INFORMATION TO USERS

This dissertation was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or “target” for pages apparently lacking from the document photographed is “Missing Page(s)”. If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in “sectioning” the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from “photographs” if essential to the understanding of the dissertation. Silver prints of “photographs” may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

University Microfilms
300 North Zeeb Road
Ann Arbor, Michigan 48106
A Xerox Education Company
NAQVI, Sarwar, 1942-
EXTREMIZATION OF TERMINALLY CONSTRAINED
CONTROL PROBLEMS.

Rice University, Ph.D., 1972
Engineering, mechanical

University Microfilms, A XEROX Company, Ann Arbor, Michigan
RICE UNIVERSITY

EXTREMINIZATION OF TERMINALLY
CONSTRAINED CONTROL PROBLEMS

by

SARWAR NAQVI

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Thesis Director's Signature:

Ho-yei Chang

Houston, Texas

May, 1972
PLEASE NOTE:

Some pages may have indistinct print.
Filmed as received.

University Microfilms, A Xerox Education Company
ABSTRACT

Extremization of Terminally
Constrained Control Problems

by

SARWAR NAQVI

In this thesis, the problem of extremizing a functional I involving the state \( x(t) \), the control \( u(t) \), and the parameter \( \pi \) is considered. The admissible state, control, and parameter are required to satisfy a vector differential constraint, a vector initial constraint, and a vector terminal constraint.

This problem is transformed into a mathematically simpler, unconstrained problem of minimizing a new functional, the performance index \( R \), which involves the state, the control, the parameter, the Lagrange multiplier \( \lambda(t) \) associated with the vector differential constraint, and the Langrange multiplier \( \mu \) associated with the vector terminal constraint. To obtain the minimum \( R = 0 \) of the performance index, a gradient algorithm is first developed. In order to achieve simplicity in programming and to bypass the explicit use of the second-order derivatives, the gradient algorithm is modified so that it becomes a pure, first-order method. For better convergence property, a conjugate-gradient algorithm is also developed.

Concerning the determination of the stepsize in these algorithms, a one-cycle cubic interpolation scheme is presented. Again, the explicit use of the second-order derivatives is avoided here.
Both the gradient algorithm and the conjugate-gradient algorithm are tested through several numerical examples. The results show that, while the gradient algorithm is slow in convergence, the conjugate-gradient algorithm displays a much better convergence characteristic.
ACKNOWLEDGEMENT

The author is indebted to Dr. H.Y. Huang for suggesting the topic and helpful discussions. He would like also to thank his wife, Patricia, for her patience and understanding during the time this thesis progressed. Last, but certainly not least, he would like to thank Mrs. Sandi Peppeard for the excellent typing of the manuscript.
To my sisters

Suraiya, Zarina, Parveen,

Nasreen, Yasmeen, Zeeba.
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>2. Statement of Problem</td>
<td>2</td>
</tr>
<tr>
<td>3. Gradient Algorithm</td>
<td>7</td>
</tr>
<tr>
<td>4. Modified Gradient Algorithm</td>
<td>15</td>
</tr>
<tr>
<td>5. Conjugate-Gradient Algorithm</td>
<td>21</td>
</tr>
<tr>
<td>6. One-Dimensional Search</td>
<td>27</td>
</tr>
<tr>
<td>7. Experimental Conditions</td>
<td>30</td>
</tr>
<tr>
<td>8. Numerical Examples</td>
<td>32</td>
</tr>
<tr>
<td>9. Computational Results</td>
<td>34</td>
</tr>
<tr>
<td>10. Conclusions</td>
<td>43</td>
</tr>
<tr>
<td>References</td>
<td>44</td>
</tr>
</tbody>
</table>
1. **Introduction**

Over the past several years, considerable work has been done on the development of gradient methods for optimal control problems with terminal constraints. Depending on the approaches to enforce the satisfaction of the terminal constraints, the gradient methods developed so far can be classified mainly into two groups: (a) methods which employ the penalty function and (b) methods which employ the restoration phase.

The approach (a) was adopted, for instance, in Refs. 1-2. This approach has the advantage of replacing the constrained optimal problem by a mathematically simpler, unconstrained optimal problem. But, the approach also has the disadvantages that no clear-cut method exists for the choice of the penalty constants and that the algorithm must be repeated several times for increasing values of the penalty constants.

The approach (b) was adopted, for instance, in Refs. 3-4. This approach has the advantages that the algorithm possesses better convergence property and that the constraints are satisfied after every restoration phase so that the values of the functional between iterations are comparable. But, this approach has the disadvantage that a two-point boundary-value problem must be solved at every iteration.

It is the purpose of this thesis to develop a technique which, on the one hand, possesses the simplicity of the penalty function approach and, on the other hand, bypasses the difficulty associated with the choice of the penalty constants. For better convergence property, a conjugate-gradient version of the new technique is also developed.
2. **Statement of Problem**

In this thesis, we consider the problem of extremizing the functional

\[ I = \int_{0}^{1} f(x, u, \pi, t) dt + g(x_1, \pi) \]

with respect to the functions \( x(t) \), \( u(t) \), and the parameter \( \pi \) which satisfy the differential equation

\[ \dot{x} - \varphi(x, u, \pi, t) = 0 \]

the initial condition

\[ x_0 - k = 0 \]

and the final condition

\[ \psi(x_1, \pi) = 0 \]

In the above equations, the functions \( f \) and \( g \) are scalar, the function \( \varphi \) is an \( n \)-vector, and the function \( \psi \) is a \( q \)-vector. The symbol \( x(t) \), an \( n \)-vector, denotes the state variable; the symbol \( u(t) \), an \( m \)-vector, denotes the control variable; and the symbol \( \pi \), a \( p \)-vector, denotes the parameter. The time \( t \), a scalar, is the independent variable; without loss of generality, the prescribed initial time is \( t = 0 \) and the prescribed final time is \( t = 1 \). The subscripts 0, 1 denote quantities at \( t = 0 \) and \( t = 1 \), respectively. Finally, the constant \( k \), an \( n \)-vector, denotes the prescribed initial value of \( x \).

Thus, at the initial point, all the components of the state variable are given. At the final point, \( q \) scalar equations are specified, where \( 0 \leq q \leq n+p \).
Problems where the final time is other than unity can be reduced to the form (1)-(4) by normalizing the time with respect to the final time and by regarding the final time, if it is free, as one of the components of the parameter \( \pi \).

2.1. **First-Order Conditions.** From calculus of variations (see, for instance, Ref. 3), it is known that the previous problem is one of the Bolza type. It can be recast as that of extremizing the **augmented functional**

\[
J = \int_0^1 F dt + G \tag{5}
\]

subject to (2)-(4). In the above expression, the functions \( F \) and \( G \) are given by

\[
F = f + \lambda^T (\dot{x} - \phi) , \quad G = g + \mu^T \psi \tag{6}
\]

where \( \lambda(t) \), an \( n \)-vector, is a variable Lagrange multiplier and \( \mu \), a \( q \)-vector, is a constant Lagrange multiplier. The superscript \( T \) denotes the transpose of a matrix.

The optimum solutions \( x(t), u(t), \pi \) must satisfy (2)-(4), the Euler equations

\[
\begin{align*}
\frac{d}{dt}F_x - F_x &= 0 , \\
F_u &= 0 , \\
\int_0^1 F \pi dt + G \pi &= 0
\end{align*} \tag{7}
\]

and the following natural condition arising from the transversality condition:

\[
(F_x)_1 \gamma + G_x \gamma = 0 \tag{8}
\]

If one introduces the Hamiltonian

\[
H = f - \lambda^T \phi \tag{9}
\]
Eqs. (7)-(8) can be rewritten as

\[ \dot{x} - H_x = 0, \quad H_u = 0, \quad \int_0^1 H_{\pi} \, dt + G_{\pi} = 0 \]  

(10)

and

\[ \lambda_1 + G_{\pi x_1} = 0 \]

(11)

Summarizing, we seek functions \( x(t), u(t), \lambda(t), \pi, \mu \) which satisfy the constraints (2)-(4) and the optimality conditions (10)-(11).

2.2. **Performance Index.** Within the context of this thesis, let the norm of a vector \( v \) be defined as

\[ N(v) = v^T v \]

(12)

We introduce a measure of the error involved in the constraints (2)-(4) defined by the functional

\[ P = \int_0^1 N(\dot{x} - \omega) \, dt + N(x_0 - k) + N(\psi) \]

(13)

and a measure of the error involved in the optimality conditions (10)-(11) defined by the functional

\[ Q = \int_0^1 N(\dot{\lambda} - H_x) \, dt + \int_0^1 N(H_u) \, dt \]

\[ + N\left( \int_0^1 H_{\pi} \, dt + G_{\pi} \right) + N(\lambda_1 + G_{\pi x_1}) \]

(14)

Therefore, the cumulative error in the entire system, or the performance index \( R \), can be defined as the sum of the error in the constraints and that
in the optimality conditions, that is,

\[ R = P + Q \] (15)

We observe that the performance index \( R \) defined by Eqs. (13)-(15) has the following properties: (a) it is equal to zero for the set of functions \( x(t), u(t), \lambda(t), \pi, \mu \) satisfying the system (2)-(4), (10)-(11) and (b) it is positive for any set of functions \( x(t), u(t), \lambda(t), \pi, \mu \) not consistent with the system (2)-(4), (10)-(11). Consequently, the performance index has an absolute minimum \( R = 0 \) for the functions \( x(t), u(t), \lambda(t), \pi, \mu \) satisfying the system (2)-(4), (10)-(11). This being the case, the problem of solving the system (2)-(4), (10)-(11) for \( x(t), u(t), \lambda(t), \pi, \mu \) can be replaced by that of minimizing the performance index \( R \) with respect to the functions \( x(t), u(t), \lambda(t), \pi, \mu \).

2.3. **Minimization of Performance Index.** The problem of minimizing the performance index \( R \) with respect to the functions \( x(t), u(t), \lambda(t), \pi, \mu \) can be treated in different ways. Here, from the viewpoint of simplicity in computation, a particular approach is adopted.

