FUNCTION SPACE METHODS IN OPTIMAL CONTROL WITH APPLICATIONS TO POWER SYSTEMS

by

Sanjay Kumar Mitter

Thesis submitted for the Degree of Doctor of Philosophy in the Faculty of Electrical Engineering, Imperial College of Science and Technology, University of London

Imperial College,
London S. W. 7

July 1965

ABSTRACT

The mathematical description of a non-linear dynamical system may take various forms. A natural description is often in terms of difference or differential equations. A general theory is presented which permits us to determine the optimal control of such a system according to some suitable performance criterion.

Necessary and sufficient conditions of optimality are derived for a large class of optimal control problems. Ideas of mathematical programming, suitably extended to function spaces as well as Classical Calculus of Variations are used for this purpose. An iterative technique in function space is presented to synthesize optimal open loop and closed loop control programmes. The method presented is formally equivalent to Newton's Method in Function Space.

For many physical systems the state space is infinite dimensional. A theory is developed to solve minimization problems in Banach Spaces. The theory is illustrated by considering examples of finite and infinite dimensional control problems.

Two applications to power systems are presented.

To my wife Adriana,

and

to my mother

ACKNOWLEDGEMENTS

I am deeply indebted to my supervisor, Prof. J. H. Westcott for his interest, support and encouragement throughout the course of this work. This research has been supported by the Central Electricity Research Laboratories, Leatherhead, England. I would like to gratefully acknowledge their interest, support and above all the freedom they have allowed me while conducting this research. In this connection I would especially like to thank Mr. E. D. Farmer of the Central Electricity Research Laboratories.

The programming was done by Mr. D. Stone. I am grateful to him for the care and interest he took in this work.

I have profited immensely from the many discussions I had with Prof. Frank Stewart of Brown University, Providence, U.S.A. during his stay at Imperial College in 1963-64. I have also profited from discussions with Dr. M. A. Arbib (now at Stanford University, U.S.A.) and Prof. I. Lefkowitz of Case Institute of Technology, U. S. A.

Discussions with members of the control group at Imperial College have been invaluable. I would particularly like to mention Dr. J. D. Pearson (now at Case

Institute of Technology, U.S.A.), Dr. J. J. Florentin (now at Brown University, U.S.A.), Mr. G. F. Bryant, Mr. J. M. C. Clark, Mr. M. D. Levine, Mr. D. Q. Mayne, Mr. S. Shapiro and Mr. T. J. Stafford.

Finally this research would never have been done without the patience and understanding of my wife.

Computing was done at the C.E.G.B. Computer Centre and the Imperial College Computer Unit.

The accurate typing was done by Miss P. A. Mills and Mrs. S. A. Thomas.

CONTENTS

CHAPTER 1

INTRODUCTION

This thesis is concerned with the optimal control of non-linear dynamic systems. We shall assume that the mathematical description of the system to be controlled is known. The results of the thesis are very general in nature and thus applicable to a wide variety of systems ranging from sampled data to distributed parameter systems.

1.1. The Control Problem

A control system consists of a dynamic process, a controlling mechanism and paths between the controller and the process as shown in Fig. 1.1

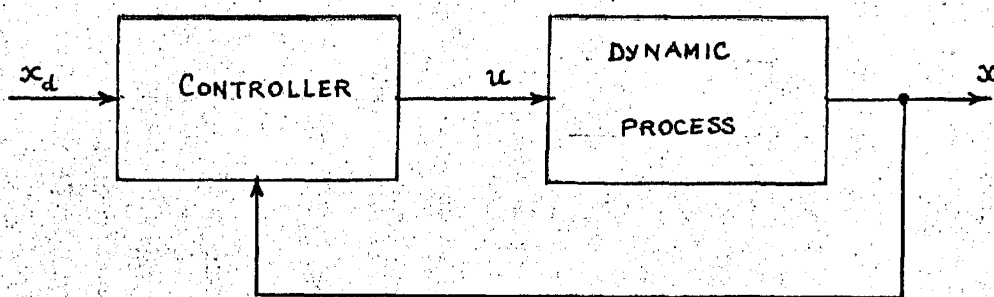


Fig. 1.1

The object of the controller is to apply inputs u(t) to the plant which will cause the plant to operate in a prescribed manner usually evidenced by the proximity of the plant outputs $x(t)$ to a desired set of outputs $x_d(t)$. The nearness of the plant outputs to the desired outputs will usually be measured by some suitable performance functional.

In addition to the plant inputs $u(t)$, there are generally external inputs which influence the dynamic behaviour of the system in an unpredictable fashion. The controller should thus effect suitable performance in the presence of disturbances. In order to do this, it may be necessary to identify the disturbances and the nature of the parameter variations. Further, for the control of a complex system, the controller will in general be a digital computer. Thus a more general model of a complex control system is shown in Fig. 2.

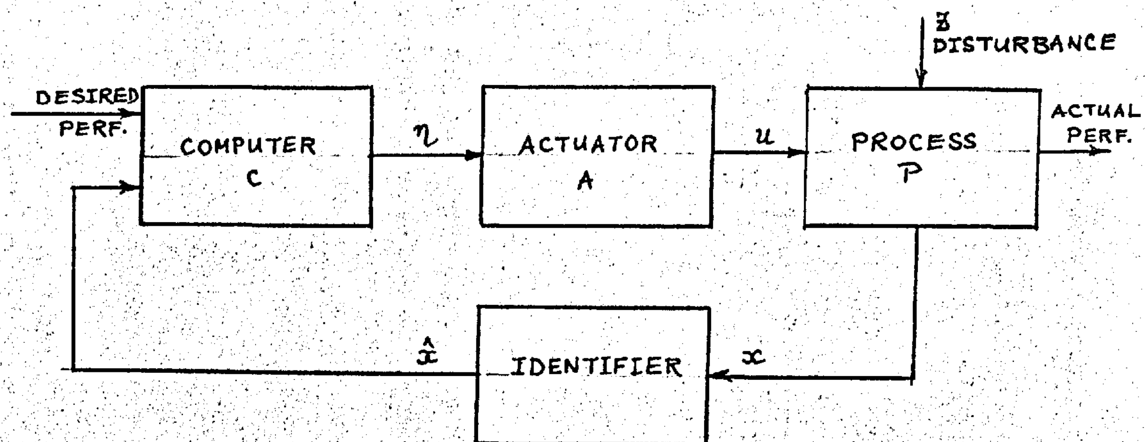


Fig. 1.2

In this thesis we shall assume that the nature of the disturbances and the parameter variations is completely determined.

1.2. The Need for Optimal Dynamic Control

Classical control theory was really concerned with the linear control of single input-single output linear time-invariant systems. This theory may be called the theory of servomechanisms. The design of the control system proceeded by using various Transform techniques. The attempts made to extend this theory to multivariable systems were not very successful. In fact methods like non-interacting control were in effect methods to reduce the multivariable system to n single variable systems.

We are however really interested in the control of complex processes. These processes are truly multivariable and in general non-linear. There is also a desire on the designer's part to obtain the best performance out of a system. Hence the need for optimal control. As a matter of fact this desire to obtain optimal control actually helps in the solution of the control problem, in the sense that a certain 'structure' is introduced into the problem. This allows a certain body of mathematics to be used. Further a firm basis for making approximations is provided, since in an actual design problem some kind of approximation will invariably have to be made.

1.3. The Concept of State[(1),(2)]

We shall generally be adopting the state description of dynamical systems, although our methods are not restricted to such a description. Here we give an intuitive idea (following Arbib) of what we mean by the 'state' of a system.

A dynamical system for us is something in which we put in certain inputs (control) and which itself puts out something at certain times. We usually think of a system S having an associated time scale T.

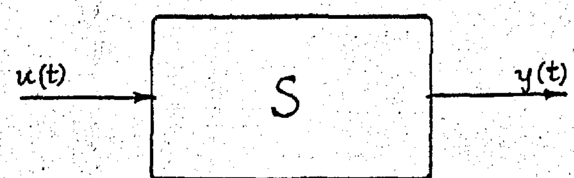


Fig. 1.3

At each instant $t \in T$, the system S receives an input u(t) and emits some output y(t). We now assume that there is a fixed set U of possible inputs and at any time $t \in T$ we choose an element u(t) from it. There is also a set of outputs Y, which includes all possible values of $y(t)$, $\forall\, t \in T$. The time scale T may be continuous or discrete (sampled) and the input and output sets will generally be linear spaces.

We would now like to arrive at the notion of a state. Now we may not be able to predict the output $y(t)$ by just knowing what the present input is. The past history of the system S may have altered S in such a way as to modify the output. In other words the output of S is a function of both the input and the past history of the system. We think of 'state' as being some attribute of the system which together with the input at that instant determines the output. But to qualify as the state of a system it must have one more property viz. that the states and inputs together suffice to determine the subsequent states.

For a large number of physical problems the state of the system is described by a set of first order nonlinear differential equations,

$$\dot{x}_i(t) = f_i(x_1(t) \ldots . x_n(t), t) \;; \quad x_i(t_o) = c_i \qquad i = 1, 2, \ldots . n \tag{1.1}$$

or in the equivalent integral equation form

$$x_i(t) = x_i(t_o) + \int_{t_o}^{t} f_i(x_1(\tau), \ldots . x_n(\tau), \tau)\, d\tau \quad i = 1, 2, \ldots . n \tag{1.2}$$

and we shall mostly be concerned with such systems. One of the reasons for adopting a differential equation description is the relative ease with which differential equations may be solved on a digital computer.

1.4. A General Formulation of the Problem

We now formulate a general optimal control problem.

We are given a dynamical system whose evolution of state is described by a non-linear operator equation given in explicit or implicit form

$$x = N(u) \tag{1.3}$$

$$\text{or } \mathbf{S}(x,u) = 0_x \tag{1.4}$$

Here x represents the state of the system and is an element of a function space $\mathcal{X}$ and u represents the control to be applied and is an element of another function space U and the non-linear operator N maps elements of the direct product $\mathcal{X} \times U$ to $\mathcal{X}$-space.

We are also given another operator equation

$$y = G(x) \tag{1.5}$$

where y is an element of the output function space Y and G is a non-linear operator which maps elements of the state space $\mathcal{X}$ to elements of the output space Y. 0_x represents the zero element of the state space $\mathcal{X}$. The problem of optimal control is to find the pair $(\hat{u},\hat{x})$

such that the functional $F(u,y)$ is a minimum. Here F is a non-linear operator from the product space $U \times Y$ to the real number space R. Since y is explicitly known in terms of x, we may write $F(u,y) = P(u,x)$ thereby eliminating y from the problem. The pair $(\hat{u},\hat{x})$ will generally be restricted to lie either in the interior of a set $\Omega \subseteq U \times \mathcal{X}$ or it may take values on the closure of the set Ω. The latter is the case when the pair $(\hat{u},\hat{x})$ has to satisfy inequality constraints.

We formulate the optimal control problem in such generality since most problems of optimal control can be cast in this form by suitably defining the underlying function spaces. This formulation thus includes problems of distributed parameter system as well as certain probabilistic systems.

The operators involved are assumed to have suitable continuity differentiability and boundedness properties.

An important sub-class of problems is:

$$\text{Minimise } P(x(t_o),u) = F(x(t_f),t_f) + \int_{t_o}^{t_f} L(x(t)u(t)t)dt \tag{1.6}$$

subject to the constraints

$$\dot{x}(t) = f(x(t),u(t),t) \; ; \; x(t_o) = c \tag{1.7}$$

$$R(x(t),u(t),t) \leq 0 \tag{1.8}$$

$$S(x(t),t) \leq 0 \tag{1.9}$$

$$G(x(t_f),t_f) = 0 \tag{1.10}$$

where x(.) and u(.) are vectors, F and L scalar-valued functions and R,S,G vector-valued functions.

For inequality constraints ≤ 0 means that each component of the vector satisfies the inequality constraint.

In order to derive the necessary conditions of optimality various continuity and differentiability assumptions have to be made. The inequality constraints also have to satisfy certain constraint qualifications. For these assumptions and conditions we refer the reader to the papers of Berkovitz(3),(4).

1.5. Review of the Available Results

At this point it seems appropriate to review some of the results that are available in the field of optimal control. We may subdivide this section into three sub-sections:

i) Existence and Uniqueness results

ii) Necessary and Sufficient Conditions of Optimality

iii) Feedback Solutions

iv) Computational Methods.

In part of this review we rely heavily on a paper by Berkovitz(5).

1.5.1. Existence and Uniqueness Results

From a mathematical standpoint, the first problem of optimal control is the problem of existence. That is, given the control problem, does there exist a lower bound for the performance functional P(t) and if so is this lower bound attained by some admissible control u^o. Closely related to this is the question of uniqueness; that is, if the minimum is attained by a control u^o, is this the only control that gives P(u) its minimum value or are there others?

The question of uniqueness is particularly important. For if there exist several ways of minimizing P(u) some of these may be easier to implement than others.

The problem that has been most studied mathematically is the time optimal control problem for linear systems, that is problems for which the system dynamics is represented by

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) \tag{1.11}$$

where A and B are $n \times n$ and $n \times m$ matrices, and it is required to bring the system from some initial state $x(t_o)$ to the origin of the x-space in minimum time. The control vector u must satisfy an amplitude constraint

$|u(t)| \leq 1$, $t \in [t_o, t^*]$ where t^* is the minimum time. A slightly different version of the same problem is to hit a moving target in minimum time.

The earliest results for this problem are due to Bushaw[6], who considered a very special problem and the first results of wider applicability are due to Bellman, Glicksberg and Gross[7] and Gamkrelidze[8]. The most general results available for this problem is due to La Salle[9]. La Salle showed that if there is an admissible control u that results in a trajectory x(t) hitting the moving target then there is an optimal control that is bang-bang i.e. u takes on only the values +1 or -1. For systems which La Salle called 'proper' he showed that this control is unique. This appears to be the first paper in which the concept of controllability and reachable sets was introduced. La Salle used a Liapunov Theorem on the range of a vector measure to prove some of his results. Extensive use of this theorem has been made in later work on Control theory. Neustadt[10] has considered the same problem with the added constraint

$$\int_{t_0}^{t_1} \varphi(u)dt \leq M$$

The most recent result for this problem is that of Halkin's[11] who proves La Salle's theorem for piecewise continuous functions.

For non-linear systems existence results have been obtained by Lee and Markus[12], Roxin[13] and Filipov[14]. Lee and Markus consider the problem of hitting a compact target set that is moving continuously over a fixed time interval. The control u is constrained to lie in a compact convex set in m-dimensional euclidean space and it contains the origin.

Their system dynamics is assumed to be of the form

$$\dot{x}(t) = g(x,t) + B(x,t)u(t) \tag{1.12}$$

and the performance functional is of the form

$$P(u,x(t_o)) = \int_{t_o}^{t_1} [L(x,t) + M(x,t)u(t)]dt \tag{1.13}$$

They show that if the set Ω of admissible control that result in the target set being hit is non-void, and if all trajectories resulting from admissible controls satisfies a uniform boundedness condition

$\|x(t)\| < \beta, \quad \beta < \infty$, then an optimal control exists.

Roxin considers the problem of minimising

$\int_{t_o}^{t_1} L(x,u,t)dt$ constrained to satisfy a non-linear differential equation $\dot{x} = f(x(t),u(t),t)$; $x(t_o) = c$. He assumes that u lies in a fixed compact set Ω in euclidean m-space. He assumes that for each (t,x), $f(t,x,u)$ maps Ω onto a convex set. He further assumes

$$\langle x(t), f(x(t),u(t),t)\rangle \leq K(\|x(t)\|^2 + 1)$$ where $\langle\ \rangle$ denotes inner product in euclidean n-space and $\|x(t)\|$ norm in euclidean n-space.

Under these assumptions, Roxin shows that the set of all points that can be reached by the system is closed. From this result one gets the existence of optimal controls.

Warga[14] suggests that a system that does not satisfy the convexity property should be relaxed by enlarging the set of allowed values of $\dot{x}(t) = f(x(t),\Omega,t)$ to the closure of the convex hull of $f(x(t),\Omega,t)$. He then shows that solutions of the relaxed problem can be uniformly approximated by solutions of the original problem.

If however the system is linear in x, that is

$$\dot{x}(t) = A(t)x(t) + \varphi(u,t)$$

and $$P(u,x(t_o)) = \int_{t_o}^{t_1} (\alpha(t)x(t) + \varphi_o(u,t))dt$$

then optimal controls exist if the set $\hat{\varphi}(\Omega,t) = (\varphi(\Omega,t), \varphi_o(\Omega,t))$ is only compact (and not necessarily convex). This result has been proved by Neustadt[15].

1.5.2. Necessary and Sufficient Conditions of Optimality

The necessary conditions of optimality may be derived using various approaches. The four main approaches are

i) Using Modifications of Classical Variational techniques

ii) Pontryagin's Maximum Principle

iii) Hamilton Jacobi or Bellman's Partial Differential Equation approach

iv) Reachable Set approach (Principle of Optimal Evolution).

Under certain strong assumptions all these results are equivalent. In the following we shall clearly indicate in what sense these results differ.

Method of Variational Calculus

The most general results appear to be those of Berkovitz(16),(17) who essentially applied the results of Bliss(18), McShane(19) and Hestenes(20) to obtain necessary conditions of optimality for the general control problem defined by relations (1.6) to (1.10). For simplicity, we assume that the constraints (1.9) and (1.10) are absent. It is convenient to define a function H by the relation,

$$H(t,x(t),u(t),\lambda_o,\lambda(t)) = \lambda_o L(x(t),u(t),t)+\langle\lambda,f(x(t),u(t),t)\rangle$$

where λ_o is a scalar and λ is an n-vector. The results of Berkovitz may be summarised in the following

<u>Theorem 1</u>. Let $\hat{u}$ be an optimal control in the class of admissible controls, let $\hat{K}$ be the corresponding trajectory and let $\hat{x}(t)$ be the function defining $\hat{K}$ on $[t_o, t_f]$. Then there exists a scalar $\lambda_o \geq 0$; n-dimensional vector $\hat{\lambda}(t)$ defined and continuous on $[t_o, t_f]$ and an r-dimensional vector $\hat{\mu}(t) \geq 0$ defined and continuous on $[t_o, t_f]$ except perhaps at values of t corresponding to corners of $\hat{K}$, where it possesses unique left and right hand limits such that the vector $(\lambda_o, \hat{\lambda}(t))$ never vanishes and such that the following conditions are fulfilled.

