\lim_{i \in K'} v_i(t) = \theta(t)$ for almost all $t \in [t_0, t_f]$, suppose that $t' \in [t_0, t_f]$ is such that $\lim_{i \in K'} v_i(t') = \theta(t') = k\pi$, with $k = 0$ or 1. It now follows from (5.18), since $\beta^k_i < 1$, and since $(\partial H/\partial u^{-1})(x_i(t'), u_i(t'), \psi(t'), t')$ is bounded for $i \in K'$, that we must also have

$$\lim_{i \in K'} v_{i+1}(t') = k\pi,$$

for almost all $t'$ such that $\lim_{i \in K'} v_i(t') = k\pi.$

By construction, $v_0(t) \in (0, \pi)$ for all $t \in [t_0, t_f]$, and hence, $\sin v_0(t') > 0$. Let $K''$ be an infinite subsequence of the positive integers defined by $K'' = K' \cup \{i(i-1) \in K'\}$. It now follows from (5.21) and (5.23) that $\{(1)^{k+1}(k\pi - v_i(t))\}_{i \in K''}$ is a strictly positive sequence which converges to zero, and hence there must exist an infinite subsequence $K''' \subset K'$ such that

$$\{(1)^{k+1}(k\pi - v_{i+1}(t')) < (1)^{k+1}(k\pi - v_i(t')) \quad \text{for all} \quad i \in K'''.$$

Note that $\{i|\lim_{i \in K'} v_{i+1}(t') \neq \lim_{i \in K'} v_i(t)'\}$ is a null set.
Combining (5.24) with (5.18), and recalling that \( \sin \nu(t') > 0 \) for all \( i \), we conclude that

\[
(-1)^k \frac{\partial H}{\partial u^{(k-1)}}(x_i(t'), u_i(t'), \psi_i(t'), t) \geq 0 \quad \text{for all } i \in K^{m}.
\]

It now follows from the continuity of \( \frac{\partial H}{\partial u^{(k-1)}} \) that

\[
(-1)^k \frac{\partial H}{\partial u^{(k-1)}}(\hat{x}(t'), \hat{u}(t'), \hat{\psi}(t'), t') \geq 0,
\]

where \( \hat{\psi}(t') = \hat{\mu}^0 \hat{p}(t') + \sum_{i=1}^{m^2} \hat{\mu}_i \hat{q}_i(t') \). This establishes (5.16a); (5.16b) can be established in a similar way.

The following result is now obvious.

**Corollary 5.27.** Let \( \{(\xi_i, u_i)\} \) be a sequence generated by the modified Algorithm 4.24 in solving problem (5.1) by means of (5.7). Then, either \( \{(\xi_i, u_i)\} \) is finite and its last element satisfies the Pontryagin-maximum-principle-in-differential-form, or \( \{(\xi_i, u_i)\} \) is infinite and every pair of points \( (\xi_i, u_i) \) which satisfies for some \( K \subset \{0, 1, 2, \ldots\} \) either (i) \( \lim_{i \to K} \| \xi_i - \xi \|_1 = 0 \), or (ii) \( \lim_{i \to K} \| \xi_i - \xi \|_1 < \infty \), also satisfies the Pontryagin-maximum-principle-in-differential-form.

### 6. Experimental results.

To illustrate the behavior of our algorithm, we have applied it to two problems. The first was,

\[
\text{min} \int_0^2 \left( \left\| x(t) - \left[ \frac{1}{2} (x_f - x_0) + x_0 \right] \right\|^2 + \| u \|^2 \right) dt,
\]

where

\[
x_0 = (10, 10, 10, 10)^T, \quad x_f = (2, 2, 2, 2)^T,
\]

with \( x(0) \in \mathbb{R}^4, u(t) \in \mathbb{R}^2 \) and

\[
\begin{pmatrix}
\dot{x}_1 \\
\dot{x}_2 \\
\dot{x}_3 \\
\dot{x}_4
\end{pmatrix} =
\begin{pmatrix}
-1 & 0 & 0 & 0 \\
0 & -2 & 0 & 0 \\
0 & 0 & -3 & 0 \\
0 & 0 & 0 & -4
\end{pmatrix}
\begin{pmatrix}
x_1 \\
x_2 \\
x_3 \\
x_4
\end{pmatrix} +
\begin{pmatrix}
5 & 0 \\
0 & 5 \\
10 & 0 \\
0 & 10
\end{pmatrix}
\begin{pmatrix}
u_1 \\
u_2
\end{pmatrix}
\]

and with the constraints

\[
\| x(0) - x_0 \|_2 \leq 1, \quad \| x(2) - x_f \|_2 \leq 1.
\]

Then, we added the constraint

\[
| u^1(t) | \leq 1, \quad | u^2(t) | \leq 1, \quad t \in [0, 2],
\]

to obtain the second problem.

The program we used was designed for general situations and did not exploit the structure of (6.1) and (6.2). The integration was performed using the Euler-Cauchy method, with an initial step size of 1/32. The algorithm was programmed to increase the integration precision on demand, according to the scheme outlined in [8]. This adaptive integration scheme was used so as to reduce computer
Fig. 1. Solution of the problems (6.1)-(6.3) and (6.1)-(6.4)

Fig. 2. Solution of the problem (6.1)-(6.3). The 5th iterate of initialization cycle is feasible and is used as first iterate of minimization cycle.
The computations were carried out on a CDC 6400 computer and were stopped after 40 iterations. (The first problem required about 20% less time than the second.) In each case, the program required fewer than 5 iterations to compute an initial feasible solution \((x_0, u(\cdot))\) and \((\xi_0, \bar{u}(\cdot))\), respectively. The results of the computation are shown graphically in the accompanying figures.

7. Conclusion. In the form stated, Algorithm 4.24 is readily implementable either by using fixed precision integration or adaptive integration schemes such as those described in [8]. Our experience has been that the adaptive precision schemes are invariably considerably faster than fixed precision schemes. Although we have not had the opportunity to confirm the following conjecture by experiment, we feel reasonably sure that our algorithm will outperform its rivals, the various penalty methods, at least in the case when the constraints are such as to cause “ridge paralysis” in the penalty methods.

In modifying Algorithm 4.24 in § 5, a substitution formula (Valentine’s trick) was used to extend Algorithm 4.24 to problems with affine instantaneous inequality constraints and, in addition, a perturbation method was added to ensure that convergence was possible only to points satisfying both a first and a second order optimality condition for the derived problem (5.7). This eliminated points which satisfy a first order condition for (5.7) but not for (5.1). Industrial experience with the substitution formula, used in conjunction with simpler algorithms, indicates that it performs quite well; our own experimental results confirm this view.

Acknowledgment. We wish to thank Dr. R. Klessig for his helpful comments, Mr. H. Mukai for preparing the computer code and performing a number of experiments for us, and Dr. H. J. Kelly for sharing with us his experience in the use of the Valentine transformation.
REFERENCES