We observe that, if the functions \( u(t), \pi, \mu \) are prescribed, the function \( x(t) \) can be obtained from the system

\[ x_o - k = 0 \; , \; \dot{x} - \varphi = 0 \] (16)

by forward integration and the function \( \lambda(t) \) can be obtained from the system

\[ \lambda_1 + G_{x_1} = 0 \; , \; \dot{\lambda} - H_x = 0 \] (17)
by backward integration. Then, the error in the constraints (13) is reduced to

$$P = N(\psi)$$  \hspace{1cm} (18)$$

the error in the optimality conditions (14) is reduced to

$$Q = \int_0^1 N(H_u) dt + N\left(\int_0^1 H \frac{dt}{\pi} + G \frac{1}{\pi}\right)$$  \hspace{1cm} (19)$$

and the performance index is given by

$$R = P + Q$$  \hspace{1cm} (20)$$

with $P, Q$ defined by (18)-(19), respectively.

With this understanding, the problem of minimizing the performance index $R$ is formulated as follows: Find the functions $u(t)$, $\pi$, $\mu$, the function $x(t)$ satisfying the system (16), and the function $\lambda(t)$ satisfying the system (17) which minimize the performance index $R$ defined by (18)-(20).
3. **Gradient Algorithm**

In accordance with the formulation stated in Section 2.3, the gradient algorithm for reducing the performance index iteratively is presented in the following.

3.1. **Nominal and Varied Functions.** Let \( u(t), \pi, \mu \) denote the nominal functions and \( x(t), \lambda(t) \) the corresponding functions obtained from Eqs. (16)-(17). Also, let \( \bar{u}(t), \bar{\pi}, \bar{\mu} \) denote the varied functions and \( \bar{x}(t), \bar{\lambda}(t) \) the corresponding functions obtained from Eqs. (16)-(17). The two systems are related by

\[
\bar{u} = u + \Delta u, \quad \bar{\pi} = \pi + \Delta \pi, \quad \bar{\mu} = \mu + \Delta \mu
\]

\[
\bar{x} = x + \Delta x, \quad \bar{\lambda} = \lambda + \Delta \lambda
\]

(21)

where \( \Delta u(t), \Delta \pi, \Delta \mu, \Delta x(t), \Delta \lambda(t) \) are the variations leading from the nominal functions \( u(t), \pi, \mu, x(t), \lambda(t) \) to the varied functions \( \bar{u}(t), \bar{\pi}, \bar{\mu}, \bar{x}(t), \bar{\lambda}(t) \), respectively. If these variations are sufficiently small, they satisfy approximately the linearized systems

\[
\Delta x = 0, \quad \Delta \lambda - \varphi_T \Delta x - \varphi_u \Delta u - \varphi_{\pi T} \Delta \pi = 0
\]

(22)

and

\[
\Delta \lambda = G_{x1}^T \Delta x + G_{\pi1}^T \Delta \pi + G_{\mu1}^T \Delta \mu = 0
\]

\[
\Delta \lambda - H_{x1}^T \Delta x - H_{\lambda x} \Delta \lambda - H_{xu} \Delta u - H_{\pi x} \Delta \pi = 0
\]

(23)

Eqs. (22) are obtained from Eqs. (16) and Eqs. (23) from Eqs. (17). In Eqs. (22)-(23), all the partial derivatives are evaluated at the nominal.

---

1 For the case of quadratic functional-linear constraint, Eqs. (22)-(23) become exact.
3.2. First Variation of Performance Index. Due to the variations
\( \Delta u(t), \Delta \pi, \Delta \mu, \Delta x(t), \Delta \lambda(t) \), the performance index defined by Eqs. (18)-(20)
has a first-order change \( \delta R \) given by

\[
\delta R = 2 \left\{ \psi^T \left( \psi_{\Delta x}^T \Delta x_1 + \psi_{\Delta \pi}^T \right) \\
+ \int_0^1 H_{\Delta x}^T H_{\Delta x} + H_{\Delta \lambda}^T H_{\Delta \lambda} + H_{\Delta u}^T H_{\Delta u} + H_{\Delta \pi}^T H_{\Delta \pi} \right) dt \\
+ \Gamma^T \left\{ \left( H_{\Delta x}^T \Delta x + H_{\Delta \lambda}^T \Delta \lambda + H_{\Delta u}^T \Delta u + H_{\Delta \pi}^T \Delta \pi \right) dt \\
+ G_{\Delta x}^T \Delta x_1 + G_{\Delta \pi}^T \Delta \pi_1 + G_{\Delta \lambda}^T \Delta \lambda_1 \right\}
\]
(24)

where the symbol \( \Gamma \) is a p-vector defined by

\[
\Gamma = \int_0^1 H_{\Delta x}^T dt + G_{\Delta \pi}
\]
(25)

Again, the function \( \psi \) and all the partial derivatives in Eqs. (24)-(25) are
evaluated at the nominal.

Since the variations \( \Delta u(t), \Delta \pi, \Delta \mu, \Delta x(t), \Delta \lambda(t) \) satisfy the systems
(22)-(23), the following relation also holds:

\[
2 \left\{ \int_0^1 \left( X_{\Delta x}^T \Delta x - \varphi_{\Delta x}^T \Delta x - \varphi_{\Delta u}^T \Delta u - \varphi_{\Delta \pi}^T \Delta \pi \right) dt \\
+ \sigma^T (\Delta \lambda_1 + G_{\Delta x}^T \Delta x_1 + G_{\Delta \pi}^T \Delta \pi_1 + G_{\Delta \lambda}^T \Delta \lambda_1) \\
+ \int_0^1 \lambda^T (\Delta \lambda - H_{\Delta x}^T \Delta x - H_{\Delta \lambda}^T \Delta \lambda - H_{\Delta u}^T \Delta u - H_{\Delta \pi}^T \Delta \pi) dt \right\} = 0
\]
(26)

where the symbols \( X(t), \sigma, \lambda(t) \) are arbitrary n-vector functions. After
integrating by parts the terms involving \( \Delta \dot{x}, \Delta \dot{\lambda} \) and using Eq. (22-1), the relation
(26) can be rewritten in the form

\[
2\left( X_1^T \Delta x_1 - \int_0^1 X^T \Delta x_1 dt - \int_0^1 X^T (\varphi_x \Delta x + \varphi_u \Delta u + \varphi_\pi \Delta \pi) dt \right) \\
+ \sigma^T (\Delta \lambda_1 + G_{x_1}^T \Delta x_1 + G_{\pi_1}^T \Delta \pi + G_{\mu}^T \Delta \mu) \\
+ \Lambda_1^T \Delta \lambda_1 - \Lambda_0^T \Delta \lambda_0 - \int_0^1 \Lambda^T \Delta \lambda dt \\
- \int_0^1 \Lambda^T (H_{xx}^T \Delta x + H_{x\lambda}^T \Delta \lambda + H_{xu}^T \Delta u + H_{x\pi}^T \Delta \pi) dt \right) = 0
\]

(27)

Because of this relationship, the left-hand side of Eq. (27) can be added to the right-hand side of Eq. (24) without changing the value of the first variation of the performance index. If this operation is performed, the first variation \( \delta R \), after rearrangement, is given by

\[
\delta R = 2 \left\{ \int_0^1 \left( -H_{ux}^T \Delta x + H_{u\lambda}^T \Delta \lambda + H_{uu}^T \Delta u + H_{u\pi}^T \Delta \pi \right) T^T du dt \\
+ \left[ \int_0^1 \left( -H_{x\pi}^T \Delta x + H_{\pi u}^T \Delta u + H_{\pi \pi}^T \Delta \pi \right) \right] T^T du dt \\
+ G_{x_1}^T \sigma + G_{\pi_1}^T \Gamma + G_{\mu}^T \psi \right\} \Delta \pi \\
+ (\psi_{x_1}^T \sigma + \psi_{\pi}^T \Gamma) T^T \Delta \mu \\
+ (X_1 + G_{x_1}^T \sigma + G_{x_1}^T \Gamma + G_{x_1}^T \psi) T^T \Delta x_1 \\
+ \left( -X - H_{xx}^T \Delta x + H_{x\lambda}^T \Delta \lambda + H_{xu}^T \Delta u + H_{x\pi}^T \Delta \pi \right) T^T dx dt \\
- \Lambda_0^T \Delta \lambda_0 + (\sigma + \Lambda_1) T^T \Delta \lambda_1 \\
+ \left[ \int_0^1 (-\Lambda + \varphi_x^T \Delta x - \varphi_u^T \Delta u - \varphi_\pi^T \Delta \pi) T^T \Delta \lambda dt \right] \right\}
\]

(28)
For Eq. (28), the following relationships have been employed:

\[ H_{\lambda x} = -\varphi_x, \quad H_{\lambda u} = -\varphi_u, \quad H_{\lambda \pi} = -\varphi_{\pi} \]  

(29)

Since the first variation \( \delta R \) given by Eq. (28) is the same for all combinations of functions \( X(t), M(t), \sigma \), one can choose these functions in a convenient way. Here, we define the function \( \Lambda(t) \) by the system

\[ \Lambda_0 = 0, \quad \dot{\Lambda} - \varphi_x T_{\lambda} \Lambda - \varphi_u T_{u} \Lambda - \varphi_{\pi} T_{\pi} \Lambda = 0 \]  

(30)

the vector \( \sigma \) by the equation

\[ \sigma = -\Lambda_1 \]  

(31)

and the function \( X(t) \) by the system

\[ X_1 + G_{x1}^T \sigma + G_{x1}^T \Lambda + G_{x1}^T \psi = 0 \]  

\[ \dot{X} + H_{xx}^T \Lambda - H_{xu}^T X - H_{xu}^T H + H_{x\pi}^T \Lambda = 0 \]  

(32)

Then, the first variation (28) is simplified to

\[ \delta R = 2 \left[ \int_0^1 A^T \Delta u dt + B^T \Delta \pi + C^T \Delta \mu \right] \]  

(33)

where \( A(t), B, C \) are given by

\[ A = -H_{ux}^T \Lambda + H_{u\lambda}^T X + H_{uu}^T H + H_{u\pi}^T \Lambda \]  