I. Along K the following equations hold:

$$\frac{d\hat{x}(t)}{dt} = f(\hat{x}(t), \hat{u}(t), t) \ ; \quad x(t_o) = c \tag{1.14}$$

$$\frac{d\hat{\lambda}(t)}{dt} = -D_x H(\hat{x}(t), \hat{\lambda}(t), \hat{u}(t), t) - (D_x R(\hat{x}(t), \hat{u}(t), t))^T \hat{\mu}(t) \tag{1.15}$$

$$D_u H \ (\hat{x}(t), \hat{u}(t), \hat{\lambda}(t), t) + [D_u R(\hat{x}(t), \hat{u}(t), t)]^T \hat{\mu} = 0 \tag{1.16}$$

$$\hat{\mu}_i(t) R_i(\hat{x}(t), \hat{u}(t), t) = 0 \ ; \quad i = 1, 2, \ldots r \tag{1.17}$$

$$\hat{\lambda}(t_f) = D_x F \ (\hat{x}(t_f), t_f) \tag{1.18}$$

(See section on Notation later in this chapter)

Along K the function H is continuous. The above equations are the Euler-Lagrange Equations of the problem.

II. For every element (t,x,u) of K and every u such that $u = u(t)$ for some admissible control u,

$$H(t,\hat{x},u,\hat{\lambda}_o,\hat{\lambda}) \geq H(t,\hat{x},\hat{u},\lambda_o,\hat{\lambda})$$

This is known as the Weierstrass Condition.

III. Let $I(t,x)$ denote the subset of indices $i = 1,\ldots.r$ such that $R_i(t,x,u) = 0$. Then at each point of K $\langle e,(D_u^2 H)e\rangle \geq 0$, for all vectors $e = \begin{pmatrix} e_1 \\ \vdots \\ e_m \end{pmatrix}$ satisfying

$$\sum_{j=1}^{m} \frac{\partial R_i}{\partial u_j} e_j = 0, \quad i \quad I.$$

If the system is normal then λ_o may be set equal to 1. The above results were proved by Berkovitz using the following slack variable technique to convert the inequality constraints to equality constraints

$$\frac{d}{dt}(\xi^2) = R(x,u,t) \; ; \quad \xi(t_o) = 0,$$ where ξ is an r-vector, and applying the theorems of Bliss and McShane.

Pontryagin's Maximum Principle

In the work of Pontryagin and his co-workers(21) the control vector u is assumed to belong to a certain class D and constrained to lie in a fixed closed subset Ω of an arbitrary Hausdorff* Space. The class D should

*A linear topological space satisfying the Hausdorff separation axiom (distinct points have disjoint neighbourhoods) is called a Hausdorff Linear Space.

satisfy the following three conditions

i) All the controls $u(t)$, $t \in [t_o, t_f]$ which belong to the class D are measurable and bounded

ii) If $u(t)$, $t \in [t_o, t_f]$ is an admissible control, v is an arbitrary point of Ω and t' and t'' are numbers such that $t_o \leq t' \leq t'' \leq t_1$, the control $u_1(t)$, $t_o \leq t \leq t_1$, defined by the formula

$$u_1(t) = \begin{cases} v \text{ for } t' \leq t \leq t'' \\ u(t) \text{ for } t_o \leq t \leq t' \text{ or } t'' \leq t \leq t_f \end{cases}$$

is also admissible

iii) If the interval $t_o \leq t \leq t_f$ is broken up by means of subdivision points into a finite number of subintervals, on each of which the control $u(t)$ is admissible then this control is also admissible in the entire interval $[t_o, t_f]$. An admissible control considered on a sub-interval is also admissible. A control obtained from an admissible control $u(t)$, $t \in [t_o, t_f]$ by a translation in time [i.e. the control $u_1(t) = u(t-\alpha)$, $t \in [t_o+\alpha, t_1+\alpha]$ is also admissible.

The only other condition they require is that the functions $\frac{\partial f_i}{\partial x_j}(x_1, \ldots x_n, u)$ $i,j = 1,2,\ldots n$ are continuous on the direct product $X \times \bar{\Omega}$, where X is the n-dimensional state space and $\bar{\Omega}$ is the closure of Ω.

For this class of problems, Pontryagin's results are

$$\frac{dx}{dt} = f(x,u,t).$$

$$\frac{d\lambda}{dt} = -H_x \tag{1.19}$$

and $$H(t,x^o,u,\lambda_o,\lambda) \leq H(t,x^o,u^o,\lambda_o,\lambda) \tag{1.20}$$

where $\lambda_o \leq 0$. It is worth mentioning that these results are obtained by considering strong variations about a trajectory. Equations (1.19) and (1.20) constitute Pontryagin's Maximum Principle.

If the set Ω is a subset of Euclidean m-space with piecewise-smooth boundary then Ω may be represented by means of a set of inequality constraints

$$R_i(u(t)) \leq 0 \qquad i = 1,2,\ldots.r$$

If the control u(t) is bounded and measurable then an almost everywhere version of Berkovitz's results may be obtained by using another theorem of McShane[22]. In this sense Berkovitz's results and Pontryagin's results are equivalent.

These arguments were modified by Gamkrelidze[23] to obtain results where the state variable was constrained to lie in a region with piecewise smooth boundary; this leads to conditions of the form

$$S(x(t),t) \leq 0$$

Berkovitz also obtained results for this case using variational arguments.

<u>The Method of Dynamic Programming (Hamilton-Jacobi Theory)</u>

Let $v(t_o,x(t_o))$ be the value of the minimum of $P(u)$ as a function of initial time and position $(t_o,x(t_o))$. Invoking his Principle of Optimality, Bellman obtained the following partial differential equation for the value function $v(t_o,x(t_o))$,

$$-\frac{\partial v}{\partial t} = \min_{u \in \Omega} [L(x(t),u(t),t) + \langle D_x v, f(x,u,t)\rangle]$$

where the minimisation is carried over the set of admissible controls . The validity of this equation has been rigorously established by Berkovitz (under certain assumptions). Under suitable assumptions on the properties of the control u as a function of t and x, he has shown that v is piecewise c^2 on an appropriate region of (t,x) space. It can further be shown under these assumptions that if $\Lambda(t,x)$ is defined to be the value at t of the vector λ associated with the optimal trajectory through (t,x) then,

$$D_x v(t,\hat{x}) = \Lambda(t,x)$$

$$\frac{\partial v}{\partial t} = -f(t,\hat{x},\hat{u}) - \langle \Lambda(t,\hat{x}), f(\hat{x},\hat{u},t)\rangle$$

where $\hat{u} = \hat{u}(t)$ is the value of the optimal control corresponding to the point $(t,\hat{x})$. Hence,

$$\frac{\partial v}{\partial t}(t,\hat{x}) + H(t,\hat{x},\hat{u},\Lambda) = 0 \qquad (1.21)$$

which is a Hamilton-Jacobi Equation.

Combining equation (1.21) with the Weierstrass Condition we may write down Bellman's equation.

An analogous viewpoint is the Hamilton-Jacobi theory as developed by Caratheodory[24]. This was resurrected by Kalman[25] and applied to the control problem.

Reachable Set Theory

A theory parallel to the theory of Pontryagin and his co-workers has been developed by Halkin[26] and Roxin[27]. Halkin's results are more general. The viewpoint of Halkin is quite different from Pontryagin's. In the following we shall indicate some of the salient points of Halkin's work.

Let us first formulate the problem following Halkin. Consider an optimal control system

$\dot{x}(t) = f(x(t),u(t)t)$, where $x(t) \in E^n$, $u(t) \in E^r$ and $t \in [0,1]$.*

We are given a class F of bounded measurable functions $[0,1] \to \Omega$, where Ω is a closed subset of E^r, such that

i) $\forall u(t) \in F$, $f(x;u(t),t)$ is measurable on $[0,1]$ and C' in x and all solutions of the differential equation are bounded.

*See the section on notation later in this chapter.

ii) $u'(.) \in F$, $u''(.) \in F$, $\tau \in [0,1] \Rightarrow u(.) \in F$

and $u(t) = u'(t)$ on $[0,\tau]$

$= u''(t)$ on $(\tau, 1]$.

An initial manifold $h_i(x) = 0$, $i = 1,2,\ldots\ldots$ and a terminal manifold $g_i(x) = 0$, $i = 1,2,\ldots m$ are given. It is assumed that the gradients of each g and h exist and are linearly independent.

The problem of optimal control is to choose the pair $(\hat{x}(.),\hat{u}(.))$ such that $g_o(x(1))$ is a minimum.

For this problem the results of Halkin are the same as that of Pontryagin viz.

<u>Maximum Principle</u>

If $(\hat{x}(.),\hat{u}(.))$ is optimal, there exists an absolutely continuous function $\hat{\lambda}(.)$ on $[0,1]$, not zero, and differentiable almost everywhere such that

i) $\langle f(\hat{x}(t),\hat{u}(t),t), \hat{\lambda}(t)\rangle \geq \langle f(\hat{x}(t),u(t),t), \hat{\lambda}(t)\rangle$ almost everywhere on $[0,1]$ and $\forall u(.) \in F$.

ii) $\hat{\lambda}(.)$ satisfies the adjoint equation

$$\dot{\hat{\lambda}}(t) = -(D_x f(\hat{x}(t),\hat{u}(t),t))^T \hat{\lambda}(t)$$

iii) There exist constants $\alpha = (\alpha_1, \ldots \alpha_l)$, $\beta = (\beta_o, \ldots \beta_m)$ such that

a) $\hat{\lambda}(0) = (D_x h(\hat{x}(0))^T \alpha$ where h is an ℓ-vector

b) $\hat{\lambda}(1) = (D_x \bar{g}(\hat{x}(1)))^T \beta$ and $\bar{g} = (g_o, g)$ is an m+1-vector.

c) $\beta_o \geq 0$

Let us mention that the satisfaction of the terminal constraints complicates the proof considerably.

As in the work of Pontryagin and his co-workers the proof of the maximum principle ultimately relies on the separation theorems on disjoint convex sets in Euclidean space. But the way the convex sets are constructed is quite different in the work of Halkin.

In the work of Halkin two sets are constructed

a) W = set of all points accessible at time $t = 1$

b) $S(x) = \{y : g(y) = 0, \ g_o(y) > g_o(x)\}$.

The following lemma is then obtained:

Lemma 1: The set W and $S(x(1))$ are disjoint.

It is also not too difficult to show that $x(1)$ lies on the boundary of W, denoted by ∂W.

A particular linearisation is now introduced. The functions h_i, g_i and f are linearised so that

$$h_i(x) \to h_i(\hat{x}(0)) + D_x h_i(\hat{x}(0))(x-\hat{x}(0))$$

$$g_i(x) \to g_i(\hat{x}(0)) + D_x g_i(\hat{x}(1))(x-\hat{x}(1))$$

$$f(x(t),u,t) \to f(\hat{x}(t),u,t) + D_x f(\hat{x}(t),u,t)(x-\hat{x}(t))$$

giving rise to a linear optimal control problem corresponding to the original non-linear problem.

The two following Lemmas can then be obtained using certain generalisations of Liapunov theorems and the Brouwer fixed point theorem.

Lemma 2. W and $S(\hat{x}(1))$ disjoint implies that $\hat{W}$ and $\hat{S}(\hat{x}(1))$ are separable by a hyperplane.

Here $\hat{W}$ and $\hat{S}$ are sets similar to W and S for the linear problem.

Lemma 3. $\hat{x}(1) \in \partial W$ implies that $\hat{x}(1) \in \partial \hat{W}$.

Once these lemmas have been proved the maximum principle can be proved without too much difficulty. We refer the reader to Halkin's original work.

Sufficiency Results

A sufficiency theorem for optimal control was given by Berkovitz along the lines of the Classical Calculus of Variations. It does not however appear to be too useful. In Chapter 4 we shall give some sufficiency results, again along the lines of the Classical Calculus of Variations. Sufficiency theorems in terms of the Hamiltonian H was stated by Rozonoer[28] for linear systems in which the terminal time t_f was fixed and $x(t_f)$ free. Slightly more general results have been given by Lee[29].

Operator Theoretic Approach

We would also like to mention the results obtained by Balakrishnan[30] for the final value problem for linear systems in which the control is required to satisfy an energy-type of inequality constraint. Balakrishnan also indicates a steepest descent method for synthesizing the optimal control. A proof of convergence of the steepest descent method is also given.

1.5.3. Feedback Solutions

One of the most important concepts in automatic control is the concept of feedback. This means that the optimal control $\hat{u}$ should be obtained in the form $\hat{u}(t) = \hat{u}(\hat{x}(t), t)$. In other words the optimal control should be a function of the instantaneous state of the system. Using variational theory or the maximum principle the optimal control is obtained in the form $u(t) = u(x(t_o), t)$ where t_o is the initial time. In order to obtain a feedback form of solution it would be necessary to solve the Euler-Lagrange equations of the problem repeatedly. In the next section we shall indicate that the solution of the Euler-Lagrange equations leads to a complicated two-point boundary value problem. We might thus conclude that in general it is not possible to obtain the optimal control in feedback form.

If the dynamics of the system is linear, that is, of the form

$\frac{dx}{dt} = A(t)x(t) + B(t)u(t)$, where $A(t)$ and $B(t)$ are matrices of suitable order, then in certain special cases the optimal control may be obtained in feedback form. Of these cases, the most important from the practical point of view is the case where

$$P(x(t_o),u) = \frac{1}{2} \int_{t_o}^{t_f} (x^T(t)Q(t)x(t) + u^T(t)R(t)u(t))dt,$$

Q and R being bounded positive definite symmetric matrices. In this case the optimal control law may be written in the form,

$\hat{u}(t) = -R^{-1}(t)B^T(t)K(t)\hat{x}(t)$, where $K(t)$ satisfies the matrix Ricalti equation

$$\dot{K}(t)+K(t)B(t)K(t)+K(t)A(t)+A^T(t)K(t)+Q(t) = 0 \ ; \quad K(t_f) = 0$$

Kalman[31] has shown that this control law is stable. He has also shown that the Ricatti equation when solved backwards in time is stable.

For linear systems, feedback control laws have been obtained for certain time optimal control problems.

Finally we would like to mention that the dynamic programming point of view enables one in principle to construct feedback control laws. However the discrete form of dynamic programming leads to the now familiar computer storage problems and in continuous form the resulting partial differential equation cannot be solved in general.

1.5.4. Computational Methods

To simplify matters let us assume that the final time t_f is fixed, there are no terminal constraints to be satisfied and there are no inequality constraints present. To synthesize the optimal control it is then necessary to solve the following equations

$$\frac{dx(t)}{dt} = f(x(t),u(t),t) \; ; \quad x(t_o) \text{ given}$$

$$\frac{d\lambda(t)}{dt} = - D_x H(x(t),\lambda(t),u(t),t)$$

$$\lambda(t_f) = D_x F(x(t_f),t_f) .$$

We note that we are faced with a non-linear two point boundary value problem. A discussion of various aspects of this two-point boundary value problem will be given in Chapter 5 and here we content ourselves by mentioning the problem.

1.6. Nature Scope and Contribution of this Thesis

As we have indicated before, this thesis is concerned with the optimal control of non-linear dynamical systems. We shall now survey the contents of the various chapters and at the end indicate why the title 'Function Space Methods in Optimal Control' was chosen.

Chapter 2 is devoted to investigating in detail the relationship between mathematical programming (linear and non-linear programming) and discrete-time optimal control systems. Essentially we have used methods of mathematical programming to derive necessary and sufficient conditions of optimality for a class of discrete-time optimal control problems. We have also shown the relationship between the Discrete Maximum Principle and our results. Various erroneous results on the Discrete Maximum Principle have been published in the literature.[32][33] Our methods clearly show that to obtain the Discrete Maximum Principle some strong form of convexity assumption has to be made. An important aspect of the theory of mathematical programming is duality. In this chapter we prove certain duality results for optimal control problems. Duality has important applications in the decomposition of large scale systems. Finally we indicate how optimal control problems could be solved computationally as problems of mathematical programming. This may prove very useful in many cases since computer programmes exist for the solution of linear and non-linear programming problems. To the best of our knowledge this is the first complete treatment showing clearly the relationship between control and programming. Previously Zadeh and Whalen[34] indicated how certain linear time-optimal and fuel optimal control problems could be solved as linear programming problems. It is also our understanding that Prof. Rosen(35) has been working on somewhat similar lines (Prof. Rosen lectured on the solution of state-constrained problems by mathematical programming methods in the Control and Programming conference held in Colorado Springs, April 1965). In a discussion of Ringlee's(36) paper in the IFAC conference in Basel, 1963, we indicated that discrete-time optimal control problems could be solved as mathematical programming problems. One of the classical works in the field of mathematical economics is due to Samuelson, Dorfman and Solow(37). Mathematical economics relies heavily on mathematical programming. We feel that there is much to be gained in applying ideas from control theory to mathematical economics and vice versa.

The original motivation for investigating the relationship between optimal control and mathematical programming was provided by certain problems in the control of power systems. Roughly speaking the problem is as follows: Given the load demand for a power system, it is required to schedule generation to meet the demand such that the total cost of generation is minimised and the constraints of the system are not violated. An alternative way of looking at the problem is : find the optimum set points of the generators, subject to certain constraints such that the total cost of generation is minimised. In Chapter 3 we proceed to solve this problem using the results of Chapter 2.

Chapter 4 might be thought of as a preliminary to Chapter 5 where we consider second order computational methods. Second order necessary conditions and sufficient conditions of optimality have so far been neglected in the literature of control. These conditions however are important in second order computational methods and in the synthesis of neighbouring optimal feedback controls. In this chapter we present a detailed treatment of the Jacobi condition and conjugate points for a class of optimal control problems. The treatment was motivated by the recent book of Gelfand and Fomin[38].

Chapter 5 is concerned with the synthesis of optimal controls using second order methods. These methods may be thought of as extensions to the gradient methods proposed by Kelley[39] and Bryson[40]. Our results overlap to some extent results obtained by Merriam[41],[42]. In a visit to the U.S.A. in April 1965, Prof. Bryson of Harvard University pointed out to me that Kelley[43] and more recently Bryson[44] have also expressed similar ideas. In his work Merriam does not consider terminal constraints. Kelley's actual computational technique is quite different from ours. Bryson's derivation of results is also different from ours. A preliminary account of our results was given in a symposium on Optimal Control held at Imperial College in April 1964[45]. In this chapter we also present a discussion of the advantages and disadvantages of various computational methods for solving optimal control problems. Computer results are also presented.

There are many problems in optimal control where the system state vector is infinite dimensional. Examples of such systems are provided by distributed parameter systems, systems with pure time delay and stochastic systems. In Chapter 7 we present a variational theory for minimization problems in function spaces.