(34-1)

\[ B = \int_0^1 (-H_{\pi\lambda}^T \Lambda + H_{\pi\lambda}^T X + H_{\pi\mu}^T H + H_{\pi\pi}^T \Lambda) dt \]  

\[ + G_{\pi\lambda}^T \sigma + G_{\pi\pi}^T \Lambda + G_{\pi\mu}^T \psi \]  

(34-2)
3.3. Variations. Let the variations $\Delta u(t)$, $\Delta \pi$, $\Delta \mu$, $\Delta x(t)$, $\Delta \lambda(t)$, which lead from the nominal to the varied functions, be expressed in the form

$$\Delta u = -\alpha A, \quad \Delta \pi = -\alpha B, \quad \Delta \mu = -\alpha C$$

$$\Delta x = -\alpha D, \quad \Delta \lambda = -\alpha E$$

where the symbol $\alpha$ is a positive scaling factor, or stepsize. Since these variations must satisfy the system (22)-(23), it follows immediately that the functions $A(t)$, $B$, $C$, $D(t)$, $E(t)$ must satisfy

$$\dot{D}_o = 0, \quad D - \phi_x^T D - \phi_u^T A - \phi_\pi^T B = 0$$

and

$$E_1 + G_{x1}^T D_1 + G_{x1}^T B + G_{x1}^T C = 0$$

$$\dot{E} - H_{xx}^T D - H_{x\lambda}^T E - H_{xu}^T A - H_{x\pi}^T B = 0$$

With $A(t)$, $B$, $C$ given by Eqs. (34), the function $D(t)$ can be determined from the system (36) by forward integration and the function $E(t)$ can be determined from the system (37) by backward integration.

3.4. Descent Property. With variations given by Eqs. (35), the first variation of the performance index (33) reduces to

$$\delta R = -2\alpha \left[ \int_0^1 A^T A dt + B^T B + C^T C \right]$$

(38)
This expression is of fundamental importance. It states that the first variation of the performance index is always negative, that is,

$$\delta R < 0$$  \hspace{1cm} (39)

Therefore, if the stepsize $\alpha$ is sufficiently small, the performance index can be reduced.

3.5. **Gradient Algorithm.** Summarizing the previous analysis, the gradient algorithm for minimizing the performance index is given as follows.

(a) For a given nominal $u(t)$, $\pi$, $\mu$, determine $x(t)$ from the system

$$x_0 = k, \quad \dot{x} = \varphi$$  \hspace{1cm} (40)

by forward integration and $\lambda(t)$ from the system

$$\lambda_1 = -G x_1, \quad \dot{\lambda} = H x$$  \hspace{1cm} (41)

by backward integration. With $u(t)$, $\pi$, $\mu$, $x(t)$, $\lambda(t)$ known, calculate $H_u(t)$, $\Gamma$, $\psi$ and evaluate the performance index with the expression

$$R = \int_0^1 H_u^T H_u \, dt + \Gamma^T \Gamma + \psi^T \psi$$  \hspace{1cm} (42)

If $R$ satisfies the inequality

$$R < \varepsilon_1$$  \hspace{1cm} (43)

where $\varepsilon_1$ is a preselected small positive quantity, the algorithm is terminated
and the solution considered achieved. If \( R \) violates Ineq. (43), proceed to Step (b).

(b) Determine \( \Lambda(t) \) from the system

\[
\Lambda_0 = 0 , \quad \dot{\Lambda} = \varphi_x^T \Lambda - \varphi_u^T H_u - \varphi_{\pi}^T \Gamma
\]  

(44)

by forward integration and \( X(t) \) from the system

\[
X_1 = -(-G_{x_1 x_1}^T \Lambda + G_{x_1 \pi}^T \Gamma + G_{x_1 u}^T \psi)
\]

\[
\cdot X = -H_{xx}^T \Lambda + H_{x\lambda}^T X + H_{xu}^T H_u + H_{x\pi}^T \Gamma
\]  

(45)

by backward integration.

(c) With \( H_u(t), \Pi, \Psi, \Lambda(t), X(t) \) known, determine \( A(t), B, C \) from the relations

\[
A = -H_{ux}^T \Lambda + H_{u\lambda}^T X + H_{uu}^T H_u + H_{u\pi}^T \Gamma
\]

\[
B = \int_0^1 (-H_{\pi x}^T \Lambda + H_{\pi\lambda}^T X + H_{\pi u}^T H_u + H_{\pi\pi}^T \Gamma) dt
\]

\[
- G_{\pi x}^T \Lambda_1 + G_{\pi\pi}^T \Gamma + G_{\pi u}^T \psi
\]

\[
C = -\psi_{x_1}^T \Lambda_1 + \psi_{\pi}^T \Gamma
\]

(46)

determine \( D(t) \) from the system

\[
D_0 = 0 , \quad \dot{D} = \varphi_x^T D + \varphi_u^T A + \varphi_{\pi}^T B
\]  

(47)
by forward integration and determine \( E(t) \) from the system

\[
E_1 = -(G_{x_1}^T \bar{x}_{1,1} + G_{\pi}^T \bar{\pi} + G_{\mu}^T \bar{\mu})
\]

\[
\dot{E} = H_{x_1}^T D + H_{x_\lambda}^T E + H_{x_u}^T A + H_{x_\mu}^T B
\]

by backward integration.

(d) Define a one-parameter family of varied functions by

\[
\bar{u} = u - \alpha A, \quad \bar{\pi} = \pi - \alpha B, \quad \bar{\mu} = \mu - \alpha C
\]

\[
\bar{x} = x - \alpha D, \quad \bar{\lambda} = \lambda - \alpha E
\]

Then, the performance index at the varied functions becomes a function of the parameter \( \alpha \) only, that is,

\[
\bar{R} = \bar{R}(\alpha)
\]

Perform a one-dimensional search on \( \bar{R}(\alpha) \) so that the optimum stepsize \( \alpha \) minimizing \( \bar{R}(\alpha) \) is determined. In this respect, a particular search technique is presented in Section 6.

(e) With the optimum \( \alpha \) known, determine the function \( \bar{u}(t), \bar{\pi}, \bar{\mu} \)

from the relations

\[
\bar{u} = u - \alpha A, \quad \bar{\pi} = \pi - \alpha B, \quad \bar{\mu} = \mu - \alpha C
\]

Return to Step (a) and start a new iteration with \( \bar{u}(t), \bar{\pi}, \bar{\mu} \) as the nominal.
4. Modified Gradient Algorithm

The gradient algorithm presented in the previous section has the following characteristics: (a) it employs two differential systems, namely, the original system (40)-(41) and the linearized systems (44)-(45) and (47)-(48); and (b) it employs explicitly not only the functions $f$, $g$, $\varphi$, $\psi$ and their first derivatives but also their second derivatives. From the viewpoint of simplicity in programming and time-saving in computation, these two characteristics are unfavorable. Both simplicity in programming and time-saving in computation can be achieved should one use only one differential system and employ explicitly only the functions $f$, $g$, $\varphi$, $\psi$ and their first derivatives.

4.1. Reduction to One Differential System. Let $u(t)$, $\pi$, $\mu$, $x(t)$, $\lambda(t)$ be a set of functions which satisfies the original differential systems (40)-(41). Also, let $\overline{u}(t)$, $\overline{\pi}$, $\overline{\mu}$, $\overline{x}(t)$, $\overline{\lambda}(t)$ be another set of functions which satisfies the same differential systems. The two sets are related by

\begin{align}
\overline{u} &= u + du \\
\overline{\pi} &= \pi + d\pi \\
\overline{\mu} &= \mu + d\mu \\
\overline{x} &= x + dx \\
\overline{\lambda} &= \lambda + d\lambda
\end{align}

(52)

where $du(t)$, $d\pi$, $d\mu$, $dx(t)$, $d\lambda(t)$ are the variations leading from the set $u(t)$, $\pi$, $\mu$, $x(t)$, $\lambda(t)$ to the set $\overline{u}(t)$, $\overline{\pi}$, $\overline{\mu}$, $\overline{x}(t)$, $\overline{\lambda}(t)$, respectively. If these variations are sufficiently small, they satisfy approximately the linearized systems

\begin{align}
dx_0 &= 0 \\
dx &= \phi_x^T dx + \phi_u^T du + \phi_{\pi}^T d\pi
\end{align}

(53)
and

\[
\frac{d\lambda}{\tau} = -(G^T x_1 x_1^T dx_1 + G^T x_1 \pi d\pi + G^T x_1 \mu d\mu) \tag{54}
\]

\[
\frac{d\lambda}{\tau} = H^T x \lambda dx + H^T x \lambda d\lambda + H^T x \mu du + H^T x \pi d\pi
\]

Now, if one chooses

\[
du = -\beta u, \quad d\pi = -\beta \pi, \quad d\mu = -\beta \mu \tag{55}
\]

\[
dx = \beta \lambda, \quad d\lambda = -\beta X
\]

where \(\beta\) is a prescribed small quantity, the systems (53)-(54) are reduced to the systems (44)-(45). Therefore, instead of determining \(\lambda(t), X(t)\), from the linearized systems (44)-(45), one can calculate them in the following way: Employing \(\bar{u}(t), \bar{\pi}, \bar{\mu}\) defined by Eqs. (52), (55), obtain \(\bar{x}(t)\) from the system (40) by forward integration and \(\bar{\lambda}(t)\) from the system (41) by backward integration. Once \(\bar{x}(t), \bar{\lambda}(t)\) are known, \(X(t) \lambda(t)\) are determined by

\[
X = (\lambda - \bar{\lambda})/\beta, \quad \lambda = -(x - \bar{x})/\beta \tag{56}
\]

On the other hand, if one chooses

\[
du = -\beta A, \quad d\pi = -\beta B, \quad d\mu = -\beta C \tag{57}
\]

\[
dx = -\beta D, \quad d\lambda = -\beta E
\]

the systems (53)-(54) are reduced to the systems (47)-(48). Therefore, instead of determining \(D(t), E(t)\) from the linearized systems (47)-(48), one can calculate them in the following way: Employing \(\bar{u}(t), \bar{\pi}, \bar{\mu}\) defined
by Eqs. (52), (57), obtain $\bar{x}(t)$ from the system (40) by forward integration and $\bar{\lambda}(t)$ from the system (41) by backward integration. Once $\bar{x}(t)$, $\bar{\lambda}(t)$ are known, $D(t)$, $E(t)$ are determined by

$$D = (x - \bar{x})/\beta, \quad E = (\lambda - \bar{\lambda})/\beta$$

(58)