These results are extensions of the results of Kuhn and Tucker in non-linear programming to infinite dimensional spaces. The theory developed is then applied to certain representative control problems. We also indicate some successive approximation methods for the solution of such problems. This chapter is mathematical in nature and uses extensively results of Functional Analysis. The relevant mathematical background is summarised in Appendix B.

From a mathematical standpoint, it would have been more logical to put the results of Chapter 7 right at the beginning and then proceed to show how some of the results of the other chapters may be considered to be special cases. We have however preferred to adopt the reverse order. The problems of Chapter 2 have been essentially solved by Ordinary Calculus type methods, whilst the problems of Chapter 7 have also been solved by Differential Calculus type methods, but Differential Calculus in Banach spaces(46),(47). In problems of optimal control, these Banach Spaces are invariably suitable function spaces.

It is clear that we could have discussed approximation methods in the abstract framework of Functional Analysis. Most of the well known methods for minimizing functions of several variables have counterparts in function spaces and well known in Functional Analysis(48),(49). However, we are not just interested in indicating a computational algorithm and proving that the procedure converges, but in actually carrying out the computations on a digital computer and providing suitable control-theoretic interpretations. In our opinion, the main contribution of the work of Kelley and Bryson in solving optimal control problems was not the use of a gradient method itself, but to show how the gradients in function space could be computed relatively easily on a digital computer. Similarly the second order methods we use in Chapter 5 are more or less well known in Functional Analysis. What we consider important is their use in the manner we have shown to solve optimal control problems. This is also our justification of giving a formal presentation of the material in that chapter. We simply indicate the way we proceed to solve the problems on a digital computer.

It should by now be clear why we have chosen the title 'Function Space Methods in Optimal Control'. Mathematically the principal difference between continuous time optimal control problems and discrete time optimal control problems is that in the former we are dealing with problems in function space rather than in Euclidean Space.

We present two applications to power systems: i) Economic Scheduling of Power Generation ii) Solution of a boiler Problem (in Chapter 6). Some of the ideas presented in this thesis could be used as a design philosophy for the optimal control of a power system.

The thesis is written in a way that most of the chapters are essentially self contained. Chapters 2 and

7 are generally mathematically rigorous. However, wherever possible, we have tried to motivate the discussion by means of examples. In these two chapters often many technical assumptions have been made. These have been included to render the subsequent mathematical results correct. The reader may omit them if he so wishes. Ideas of mathematical programming have only recently been used in the field of engineering. Hence some of the results have been summarised in Appendix A. Appendix B, to some extent, covers the mathematical background necessary to read Chapter 7. The thesis could also be read in the following order: Chapter 4, Chapter 5, Chapter 6, Chapter 2, Chapter 3, Chapter 7.

1.7. Notation

We shall try to use a uniform notation throughout this thesis as far as possible.

Set theoretic notation is often used. A set is a collection of objects. The set X is written as $\{x_k\}$ where x_k represents the element of the set. The notation $x_k \in X$ means that x_k is an element of the set X. When a set X represents a set of point x having a particular property P(x), we write this as

$$X = \{x : P(x)\}$$

Thus the set of points

$$X = \{(x_1, x_2) : x_1^2 + x_2^2 \leq 1\}$$

represent all points which lie inside and on the circumference of the circle with radius unity and centre at the origin.

A <u>space</u> X is a set X together with a given topology on X. For our purposes, a space X is a set X with which we can associate the notion of distance function d(x,y). Thus in three dimensional E^3 we have the usual 'Euclidean distance'

$$d(x,y) = \sqrt{(x_1-y_1)^2 + (x_2-y_2)^2 + (x_3-y_3)^2}$$

In most of our work we shall be concerned with finite dimensional Euclidean spaces. When we write $x(.) \in E^n$, we mean x(.) is an n-vector.

A function or a mapping is to be thought of as a relation between two sets X and Y, say. We shall often write $f : X \rightarrow Y$. We mean given an element $x \in X$, there exists a $y \in Y$, such that $y = f(x)$. The function or mapping is to be clearly distinguished from its value.

We shall also use vector-matrix notation extensively. If we want to refer to a particular component of a vector x, this will be denoted by x^j, say. Capital letters will be used for matrices. The adjoint (transposed) matrix will be written as A^T, say. Dot indicates differentiation with respect to time; thus $\dot{x} = \frac{dx}{dt}$.

$\langle .,. \rangle_{E^n}$ represents inner product in n-dimensional Euclidean space. Thus $\langle x,y \rangle_{E^n} = \sum_{i=1}^{n} x^i y^i$. Usually we shall drop the suffix E^n.

Consider the scalar-valued function

$$f(x,u) = f(x',\ldots.x^n,u',\ldots.u^m)$$

$$D_x f(\bar{x},\bar{u}) = \left(\frac{\partial f}{\partial x'},\ldots,\frac{\partial f}{\partial x^n}\right)_{\substack{x=\bar{x}\\ u=u}}$$ i.e. a row vector.

Similarly $D_u f(\bar{x},\bar{u}) = \left(\frac{\partial f}{\partial u'},\ldots.,\frac{\partial f}{\partial u^m}\right)_{\substack{x=\bar{x}\\ u=u}}$

For a vector-valued function f(x,u) where f is an n-vector

$$D_x f(\bar{x},\bar{u}) = \begin{pmatrix} \frac{\partial f'}{x'} & \ldots\ldots & \frac{\partial f'}{x^n} \\ \frac{\partial f^n}{\partial x'} & \ldots\ldots & \frac{\partial f^n}{\partial x^n} \end{pmatrix}_{\substack{x=\bar{x}\\ u=u}}$$ i.e. an n × n matrix.

For the scalar valued function f(x,u), the total differential

$$Df(x,u).(\xi,\eta) = D_x f(x,u).\xi + D_u f(x,u).\eta$$

$$= \langle D_x f(x,u),\xi + \langle D_u f(x,u),\eta\rangle$$

ξ and η are the increments of x and u.

$$D^2_{xu} f(\overline{x},\overline{u}) = \begin{pmatrix} \frac{\partial^2 f}{u^1 \quad x^1} & \cdots\cdots & \frac{\partial^2 f}{u^m \quad x^1} \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial u^1 \partial x^n} & \cdots\cdots & \frac{\partial^2 f}{\partial u^m \partial x^n} \end{pmatrix}$$

CHAPTER 2

OPTIMAL CONTROL AND MATHEMATICAL PROGRAMMING

1. Introduction

Sampled-data control systems have attracted a certain amount of attention in the literature of automatic control(50). Recently it has been recognized that state-space and optimal control ideas allow us to investigate such systems in a unified and often very much simplified fashion. In this chapter we consider a non-linear discrete-time optimal control problem and proceed to solve it using methods of linear and non-linear programming. The basic theorem we use is the Kuhn-Tucker theorem which is an extension of the Lagrange Multiplier rule of Ordinary Calculus. We have summarized the necessary mathematical results in Appendix A.

The chapter may be divided into 9 sections. In Section 2 we state the problem. In Section 3 we investigate necessary and sufficient conditions of optimality. In Section 4 we indicate the relationship of our results with the discrete-maximum principle. Section 6 is devoted to duality results of optimal control problems. In Section 7 we make some comments on the discrete maximum principle and sufficiency results.

An important class of problems in control are linear control problems with a quadratic performance criterion. In Section 8 we solve such a problem using mathematical programming methods. In Section 9 we indicate how we might use mathematical programming methods for the computational solution of optimal control problems.

Before proceeding to the general theory, it is worth considering a simple example to motivate the subsequent development.

Consider a 1-stage one dimensional control problem:

$x_1 = a\,x_o + b\,u_o$, where x is the state variable and u the control variable and x_o is given.

It is required to minimise

$$\phi(x_1,u_o) = \tfrac{1}{2}(x_1^2 + u_o^2)$$

subject to the inequality constraints

$$\alpha \leq u_o \leq \beta.$$

Let us first assume that the inequality constraint is absent. Let the optimal u_o and x_1 be $\hat{u}_o$ and $\hat{x}_1$. Then we could obtain $\hat{u}_o$ and $\hat{x}_1$ by using the Lagrange multiplier rule of ordinary calculus. Thus, form

$$L(x_o,x_1,u_o,\hat{\lambda}_1) = \tfrac{1}{2}(x_1^2+u_o^2) - \hat{\lambda}_1(x_1 - ax_o - bu_o)$$

Differentiating with respect to u_o and x_1, and equating the derivatives to zero, we get

$$\frac{\partial L}{\partial x_1} = x_1 - \hat{\lambda}_1 = 0$$

$$\frac{\partial L}{\partial u_o} = u_o + b\,\hat{\lambda}_1 = 0$$

Thus $\hat{x}_1 = \hat{\lambda}_1$

$\hat{u}_o = -\,b\,\hat{\lambda}_1$ and hence $\hat{u}_o = -\,b(ax_o + b\hat{u}_o)$. Therefore,

$$\hat{u}_o = \frac{-ab}{1+b^2}\,x_o$$

We note that the control law is linear feedback law . When the saturation type inequality constraints, $\alpha \leq u_o \leq \beta$ are present, we can no longer apply the Lagrange multiplier rule directly. However the Kuhn-Tucker theorem of non-linear programming can now be used to obtain the requisite solution.

In this case, we have to form

$$L(x_o, x_1, u_o, \hat{\lambda}_1, \hat{\mu}_1, \hat{\nu}_1) = \tfrac{1}{2}(x_1{}^2 + u_o{}^2) - \hat{\lambda}_1(x_1 - ax_o - bu_o)$$

$$+ \hat{\mu}_1(u_o - \beta) + \hat{\nu}_1(\alpha - u_o).$$

Differentiating with respect to x_1 and u_o,

$$\frac{\partial L}{\partial x_1} = x_1 - \hat{\lambda}_1 = 0$$

$$\frac{\partial L}{\partial u_o} = u_o + b\hat{\lambda}_1 + \hat{\mu}_1 - \hat{\nu}_1 = 0$$

Hence $u_o + b(ax_o+bu_o) + \hat{\mu}_1 - \hat{\nu}_1 = 0$, and finally

$$\hat{u}_o = \frac{-ab}{1+b^2} x_o + \frac{\hat{\nu}_1-\hat{\mu}_1}{1 + b^2}$$

From the Kuhn-Tucker theorem we also obtain,

$$\hat{\mu}_1 \geq 0 \quad \hat{\nu}_1 \geq 0$$

$$\hat{\mu}_1(\hat{u}_o-\beta) = 0 \text{ and } \hat{\nu}_1(\alpha - \hat{u}_o) = 0$$

Clearly if $\alpha < u_o < \beta$, $u_o = \frac{-ab}{1+b^2} x_o$. This is the same solution as we obtained by the Lagrange Multiplier rule.

In the following pages we generalise these ideas to solve non-linear multivariable discrete-time optimal control problems. The various assumptions we have made are to ensure the existence and sometimes uniqueness of the various multipliers and to exclude certain pathological cases where the Kuhn-Tucker theorem does not hold.

2. Problem Statement

In this chapter the state vector will be an element x of a Euclidean Space E^n, the control vector will be an element u of a Euclidean Space E^m and time will assume the discrete values 0, 1, 2,....,k. The evolution of the system will be described by the difference equations,

$$x_{i+1} = x_i + f_i(x_i,u_i); \quad x_i(0) = c \text{ given } i = 0,1,2,\ldots,k-1\ldots \tag{2.1}$$

A certain subset $\Omega \subseteq E^{n+m}$ is given and for each i, i = 0,1,2 k-1, the pair (x_i,u_i) is constrained to belong to this set. The set Ω is defined by the inequalities,

$$R_i(x_i,u_i) \leq 0\ldots\ldots \tag{2.2}$$

R_i is a mapping from E^{n+m} to E^p. For each i = 0,1.2,...k-1 we shall assume that the vector functions f_i and R_i satisfy the following conditions:

a) The vector-valued functions f_i and R_i are defined on Ω and for every $(x,u) \in \Omega$, f_i and R_i are twice continuously differentiable with respect to x and u

b) The mapping

$$\begin{pmatrix} I & & & & \vdots & & \\ -(I+D_{x_1}f_1) & I & & & \vdots & D_{u_0}f_0 & \\ & -(I+D_{x_2}f_2) & I & & \vdots & & \\ & & \ddots & & \vdots & & \ddots \\ & & & -(I+D_{x_{k-1}}f_{k-1})\ I & \vdots & & D_{u_{k-1}}f_{k-1} \end{pmatrix}$$

is on to. In effect this assumption is a local controllability assumption for the k-stage problem.

c) The inverse of the mapping

$$\begin{pmatrix} I & & & & \\ -(I+D_{x_1}f_1) & I & & & \\ & -(I+D_{x_2}f_2) & I & & \\ & & \ddots & \ddots & \\ & & & -(I+D_{x_{k-1}}f_{k-1}) & I \end{pmatrix}$$

exists.

We further require that the terminal point x_k must lie on the manifold defined by

$$g(x_k) = 0 \tag{2.3}$$

where g is a twice continuously differentiable mapping from $E^n \to E^q$, $q \leq n$. We assume that g satisfies the following condition

d) The matrix Dg_{x_k} is non-singular.

The inequality constraints $R(x,u)$ must satisfy the following constraint qualification.

e) Let $\{R^i(x,u)\}$ $i = 0,1,2,\ldots..r$ be the subset of the constraints R^i for which each (x,u) lies on the boundary of the constraint set. We shall say that the function $R^i : E^{n+m} \to E^r$ is regular at a point $(\bar{x},\bar{u}) \in E^{n+m}$, if and only if for every $(\xi,\eta) \in E^{n+m}$, $(\xi,\eta) \neq 0$ such that the equality $(x,u) = (\bar{x}+\xi, \bar{u}+\eta)$ implies the inequality,

$$R^i(x,u) = R^i(\bar{x},\bar{u}) + D_x R^i(\bar{x},\bar{u}).(\xi,0) + D_u R^i(\bar{x},\bar{u}).(0,\eta) \leq 0$$

there exists a function $\psi : [0,1] \to E^{n+m}$ with the following properties,

i) $D\psi(t).\tau$ exists for $0 \leq t \leq 1$

ii) $(\bar{x},\bar{u}) = \psi(0)$

iii) $R^i[\psi(t)] \leq 0, \quad 0 \leq t \leq 1$

iv) $(\xi,\eta) = D\psi(0)$.

Two sequences $\hat{u} = (\hat{u}_o, \ldots \hat{u}_{k-1})$ and $\hat{x} = (\hat{x}_1, \ldots \hat{x}_k)$ are said to be optimal if they satisfy the conditions

(A) $x_{i+1} = x_i + f_i(x_i,u_i)$; $x_i(0) = c \qquad i = 0,1,2,\ldots k-1$ (2.4)

(B) $R_i(x_i,u_i) \leq 0 \qquad i = 0,1,2,\ldots k-1$ (2.5)

(C) $g(x_k) = 0$ (2.6)

and minimises

$$h(x_k) + \sum_{i=o}^{k-1} \phi(x_i,u_i) \qquad (2.7)$$

We assume

f) The scalar-valued function h is twice continuously differentiable with respect to x_k and the scalar-valued function ϕ is twice continuously differentiable with respect to x and u for every (x,u) .

3. Necessary and Sufficient Conditions of Optimality

Under the assumptions we have made we can deduce the necessary conditions of optimality by applying the Kuhn-Tucker Theorem (see Appendix A) to this problem.

We form the Lagrangian,

$$L(\hat{x},\hat{u},\hat{\lambda},\hat{\mu},\hat{\vartheta}) = \langle \hat{\vartheta}, g(\hat{x}_k)\rangle + h(\hat{x}_k) + \sum_{i=o}^{k-1} [\phi(\hat{x}_i,\hat{u}_i) - \langle\hat{\lambda}_{i+1}, \hat{x}_{i+1} - \hat{x}_i - f_i(\hat{x}_i,\hat{u}_i)\rangle + \langle\hat{\mu}_{i+1}, R_i(\hat{x}_i,\hat{u}_i)\rangle] \tag{2.8}$$

where $\hat{\lambda}_i$ is an element of E^n, $\hat{\mu}_i$ is an element of E^p and $\hat{\vartheta}$ is an element of E^q.

Differentiating L with respect to $\hat{x}_i$, $i = 1,2,\ldots,k-1$ and $\hat{u}_i$, $i = 0,1,\ldots k-1$ and equating $D_{x_i}L(\hat{x}_i,\hat{u}_i)$ and $D_{u_i}L(\hat{x}_i,\hat{u}_i)$ to zero we obtain,

$$\lambda_i + \lambda_{i+1} + D_x\phi(\hat{x}_i,\hat{u}_i) + (D_x f_i(\hat{x}_i,\hat{u}_i))^T\hat{\lambda}_{i+1} + (D_x R_i(\hat{x}_i,\hat{u}_i))^T\hat{\mu}_{i+1} = 0 \tag{2.9}$$

for $i = 1,2,\ldots k-1$

$$D_{u_i}\phi(\hat{x}_i,\hat{u}_i) + (D_{u_i} f_i(\hat{x}_i,\hat{u}_i))^T\hat{\lambda}_{i+1} + (D_{u_i} R_i(\hat{x}_i,\hat{u}_i))^T\hat{\mu}_{i+1} = 0 \tag{2.10}$$

for $i = 0,1,\ldots k-1$.

Differentiating with respect to $\hat{x}_k$, we obtain,

$$\hat{\lambda}_k = (Dg(\hat{x}_k))^T\hat{\nu} + Dh(\hat{x}_k) \tag{2.11}$$

We also obtain that the optimal u_i and x_i satisfy

$$\begin{aligned} &\hat{\mu}_{i+1} \geq 0 \\ &< \hat{\mu}_{i+1}, R_i(\hat{x}_i,\hat{u}_i)> = 0 \end{aligned} \tag{2.12}$$

for $i = 0,1,2,\ldots.k-1$.

Under the assumptions we have made the $\hat{\lambda}$'s are unique.

In order to prove <u>sufficiency</u> we have to impose further restrictions. We shall assume that

g) $f_i(x_i,u_i) = A_i x_i + B_i u_i$ (2.13)

where A_i and B_i are $n \times n$ and $n \times m$ matrices for each i. For the matrices A_i and B_i we further assume $\|A_i\| < \infty$ and $\|B_i\| < \infty$ where

$\|A_i\|^2 = \text{trace}\,[A_i A_i^T]$ and $\|B_i\|^2 = \text{trace}\,[B_i B_i^T]$ for each i.

h) $g(x_k) = Mx_k$ (2.14)

where M is an $n \times q$ matrix, and $\|M\| < \infty$

i) ϕ is strictly convex with respect to u_i for each i and convex with respect to x_i for each i

h is convex with respect to x_k

R_i is convex with respect to x_i and u_i for each i.