4.2. **Approximation for $A(t)$, $B$, $C$**. With the linearized systems (44)-(45) and (47)-(48) bypassed, the explicit use of second derivatives appears only in Eqs. (46) for the determination of $A(t)$, $B$, $C$. To avoid this explicit use of second derivatives, one observes that, for sufficiently small variations $d\bar{u}(t)$, $d\bar{\pi}$, $d\bar{\mu}$, $dx(t)$, $d\lambda(t)$, the following first-order approximations hold:

$$\bar{H}_u - H_u = H_u^T dx + H_{u\lambda}^T d\lambda + H_{uu}^T du + H_{u\pi}^T d\pi$$

$$\bar{\Gamma} - \Gamma = \int_0^1 (H_{\pi x}^T dx + H_{\pi \lambda}^T d\lambda + H_{\pi u}^T du + H_{\pi \pi}^T d\pi)dt$$

$$+ G_{\pi x 1}^T dx + G_{\pi \pi}^T d\pi + G_{\pi \mu}^T d\mu$$

(59)

$$\bar{\psi} - \psi = \psi_{x 1}^T dx + \psi_{\pi}^T d\pi$$

where the symbols $\bar{H}_u(t)$, $\bar{\Gamma}$, $\bar{\psi}$ denote the functions $H_u(t)$, $\Gamma$, $\psi$ evaluated at $\bar{u}(t)$, $\bar{\pi}$, $\bar{\mu}$, $\bar{x}(t)$, $\bar{\lambda}(t)$. If one chooses the variations defined by Eqs. (55), Eqs. (59) reduce to

$$\frac{(H_u - \bar{H}_u)}{\beta} = -H_{u x}^T \Lambda + H_{u \lambda}^T X + H_{u u}^T H_u + H_{u \pi}^T \Gamma$$

(60-1)

$$\frac{(\Gamma - \bar{\Gamma})}{\beta} = \int_0^1 (-H_{\pi x}^T \Lambda + H_{\pi \lambda}^T X + H_{\pi u}^T H_u + H_{\pi \pi}^T \Gamma)dt$$

$$- G_{\pi x 1}^T \Lambda + G_{\pi \pi}^T \Gamma + G_{\pi \mu}^T \psi$$

(60-2)
\[
(\psi - \bar{\psi})/\beta = -\psi^T_{x_1} + \psi^T_{\pi}
\]

Comparing Eqs. (46) and Eqs. (60), one obtains the relations

\[
A = (H_u - \bar{H}_u)/\beta , \quad B = (\bar{\Gamma} - \bar{\Gamma})/\beta , \quad C = (\psi - \bar{\psi})/\beta
\]

Therefore, instead of determining \(A(t), B, C\) from Eqs. (46), one can calculate them in the following way: Employing \(\bar{u}(t), \bar{\pi}, \bar{\mu}, \bar{x}(t), \bar{\lambda}(t)\) defined by Eqs. (52), (55), evaluate \(\bar{H}_u, \bar{\Gamma}, \bar{\psi}\). Once \(\bar{H}_u, \bar{\Gamma}, \bar{\psi}\) are known, \(A(t), B, C\) are determined by Eqs. (61).

4.3. Modified Gradient Algorithm. Summarizing the previous analysis, the modified gradient algorithm for minimizing the performance index is given as follows.

(a) For a given nominal \(u(t), \pi, \mu\), determine \(x(t)\) from the system

\[
x_0 = k , \quad \dot{x} = \phi
\]

by forward integration and \(\lambda(t)\) from the system

\[
\lambda_1 = -C_{x_1} , \quad \dot{\lambda} = H_x
\]

by backward integration. With \(u(t), \pi, \mu, x(t), \lambda(t)\) known, calculate \(H_u(t), \Gamma, \psi\) and evaluate the performance index with the expression

\[
R = \int_0^1 H_u^T H_u dt + \Gamma^T \Gamma + \psi^T \psi
\]

If \(R\) satisfies the inequality

\[
R \leq \epsilon_1
\]
where $\epsilon_1$ is a preselected small positive quantity, the algorithm is terminated and the solution considered achieved. If $R$ violates Ineq. (65), proceed to Step (b).

(b) Calculate $\tilde{u}(t)$, $\tilde{\pi}$, $\tilde{\mu}$ with the relations

$$\tilde{u} = u - \beta H_u \quad , \quad \tilde{\pi} = \pi - \beta \Gamma \quad , \quad \tilde{\mu} = \mu - \beta \psi$$

(66)

With $\tilde{u}(t)$, $\tilde{\pi}$, $\tilde{\mu}$ known, determine $\tilde{x}(t)$ from the system (62) by forward integration and $\tilde{\lambda}(t)$ from the system (63) by backward integration. Calculate $\tilde{H}_u(t)$, $\tilde{\Gamma}$, $\tilde{\psi}$ and determine $X(t)$, $M(t)$ from the relations

$$X = (\lambda - \tilde{\lambda})/\beta \quad , \quad \Lambda = -(x - \tilde{x})/\beta$$

(67)

and $A(t)$, $B$, $C$ from the relations

$$A = (H_u - \tilde{H}_u)/\beta \quad , \quad B = (\Gamma - \tilde{\Gamma})/\beta \quad , \quad C = (\psi - \tilde{\psi})/\beta$$

(68)

(c) Calculate $\bar{u}(t)$, $\bar{\pi}$, $\bar{\mu}$ with the relations

$$\bar{u} = u - \beta A \quad , \quad \bar{\pi} = \pi - \beta B \quad , \quad \bar{\mu} = \mu - \beta C$$

(69)

With $\bar{u}(t)$, $\bar{\pi}$, $\bar{\mu}$ known, determine $\bar{x}(t)$ from the system (62) by forward integration and $\bar{\lambda}(t)$ from the system (63) by backward integration.

Calculate $D(t)$, $E(t)$ with the relations

$$D = (x - \bar{x})/\beta \quad , \quad E = (\lambda - \bar{\lambda})/\beta$$

(70)

(d) Define a one-parameter family of varied functions by
\[ \tilde{u} = u - \alpha A \quad , \quad \tilde{\pi} = \pi - \alpha B \quad , \quad \tilde{\mu} = \mu - \alpha C \quad (71) \]

\[ \tilde{x} = x - \alpha D \quad , \quad \tilde{\lambda} = \lambda - \alpha E \]

Then, the performance index at the varied functions becomes a function of the parameter \( \alpha \) only, that is,

\[ \tilde{R} = \tilde{R}(\alpha) \quad (72) \]

Perform a one-dimensional search on \( \tilde{R}(\alpha) \) so that the optimum stepsize \( \alpha \) minimizing \( \tilde{R}(\alpha) \) is determined. A particular search technique is presented in Section 6.

(e) With the optimum \( \alpha \) known, determine the function \( \tilde{u}(t), \tilde{\pi}, \tilde{\mu} \) from the relations

\[ \tilde{u} = u - \alpha A \quad , \quad \tilde{\pi} = \pi - \alpha B \quad , \quad \tilde{\mu} = \mu - \alpha C \quad (73) \]

Return to Step (a) and start a new iteration with \( \tilde{u}(t), \tilde{\pi}, \tilde{\mu} \) as the nominal.
5. **Conjugate-Gradient Algorithm**

Since gradient methods, in general, are notorious for their slow convergence, the conjugate-gradient algorithm for minimizing the performance index is also presented here in an attempt to improve the convergence property.\(^2\)

5.1. **Variations.** While the variations \(\Delta u(t), \Delta \pi, \Delta \mu, \Delta x(t), \Delta \lambda(t)\) for the gradient algorithm are given by Eqs. (35), the variations for the conjugate-gradient algorithm are given by

\[
\begin{align*}
\Delta u &= -\alpha a, \quad \Delta \pi = -\alpha b, \quad \Delta \mu = -\alpha c \\
\Delta x &= -\alpha d, \quad \Delta \lambda = -\alpha e
\end{align*}
\](74)

where the functions \(a(t), b, c, d(t), e(t)\) are defined by

\[
\begin{align*}
a &= A + \gamma \hat{a} \quad b &= B + \gamma \hat{b} \quad c &= C + \gamma \hat{c} \\
d &= D + \gamma \hat{d} \quad e &= E + \gamma \hat{e}
\end{align*}
\](75)

In Eqs. (75), the symbols \(\hat{a}(t), \hat{b}, \hat{c}, \hat{d}(t), \hat{e}(t)\) denote the functions \(a(t), b, c, d(t), e(t)\) used in the previous iteration. Also, in Eqs. (75), the scalar quantity \(\gamma\) is given either by

\[
\gamma = 0
\](76)

at the starting iteration or by

\(^2\)The conjugate-gradient algorithm in this section is a direct extension of those given in Refs. 2 and 6.
\[
\gamma = \frac{1}{\int_0^1 A^T A + B^T B + C^T C}
\]

(77)

at any other iteration. Again, in Eq. (77), the symbols \(\hat{A}(t)\), \(\hat{B}\), \(\hat{C}\) denote the functions \(A(t)\), \(B\), \(C\) used in the previous iteration.

5.2. **Restart.** When the variations (74) are used, the first variation of the performance index (33) becomes

\[
\delta R = -2\alpha \left[ \int_0^1 A^T \dot{a} dt + B^T b + C^T c \right]
\]

(78)

In order to have the descent property, that is

\[
\delta R < 0
\]

(79)

for the positive stepsizen \(\alpha\), the inequality

\[
\int_0^1 A^T \dot{a} dt + B^T b + C^T c > 0
\]

(80)

must be satisfied.