We shall now prove that under assumptions a) - i) if the sequences $\hat{x}_1,\ldots\hat{x}_k$ and $\hat{u}_o,\ldots,\hat{u}_{k-1}$ satisfy equations (2.9) - (2.12) then

$$h(\hat{x}_k)+\sum_{i=o}^{k-1}\phi(\hat{x}_i,\hat{u}_i)\le h(x_k)+\sum_{i=o}^{k-1}\phi(x_i,u_i),\ \text{for all } (x_i,u_i)$$

satisfying equation (2.13) and belonging to the constraint set .

Let

$$\hat{L} = L(\hat{x}_1,\ldots\hat{x}_k,\hat{u}_o,\ldots\hat{u}_{k-1},\hat{\lambda}_1,\ldots\hat{\lambda}_k,\hat{\mu}_1,\ldots\hat{\mu}_k,\hat{\vartheta}) \qquad (2.15)$$

$$\bar{L} = L(x_1,\ldots x_k,u_o,\ldots u_{k-1},\hat{\lambda}_1,\ldots\hat{\lambda}_k,\hat{\mu}_1\ldots\hat{\mu}_k,\hat{\vartheta}) \qquad (2.16)$$

From (2.8) and (2.16), using the fact that ϕ, h and R_i are convex in x and u, we get

$$L = \sum_{i=o}^{k-1} [\phi(x_i,u_i)-\langle\hat{\lambda}_{i+1},x_{i+1}-x_i-A_ix_i-B_iu_i\rangle+\langle\hat{\mu}_{i+1},R_i(x_i,u_i)\rangle]$$

$$+ h(x_k) + \langle\hat{\vartheta}, Mx_k\rangle$$

$$\ge \sum_{i=o}^{k-1} [\phi(\hat{x}_i,\hat{u}_i)+\langle D_{x_i}\phi(\hat{x}_i,\hat{u}_i),x_i-\hat{x}_i\rangle+\langle D_{u_i}\phi(\hat{x}_i,\hat{u}_i),u_i-\hat{u}_i\rangle$$

$$- \langle\hat{\lambda}_{i+1},x_{i+1}-x_i-A_ix_i-B_iu_i\rangle + \langle\hat{\mu}_{i+1},R_i(\hat{x}_i,\hat{u}_i)\rangle$$

$$+ \langle\hat{\mu}_{i+1},D_{x_i}R_i(\hat{x}_i,\hat{u}_i)(x_i-x_i)+D_{u_i}R_i(\hat{x}_i,\hat{u}_i)(u_i-\hat{u}_i)\rangle]$$

$$+ h(\hat{x}_k)+\langle Dh(x_k),x_k-\hat{x}_k\rangle+\langle\hat{\vartheta},M(x_k-\hat{x}_k)\rangle+\langle\hat{\vartheta},M\hat{x}_k\rangle$$

$$= \hat{L} + \sum_{i=1}^{k-1} [\langle\hat{\lambda}_i,x_i-\hat{x}_i\rangle-\langle\hat{\lambda}_{i+1},x_{i+1}-\hat{x}_{i+1}]+\langle\hat{\lambda}_k,x_k-\hat{x}_k\rangle \qquad (2.17)$$

In obtaining the inequality (2.17) we have used equations (2.9), (2.10) and (2.11).

But $\sum_{i=1}^{k-1} [\langle\hat{\lambda}_i, x_i-\hat{x}_i\rangle - \langle\hat{\lambda}_{i+1}, x_{i+1}-\hat{x}_{i+1}\rangle] + \langle\hat{\lambda}_k, x_k-\hat{x}_k\rangle = 0$

Hence $\bar{L} \geq \hat{L}$.

Using (2.12) we therefore get the desired result.

4. Relationship with the Discrete Minimum Principle

Recently there has been a lot of interest in the Discrete Maximum Principle and it has been shown that the Discrete Maximum (Minimum) Principle does not hold in general for non-linear systems. We shall now obtain the Discrete Minimum Principle for a restricted class of systems from the results we have obtained in Section 3.

We assume that assumptions (a) to (i) hold.

Let us define the function

$$H(x_i,u_i,\lambda_{i+1},\mu_{i+1}) = \phi(x_i,u_i)+\langle\lambda_{i+1},A_iX_i+B_iu_i\rangle+\langle\mu_{i+1},R_i(x_i,u_i)\rangle$$

$$i = 0,1,2,\ldots k-1 \qquad (2.18)$$

We shall show that

$$H(\hat{x}_i,\hat{u}_i,\hat{\lambda}_{i+1},\hat{\mu}_{i+1}) \leq H(\hat{x}_i,u_i,\hat{\lambda}_{i+1},\hat{\mu}_{i+1}) \quad i = 0,1,2,\ldots k-1.$$

Let us first note that using (2.18), equations (2.9) and (2.10) may be re-written as

$$\hat{\lambda}_i = \hat{\lambda}_{i+1}+D_{x_i}H(\hat{x}_i,\hat{u}_i,\hat{\lambda}_{i+1},\hat{\mu}_{i+1}) \qquad (2.19)$$

$$D_{u_i} H(\hat{x}_i, \hat{u}_i, \hat{\lambda}_{i+1}, \hat{\mu}_{i+1}) = 0 \qquad (2.\quad)$$

In view of the assumptions we have made H is convex in u_i.

Hence

$$H(\hat{x}_i, u_i, \hat{\lambda}_{i+1}) \geq H(\hat{x}_i, \hat{u}_i, \hat{\lambda}_{i+1}, \hat{\mu}_{i+1}) + \langle D_{u_i} H(\hat{x}_i, \hat{u}_i, \hat{\lambda}_{i+1}, \hat{\mu}_{i+1}), u_i - \hat{u}_i \rangle \qquad (2.21)$$

from which the result easily follows.

5. Summary of Results of Sections 3 and 4

It is perhaps worth summarizing the results of Sections 3 and 4.

Proposition 2.1. (Necessary Conditions) Let $\hat{x} = (\hat{x}_1, \ldots, \hat{x}_k)$ and $\hat{u} = (\hat{u}_o, \ldots \hat{u}_{k-1})$ be two sequences which are optimal for the problem formulated in Section 2 and let $\hat{x}$ and $\hat{u}$ satisfy assumptions a) - f). Then there exists a unique sequence of non-zero vectors $\hat{\lambda} = (\hat{\lambda}_1, \ldots \hat{\lambda}_k)$ and a sequence of non-negative vectors $\hat{\mu} = (\hat{\mu}_1, \ldots \hat{\mu}_k)$ and a vector $\hat{\nu}$ such that

$$\hat{\lambda}_i = \hat{\lambda}_{i+1} + D_{x_i}\phi(\hat{x}_i, \hat{u}_i) + (D_{x_i} f_i(\hat{x}_i, \hat{u}_i))^T \hat{\lambda}_{i+1} + (D_{x_i} R_i(\hat{x}_i, \hat{u}_i))^T \hat{\mu}_{i+1} = 0$$

$$i = 1, 2, \ldots k-1$$

$$\hat{\lambda}_k = (Dg(\hat{x}_k))^T \hat{\nu} + Dh(\hat{x}_k)$$

$$D_{u_i}\phi(\hat{x}_i, \hat{u}_i) + (D_{u_i} f_i(\hat{x}_i, \hat{u}_i))^T \hat{\lambda}_{i+1} + (D_{u_i} R_i(\hat{x}_i, \hat{u}_i))^T \hat{\mu}_{i+1} = 0$$

$$i = 0, 1, 2, \ldots k-1$$

$$\hat{\mu}_{i+1} \geq 0 \qquad\qquad i = 0,1,2,\ldots k-1.$$
$$\langle \hat{\mu}_{i+1}, R_i(\hat{x}_i,\hat{u}_i)\rangle = 0$$

Proposition 2.2. (Sufficient Conditions) Let the system be linear and let all the hypotheses of Theorem 2.1 be fulfilled. Further let assumptions g) - i) of Section 2 be satisfied. Then the conditions of Proposition 2.1 are also sufficient.

Proposition 2.3. (Discrete Minimum Principle) Let all the assumptions of propositions 2.1 and 2.2 be fulfilled. Define the Hamiltonian function as

$$H(x_i,u_i,\lambda_{i+1},\mu_{i+1}) = \phi(x_i,u_i)+\langle\lambda_{i+1},A_ix_i+B_iu_i\rangle+\langle\mu_{i+1},R_i(x_i,u_i)\rangle$$
$$i = 0,1,2,\ldots k-1$$

Then a set of necessary and sufficient conditions for $\hat{u}$ and $\hat{x}$ to be optimal are

$$\hat{\lambda}_i = \hat{\lambda}_{i+1}+D_{x_i}H(\hat{x}_i,\hat{u}_i,\hat{\lambda}_{i+1},\hat{\mu}_{i+1}) \qquad i = 1,2,\ldots k-1$$
$$\hat{\lambda}_k = (Dg(\hat{x}_k))^T + Dh(\hat{x}_k)$$
$$H(\hat{x}_i,\hat{u}_i,\hat{\lambda}_{i+1},\hat{\mu}_{i+1}) \leq H(\hat{x}_i,u_i,\hat{\lambda}_{i+1},\hat{\mu}_{i+1})$$

6. Some Comments

In showing that under certain assumptions the necessary conditions of optimality are also sufficient and in obtaining the Discrete Minimum Principle we have relied heavily on convexity. It is clear that a set of sufficient conditions for non-linear systems could be obtained by further assuming

j) $f_i(x_i,u_i)$ is convex in x_i and u_i for $i = 0,1,2,\ldots k-1$

$g(x_k)$ is convex in x_k

and k) $\hat{\lambda}_i \geq 0$ (i.e. each component of $\lambda_i \geq 0$) for

$i = 1,\ldots\ldots k$

$\hat{\vartheta} \geq 0$ (i.e. each component of $\hat{\vartheta} \geq 0$),

since $\mathcal{L}$ will then be convex in x and u.

The Discrete Minimum principle will also clearly hold. Unfortunately it does not seem very easy to obtain conditions under which $\hat{\lambda}_i$ will be greater than zero, where $\hat{\lambda}_i$ satisfies equation (2.9). We can however obtain sufficient conditions (possibly very restrictive) for the solution of equation (2.9) to be positive. The conditions are suggested by certain results in Bellman[51]. We investigate this in the next chapter when we consider continuous time dynamic systems.

In optimal control problems constraints of the form

$$S_i(x_i) \leq 0, \quad i = 1,2,\ldots k \qquad (2.22)$$

are often present. In the literature these constraints are known as state variable constraints. These constraints may be converted into constraints of the form $R_i(x_i,u_i)$ which we have considered, by writing

$$S_i(x_i) = S_i(x_{i-1}+f_{i-1}(x_{i-1},u_{i-1}))$$

$$= T_i(x_{i-1},u_{i-1}) \qquad i = 1,2,\ldots.k.$$

This corresponds to the intuitive idea that to satisfy a state constraint at time instant i, the control at time instant i-1 (at least) must be suitably chosen.

7. Duality

One of the most important aspects of the theory of linear and non-linear programming is duality[52],[53]. For continuous time optimal control problems the importance of duality was demonstrated by Pearson[54]. Duality theory also has important application in decomposing large scale systems[55]. In this section we develop a duality theory for a class of discrete-time optimal control problems using mathematical programming methods. The basic idea is to construct a maximisation problem corresponding to the given minimisation problem such that the value of the optimal performance function of the two problems is the same.

Let us first define the two problems.

Primal Problem. Choose two sequences $\hat{u} = (\hat{u}_o,\ldots.\hat{u}_{k-1})$ and $x = (\hat{x}_1,\ldots.,\hat{x}_k)$ such that the following conditions are satisfied

$$x_{i+1} = x_i + A_i x_i + B_i u_i \; ; \quad x_o = c \text{ given} \qquad (2.23)$$

for $i = 0,1,2,\ldots k-1$

$$R_i(x_i,u_i) \leq 0 \tag{2.24}$$

$i = 0,1,2,\ldots k-1$

$$M\, x_k = 0 \tag{2.25}$$

and the following performance function

$$\check{L} = h(x_k) + \sum_{i=o}^{k-1} \phi(x_i,u_i) \tag{2.26}$$

is minimised.

<u>II Dual Problem</u>. Choose sequences $\hat{u} = (\hat{u}_o,\ldots\hat{u}_{k-1})$, $\hat{x} = (\hat{x}_1,\ldots\hat{x}_k)$, $\hat{\lambda} = (\hat{\lambda}_1,\ldots\hat{\lambda}_k)$ and $\mu = (\hat{\mu}_1,\ldots\hat{\mu}_k)$ and the vector $\hat{\nu}$ such that the following conditions are satisfied:

$$\lambda_i = \lambda_{i+1} + D_{x_i}\phi(x_i,u_i) + A_i^T\lambda_{i+1} + (D_{x_i}R_i(x_i,u_i))^T\mu_{i+1} \;, \quad i = 1,2,\ldots k-1 \tag{2.27}$$

$$\lambda_k = M^T\nu + Dh(x_k) \tag{2.28}$$

$$D_{u_i}\phi(x_i,u_i) + B_i^T\lambda_{i+1} + (D_{u_i}R_i(x_i,u_i))^T\mu_{i+1} = 0 \tag{2.29}$$

$i = 0,1,2,\ldots k-1$

$$\mu_{i+1} \geq 0, \quad i = 0,1,\ldots k-1 \tag{2.30}$$

and the following performance function

$$\hat{L} = h(x_k)+\langle x_k,\lambda_k - M^T\nu - Dh(x_k)\rangle$$

$$+\sum_{i=o}^{k-1}[\phi(x_i,u_i)-\langle\lambda_{i+1},X_{i+1}-x_i-A_ix_i-B_iu_i\rangle+\langle\mu_{i+1},R_i(x_i,u_i)\rangle]$$

$$+\sum_{i=1}^{k-1}[\langle x_i,\lambda_i-\lambda_{i+1}-D_{x_i}\phi(x_i,u_i)-A_i{}^T\lambda_{i+1}-(D_{x_i}R_i(x_i,u_i))^T\mu_{i+1}\rangle]$$

is maximised. (2.31)

We shall assume that assumptions a) - i) are satisfied by the Primal problem. We then have the following Duality result:

If $\hat{u}$ and $\hat{x}$ minimize (2.25) subject to the constraints (2.22) - (2.24) then there exist sequences $\hat{\lambda}$, $\hat{\mu}$ and a vector $\hat{\nu}$ such that $\hat{u}$, $\hat{x}$, $\hat{\lambda}$, $\hat{\mu}$ and $\hat{\nu}$ maximise (2.31) subject to (2.24) - (2.30) and $\check{L} = \hat{L}$.

The proof of this result is quite simple. We know from the results in Section 3 that for the primal problem there exist sequences $\hat{\lambda}$, $\hat{\mu}$ and a vector $\hat{\nu}$ such that equations (2.27) - (2.30) are satisfied. We also have

$$\sum_{i=o}^{k-1}[\langle\hat{\mu}_{i+1}, R_i(\hat{x}_i,\hat{u}_i)\rangle] = 0.$$

It is easy to see that $\check{L} = \hat{L}$. We shall now show that $\hat{u}$, $\hat{x}$, $\hat{\lambda}$, $\hat{\mu}$ and $\hat{\nu}$ indeed maximise L. The proof is by contradiction. Let us assume that there exists sequences

u, x, λ, μ and a vector ν such that

$$\hat{L}(u,x,\lambda,\mu,\nu) > \hat{L}(\hat{u},\hat{x},\hat{\lambda},\hat{\mu},\hat{\nu}).$$

We first note using (2.23) - (2.30)

$$\hat{L}(\hat{u},\hat{x},\hat{\lambda},\hat{\mu},\hat{\nu}) = h(\hat{x}_k)+\sum_{i=o}^{k-1}[\phi(\hat{x}_i,\hat{u}_i)+\langle\hat{\mu}_{i+1},R_i(\hat{x}_i,\hat{u}_i)\rangle]$$

$$\geq h(\hat{x}_k)+\sum_{i=o}^{k-1}[\phi(\hat{x}_i,\hat{u}_i)+\langle\mu_{i+1},R_i(\hat{x}_i,\hat{u}_i)\rangle]$$

since $\sum_{i=o}^{k-1}\langle\hat{\mu}_{i+1},R_i(\hat{x}_i,\hat{u}_i)\rangle = 0$ and $\mu_{i+1} \geq 0$, $R_i(\hat{x}_i,\hat{u}_i) \leq 0$.

Hence $\hat{L}(u,x,\lambda,\mu,\nu) > \hat{L}(\hat{u},\hat{x},\hat{\lambda},\hat{\mu},\hat{\nu}) \geq \hat{L}(\hat{u},\hat{x},\hat{\lambda},\mu,\hat{\nu})$.

In view of our assumptions L is a convex function of u and x.

Hence,

$$\hat{L}(\hat{u},\hat{x},\hat{\lambda},\mu,\hat{\nu}) = h(\hat{x}_k) + \langle\hat{x}_k,\hat{\lambda}_k-M^T\hat{\nu}-Dh(\hat{x}_k)\rangle$$

$$+\sum_{i=o}^{k-1}[\phi(\hat{x}_i,\hat{u}_i)-\langle\hat{\lambda}_{i+1},\hat{x}_{i+1}-\hat{x}_i-A_i\hat{x}_i-B_i\hat{u}_i\rangle+\langle\mu_{i+1},R_i(\hat{x}_i,\hat{u}_i)\rangle]$$

$$+\sum_{i=1}^{k-1}[\langle\hat{x}_i,\hat{\lambda}_i-\hat{\lambda}_{i+1}-D_{x_i}\phi(\hat{x}_i,\hat{u}_i)-A_i^T\hat{\lambda}_{i+1}-(D_{x_i}R_i(\hat{x}_i,\hat{u}_i))^T\mu_{i+1}\rangle]$$

Using the convexity of L, that is writing

$$\phi(\hat{x}_i,\hat{u}_i)\geq\phi(x_i,u_i)+\langle D_{u_i}\phi(x_i,u_i),\hat{u}_i-u_i\rangle+\langle D_{x_i}\phi(x_i,u_i),\hat{x}_i-x_i\rangle$$

etc. and simplifying, we obtain

$$\hat{L}(\hat{u},\hat{x},\hat{\lambda},\mu,\hat{\vartheta}) \geq L(u,x,\lambda,\mu,\vartheta)$$

$$+ \sum_{i=0}^{k-1} [\langle D_{u_i}\phi(x_i,u_i)+B_i^T\lambda_{i+1}+(D_{u_i}R_i(x_i,u_i)^T\mu_{i+1},\hat{u}_i-u_i\rangle]$$

But $\hat{L}(u,x,\lambda,\mu,\vartheta) > \hat{L}(\hat{u},\hat{x},\hat{\lambda},\mu,\hat{\vartheta})$.

Hence $\sum_{i=0}^{k-1} [\langle D_{u_i}\phi(x_i,u_i)+B_i^T\lambda_{i+1}+(D_{u_i}R_i(x_i,u_i))^T\mu_{i+1},\hat{u}_i-u_i\rangle]<0$

which contradicts (2.29).

We have thus proved our duality result.

Unfortunately these Duality results are in general <u>not true for non-linear systems</u>. The reason for this is that for non-linear systems L will in general not be a convex function of u.

It is worth noting that in the Dual problem the inequality constraints take a particularly simple form, namely they are just non-negativity constraints.

There is a certain resemblance between Duality results of mathematical programming and Kalman's theorem on the Duality between optimal filtering and optimal regulation. We have not found any deep connection as yet.

8. Linear Optimal Control Problem with a Quadratic Performance Function and an Amplitude Constraint

We now apply the theory we have developed to the following problem:

Find sequences $\hat{x} = (\hat{x}_1, \ldots. \hat{x}_k)$ and $\hat{u} = (\hat{u}_o, \ldots, \hat{u}_{k-1})$ which satisfy the system equation

$$x_{i+1} = x_i + Ax_i + Bu_i, \quad x_o = c \text{ (given)} \quad i = 0,1,2,\ldots k-1 \tag{2.32}$$

where A and B are constant $n \times n$ and $n \times m$ matrices, and the inequality constraints

$$|u_i| \leq 1, \quad i = 0,1,2,\ldots.k-1 \tag{2.33}$$

and which minimize the performance function

$$\frac{1}{2} \sum_{i=o}^{k-1} [\langle x_{i+1}, Px_{i+1}\rangle + \langle u_i, Ru_i\rangle] \tag{2.34}$$

where P is an $n \times n$ constant positive semidefinite symmetric matrix and R is a constant $m \times m$ positive definite symmetric matrix.

To solve the problem we form the Lagrangian

$$L(x,u,\lambda,\mu,\nu) = \sum_{i=o}^{k-1} [\tfrac{1}{2}\langle x_{i+1}, Px_{i+1}\rangle + \tfrac{1}{2}\langle u_i, Ru_i\rangle$$

$$- \langle \lambda_{i+1}, x_{i+1} - x_i - Ax_i - Bu_i\rangle + \langle \mu_{i+1}, u_i - 1\rangle + \langle \nu_{i+1}, -u_i - 1\rangle] \tag{2.35}$$

Then an application of the results in Section 3 yields the following necessary and sufficient conditions of optimality

$$\hat{\lambda}_i = \hat{\lambda}_{i+1} - P\hat{x}_i + A^T\hat{\lambda}_{i+1} \; ; \quad i = 1,2,\ldots k-1 \qquad (2.36)$$

$$\hat{\lambda}_k = P\hat{x}_k \qquad (2.37)$$

$$R\hat{u}_i + B^T\hat{\lambda}_{i+1} + \hat{\mu}_{i+1} - \hat{\nu}_{i+1} = 0 \; ; \quad i = 0,1,\ldots k-1 \qquad (2.38)$$

$$\mu_{i+1} \geq 0, \quad \nu_{i+1} \geq 0 \quad i = 0,1,2,\ldots k-1 \qquad (2.39)$$

$$\langle \hat{\mu}_{i+1}, \hat{u}_i - 1 \rangle = 0 \; ; \quad \langle \hat{\nu}_{i+1}, -\hat{u}_i - 1 \rangle = 0 \qquad (2.40)$$

Let us now assume that the inequality constraints (2.33) were not operative. The solution of this linear problem with a quadratic performance function is well known. In particular it is known that the optimal control law is a linear feedback law.

In view of this, it is worth considering whether we can obtain the solution of the constrained problem from the known solution of the unconstrained one.

Let us first make the observation that since $\mu_{i+1} \geq 0$ and $\langle \hat{\mu}_{i+1}, \hat{u}_i - 1 \rangle = 0$, if $\hat{u}_i - 1 < 0$ we must have $\hat{\mu}_{i+1} = 0$. The same observation is true for $\hat{\nu}_{i+1}$.

Let $\bar{x} = (\bar{x}_1, \ldots \bar{x}_k)$, $\bar{u} = (\bar{u}_o, \ldots \bar{u}_{k-1})$, $\bar{\lambda} = (\bar{\lambda}_1, \ldots \bar{\lambda}_k)$ be the solution of the unconstrained problem, and let $j \leq k-1$ be the first instant of time when some components of $\bar{u}_i$ satisfy $\bar{u}_i - 1 \leq 0$ or $-\bar{u}_i - 1 \leq 0$. Then we set

$$\hat{u}_i = \bar{u}_i \quad , \quad i < j$$

$$\hat{x}_i = \bar{x}_i \quad , \quad i < j$$

$$\hat{\lambda}_i = \bar{\lambda}_i \quad , \quad i \leq j+1$$

$$\hat{\mu}_i = \hat{\nu}_i = 0 \quad , \quad i \leq j$$

i.e. for time $i \leq j$ the solution of the constrained and unconstrained parts is the same.

Let $\bar{u}_i^k$ be the k^{th} component of $\bar{u}_i$. We shall define three index sets,

$$I_1 = \{k : |\bar{u}_j^k| < 1\}$$

$$I_2 = \{k : \bar{u}_j^k - 1 = 0\}$$

$$I_3 = \{k : -\bar{u}_j^k - 1 = 0\}$$

If $k \in I_1$ set $\hat{u}_j^k = \bar{u}_j^k$, $\hat{\mu}_{j+1}^k = \hat{\nu}_{j+1}^k = 0$

If $k \in I_2$ set $\hat{u}_j^k = 1$, $\hat{\nu}_{j+1}^k = 0$

If $k \in I_3$ set $\hat{u}_j^k = -1$, $\hat{\mu}_{j+1}^k = 0$

Set $\hat{x}_{j+1} = \hat{x}_j + A\hat{x}_j + B\hat{u}_j$

We however have to guarantee that we can determine

$$\hat{\mu}_{i+1}^k \geq 0 \ , \quad k \in I_2 \text{ and } \hat{\nu}_{j+1}^k \geq 0 \ , \ k \in I_3.$$

from equation (2.38). If we can do this we can proceed to the next step.

9. Computational Considerations

Many problems of optimal control can be reduced to problems of linear and non-linear programming and then solved using standard techniques. We give some examples below.

Example 1. Consider the following discrete-time dynamic system

$$\begin{bmatrix} x^1_{i+1} \\ x^2_{i+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x^1_i \\ x^2_i \end{bmatrix} + \begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix} [u_i + d_i]$$

$$i = 0,1,2,\ldots k-1 \qquad (2.41)$$

Let the initial conditions be $x^1(0) = x^2(0) = 0$. d_i is to be thought of as a known bounded disturbance. We consider the problem of finding a sequence $\hat{u} = (\hat{u}_o, \ldots \hat{u}_{k-1})$, $|u_i| \leq 1$ $i = 0,1,\ldots k-1$ such that

$$\phi(u) = \max_{i \leq j \leq 2} \max_{i \leq i \leq k} |x^j_i (u,d)|$$

is minimised.

This is a typical minimax problem. From the control point of view, this is the problem of minimising the maximum deviation of the system state from the equilibrium state due to the presence of a disturbance.

Let us first note that the solution of equation

(2.41) may be written as

$$\begin{bmatrix} x^1{}_i \\ x^2{}_i \end{bmatrix} = \begin{bmatrix} 1 & i \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x^1{}_o \\ x^2{}_o \end{bmatrix} . + \sum_{\ell=1}^{i} \begin{bmatrix} (\frac{1}{2}+i-\ell) \\ 1 \end{bmatrix} (u_{\ell-1}+d_{\ell-1}) \qquad (2.42)$$

To fix ideas let $k = 2$ and $d_o = 2$, $d_1 = -2$.

Then from (2.42)

$$x^1{}_1 = \frac{1}{2} u_o + 1$$

$$x^2{}_1 = u_o + 2$$

$$x^1{}_2 = \frac{3}{2} u_o + \frac{1}{2} u_1 + 2$$

$$x^2{}_2 = u_o + u_1.$$

Introduce a scalar $c \geq 0$ such that

$$\left| x_i^j \ (u,d) \right| \leq c$$

Hence our original problem reduces to

Minimise c

Subject to $c \geq 0$

$$\frac{1}{2} u_o + 1 \leq c$$

$$u_o + 2 \leq c$$

$$\frac{3}{2} u_o + \frac{1}{2} u_1 + 2 \leq c$$

$$u_o + u_1 \leq c$$

$$\frac{1}{2} u_o + 1 \geq -c$$

$$u_o + 2 \geq -c$$

$$\frac{3}{2} u_o + \frac{1}{2} u_1 + 2 \geq -c$$

$$u_o + u_1 \geq -c$$

$$u_o \leq 1$$

$$u_o \geq -1$$

$$u_1 \leq 1$$

$$u_1 \geq -1$$

This is now a problem in Linear Programming and can be solved by standard techniques. Some typical results are shown in Fig. 2.1.

Example 2

Consider the discrete time system

$$x_{i+1} = A_i x_i + B_i u_i \quad ; \quad i = 0,1,2,\dots k-1 \tag{2.43}$$

$x_o = c$ given

Consider the problem of minimizing

$$\sum_{i=o}^{k-1} [\langle x_{i+1}, Px_{i+1}\rangle + \langle u_i, Ru_i\rangle] \tag{2.44}$$

subject to the constraints

$$|u_i| \leq 1 \quad i = 0,1,2,\dots k-1 \tag{2.45}$$

$$|x_i| \leq 1 \quad i = 1,2,\dots k \tag{2.46}$$

The solution of equation (2.43) may be represented as

$$x_i = \Phi_i x_o + \sum_{j=1}^{i} \Phi(i)\, \Phi^{-1}(j) B_{j-1} u_{j-1} \quad i = 1,2,\dots,k \tag{2.47}$$

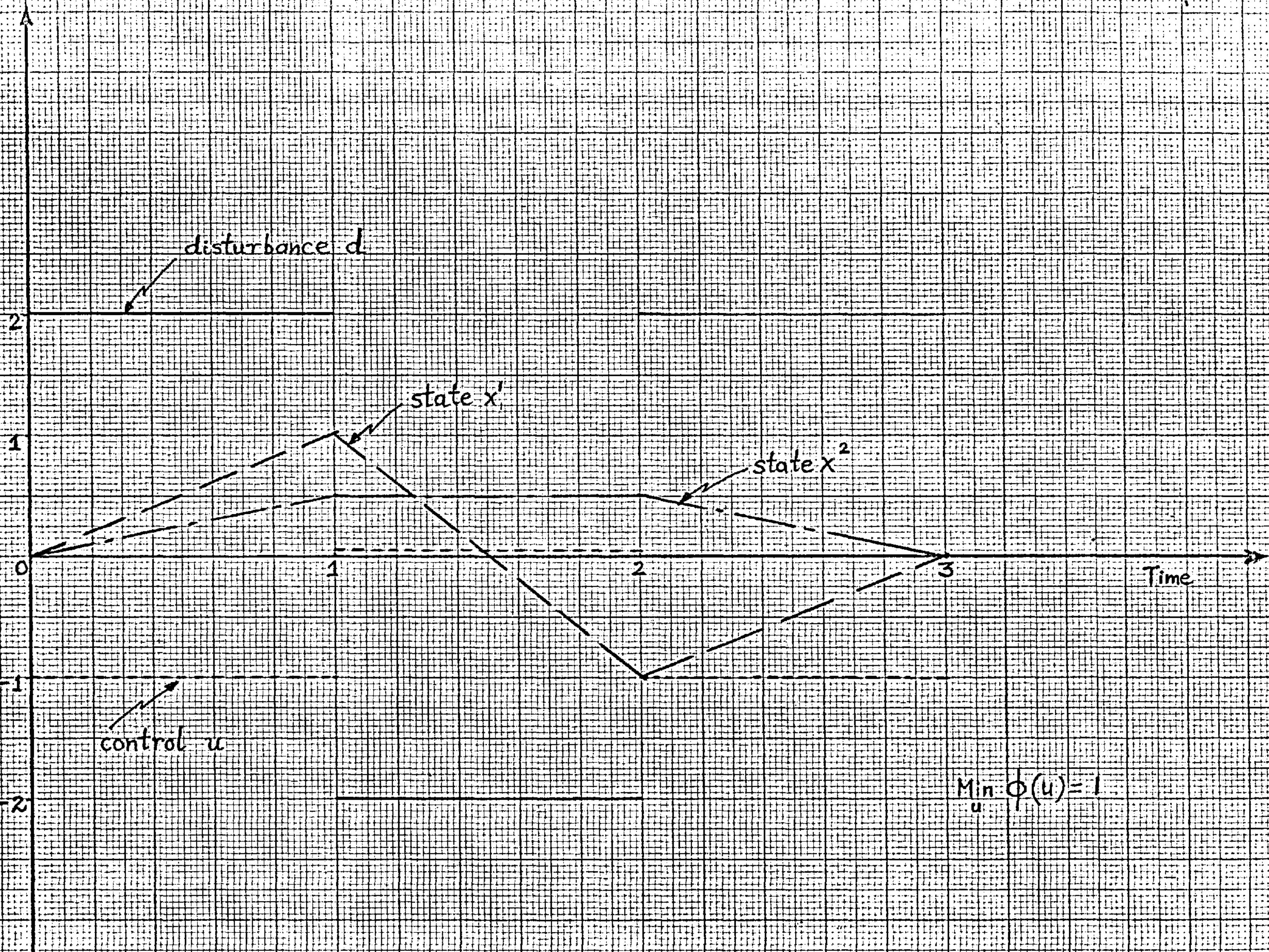


Fig 2.1 Results of 3-stage L.P. Problem

where $\phi_i = A_{i-1} A_{i-2} \ldots . A_o$

If we substitute (2.47) into (2.44) and (2.46) we have a non-linear programming problem in u-space. Let us mention however that if the number of control variables and the number of stages are large then the resulting non-linear programming problem will also be large. If the constrains are linear then library programs exist (Rosen's Gradient Projection Method)[56] to solve such problems.

Another powerful method of solving non-linear programming problems is the method of feasible directions due to Zoutendijk[57].

We shall not consider in detail the various existing computational methods for the solution of programming problems, but shall content ourselves with introducing a new primal-dual algorithm for solving a class of quadratic programming problems.

We are interested in solving the following problem: Find sequences $u = (u_o, \ldots . u_{k-1})$, $x = (x_1, \ldots . x_k)$ such that

$$\sum_{i=o}^{k-1} [\tfrac{1}{2}\langle x_{i+1}, Px_{i+1}\rangle + \tfrac{1}{2}\langle u_i, Ru_i\rangle]$$ is minimised subject

to
$$x_{i+1} = Ax_i + Bu_i , \quad i = 0,1,2,\ldots . k-1 \qquad (2.48)$$

$$x_o = c \text{ given}$$

$$Cx_i + Du_i \leq 0 \ , \quad i = 0,1,2,\ldots k-1 \qquad (2.49)$$

We assume that P is an n × n constant positive semi-definite matrix and R is an m × m constant positive definite matrix. We shall also assume that the usual assumptions on the system equations and the inequality constraints are satisfied.

From Section 6, the corresponding Dual problem is: find sequences $u = (u_o,\ldots u_{k-1})$, $x = (x_1,\ldots x_k)$, $\lambda = (\lambda_1,\ldots\lambda_k)$ and $\mu = (\mu_1,\ldots,\mu_k)$ such that

$$\sum_{i=o}^{k-1} [\tfrac{1}{2}\langle x_{i+1}, Px_{i+1}\rangle + \tfrac{1}{2}\langle u_i, Ru_i\rangle - \langle\lambda_{i+1}, x_{i+1}-Ax_i-Bu_i\rangle + \langle\mu_{i+1}, cx_i+Du_i\rangle]$$

$$+ \sum_{i=1}^{k-1} (\langle x_i, \lambda_i - A^T\lambda_{i+1} - Px_i - c^T\mu_{i+1}\rangle) + \langle x_k, \lambda_k\rangle$$

subject to $\lambda_i = A^T\lambda_{i+1} + c^T\mu_{i+1} + Px_i \quad i = 1,2,\ldots k-1$

$$\lambda_k = Px_k \qquad (2.50)$$

$$Ru_i + B^T\lambda_{i+1} + D^T\mu_{i+1} = 0 \quad i = 0,1,\ldots k-1 \qquad (2.51)$$

$$\mu_{i+1} \geq 0 \quad i = 0,1,\ldots k-1 \qquad (2.52)$$

The dual integrand can be simplified and written as:

$$= \langle x_k, Px_k\rangle + \langle x_o, A^T\lambda_1 + c^T\mu_1\rangle$$

$$- \sum_{i=o}^{k-1} [\tfrac{1}{2}\langle x_{i+1}, Px_{i+1}\rangle + \tfrac{1}{2}\langle R^{-1}(B^T\lambda_{i+1}+D^T\mu_{i+1}), (B^T\lambda_{i+1}+D^T\mu_{i+1})\rangle]$$

In the above we have eliminated u_i by using equation (2.51). In the dual problem let us now regard x and μ as the control variables and λ as the state variable.

The Algorithm

i) Let x_o = c the given initial condition of the primal problem

Guess $x_1, \ldots x_k$ and $\mu_1, \ldots \mu_k$, each $\mu \geq 0$

ii) Solve the recurrence equation (2.50) backwards.

iii) Let us adjoin the system equations (2.50) to the dual integrand by means of multiplier $y_1, \ldots, y_{k-1}$ obtaining

$$\langle x_k, Px_k\rangle + \langle x_o, A^T\lambda_1 + c^T\mu_1\rangle$$

$$- \sum_{i=o}^{k-1} [\tfrac{1}{2}\langle x_{i+1}, Px_{i+1}\rangle + \tfrac{1}{2}\langle R^{-1}(B^T\lambda_{i+1}+D^T\mu_{i+1}), (B^T\lambda_{i+1}+D^T\mu_{i+1})\rangle]$$

$$+ \sum_{i=1}^{k-1} [\langle y_i, -\lambda_i + A^T\lambda_{i+1} + c^T\mu_{i+1} + Px_i\rangle]$$

$$+ \langle y_k, \lambda_k\rangle$$

Solve the recurrence equations

$$y_o = x_o$$

$$y_1 = Ay_o - BR^{-1}[B^T(\lambda_1)_{old} + D^T(\mu_1)_{old}]$$

$$y_{i+1} = Ay_i - BR^{-1}[B^T(\lambda_{i+1})_{old} + D^T(\mu_{i+1})_{old}] \; ; \quad i = 1,2,\ldots.k-1$$

simultaneously calculating

$$(\mu_{i+1})_{new} = (\mu_{i+1})_{old} + Max(0, Cy_i - DR^{-1}D^T(\mu_{i+1})_{old} - DR^{-1}B^T(\lambda_{i+1})_{old})$$

$$i = 0,1,2,\ldots.k-1$$

Choose

$$(x_k)_{new} = (x_k)_{old} + \varepsilon P(x_k)_{old}$$

$$(x_i)_{new} = (x_i)_{old} + \varepsilon[P^T y_i - P(x_i)_{old}] \quad i = 1,2,\ldots.k-1 \; ,$$

where ε is a small positive number.