In the light of definitions (75), Ineq. (80) can be expressed as

\[
\left[ \int_0^1 A^T \dot{a} dt + B^T b + C^T c \right] + \gamma \left[ \int_0^1 A^T \ddot{a} dt + B^T \ddot{b} + C^T \ddot{c} \right] > 0
\]

(81)

If the problem is of quadratic functional-linear constraint type and an exact one-dimensional search is performed for the determination of the stepsize, the following property holds:

\[
\int_0^1 A^T \ddot{a} dt + B^T \ddot{b} + C^T \ddot{c} = 0
\]

(82)
Therefore, Ineq. (81) reduces to
\[
\int_0^1 A^T \mathbf{d}t + B^T \mathbf{B} + C^T \mathbf{C} > 0
\]
(83)
and is satisfied. On the other hand, if the problem is not of quadratic functional-linear constraint type and/or the one-dimensional search on the stepsize is not exact, the property (82) does not hold and Ineq. (81) may be violated. This being the case, the conjugate-gradient algorithm must be restarted. The restart of the algorithm is effected by setting
\[
\gamma = 0
\]
(84)
It is seen that, with \( \gamma = 0 \), the functions \( a(t), b, c \) reduce to \( A(t), B, C \), respectively, and the satisfaction of Ineq. (80) follows immediately.

5.3. **Conjugate-Gradient Algorithm.** For conciseness, the conjugate-gradient algorithm for minimizing the performance index, using only the original differential system, the function \( f, g, \varphi, \psi \) and their first derivatives, is given as follows.

(a) For a given nominal, \( u(t), \pi, \mu \), determine \( x(t) \) from the system
\[
x_0 = k \quad , \quad \dot{x} = \varphi
\]
(85)
by forward integration and \( \lambda(t) \) from the system
\[
\lambda_1 = -G_x \quad , \quad \dot{\lambda} = H_x
\]
(86)
by backward integration. With \( u(t), \pi, \mu, x(t), \lambda(t) \) known, calculate \( H_u(t), \Gamma, \psi \) and evaluate the performance index with the expression
\[ R = \int_{0}^{1} H_u T \, u \, dt + \Gamma^T \Gamma + \psi^T \psi \]  

(87)

If \( R \) satisfies the inequality

\[ R \leq \varepsilon_1 \]  

(88)

where \( \varepsilon_1 \) is a preselected small positive quantity, the algorithm is terminated and the solution considered achieved. If \( R \) violates Ineq. (88), proceed to Step (b).

(b) Calculate \( \bar{u}(t), \bar{\pi}, \bar{\mu} \) with the relations

\[ \bar{u} = u - \beta H_u, \quad \bar{\pi} = \pi - \beta \Gamma, \quad \bar{\mu} = \mu - \beta \psi \]  

(89)

With \( \bar{u}(t), \bar{\pi}, \bar{\mu} \) known, determine \( \bar{x}(t) \) from the system (85) by forward integration and \( \bar{\lambda}(t) \) from the system (86) by backward integration. Calculate \( \bar{H}_u(t), \bar{\Gamma}, \bar{\psi} \) and determine \( X(t), \Lambda(t) \) from the relations

\[ X = (\lambda - \bar{\lambda}) / \beta, \quad \Lambda = -(x - \bar{x}) / \beta \]  

(90)

and \( A(t), B, C \) from the relations

\[ A = (H_u - \bar{H}_u) / \beta, \quad B = (\Gamma - \bar{\Gamma}) / \beta, \quad C = (\psi - \bar{\psi}) / \beta \]  

(91)

(c) At the starting or restarting points, define

\[ \gamma = 0 \]  

(92)

For any other point, define
\[ \gamma = \frac{\int_0^1 A^T A dt + B^T B + C^T C}{\int_0^1 \hat{A}^T \hat{A} dt + \hat{B}^T \hat{B} + \hat{C}^T \hat{C}} \]  

(93)

Calculate \( a(t), b, c \) with the relations

\[ a = A + \gamma \hat{a}, \quad b = B + \gamma \hat{b}, \quad c = C + \gamma \hat{c} \]  

(94)

(d) Calculate \( \tilde{u}(t), \tilde{\pi}, \tilde{\mu} \) with the relations

\[ \tilde{u} = u - \beta a, \quad \tilde{\pi} = \pi - \beta b, \quad \tilde{\mu} = \mu - \beta c \]  

(95)

With \( \tilde{u}(t), \tilde{\pi}, \tilde{\mu} \) known, determine \( \tilde{x}(t) \) from the system (85) by forward integration and \( \tilde{\lambda}(t) \) from the system (86) by backward integration. Calculate \( d(t), e(t) \) with the relations

\[ d = (x - \tilde{x})/\beta, \quad e = (\lambda - \tilde{\lambda})/\beta \]  

(96)

(e) Define a one-parameter family of varied functions by

\[ \tilde{u} = u - \alpha a, \quad \tilde{\pi} = \pi - \alpha b, \quad \tilde{\mu} = \mu - \alpha c \]  

\[ \tilde{x} = x - \alpha d, \quad \tilde{\lambda} = \lambda - \alpha e \]  

(97)

Then, the performance index at the varied functions becomes a function of the parameter \( \alpha \) only, that is

\[ \tilde{K} = \tilde{K}(\alpha) \]  

(98)

Perform a one-dimensional search on \( \tilde{K}(\alpha) \) so that the optimum stepsize
α minimizing \( \hat{R}(a) \) is determined. A particular search technique is presented in Section 6.

(f) With the optimum \( \alpha \) known, determine the function \( \bar{u}(t) \), \( \bar{\pi} \), \( \bar{\mu} \) from the relations

\[
\bar{u} = u - \alpha a \quad \bar{\pi} = \pi - \alpha b \quad \bar{\mu} = \mu - \alpha c
\]

(99)

Return to Step (a) and start a new iteration with \( \bar{u}(t) \), \( \bar{\pi} \), \( \bar{\mu} \) as the nominal.
6. One-Dimensional Search

In the algorithms presented in Sections 3-5, a one-dimensional search for the optimal stepsize \( \alpha \) minimizing the function \( \tilde{R}(\alpha) \) is needed. Preferably, the search must be performed until the condition

\[ \tilde{R}'(\alpha) = 0 \]  

is satisfied. In Eq. (100), the symbol \( \tilde{R}'(\alpha) \) denotes the derivative of \( \tilde{R}(\alpha) \) with respect to \( \alpha \). However, this exact search, in general, is time-consuming. Therefore, we renounce the use of the exact search and employ a one-cycle cubic interpolation described below.

(a) At \( \alpha = 0 \), the quantities \( \tilde{R}(0) \), \( \tilde{R}_\alpha(0) \) are given by

\[ \tilde{R}(0) = R \quad , \quad \tilde{R}_\alpha(0) = -2\left[ \int_0^1 A^T \text{ad}t + B^T b + C^T c \right] \]  

Calculate \( \alpha_1 \) by the relation

\[ \alpha_1 = -\frac{2\tilde{R}(0)}{\tilde{R}_\alpha(0)} \]  

and evaluate \( \tilde{R}(\alpha_1) \). If the inequality

\[ \tilde{R}(\alpha_1) > \tilde{R}(0) \]  

holds, proceed to Step (b). If Ineq. (103) is violated, proceed to Step (c).

(b) Calculate \( \alpha_2 \) by the relation

\[ \alpha_2 = \alpha_1 / 2 \]
and evaluate $\tilde{R}(\alpha_2)$. If the inequality

$$\tilde{R}(\alpha_2) < \tilde{R}(0)$$  \hspace{1cm} (105)

holds, proceed to Step (d). On the other hand, if Ineq. (105) is violated, set

$$\alpha_1 = \alpha_2 \ , \ \tilde{R}(\alpha_1) = \tilde{R}(\alpha_2)$$  \hspace{1cm} (106)

and return to the beginning of Step (b).

(c) Calculate $\alpha_2$ by the relation

$$\alpha_2 = 2\alpha_1$$  \hspace{1cm} (107)

and evaluate $\tilde{R}(\alpha_2)$. If the inequality

$$\tilde{R}(\alpha_2) > \tilde{R}(\alpha_1)$$  \hspace{1cm} (108)

holds, proceed to Step (d). On the other hand, if Ineq. (108) is violated, set

$$\alpha_1 = \alpha_2 \ , \ \tilde{R}(\alpha_1) = \tilde{R}(\alpha_2)$$  \hspace{1cm} (109)

and return to the beginning of Step (c).

(d) Calculate the optimal stepsize $\alpha$ with the formula

$$\alpha = \left\{ \sqrt{W^2 - 3Z\tilde{R}(0)} - W \right\} / 3Z$$  \hspace{1cm} (110)

where $W$, $Z$ are given by
\[ W = \frac{\alpha_1 \alpha_2}{\alpha_2 - \alpha_1} \left[ \frac{\tilde{R}(\alpha_1) - \tilde{R}(0) - \tilde{R}_\alpha(0)\alpha_1}{\alpha_1^3} - \frac{\tilde{R}(\alpha_2) - \tilde{R}(0) - \tilde{R}_\alpha(0)\alpha_2}{\alpha_2^3} \right] \]

\[ Z = \frac{-1}{\alpha_2 - \alpha_1} \left[ \frac{\tilde{R}(\alpha_1) - \tilde{R}(0) - \tilde{R}_\alpha(0)\alpha_1}{\alpha_1^2} - \frac{\tilde{R}(\alpha_2) - \tilde{R}(0) - \tilde{R}_\alpha(0)\alpha_2}{\alpha_2^2} \right] \]
7. **Experimental Conditions**

The modified gradient algorithm in Section 4 and the conjugate-gradient algorithm in Section 5 were tested through several numerical examples using the Rice University Burroughs B-5500 Computer. These algorithms were programmed in FORTRAN IV and in double-precision arithmetic.

For integration of differential systems, the Hamming's modified predictor-corrector method with a special Range-Kutta starting routine was employed (Ref. 7). For integration of definite integrals, the Simpson's rule was used. For both cases, the interval of integrations was divided into 20 steps. Additional pertinent experimental conditions are given in the following.