It can be shown without too much difficulty that the process converges.

The reason for solving the dual problem rather than the primal is the simplicity of the inequality constraints in the dual problem.

CHAPTER 3

ECONOMIC SCHEDULING OF POWER GENERATION

3.1 Introduction

A typical computer control system for controlling two power system areas is shown in Fig. 3.1. The symbols used are indicated in Table 3.2. Note that the design of the control system is conventional except for the use of a digital computer. One of the computations that the digital computer has to perform is to find the optimum set points $P_{s1},....P_{sn}$. This problem is known as the economic scheduling problem in the Power Systems literature. This is because the criterion for optimality is an economic one.

A typical load demand curve for a power system is shown in Fig. 3.3. In general, each generator has to satisfy maximum and minimum power limit constraints and maximum and minimum rate of rise of generation constraints. It is thus, in general, not possible to do the scheduling calculation at a particular time instant independently, without considering the load demand at a later instant. If this is done it might result in the load demand at a later instant not being met (due to the rate of rise constraints). We thus have a dynamic allocation problem to solve. In fact it is an optimal control problem.

3.2 Problem Statement

Let us consider an N-node network. At node i let

$p^i(t)$ be the active power at time t

$c^i(t)$ be the local active power demand at time t.

This is assumed to be known.

The cost of production L is a function of the active powers:

$$L = L(p^1(t), \ldots p^n(t)).$$

The total losses in the network g is a function of the power injected into the network, i.e.

$$g = g(p^1(t) - c^1(t), \ldots, p^n(t) - c^n(t))$$

Since the load demand has to be met we must satisfy,

$$h(p^1(t), \ldots, p^n(t), c^1(t), \ldots c^n(t)) = g - \sum_i (p^i(t) - c^i(t)) = 0 \tag{3.1}$$

Using Vector notation, the dynamic allocation problem to be solved is,

$$\underset{p(.)}{\text{Minimise}} \int_0^{t_f} L(p(t))dt \tag{3.2}$$

$$\text{subject to } h(p(t), c(t)) = 0 \tag{3.3}$$

$$\alpha \leq p(t) \leq \beta \tag{3.4}$$

$$\gamma \leq \dot{p}(t) \leq \delta \tag{3.5}$$

It is assumed that p(0) is known.

Let us introduce the variables $x(t) = p(t)$ and $u(t) = \dot{p}(t)$. We can then rewrite the above problem as an optimal control problem,

$$\underset{u(.)}{\text{Minimise}} \int_0^{t_f} L(x(t))dt \tag{3.6}$$

$$\text{subject to } \dot{x}(t) = u(t) \;; \quad x(0) \text{ known} \tag{3.7}$$

$$h(x(t),c(t)) = 0 \tag{3.8}$$

$$\alpha \leq x(t) \leq \beta \tag{3.9}$$

$$\gamma \leq u(t) \leq \delta \tag{3.10}$$

u is now regarded as the control variable and x the state variable.

It is also convenient to write the discrete version of the problem,

$$\underset{\{u_i\}}{\text{Minimise}} \sum_{i=1}^{K} L(x_i) \tag{3.11}$$

$$x_{i+1} = x_i + u_i \;; \quad i = 0,1,2,\ldots k-1, \; x_o \text{ known} \tag{3.12}$$

$$h(x_i,c_i) = 0 \qquad i = 1,2,\ldots\ldots k \tag{3.13}$$

$$\alpha \leq x_i \leq \beta \quad , \; i = 1,2,\ldots\ldots k \tag{3.14}$$

$$\gamma \leq u_i \leq \delta \quad , \; i = 0,1,2,\ldots k-1 \tag{3.15}$$

Subscript represents time as usual.

For some considerations later, we might view the discrete problem as a network

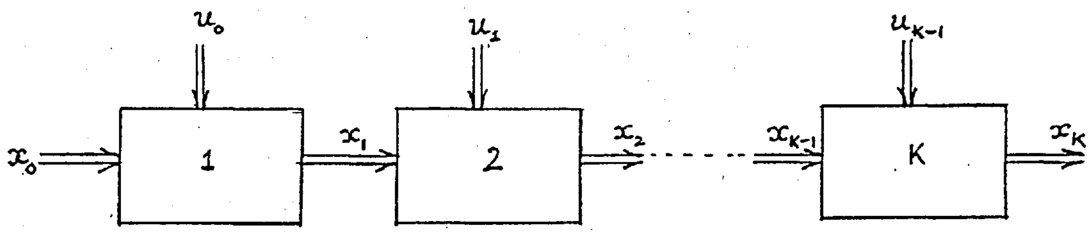


Fig. 3.4

3.3 Two Methods of Solution

Now that we have formulated the problem as a discrete optimal control problem, we could use the theory we have developed in the previous chapter to solve the problem. We can also use the computational methods we have indicated in the previous chapter. We shall now consider various methods of solving this particular problem in a little more detail.

3.3.1. Non-linear Programming Solution

One method of solution would be reduction to a vast non-linear programming problem in u-space. For example, the equations (3.12) could be written as

$$x_1 = x_o + u_o$$

$$x_2 = x_o + u_o + u_1$$

$$x_3 = x_o + u_o + u_1 + u_2$$

$$\cdots\cdots\cdots\cdots\cdots$$

$$x_k = x_o + u_o + \cdots + u_{k-1}$$

The constraints (3.14) could similarly be written in terms of u. If we now define the vector

$U = (u_o,\ldots,u_{k-1})^T$, it is clear that the problem has been reduced to

$$\underset{U}{\text{Minimise}}\ F(U)$$

$$G(U) = 0$$

$$AU \leq D,$$

where F and G are suitable functions, A a suitable matrix and D a suitable vector. Rosen's Gradient Projection method can then be applied.

When the number of generators is large and the number of stages is large, the resulting non-linear programming problem will also be very large.

3.3.2. Decomposition Technique

The main difficulty in solving the economic scheduling problem is the dimensionality of the state and control vectors. Hence some method of decomposing the problem into sub-problems, solving the sub-problems separately and co-ordinating the solutions to obtain the solution of the integrated problem seems to be attractive. Such a technique was proposed by Lasdon[83] to solve non-linear programming problems.

We shall illustrate this technique with a two-stage scheduling problem. To be mathematically correct, we have to assume that the constraints,

$$h(x_i, c_i) = 0$$

are linear. These constraints will be linear if the power losses are neglected or suitably linearised.

From (3.11) to (3.14) the two-stage problem to be solved is,

$$\underset{u_o, u_1}{\text{Minimise}} \; (L(x_1) + L(x_2))$$

subject to

$$x_1 = x_o + u_o$$

$$x_2 = x_1 + u_1$$

$$\alpha \leq x_o + u_o \leq \beta$$

$$\alpha \leq x_1 + u_1 \leq \beta$$

$$\gamma \leq u_o \leq \delta$$

$$\gamma \leq u_1 \leq \delta$$

$$h(x_1, c_1) = 0$$

$$h(x_2, c_2) = 0$$

Note that this problem is linear in u and we would have to apply the Discrete Minimum Principle to obtain the optimal control. It is also to be expected that the solution will be 'bang-bang' i.e. either the state variables will be at their limits and/or the control variables will be at their limits.

A 'bang-bang' type of solution for the control variables may be undesirable in practice and the constraints on u_o and u_1 may be approximated by soft constraints. Thus it may be more reasonable to meet the demand in the least square sense. With these two assumptions, let the modified cost function and problem be

$$\underset{\{u_o,u_1\}}{\text{Minimise}}\ \phi(u_o,u_1,x_1,x_2) = \phi_1(u_o,x_1) + \phi_2(u_1,x_2)$$

$$\text{subject to}\quad x_1 = x_o + u_o$$

$$x_2 = x_1 + u_1$$

$$x_o + u_o - \beta \leq 0$$

$$\alpha - x_o - u_o \leq 0$$

$$x_1 + u_1 - \beta \leq 0$$

$$\alpha - x_1 - u_1 \leq 0$$

From the duality theory developed in Chapter 2, the Dual problem is easily formed.

Consider the <u>sub-problems</u>:

a) Minimise $\phi_1(u_o,x_1) - \langle\lambda_1,x_1\rangle$

$$\text{subject to } x_o + u_o - \beta \leq 0$$

$$\alpha - x_o - u_o \leq 0$$

b) Minimise $\phi_2(u_1,x_2) - \langle\lambda_2,x_2\rangle + \langle\lambda_1,x_1\rangle$

$$\text{subject to } x_1 + u_1 - \beta \leq 0$$

$$\alpha - x_1 - u_1 \leq 0$$

(Essentially we have broken the connection between the two stages). Sub-problems a) and b) are solved for guessed values of λ_1 and λ_2, thereby obtaining u_o,u_1,x_1 and x_2 as functions of λ_1 and λ_2. We also obtain multipliers corresponding to the inequality constraints in the sub-problems. These serve as the μ's in the dual. It can be proved that the correct values of λ's and μ's are obtained when the dual problem is maximised. Therefore a steepest descent algorithm could be used to update the λ's and μ's.

This is a heuristic and highly simplified description of the decomposition method. A complete description is beyond the scope of the present thesis.

3.4 Example

A special case of the problem arises when the cost function is linear and the power losses are neglected. This problem is

$$\text{Minimise} \quad \sum_{i=o}^{k} \langle a, x_i \rangle$$

$$\text{subject to } x_{i+1} = x_i + u_i \quad ; \quad i = 0,1,2,\ldots.k-1, \quad x_o \text{ known}$$

$$\sum_{j=1}^{N} x_i^j = c_i \ , \quad i = 1,2,\ldots.k$$

$$\alpha \leq x_i \leq \beta \qquad i = 1,2,\ldots.k$$

$$\gamma \leq u_i \leq \delta \qquad i = 0,1,\ldots.k-1$$

This problem is completely linear and can be solved as a linear programming problem. Since there are a large number of inequality constraints it is more convenient to solve the Dual Linear Program.

A typical problem was solved using the L.P.90 programme. The data for the problem (provided by the Central Electricity Research Laboratories) is shown in Table 3.4. The solution of the problem is shown in Fig. 3.5. The solution was checked using the Primal-Dual Algorithm presented in Chapter 2 (by hand calculation). The solution exhibits the 'bang-bang' property indicated earlier. It may be observed that one or two cheap sets depart from their upper limits. This is to satisfy a constraint (a new set is being synchronised). Thus the fault may be said to lie in the plant ordering.

From the experience gained in solving this problem it would appear that a simple programme to solve the scheduling problem could be written. This programme would contain some minor modifications to the so-called 'merit-order' scheduling.

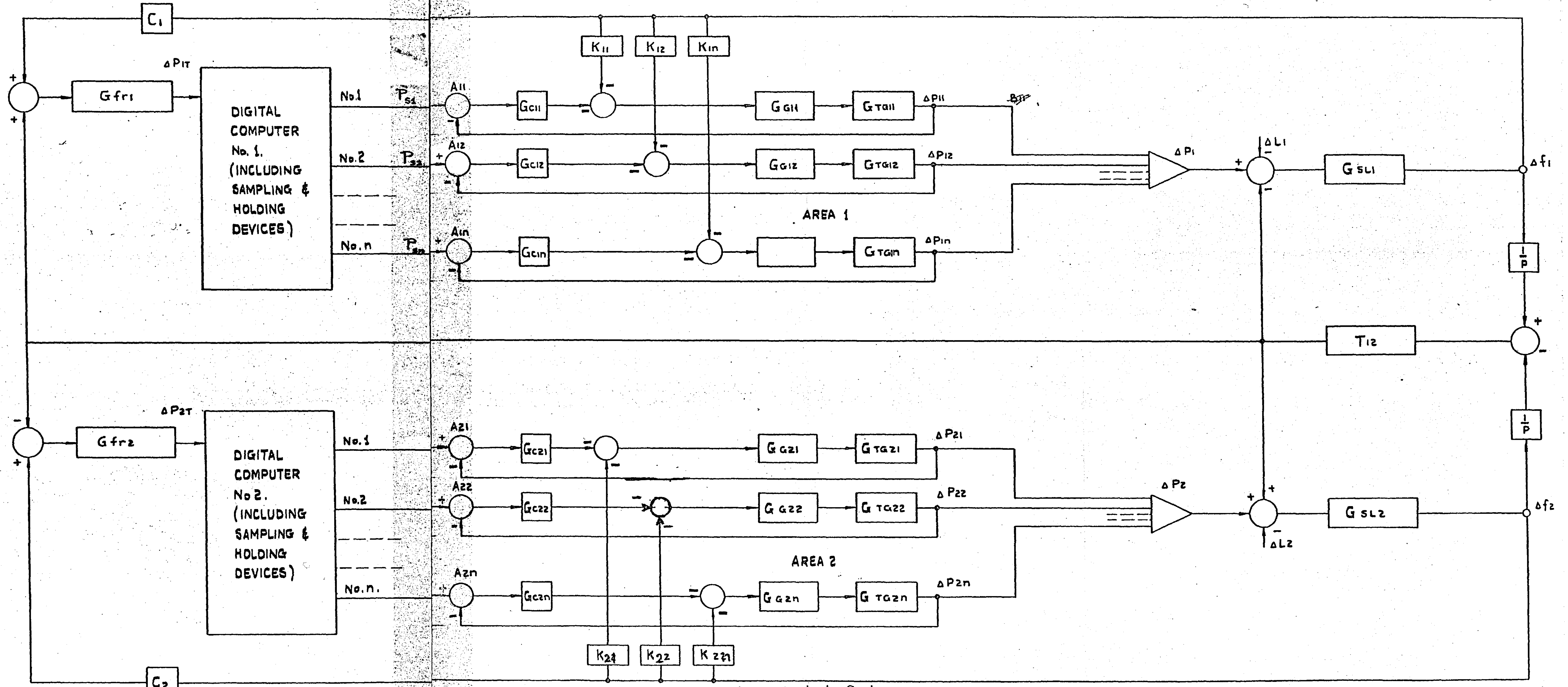


Fig.3.1 Typical Computer Control System

Symbols for Fig. 3.1

L_1 = incremental load in Area 1.

L_2 = incremental load in Area 2.

$G_{SL1} = \frac{A_1}{1+d_1 s}$ System lag in Area 1 - A_1, d_1 are constants

$G_{SL2} = \frac{A_2}{1+d_{2s}}$ System lag in Area 2

G_{G11}, G_{G12}, etc., = Dynamic response of governor actuators.

G_{TG11} = Transfer function of - Generator - Turbine Unit 1 in Area 1.

G_{TG12} = Transfer function of - Generator - Turbine Unit 2 in Area 1

G_{TG1n} = Transfer function of Generator - Turbine Unit in Area 1.

G_{TG21} = Transfer function of Generator - Turbine Unit 1 in Area 2.

G_{TG22} = Transfer function of Generator - Turbine Unit 2 in Area 2.

G_{TGn2} = Transfer function of Generator - Turbine Unit n in Area 2.

G_{fr1} = Transfer function of the frequency regulator in Area 1

G_{fr2} = Transfer function of the frequency regulator in Area 2

T_{12} = Synchronizing torque coefficient of tie line

K_{11}, K_{12}, etc. Gains of the frequency error feedback to be applied to turbine in Area 1.

K_{21}, K_{22}, etc. Gains of the frequency error feedback to be applied to the turbine in Area 2.

G_{C11}, G_{C12},		Compensation networks to be designed to suit individual generator characteristics.
G_{C21}, G_{C22}		
f1	=	frequency error in Area 1
f2	=	frequency error in Area 2

TABLE 3.2

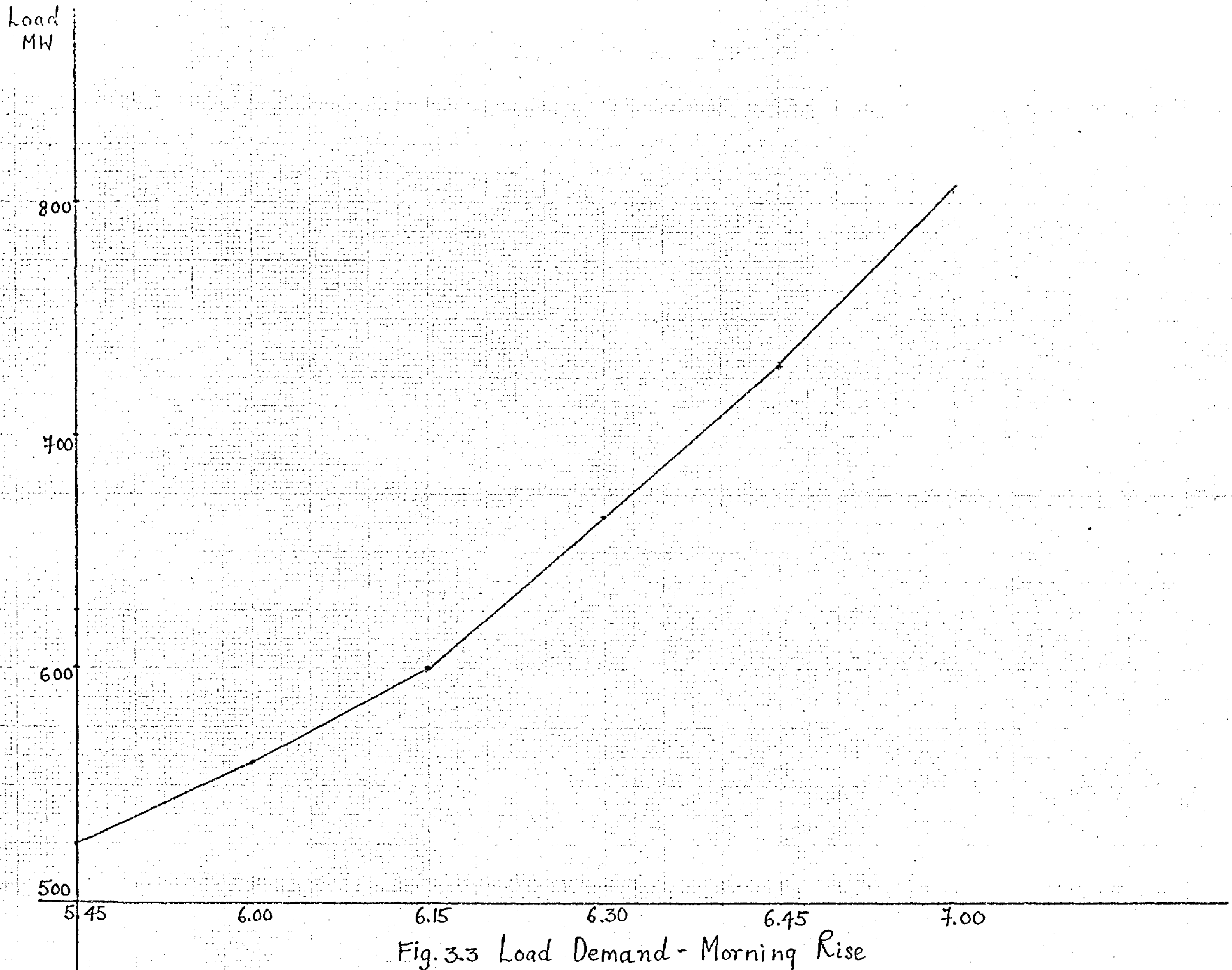


Fig. 3.3 Load Demand - Morning Rise

Data for Problem

List of Gens in Order of Merit	P	p	M	m	W	d	c	Initial State	Initial output g(o)
RICH 1	120	50	4	4	20	5	1.50	AV	120
BELV 5	120	75	12	12	80	7	1.74	AV	120
LITT 9	60	25	3	3	0	0	1.97	AV	60
LITT 6	60	25	3	3	0	0	2.00	AV	60
LITT	60	25	3	3	0	0	2.00	DS	
NFLT 2	120	45	6	5	12	5	2.23	AV	120
NFLT 3	120	45	6	5	12	5	2.23	DS	
BR.B 1	56	28	3	3	15	10	2.40	DS	
BR.B 2	56	28	3	3	15	10	2.43	DS	
BR.B 5	60	25	3	3	15	10	2.45	AV	25
BR.A 1	50	20	7	7	0	0	3.20	AV	20
Total Max Capacity	842							Total	525

P Maximum permissible output in MW

p Minimum permissible output in MW

M Maximum permissible rate of increase in MW/min

m Minimum permissible rate of decrease in MW/min

W Steady output when warming up in MW

d Time required to warm up in mins.

c Cost in £/MWhr

Loads

Time	MW
5.45	525
6.00	560
6.15	600
6.30	665
6.45	730
7.00	805

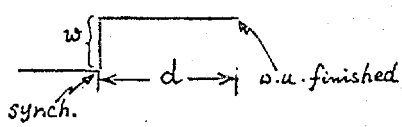
Synchronising times

BR.B1	5.50
LITT 8	6.10
NFLT	6.25
BR.B 2	6.25

Initial States:

AV - 'Available' i.e. warmed up and limited only by P,p,M and m.

DS - due to be synchronised i.e. to be synchronised at the time given in the list of sync. times. On syn. chronisings we assume the generator does



i.e. it holds a steady output W MW for d mins.

(This W/u pattern is very much simplified and it would not be too unreasonable to ignore it in preliminary calculations and assume that W = d = 0).

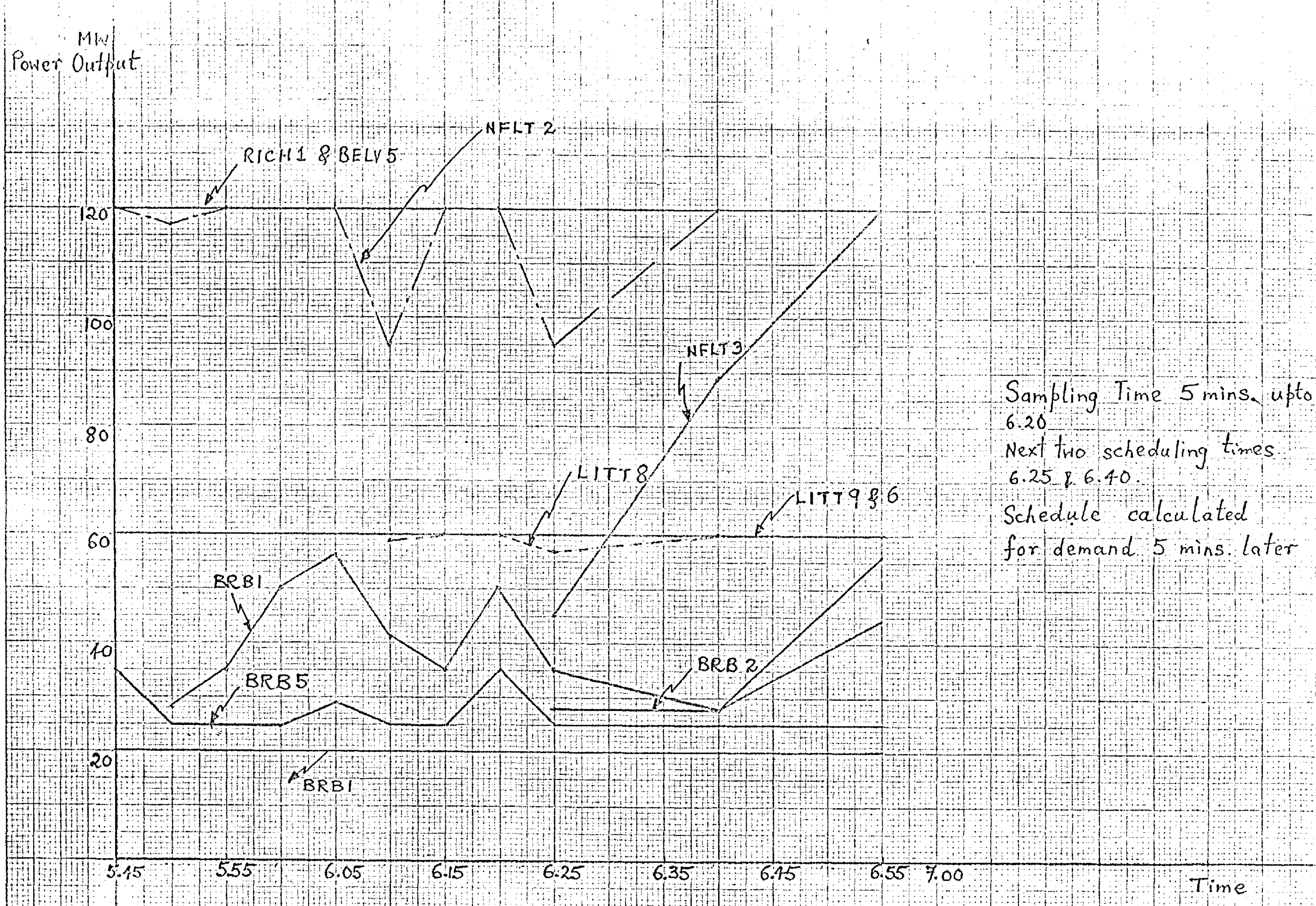


Fig.3.5 Typical Load Schedule For Morning Rise

CHAPTER 4

SECOND ORDER NECESSARY CONDITIONS AND SUFFICIENT CONDITIONS FOR A CLASS OF OPTIMAL CONTROL PROBLEMS

4.1 Introduction.

In the last two chapters we have been considering discrete time optimal control problems and solving them using mathematical programming methods. In this chapter and in Chapter 5 we want to consider a class of continuous time optimal control problems. For this class of problem, the first order necessary conditions of optimality are well known. However, not much attention has been paid to second order necessary conditions and sufficient conditions. It is our purpose to do this in this chapter.

In order to do this we have to investigate the second variation of the performance functional and the so-called Accessory minimization problem. In the literature of the Classical Calculus of Variations, the second order necessary conditions are the Legendre condition and the Jacobi condition.

Both these conditions are of great importance in second order successive approximation schemes such as the Second Variation Method and Newton's Method, and in the design of neighbouring optimal feedback control. In fact if the nominal trajectory is not sufficiently near the optimal trajectory these conditions will not be satisfied and the approximation procedure will not converge. Examination of these conditions suggests some modifications to the second variation method so that the procedure may be made to converge for any nominal trajectory. These details are presented in the next chapter.

It should be mentioned that the importance of the Jacobi condition has been recognised by Merriam.[58] He, however, does not present any detailed analysis. Our treatment is motivated by the recent book of Gelfand and Fomin.[59]

4.2 Problem Statement.

We consider the following Bolza Problem.

Minimize $P(x(t_o),u) = F(x(t_f),t_f) + \int_{t_o}^{t_f} L(x(t),u(t),t)\,dt$ (4.1)

subject to the constraints,

$$\frac{dx}{dt}(t) = f(x(t),u(t),t); \; x(t_o) = c \text{ , given} \qquad (4.2)$$

$$G(x(t_f),t_f) = 0; \quad t_f \text{ specified} \qquad (4.3)$$

where $x(.) \in E^n$, $u(t) \in E^m$ and $f: E^{n+m+1} \rightarrow E^n$ and $G: E^{n+1} \rightarrow E^p$, $p \leq n$.

We shall assume

i) All functions are twice continuously differentiable with respect to their arguments in the interval

$[t_o t_f]$.

ii) u(.) and x(.) belong to bounded open regions Ω and B of E^m and E^n. If $u(t) \in \Omega$ for all $t \in [t_o, t_f]$ then u is an <u>admissible</u> control.

iii) The matrix $D_x G(x(t_f), t_f)$ is non-singular.

iv) The system is locally completely controllable in the interval $(t_o, t_f]$ along any trajectory $\bar{x}$ corresponding to an admissible control $\bar{u}$; that is, for the linearized system

$$\frac{d}{dt}(x(t)-\bar{x}(t)) = D_x f(\bar{x}(t),\bar{u}(t),t)(x(t)-\bar{x}(t)) +$$

$$D_u f(\bar{x}(t),\bar{u}(t),t)(u((t)-\bar{u}(t)) \quad x(t_o)- \bar{x}(t_o) = 0$$

we have

$$\int_{t_o}^{t} \Phi(t,\tau) D_u f(\bar{x}(\tau),\bar{u}(\tau),\tau)(D_u f(\bar{x}(\tau),\bar{u}(\tau),\tau))^T \Phi^T(t,\tau) d\tau > 0 \qquad (4.4)$$

for all $t \in [t_o, t_f]$, where $\Phi(t, t_o)$ is the solution of

$$\frac{d\Phi(t_1 t_o)}{dt} = D_x f(\bar{x}(t),\bar{u}(t),t)\Phi(t,t_o);$$

$$\Phi(t\ ,t_o) = I \qquad (4.5)$$

In the following, to simplify the notation, we shall often write x when we really want to write x(t).

4.3 <u>First Order Necessary Conditions</u>.

For the Bolza problem we have formulated the first order necessary conditions may be derived in the usual way. The constraints are adjoined to the performance functionals by means of Lagrange multipliers $(\lambda_o, \lambda(t)) \neq (0,0)$ where $\lambda_o \geq 0$ is a constant and $\lambda(t)$ an n-vector, and ν is a p-vector.

$$\mathcal{L} = F(x(t_f),t_f) + \langle \nu, G(x(t_f),t_f) \rangle + \int_{t_o}^{t_f} [\lambda_o L(x,u,t) - \langle \lambda, \dot{x} - f(x,u,t) \rangle] dt$$

Define the Hamiltonian Function,

$$H(x,u.\lambda,t) = \lambda_o L(x,u,t) + \langle \lambda, f(x,u,t) \rangle \tag{4.6}$$

Then equating the first variation to zero and performing integration by parts, we obtain that the optimal trajectory and control satisfy

$$\dot{x} = f(x,u,t) = D_\lambda H(x,u,\lambda,t); \quad x(t_o) = c \tag{4.7}$$

$$\dot{\lambda} = -D_x H(x,u,\lambda,t); \quad \lambda(t_f) = D_x F(x(t_f),t_f) + D_x G(x(t_f),t_f)^T \nu \tag{4.8}$$

$$G(x(t_f),t_f) = 0 \tag{4.9}$$

$$D_u H(x,u,\lambda,t) = 0 \tag{4.10}$$

4.4 <u>Controllability and Normality</u>.

In Classical Calculus of Variations, if the problem if normal, we may set $\lambda_o = 1$ and this defines a unique set of multipliers λ.[*] We shall show that in view of our controllability assumption, the Accessory Minimization Problem is normal and hence our original problem is normal.[60]

The Accessory Minimization Problem is defined as[61]

Minimise $\frac{1}{2}\delta^2 P = \frac{1}{2}\langle\delta x(t_f), D_x^2\psi(x(t_f),t_f).x(t_f)\rangle$

$$+ \frac{1}{2}\int_{t_o}^{t_f} \langle D_u^2 H.\delta u, \delta u\rangle + \langle D_x^2 H.\delta x, \delta x\rangle + 2\langle D_{ux}^2 H\delta x, \delta u\rangle]dt \,\# \tag{4.11}$$

subject to $\delta\dot{x} = D_x f.\delta x + D_u f.\delta u; \quad \delta x(t_o) = 0$ (4.12)

$$D_x G(x(t_f),t_f)\delta x(t_f) = 0 \tag{4.13}$$

where $\psi = F + \langle\nu, G\rangle$.

<u>Proposition 1.</u>

<u>If the system is locally completely controllable in $[t_o, t_f]$ then the accessory minimization problem is normal and hence the original problem is normal.</u>

---

[*]<u>Def.</u>: A problem is said to be <u>abnormal</u> if we can find a set of λ's with λ_o=0 which satisfy equations (4.7)-(4.10).

#We have dropped the arguments of $D_u^2 H$ etc. They are calculated along the optimal control and trajectory. Also $\delta x = x-\hat{x}$ etc. where $\hat{x}(.)$ is the optimal trajectory.

<u>Proof</u>: We shall prove the above proposition by contradiction. Let $(\delta x, \delta\lambda, \delta u)$ be the optimal trajectory and the optimal control for the accessory minimization problem. Hence they satisfy the Euler-Lagrange equations of the problem. These are

$$\delta\dot{x} = D_x f.\delta x + D_u f.\delta u; \quad \delta x(t_o) = 0 \tag{4.14}$$

$$\delta\dot{\lambda} = -\lambda_o(D_x^2 H.\delta x + D_{xu}^2 H.\delta u) - (D_x f)^T.\delta\lambda \tag{4.15}$$

$$\lambda_o(D_u^2 H.\delta u + D_{ux}^2 H.\delta x) + (D_u f)^T.\delta\lambda = 0 \tag{4.16}$$

$$D_x G(x(t_f),t_f).\delta x(t_f) = 0 \tag{4.17}$$

$$\delta\lambda(t_f) = D_x^2 \Psi(x(t_f).t_f).\delta x(t_f) + (D_x G(x(t_f),t_f))^T.\delta\nu \tag{4.18}$$

Let us assume that the system is abnormal. Hence $\lambda_o = 0$. Hence from (4.15) and (4.16) we get,

$$\delta\dot{\lambda} = -(D_x f)^T.\delta\lambda \tag{4.19}$$

$$(D_u f)^T.\delta\lambda = 0 \tag{4.20}$$

Solving (4.19), we have

$\delta\lambda(t) = \Phi^T(t_f,t)\delta\lambda(t_f)$, where $\Phi(t,t_f)$ is the solution of $\frac{d\Phi}{dt}(t,t_f) = D_x f.\Phi(t,t_f)$; $\Phi(t_f,t_f) = I$ and therefore from (4.20),

$$(D_u f)^T \Phi^T(t_f,t)\delta\lambda(t_f) = 0 \tag{4.21}$$

But condition (4.21) expresses the fact that the rows of the matrix $\Phi(t_f,t)D_u f$ are linearly dependent. But a necessary and sufficient condition for the rows of $\Phi(t_f, t)D_u f$ to be linearly dependent is[62] that the Grammian matrix $\int_{t_o}^{t_f} \Phi(t_f, t)D_u f(D_u f)^T \Phi^T(t_f, t)dt = 0.$

This, however, contradicts our controllability assumption. Hence the accessory minimisation problem is normal and therefore by a theorem of Bliss the original problem is normal.

4.5 Second Order Necessary Conditions and Sufficient Conditions of Optimality.

For the Bolza problem we have formulated, the second variation is given by,

$$\delta^2\mathcal{L} = \langle\delta x(t_f), D_x^2\psi(x(t_f), t_f).\delta x(t_f)\rangle + \int_{t_o}^{t_f}[\langle D_x^2 H.\delta x, \delta x\rangle + \langle D_u^2 H.\delta u, \delta u\rangle + 2\langle D_{ux}^2 H.\delta x, \delta u\rangle]dt \quad (4.22)$$

It is well known in the Calculus of Variations that a necessary condition for a <u>weak minimum</u> of the Bolza problem is $\delta^2\mathcal{L} \geq 0$ and a set of sufficient conditions for a weak minimum is $\delta\mathcal{L} = 0$, $\delta^2\mathcal{L} > 0$ for all $\delta x \neq 0$, $\delta u \neq 0$, sufficiently small, where δx and δu are related by

$$\delta\dot{x} = D_x f.\delta x + D_u f.\delta u; \quad \delta x(t_o) = 0 \quad (4.23)$$

It is also well known that a necessary condition for $\delta^2\mathcal{L} \geq 0$, is $D_u^2 H \geq 0$. This is known as the weak form of Legendre's condition. In the following we shall assume that the strong form of Legendre's condition is satisfied, namely, $D_u^2 H > 0$. We then obtain the following:

Proposition 2: Let $D_u^2 H > 0$ and let $K(.)$ be an arbitrary positive definite symmetric matrix which satisfies the matrix Riccati equation.

$$\dot{K}(t) + K(t)B(t)K(t) + K(t)A(t) + A^T(t)K(t) + C(t) = 0 \qquad (4.24)$$

with the boundary condition $K(t_f) = D_x^2 \psi(x(t_f), t_f)$, where

$$\begin{aligned} A &= D_x f - D_u f . (D_u^2 H)^{-1} . D_{ux}^2 H \\ B &= -D_u f . (D_u^2 H)^{-1} . (D_u f)^T \\ C &= D_x^2 H - D_{xu}^2 H . (D_u^2 H)^{-1} . D_{ux}^2 H \end{aligned} \qquad (4.25)$$

Let the solution of this equation be defined everywhere in the interval $[t_o, t_f]$. Then $\delta^2 \mathcal{L} > 0$, for all $\delta x \neq 0$, $\delta u \neq 0$, sufficiently small where $\delta^2 \mathcal{L}$ is given by (4.22) and δx and δu are related by (4.23).