7.1. **Starting Nominal.** For all the examples, the nominal functions $u(t), \pi, \mu$ employed for starting an algorithm were arbitrarily chosen as

$$u = 0 \quad \pi = 0 \quad \mu = 0$$  \hspace{1cm} (112)

7.2. **Convergence Criterion.** The conjugate-gradient algorithm was continued until the performance index satisfied the stopping condition for convergence (88), namely,

$$R \leq \varepsilon_1$$  \hspace{1cm} (113)

For all the experiments, the quantity $\varepsilon_1$ was taken to be

$$\varepsilon_1 = 10^{-5}$$  \hspace{1cm} (114)

For the modified gradient algorithm, computation was not stopped by the convergence criterion (113)-(114). Instead, it was stopped at the $N^{th}$
iteration, where $N$ is the number of iterations for convergence using the conjugate-gradient algorithm.

7.3. **Small Step $\beta$.** The small step $\beta$ in Eqs. (66)-(70) for the modified gradient algorithm and in Eqs. (89)-(91), (95)-(96) for the conjugate-gradient algorithm was defined by

$$\beta = \frac{\varepsilon_2}{\sqrt{R}}$$

(115)

where the positive quantity $\varepsilon_2$ was taken to be

$$\varepsilon_2 = 10^{-8}$$

(116)
8. **Numerical Examples**

In this section, the numerical examples tested are given. For simplicity, all the symbols employed here are scalar quantities.

**Example 8.1.** Minimize the function (Ref. 2)

\[
I = \int_0^1 (x_1^2 + x_2^2 + 0.005u^2) \, dt
\]  

(117)

subject to the differential constraints

\[
\dot{x}_1 = x_2, \quad \dot{x}_2 = -x_2 + u
\]  

(118)

and the boundary conditions

\[
x_1(0) = 0, \quad x_2(0) = -1
\]  

(119)

This example is not constrained at the terminal point. It is used here simply to show that the algorithms in this thesis can be applied to the terminally unconstrained optimal problems, as well.

**Example 8.2.** Minimize the functional

\[
I = \int_0^1 (x^2 + u^2) \, dt + \pi^2
\]  

(120)

subject to the differential constraint

\[
\dot{x} = u
\]  

(121)

and the boundary conditions

\[
x(0) = 1, \quad x(1) = 2\pi
\]  

(122)
Example 8.3. Minimize the functional

\[ I = \int_0^1 [(x_1 + x_2)^2 + u^2] \, dt + \pi_1^2 + 3\pi_2^2 + x_1^2(1) + 4x_2^2(1) \]  \hspace{1cm} (123)

subject to the differential constraints

\[ \dot{x}_1 = 2x_1 - x_2 + u + 2\pi_1 \]  \hspace{1cm} (124)
\[ \dot{x}_2 = x_1 - 4x_2 - u + \pi_2 \]

and the boundary conditions

\[ x_1(0) = 0, \quad x_2(0) = 0 \]
\[ x_1(1) + x_2(1) + \pi_1 + \pi_2 + 1 = 0 \]  \hspace{1cm} (125)
\[ 2x_1(1) + x_2(1) - \pi_1 + 2\pi_2 + 2 = 0 \]
\[ x_1(1) - x_2(1) + 2\pi_1 + 3\pi_2 + 3 = 0 \]

Example 8.4. Minimize the functional (Ref. 8)

\[ I = \int_0^1 x^2 u^2 \, dt \]  \hspace{1cm} (126)

subject to the differential constraint

\[ \dot{x} = ux^2 \]  \hspace{1cm} (127)

and the boundary conditions

\[ x(0) = 1, \quad x(1) = e \]  \hspace{1cm} (128)

where e is the base of natural logarithms.
9. Computational Results

Using the conjugate-gradient algorithm, convergence was achieved for all the four numerical examples of Section 8. The number of iterations for convergence is 11, 4, 22, 16 for Examples 8.1-8.4, respectively. The converged solutions of these examples are presented in Tables 1-4 and the sequences of the values of I, P, Q, R are given in Tables 5-8. In Tables 5-8, the symbol i denotes the iteration number.

Using the modified gradient algorithm, the computation was stopped at 11th, 4th, 22th, 16th iteration for Examples 8.1-8.4, respectively. The convergence criterion (113)-(114) was not met at these points. For the purpose of comparison, the convergence history of the performance index R is also given in Tables 5-8.
Table 1. Converged Solution (Example 8.1, \(N = 11\)).

<table>
<thead>
<tr>
<th>(t)</th>
<th>(x_1)</th>
<th>(x_2)</th>
<th>(u)</th>
<th>(\lambda_1)</th>
<th>(\lambda_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.0000</td>
<td>-1.0000</td>
<td>12.6477</td>
<td>0.1023</td>
<td>0.1381</td>
</tr>
<tr>
<td>0.1</td>
<td>-0.0517</td>
<td>-0.1987</td>
<td>3.5741</td>
<td>0.0958</td>
<td>0.0319</td>
</tr>
<tr>
<td>0.2</td>
<td>-0.0599</td>
<td>-0.0156</td>
<td>0.6493</td>
<td>0.0844</td>
<td>0.0081</td>
</tr>
<tr>
<td>0.3</td>
<td>-0.0595</td>
<td>0.0190</td>
<td>0.2093</td>
<td>0.0723</td>
<td>0.0017</td>
</tr>
<tr>
<td>0.4</td>
<td>-0.0568</td>
<td>0.0318</td>
<td>0.0959</td>
<td>0.0607</td>
<td>0.0005</td>
</tr>
<tr>
<td>0.5</td>
<td>-0.0535</td>
<td>0.0307</td>
<td>-0.0441</td>
<td>0.0497</td>
<td>0.0016</td>
</tr>
<tr>
<td>0.6</td>
<td>-0.0509</td>
<td>0.0203</td>
<td>-0.0947</td>
<td>0.0392</td>
<td>0.0025</td>
</tr>
<tr>
<td>0.7</td>
<td>-0.0494</td>
<td>0.0104</td>
<td>-0.0684</td>
<td>0.0292</td>
<td>0.0023</td>
</tr>
<tr>
<td>0.8</td>
<td>-0.0487</td>
<td>0.0043</td>
<td>-0.0433</td>
<td>0.0194</td>
<td>0.0015</td>
</tr>
<tr>
<td>0.9</td>
<td>-0.0485</td>
<td>0.0002</td>
<td>-0.0339</td>
<td>0.0097</td>
<td>0.0006</td>
</tr>
<tr>
<td>1.0</td>
<td>-0.0486</td>
<td>-0.0018</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>
Table 2. Converged Solution (Example 8.2, N = 4).

\[ \pi = 0.2721 \]
\[ \mu = 0.2721 \]

<table>
<thead>
<tr>
<th>t</th>
<th>x</th>
<th>u</th>
<th>\lambda</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>1.0000</td>
<td>-0.8512</td>
<td>-1.6995</td>
</tr>
<tr>
<td>0.1</td>
<td>0.9198</td>
<td>-0.7538</td>
<td>-1.5077</td>
</tr>
<tr>
<td>0.2</td>
<td>0.8489</td>
<td>-0.6649</td>
<td>-1.3310</td>
</tr>
<tr>
<td>0.3</td>
<td>0.7865</td>
<td>-0.5832</td>
<td>-1.1676</td>
</tr>
<tr>
<td>0.4</td>
<td>0.7320</td>
<td>-0.5077</td>
<td>-1.0158</td>
</tr>
<tr>
<td>0.5</td>
<td>0.6848</td>
<td>-0.4372</td>
<td>-0.8742</td>
</tr>
<tr>
<td>0.6</td>
<td>0.6444</td>
<td>-0.3711</td>
<td>-0.7414</td>
</tr>
<tr>
<td>0.7</td>
<td>0.6105</td>
<td>-0.3085</td>
<td>-0.6160</td>
</tr>
<tr>
<td>0.8</td>
<td>0.5826</td>
<td>-0.2488</td>
<td>-0.4968</td>
</tr>
<tr>
<td>0.9</td>
<td>0.5606</td>
<td>-0.1915</td>
<td>-0.3826</td>
</tr>
<tr>
<td>1.0</td>
<td>0.5443</td>
<td>-0.1360</td>
<td>-0.2721</td>
</tr>
</tbody>
</table>
Table 3. Converged Solution (Example 8.3, N = 22).

\[ \pi_1 = -0.0360, \quad \pi_2 = -0.8735 \]
\[ \mu_1 = -3.8159, \quad \mu_2 = 1.0856, \quad \mu_3 = 2.6572 \]

<table>
<thead>
<tr>
<th>( t )</th>
<th>( x_1 )</th>
<th>( x_2 )</th>
<th>( u )</th>
<th>( \lambda_1 )</th>
<th>( \lambda_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.4045</td>
<td>0.7880</td>
<td>-0.0199</td>
</tr>
<tr>
<td>0.1</td>
<td>0.0366</td>
<td>-0.0988</td>
<td>0.2938</td>
<td>0.6370</td>
<td>0.0499</td>
</tr>
<tr>
<td>0.2</td>
<td>0.0777</td>
<td>-0.1529</td>
<td>0.1863</td>
<td>0.5004</td>
<td>0.1272</td>
</tr>
<tr>
<td>0.3</td>
<td>0.1201</td>
<td>-0.1768</td>
<td>0.0764</td>
<td>0.3814</td>
<td>0.2274</td>
</tr>
<tr>
<td>0.4</td>
<td>0.1608</td>
<td>-0.1798</td>
<td>-0.0464</td>
<td>0.2784</td>
<td>0.3703</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1949</td>
<td>-0.1675</td>
<td>-0.1982</td>
<td>0.1859</td>
<td>0.5818</td>
</tr>
<tr>
<td>0.6</td>
<td>0.2151</td>
<td>-0.1423</td>
<td>-0.4018</td>
<td>0.0950</td>
<td>0.8978</td>
</tr>
<tr>
<td>0.7</td>
<td>0.2098</td>
<td>-0.1046</td>
<td>-0.6882</td>
<td>-0.0076</td>
<td>1.3674</td>
</tr>
<tr>
<td>0.8</td>
<td>0.1604</td>
<td>-0.0528</td>
<td>-1.1004</td>
<td>-0.1406</td>
<td>2.0588</td>
</tr>
<tr>
<td>0.9</td>
<td>0.0384</td>
<td>0.0165</td>
<td>-1.6987</td>
<td>-0.3304</td>
<td>3.0658</td>
</tr>
<tr>
<td>1.0</td>
<td>-0.1990</td>
<td>0.1088</td>
<td>-2.5684</td>
<td>-0.6143</td>
<td>4.5171</td>
</tr>
</tbody>
</table>
Table 4. Converged Solution (Example 8.4, N = 16).

\[ \mu = -0.7358 \]

<table>
<thead>
<tr>
<th>t</th>
<th>x</th>
<th>u</th>
<th>\lambda</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>1.0000</td>
<td>0.9942</td>
<td>2.0007</td>
</tr>
<tr>
<td>0.1</td>
<td>1.1050</td>
<td>0.9067</td>
<td>1.8102</td>
</tr>
<tr>
<td>0.2</td>
<td>1.2215</td>
<td>0.8189</td>
<td>1.6379</td>
</tr>
<tr>
<td>0.3</td>
<td>1.3498</td>
<td>0.7399</td>
<td>1.4820</td>
</tr>
<tr>
<td>0.4</td>
<td>1.4917</td>
<td>0.6705</td>
<td>1.3409</td>
</tr>
<tr>
<td>0.5</td>
<td>1.6487</td>
<td>0.6063</td>
<td>1.2133</td>
</tr>
<tr>
<td>0.6</td>
<td>1.8220</td>
<td>0.5491</td>
<td>1.0978</td>
</tr>
<tr>
<td>0.7</td>
<td>2.0138</td>
<td>0.4962</td>
<td>0.9933</td>
</tr>
<tr>
<td>0.8</td>
<td>2.2255</td>
<td>0.4495</td>
<td>0.8988</td>
</tr>
<tr>
<td>0.9</td>
<td>2.4596</td>
<td>0.4066</td>
<td>0.8132</td>
</tr>
<tr>
<td>1.0</td>
<td>2.7182</td>
<td>0.3677</td>
<td>0.7358</td>
</tr>
</tbody>
</table>
Table 5. Convergence History (Example 8.1).

<table>
<thead>
<tr>
<th>i</th>
<th>I</th>
<th>P</th>
<th>Q</th>
<th>R</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$0.60042 \times 10^0$</td>
<td>0.</td>
<td>$0.40 \times 10^0$</td>
<td>$0.40 \times 10^0$</td>
<td>$0.40 \times 10^0$</td>
</tr>
<tr>
<td>1</td>
<td>$0.19086 \times 10^0$</td>
<td>0.</td>
<td>$0.17 \times 10^{-1}$</td>
<td>$0.17 \times 10^{-1}$</td>
<td>$0.17 \times 10^{-1}$</td>
</tr>
<tr>
<td>2</td>
<td>$0.10594 \times 10^0$</td>
<td>0.</td>
<td>$0.22 \times 10^{-2}$</td>
<td>$0.22 \times 10^{-2}$</td>
<td>$0.28 \times 10^{-2}$</td>
</tr>
<tr>
<td>3</td>
<td>$0.80209 \times 10^{-1}$</td>
<td>0.</td>
<td>$0.41 \times 10^{-3}$</td>
<td>$0.41 \times 10^{-3}$</td>
<td>$0.22 \times 10^{-2}$</td>
</tr>
<tr>
<td>4</td>
<td>$0.80141 \times 10^{-1}$</td>
<td>0.</td>
<td>$0.41 \times 10^{-3}$</td>
<td>$0.41 \times 10^{-3}$</td>
<td>$0.17 \times 10^{-2}$</td>
</tr>
<tr>
<td>5</td>
<td>$0.72029 \times 10^{-1}$</td>
<td>0.</td>
<td>$0.83 \times 10^{-4}$</td>
<td>$0.83 \times 10^{-4}$</td>
<td>$0.14 \times 10^{-2}$</td>
</tr>
<tr>
<td>6</td>
<td>$0.72010 \times 10^{-1}$</td>
<td>0.</td>
<td>$0.83 \times 10^{-4}$</td>
<td>$0.83 \times 10^{-4}$</td>
<td>$0.11 \times 10^{-2}$</td>
</tr>
<tr>
<td>7</td>
<td>$0.70456 \times 10^{-1}$</td>
<td>0.</td>
<td>$0.45 \times 10^{-4}$</td>
<td>$0.45 \times 10^{-4}$</td>
<td>$0.10 \times 10^{-2}$</td>
</tr>
<tr>
<td>8</td>
<td>$0.70240 \times 10^{-1}$</td>
<td>0.</td>
<td>$0.38 \times 10^{-4}$</td>
<td>$0.38 \times 10^{-4}$</td>
<td>$0.85 \times 10^{-3}$</td>
</tr>
<tr>
<td>9</td>
<td>$0.69858 \times 10^{-1}$</td>
<td>0.</td>
<td>$0.15 \times 10^{-4}$</td>
<td>$0.15 \times 10^{-4}$</td>
<td>$0.73 \times 10^{-3}$</td>
</tr>
<tr>
<td>10</td>
<td>$0.69857 \times 10^{-1}$</td>
<td>0.</td>
<td>$0.15 \times 10^{-4}$</td>
<td>$0.15 \times 10^{-4}$</td>
<td>$0.64 \times 10^{-3}$</td>
</tr>
<tr>
<td>11</td>
<td>$0.69616 \times 10^{-1}$</td>
<td>0.</td>
<td>$0.77 \times 10^{-5}$</td>
<td>$0.77 \times 10^{-5}$</td>
<td>$0.56 \times 10^{-3}$</td>
</tr>
</tbody>
</table>
Table 6. Convergence History (Example 8.2).

<table>
<thead>
<tr>
<th>i</th>
<th>I</th>
<th>P</th>
<th>Q</th>
<th>R</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.10000 x 10^1</td>
<td>0.10 x 10^1</td>
<td>0.13 x 10^1</td>
<td>0.23 x 10^1</td>
<td>0.23 x 10^1</td>
</tr>
<tr>
<td>1</td>
<td>0.80026 x 10^0</td>
<td>0.10 x 10^0</td>
<td>0.30 x 10^0</td>
<td>0.40 x 10^0</td>
<td>0.40 x 10^0</td>
</tr>
<tr>
<td>2</td>
<td>0.78884 x 10^0</td>
<td>0.93 x 10^-1</td>
<td>0.30 x 10^-1</td>
<td>0.12 x 10^0</td>
<td>0.16 x 10^0</td>
</tr>
<tr>
<td>3</td>
<td>0.84875 x 10^0</td>
<td>0.25 x 10^-4</td>
<td>0.13 x 10^-2</td>
<td>0.14 x 10^-2</td>
<td>0.84 x 10^-1</td>
</tr>
<tr>
<td>4</td>
<td>0.84979 x 10^0</td>
<td>0.87 x 10^-10</td>
<td>0.74 x 10^-6</td>
<td>0.74 x 10^-6</td>
<td>0.45 x 10^-1</td>
</tr>
</tbody>
</table>
Table 7. Convergence History (Example 8.3).

<table>
<thead>
<tr>
<th>i</th>
<th>I</th>
<th>P</th>
<th>Q</th>
<th>R</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.</td>
<td>$0.14 \times 10^2$</td>
<td>0.</td>
<td>$0.14 \times 10^2$</td>
<td>$0.14 \times 10^2$</td>
</tr>
<tr>
<td>1</td>
<td>$0.74861 \times 10^{-3}$</td>
<td>$0.13 \times 10^2$</td>
<td>$0.20 \times 10^0$</td>
<td>$0.13 \times 10^2$</td>
<td>$0.13 \times 10^2$</td>
</tr>
<tr>
<td>2</td>
<td>$0.61218 \times 10^0$</td>
<td>$0.29 \times 10^1$</td>
<td>$0.33 \times 10^1$</td>
<td>$0.62 \times 10^1$</td>
<td>$0.13 \times 10^2$</td>
</tr>
<tr>
<td>3</td>
<td>$0.10471 \times 10^1$</td>
<td>$0.10 \times 10^1$</td>
<td>$0.23 \times 10^1$</td>
<td>$0.34 \times 10^1$</td>
<td>$0.13 \times 10^2$</td>
</tr>
<tr>
<td>4</td>
<td>$0.14719 \times 10^1$</td>
<td>$0.77 \times 10^0$</td>
<td>$0.23 \times 10^{-1}$</td>
<td>$0.80 \times 10^0$</td>
<td>$0.13 \times 10^2$</td>
</tr>
<tr>
<td>5</td>
<td>$0.14878 \times 10^1$</td>
<td>$0.64 \times 10^0$</td>
<td>$0.82 \times 10^{-1}$</td>
<td>$0.73 \times 10^0$</td>
<td>$0.13 \times 10^2$</td>
</tr>
<tr>
<td>6</td>
<td>$0.14812 \times 10^1$</td>
<td>$0.59 \times 10^0$</td>
<td>$0.10 \times 10^0$</td>
<td>$0.70 \times 10^0$</td>
<td>$0.12 \times 10^2$</td>
</tr>
<tr>
<td>7</td>
<td>$0.15590 \times 10^1$</td>
<td>$0.48 \times 10^0$</td>
<td>$0.50 \times 10^{-1}$</td>
<td>$0.53 \times 10^0$</td>
<td>$0.12 \times 10^2$</td>
</tr>
<tr>
<td>8</td>
<td>$0.15561 \times 10^1$</td>
<td>$0.48 \times 10^0$</td>
<td>$0.49 \times 10^{-1}$</td>
<td>$0.53 \times 10^0$</td>
<td>$0.12 \times 10^2$</td>
</tr>
<tr>
<td>9</td>
<td>$0.16903 \times 10^1$</td>
<td>$0.35 \times 10^0$</td>
<td>$0.16 \times 10^{-1}$</td>
<td>$0.37 \times 10^0$</td>
<td>$0.12 \times 10^2$</td>
</tr>
<tr>
<td>10</td>
<td>$0.16894 \times 10^1$</td>
<td>$0.35 \times 10^0$</td>
<td>$0.16 \times 10^{-1}$</td>
<td>$0.37 \times 10^0$</td>
<td>$0.12 \times 10^2$</td>
</tr>
<tr>
<td>11</td>
<td>$0.16781 \times 10^1$</td>
<td>$0.35 \times 10^0$</td>
<td>$0.13 \times 10^{-1}$</td>
<td>$0.37 \times 10^0$</td>
<td>$0.11 \times 10^2$</td>
</tr>
<tr>
<td>12</td>
<td>$0.16780 \times 10^1$</td>
<td>$0.35 \times 10^0$</td>
<td>$0.13 \times 10^{-1}$</td>
<td>$0.37 \times 10^0$</td>
<td>$0.11 \times 10^2$</td>
</tr>
<tr>
<td>13</td>
<td>$0.16914 \times 10^1$</td>
<td>$0.34 \times 10^0$</td>
<td>$0.20 \times 10^{-1}$</td>
<td>$0.36 \times 10^0$</td>
<td>$0.11 \times 10^2$</td>
</tr>
<tr>
<td>14</td>
<td>$0.16912 \times 10^1$</td>
<td>$0.34 \times 10^0$</td>
<td>$0.20 \times 10^{-1}$</td>
<td>$0.36 \times 10^0$</td>
<td>$0.11 \times 10^2$</td>
</tr>
<tr>
<td>15</td>
<td>$0.28219 \times 10^1$</td>
<td>$0.87 \times 10^{-2}$</td>
<td>$0.42 \times 10^{-1}$</td>
<td>$0.50 \times 10^{-1}$</td>
<td>$0.11 \times 10^2$</td>
</tr>
<tr>
<td>16</td>
<td>$0.28676 \times 10^1$</td>
<td>$0.60 \times 10^{-2}$</td>
<td>$0.37 \times 10^{-1}$</td>
<td>$0.43 \times 10^{-1}$</td>
<td>$0.11 \times 10^2$</td>
</tr>
<tr>
<td>17</td>
<td>$0.29816 \times 10^1$</td>
<td>$0.17 \times 10^{-2}$</td>
<td>$0.23 \times 10^{-1}$</td>
<td>$0.25 \times 10^{-1}$</td>
<td>$0.11 \times 10^2$</td>
</tr>
<tr>
<td>18</td>
<td>$0.31230 \times 10^1$</td>
<td>$0.18 \times 10^{-2}$</td>
<td>$0.22 \times 10^{-2}$</td>
<td>$0.41 \times 10^{-2}$</td>
<td>$0.10 \times 10^2$</td>
</tr>
<tr>
<td>19</td>
<td>$0.31241 \times 10^1$</td>
<td>$0.15 \times 10^{-2}$</td>
<td>$0.23 \times 10^{-2}$</td>
<td>$0.38 \times 10^{-2}$</td>
<td>$0.10 \times 10^2$</td>
</tr>
<tr>
<td>20</td>
<td>$0.31323 \times 10^1$</td>
<td>$0.81 \times 10^{-3}$</td>
<td>$0.61 \times 10^{-3}$</td>
<td>$0.14 \times 10^{-2}$</td>
<td>$0.10 \times 10^2$</td>
</tr>
<tr>
<td>21</td>
<td>$0.31382 \times 10^1$</td>
<td>$0.46 \times 10^{-3}$</td>
<td>$0.70 \times 10^{-3}$</td>
<td>$0.11 \times 10^{-2}$</td>
<td>$0.10 \times 10^2$</td>
</tr>
<tr>
<td>22</td>
<td>$0.31652 \times 10^1$</td>
<td>$0.47 \times 10^{-6}$</td>
<td>$0.18 \times 10^{-5}$</td>
<td>$0.23 \times 10^{-5}$</td>
<td>$0.10 \times 10^2$</td>
</tr>
</tbody>
</table>
Table 8. Convergence History (Example 8.4).

<table>
<thead>
<tr>
<th>i</th>
<th>I</th>
<th>P</th>
<th>Q</th>
<th>R</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.</td>
<td>$0.29 \times 10^1$</td>
<td>0.</td>
<td>$0.29 \times 10^1$</td>
<td>$0.29 \times 10^1$</td>
</tr>
<tr>
<td>1</td>
<td>$0.51968 \times 10^{-1}$</td>
<td>$0.21 \times 10^1$</td>
<td>$0.33 \times 10^0$</td>
<td>$0.24 \times 10^1$</td>
<td>$0.24 \times 10^1$</td>
</tr>
<tr>
<td>2</td>
<td>$0.36138 \times 10^0$</td>
<td>$0.82 \times 10^0$</td>
<td>$0.68 \times 10^0$</td>
<td>$0.15 \times 10^1$</td>
<td>$0.21 \times 10^1$</td>
</tr>
<tr>
<td>3*</td>
<td>$0.79858 \times 10^0$</td>
<td>$0.85 \times 10^{-1}$</td>
<td>$0.29 \times 10^0$</td>
<td>$0.38 \times 10^0$</td>
<td>$0.18 \times 10^1$</td>
</tr>
<tr>
<td>4</td>
<td>$0.92280 \times 10^0$</td>
<td>$0.11 \times 10^{-1}$</td>
<td>$0.22 \times 10^{-1}$</td>
<td>$0.33 \times 10^{-1}$</td>
<td>$0.15 \times 10^1$</td>
</tr>
<tr>
<td>5</td>
<td>$0.96356 \times 10^0$</td>
<td>$0.24 \times 10^{-2}$</td>
<td>$0.52 \times 10^{-2}$</td>
<td>$0.77 \times 10^{-2}$</td>
<td>$0.12 \times 10^1$</td>
</tr>
<tr>
<td>6</td>
<td>$0.98193 \times 10^0$</td>
<td>$0.60 \times 10^{-3}$</td>
<td>$0.19 \times 10^{-2}$</td>
<td>$0.25 \times 10^{-2}$</td>
<td>$0.11 \times 10^1$</td>
</tr>
<tr>
<td>7</td>
<td>$0.99511 \times 10^0$</td>
<td>$0.47 \times 10^{-4}$</td>
<td>$0.79 \times 10^{-3}$</td>
<td>$0.83 \times 10^{-3}$</td>
<td>$0.83 \times 10^0$</td>
</tr>
<tr>
<td>8</td>
<td>$0.99886 \times 10^0$</td>
<td>$0.25 \times 10^{-5}$</td>
<td>$0.46 \times 10^{-3}$</td>
<td>$0.46 \times 10^{-3}$</td>
<td>$0.66 \times 10^0$</td>
</tr>
<tr>
<td>9</td>
<td>$0.99997 \times 10^0$</td>
<td>$0.39 \times 10^{-8}$</td>
<td>$0.36 \times 10^{-3}$</td>
<td>$0.36 \times 10^{-3}$</td>
<td>$0.45 \times 10^0$</td>
</tr>
<tr>
<td>10</td>
<td>$0.99965 \times 10^0$</td>
<td>$0.52 \times 10^{-6}$</td>
<td>$0.26 \times 10^{-3}$</td>
<td>$0.26 \times 10^{-3}$</td>
<td>$0.32 \times 10^0$</td>
</tr>
<tr>
<td>11</td>
<td>$0.99884 \times 10^0$</td>
<td>$0.24 \times 10^{-5}$</td>
<td>$0.15 \times 10^{-3}$</td>
<td>$0.15 \times 10^{-3}$</td>
<td>$0.19 \times 10^0$</td>
</tr>
<tr>
<td>12</td>
<td>$0.99852 \times 10^0$</td>
<td>$0.34 \times 10^{-5}$</td>
<td>$0.81 \times 10^{-4}$</td>
<td>$0.85 \times 10^{-4}$</td>
<td>$0.12 \times 10^0$</td>
</tr>
<tr>
<td>13</td>
<td>$0.99836 \times 10^0$</td>
<td>$0.52 \times 10^{-5}$</td>
<td>$0.32 \times 10^{-4}$</td>
<td>$0.37 \times 10^{-4}$</td>
<td>$0.70 \times 10^{-1}$</td>
</tr>
<tr>
<td>14</td>
<td>$0.99920 \times 10^0$</td>
<td>$0.12 \times 10^{-5}$</td>
<td>$0.22 \times 10^{-4}$</td>
<td>$0.23 \times 10^{-4}$</td>
<td>$0.40 \times 10^{-1}$</td>
</tr>
<tr>
<td>15</td>
<td>$0.99985 \times 10^0$</td>
<td>$0.12 \times 10^{-7}$</td>
<td>$0.13 \times 10^{-4}$</td>
<td>$0.13 \times 10^{-4}$</td>
<td>$0.21 \times 10^{-1}$</td>
</tr>
<tr>
<td>16</td>
<td>$0.99997 \times 10^0$</td>
<td>$0.79 \times 10^{-9}$</td>
<td>$0.74 \times 10^{-5}$</td>
<td>$0.74 \times 10^{-5}$</td>
<td>$0.12 \times 10^{-1}$</td>
</tr>
</tbody>
</table>

The asterisk * indicates the iteration at which the conjugate-gradient algorithm restarted.
10. Conclusion

In this thesis, the problem of extremizing a functional $I$ involving the state $x(t)$, the control $u(t)$, and the parameter $\pi$ is considered. The admissible state, control, and parameter are required to satisfy a vector differential constraint, a vector initial constraint, and a vector terminal constraint.

This problem is transformed into a mathematically simpler, unconstrained problem of minimizing a new functional, the performance index $R$, which involves the state, the control, the parameter, the Lagrange multiplier $\lambda(t)$ associated with the vector differential constraint, and the Lagrange multiplier $\mu$ associated with the vector terminal constraint. To obtain the minimum $R = 0$ of the performance index, a gradient algorithm is first developed. In order to achieve simplicity in programming and to bypass the explicit use of the second-order derivatives, the gradient algorithm is modified so that it becomes a pure, first-order method. For better convergence property, a conjugate-gradient algorithm is also developed.

Concerning the determination of the stepsize in these algorithms, a one-cycle cubic interpolation scheme is presented. Again, the explicit use of the second-order derivatives is avoided here.

Both the gradient algorithm and the conjugate-gradient algorithm are tested through several numerical examples. The results show that, while the gradient algorithm is slow in convergence, the conjugate-gradient algorithm displays a much better convergence characteristic.
References