Proof:

$$\int_{t_o}^{t_f} \left(\frac{d}{dt} \langle K\delta x, \delta x\rangle\right) dt = \langle K\delta x, \delta x\rangle \Big|_{t_o}^{t_f} = \langle D_x^2 \psi(x(t_f)\ \delta x(t_f), \delta x(t_f)\rangle \qquad (4.26)$$

Hence, rearranging the expression for $\delta^2 \mathcal{L}$ we may write it as

$$\delta^2\mathcal{L} = \int_{t_o}^{t_f}(\langle D_u^2H.[\delta u + (D_u^2H)^{-1}.D_{ux}^2H.\delta x], \ \delta u + (D_u^2H)^{-1}.D_{ux}^2H.\delta x\rangle$$

$$+ \langle[D_x^2H - D_{xu}^2H.(D_u^2H)^{-1}.D_{ux}^2H].\delta x, \delta x\rangle +$$

$$\frac{d}{dt}\langle K\delta x, \delta x\rangle)dt \qquad (4.27)$$

But

$$\frac{d}{dt}\langle K\delta x, \delta x\rangle = \langle[\dot{K} + K.D_xf + (D_xf)^T.K]\delta x, \delta x\rangle +$$

$$2\langle(D_uf)^T.K.\delta x, \delta x\rangle \qquad (4.28)$$

From (4.27) and (4.28),

$$\delta^2\mathcal{L} = \int_{t_o}^{t_f}(\langle D_u^2H.[\delta u + (D_u^2H)^{-1}(D_{ux}^2H + (D_uf)^TK)\delta x],$$

$$\delta u + (D_u^2H)^{-1}(D_{ux}^2H + (D_uf)^TK)\delta x\rangle)dt$$

$$+ \int_{t_o}^{t_f}\langle(\dot{K} + KBK + KA + A^TK + C)\delta x, \delta x\rangle dt \qquad (4.29)$$

Since $D_u^2H > 0$ and K satisfies (4.24), we have

$$\delta^2\mathcal{L} \geq 0 \ .$$

$\delta^2\mathcal{L}$ can be zero if and only if

$$\delta u + (D_u^2H)^{-1}[D_{ux}^2H + (D_uf)^TK]\delta x = 0 \qquad (4.30)$$

But this is impossible, since from (4.23) and (4.30),

$$\delta\dot{x} = [D_xf - D_uf(D_u^2H)^{-1}(D_{ux}^2H + (D_uf)^TK)]\delta x; \ \delta x(t_o) = 0$$

implies $\delta x = 0$ everywhere in $[t_o, t_1]$ which contradicts our assumption. Hence the proposition.

The matrix Riccati equation (4.24) is precisely the same equation as that obtained by Merriam and Kalman when solving the linear-quadratic cost problem. We note that the accessory minimization problem is also a linear quadratic cost problem, but it is being used here for a different purpose. The control-theoretic significance of the Riccati equation will be brought into evidence later in this section.

If we combine the results of sections (4.3), (4.4) and (4.5), it is easily seen that we have obtained a set of sufficient conditions for the Bolza problem, namely,

A set of sufficient conditions for $\hat{x}$, $\hat{u}$ and $\hat{\lambda}$ to furnish a weak relative minimum for the Bolza problem formulated in Section 4.2 is that they satisfy

i) The Euler-Lagrange Equations,

ii) $D_u^2 H > 0$,

iii) The solution of the matrix Riccati equation (4.24) be defined everywhere in the interval $[t_o, t_f]$.

The last condition is equivalent to the Jacobi Condition of the Calculus of Variations.

We may also define a conjugate point in the following way: The time instant $t = t_c$ at which the solution of the Riccati equation becomes unbounded is called a point <u>conjugate</u> to the point $t = t_f$.

Let us now illustrate some of these ideas with an example from mechanics.[63]

Consider a simple harmonic oscillator, i.e. a particle of mass m oscillating about an equilibrium position under the action of an elastic restoring force. The particle has kinetic energy,

$$T = \tfrac{1}{2}m\dot{x}^2$$

and potential energy

$$U = \tfrac{1}{2}\alpha x^2$$

so that the action is

$$\tfrac{1}{2}\int_{t_o}^{t_f}(m\dot{x}^2 - \alpha x^2)dt \text{ which is to be minimised.}$$

We assume $x(t_o) = 0$. Let us introduce the variable u by means of the differential equation

$$\dot{x} = u; \quad x(t_o) = 0.$$

Our problem then is to minimise,

$$\tfrac{1}{2}\int_{t_o}^{t_f}(mu^2 - \alpha x^2)dt$$

subject to $\dot{x} = u; \quad x(t_o) = 0.$

The Hamiltonian is

$$H = \tfrac{1}{2}(mu^2 - \alpha x^2) + \lambda u, \text{ whence}$$

$$H_u = mu + \lambda = 0, \text{ giving } u = -\frac{m}{\lambda}$$

$$H_x = -\alpha x, \text{ giving } \dot{\lambda} = \alpha x; \quad \lambda(t_f) = 0$$

Also, $H_{uu} = m > 0$ and $H_{xx} = -\alpha$.

The Riccati equation is,

$$\dot{K} - \frac{K^2}{m} - \alpha = 0; \quad K(t_f) = 0.$$

Introduce the transformation,

$$K = -\frac{\dot{R}}{R} m$$

The Riccati equation becomes,

$$m\ddot{R} + \alpha R = 0.$$

The solution of this equation is given by

$R = C \sin(\omega t + \theta)$, where $\omega = \sqrt{\frac{\alpha}{m}}$ and C and θ are constants. Differentiating,

$$\dot{R} = C\,\omega \cos(\omega t + \theta), \text{ and hence,}$$

$$K = -\omega m \cot(\omega t + \theta).$$

Since $K(t_f) = 0$, we get,

$$\omega t_f + \theta = \frac{\pi}{2} \text{ and therefore } \theta = \frac{\pi}{2} - \omega t_f.$$

Hence, $K = -\omega m \cot(\frac{\pi}{2} + \omega(t - t_f))$

$$= -\omega m \tan(\omega(t_f - t)).$$

If $\omega(t_f - t) = \frac{\pi}{2}$, i.e., $t = t_f - \frac{\pi}{2\omega}$ then $K \rightarrow \infty$. Therefore if $t_f \geq t_o + \frac{\pi}{2}$, the Jacobi condition is violated.

So far we have not said anything about the necessity of the Jacobi condition for the second variation to be non-negative. It turns out that the Jacobi condition is also necessary.

4.6 Relationship with Dynamic Programming.

For the simple variational problem of minimising $\int_{t_o}^{t_1} L(x, \dot{x}, t)dt$ subject to $x(t_o) = a$, $x(t_1) = b$, Dreyfus(64) has obtained a matrix Ricatti equation for $D_x^2 v$, where v is the optimal return function. In order to derive this equation it is necessary to assume that $D_x^2 v$ is continuous. $D_x^2 v$ being a matrix, the continuity assumption implies that $D_x^2 v$ is bounded everywhere in $[t_o, t_f]$. Dreyfus also shows that the boundedness of the solution of the Ricatti equation is equivalent to the Jacobi condition of the Calculus of Variations.

Essentially the same arguments are valid for the Bolza problem. If $v(x(t_o), t_o)$ is the optimal return function, then invoking the Principle of Optimality, we obtain Bellman's partial differential equation,

$$-D_t v = \min_{u \in \Omega} \left[L(x,u,t) + \langle D_x v, f(x,u,t) \rangle \right] \tag{4.31}$$

If $D_x v = \lambda(x,t)$, then by taking characteristics of the partial differential equation, we obtain the usual canonical equations,

$$\dot{x} = f(x,u,t) = D_\lambda H$$

$$\dot{\lambda} = -D_x H$$

If we assume $D_x^2 v$ is continuous, then by straightforward differentiation the matrix Ricatti equation (4.24) can be obtained for $D_x^2 v$.

We can now see an interesting relationship between Dynamic Programming and Calculus of Variations. In Classical Calculus of Variations, the satisfaction of the Euler-Lagrange equations, the Legendre condition and the Jacobi condition are sufficient to embed the optimal trajectory in an <u>extremal field</u>. On the other hand, in the Dynamic Programming formulation we start by embedding the trajectory in a <u>field</u>(65) which by what we have shown is the extremal field of the Calculus of Variations. Then, as has been shown by Dreyfus(66) we can obtain the usual relations of the Calculus of Variations.

4.7 <u>Neighbouring Optimal Feedback Controls</u>.

For the simple variational problem, Dreyfus has also indicated how the Jacobi condition can be used to generate neighbouring solutions. We now consider <u>neighbouring optimal feedback control</u> for our Bolza problem. Assume there are no terminal constraints.

Supposing that the optimal control and optimal trajectory has been obtained by some method, it is desired to compute the optimal control and optimal trajectory for a slightly perturbed initial condition $x(t_o) + \delta x(t_o)$. The neighbouring optimal feedback control problem[67] is

$$\text{Minimise } \langle\lambda(t_o),\delta x(t_o)\rangle + \tfrac{1}{2}\langle\delta x(t_f),D_x^2F(x(t_f),t_f)\delta x(t_f)\rangle$$
$$+ \tfrac{1}{2}\int_{t_o}^{t_f}[\langle D_u^2H.\delta u,\delta u\rangle + \langle D_x^2H.\delta x,\delta x\rangle + 2\langle D_{ux}^2H.\delta x,\delta u\rangle]dt$$

subject to $\delta\dot{x} = D_xf.\delta x + D_uf.\delta u$; $\delta x(t_o)$ given.

Let us assume that the Ricatti equation (4.24) has a bounded solution everywhere in $[t_o,t_f]$. Then exactly as we did in Proposition 2, we may write the performance functional for the neighbouring optimal problem as

$$\langle\lambda(t_o),\ \delta x(t_o)\rangle$$
$$+ \int_{t_o}^{t_f}\langle D_u^2H.\ \ \delta u + (D_u^2H)^{-1}(D_{ux}^2H + (d_uf)^TK)\delta x\ \ ,$$
$$\delta u + (D_u^2H)^{-1}(D_{ux}^2H + (D_uf)^TK)\delta x\rangle dt$$

The minimum value of this performance functional is obtained when the integral is zero (we are assuming $D_u^2H>0$)

We must then have

$$\delta u = (D_u^2 H)^{-1}[D_{ux}^2 H + (D_u f)^T K]\delta x$$

This is, of course, the linear feedback law which we expected to obtain. If, however, the Riccati equation does not have a bounded solution everywhere in the interval of interest, then the feedback gain tends to infinity and we cannot synthesize linear feedback control.

4.8 The Accessory Minimization Problem.

The accessory minimization problem has been defined in Section 4.4 and its Euler-Lagrange equations are given by equations (4.14)-(4.16). Since the problem is normal these equations may be re-written as

$$\delta\dot{x} = A(t)\delta x + B(t)\delta\lambda; \quad \delta x(t_o) = 0 \tag{4.32}$$

$$\delta\dot{\lambda} = -C(t)\delta x - A^T(t)\delta\lambda; \tag{4.33}$$

$$\delta\lambda(t_f) = D_x^2\psi(x(t_f), t_f, \ ).\delta x(t_f) + (D_x G(x(t_f), t_f))^T\delta\nu$$

where the definition of A, B and C are given by (4.25).

Let

$$\Phi(t,t_o) = \begin{pmatrix} \Phi_{11}(t,t_o) & \Phi_{12}(t,t_o) \\ \Phi_{21}(t,t_o) & \Phi_{22}(t,t_o) \end{pmatrix} \quad \text{be the}$$

transition matrix of the linear system (4.32)-(4.33).

It can now be directly verified that

$$\delta\lambda(t) = \Phi_{22}(t)(\Phi_{22}(t_f))^{-1}D_x^2\psi(t_f)\delta x(t_f)$$
$$+ \Phi_{22}(t)(\Phi_{22}(t_f))^{-1}(D_x G(t_f))^T\delta\nu \tag{4.34}$$

where

$$\dot{\Phi}_{22}(t,t_o) = -C(t)\,\Phi_{12}(t,t_o) - A^T(t)\,\Phi_{22}(t,t_o) \qquad (4.35)$$

$$\dot{\Phi}_{22}(t,t_o) = A(t)\,\Phi_{12}(t,t_o) + B(t)\,\Phi_{22}(t,t_o) \qquad (4.36)$$

with suitable boundary conditions

and that $K(t) = \Phi_{22}(t)\Phi_{22}^{-1}(t_f)D_x^2\psi(t_f)\Phi_{12}(t_f)\Phi_{12}^{-1}(t)$ satisfies the matrix Riccati equation (4.25). Clearly if $\Phi_{12}(t,t_o)$ becomes singular at a point $t = t_c$, the solution of the Riccati equation will be undefined at that point. Thus a point $t = t_c$ at which $\Phi_{12}(t,t_o)$ becomes singular is a point conjugate to the point $t = t_f$.

Let us now indicate the rationale behind investigating the Accessory Minimization Problem. We know that if $(\hat{x}, \hat{u})$ is optimal, the second variation must be non-negative for all non-trivial variations $\delta x, \delta u$ sufficiently small. This naturally leads to investigating the minimum value of the second variation and the accessory minimization problem. We note that for the second variation to be non-negative over the interval $[t_o,t_f]$, it must be non-negative over every compact sub-interval $[t_1,t_2]$ of

$[t_o, t_f]$.

We shall now prove some simple results concerning the Accessory Minimization Problem.

From (4.16),

$$\delta u = -(D_u^2 H)^{-1}[D_{ux}^2 H.\delta x + (D_u f)^T.\delta\lambda] \tag{4.37}$$

Substituting in (4.11), we get

$$\tfrac{1}{2}\delta^2 P = \tfrac{1}{2}\langle\delta x(t_f), D_x^2\psi(x(t_f), t_f, \nu).\delta x(t_f)\rangle + \tfrac{1}{2}\int_{t_o}^{t_f}[\langle C\delta X, \delta x\rangle - \langle\delta\lambda, B\delta\lambda\rangle]dt \tag{4.38}$$

<u>Proposition 3</u>.

<u>For any $(\delta x, \delta\lambda)$ which satisfies equation (4.32) and (4.33), $\delta^2 P$ given by equation (4.38) has a value equal to zero</u>.

<u>Proof</u>.

$$\int_{t_o}^{t_f} \frac{d.}{dt}\langle\delta\lambda, \delta x\rangle = \langle\delta\lambda(t_f), \delta x(t_f)\rangle - \langle\delta\lambda(t_o), \delta x(t_o)\rangle$$

$$= \langle D_x^2\psi(x(t_f), t_f, \ )\delta x(t_f, \delta x(t_f)\rangle \tag{4.39}$$

Hence from (4.38) and (4.39)

$$\delta^2 P = \int_{t_o}^{t_f}[\langle C\delta x, \delta x\rangle - \langle\delta\lambda, B\delta\lambda\rangle + \frac{d.}{dt}\langle\delta\lambda, \delta x\rangle]dt$$

$$= \int_{t_o}^{t_f}[\langle C\delta x, \delta x\rangle - \langle\delta\lambda, B\delta\lambda\rangle + \langle -C\delta x - A^T\delta\lambda, \delta x\rangle + \langle\delta\lambda, A\delta x + B\delta\lambda\rangle]dt$$

$$= 0.$$

Hence the proposition.

Let T be the interval $[t_o, t_f]$ and $T' = [t_1, t_2]$ be a compact sub-interval of T. Since the accessory system is completely controllable in T, it is also completely controllable in T'. We now have;

Proposition 4.

Whenever $(\delta x, \delta\lambda) = (0, \delta\lambda)$ is a solution of the Accessory system (4.32)-(4.33) on some sub-interval, then also $\delta\lambda = 0$ on this sub-interval.

Proof. We shall prove the proposition by contradiction. Let $t' \in T'$ and assume $\delta\lambda(t') \neq 0$. From (4.32),

$$\delta x(t') = \int_{t_1}^{t'} \Phi(t', \tau) B(\tau) \delta\lambda(\tau) d\tau$$

$$= \left[\int_{t_1}^{t'} \Phi(t', \tau) B(\tau) \Phi^T(t', \tau) d\tau\right] \delta\lambda(t').$$

Since the system has been assumed to be completely controllable and since $D_u^2 H > 0$, the matrix within the bracket is positive definite. Hence $\delta\lambda(t') = 0$, which proves the proposition.

Let us now define a conjugate point. Two distinct points t_1 and t_2 belonging to T are said to be mutually conjugate with respect to the system (4.32) and (4.33) if there exists a solution $(\delta x, \delta\lambda)$ of the system with

$\delta x \neq 0$ on the sub-interval with end-points t_1 and t_2 while $\delta x(t_1) = \delta x(t_2) = 0$.

The system (4.32) and (4.33) is said to be <u>non-oscillatory</u> on the sub-interval T' if no two distinct points of this sub-interval are mutually conjugate.

With these definitions and propositions in hand, we can apply certain theorems of Reid[75] to obtain the following proposition:

<u>Proposition 5: For a system which is locally completely controllable, $(\hat{u}, \hat{x})$ is optimal if and only if they</u> satisfy

i) <u>Euler-Lagrange Equations</u> (4.7)-(4.10)

ii) <u>$D_u^2 H(u, x, \lambda, t) > 0$ (Strengthened Legendre condition,</u>

iii) <u>Accessory system (4.32)-(4.33) is non-oscillatory on every sub-interval $[t_1, t_2]$ of $[t_o, t_f]$.</u>

4.9 <u>Sufficiency Results using Convexity Arguments</u>.

So far, we have considered problems in which there were no inequality constraints present. We shall now give sufficiency theorems for a general class of non-linear optimal control problems in which there are inequality constraints present. The methods we use are very similar to those used in Chapter 2.

We shall consider the same problem as that defined in section 4.2, but there will be the added inequality constraint

$$h(x(t),u(t)) \leq 0, \qquad h: E^{n+m} \rightarrow E^{q}, \; q \leq m$$

present. Let h be continuously differentiable with respect to x and u.

Let $I = \{i: h_i(x(t),u(t)) = 0\}$

and let $\bar{h}(x(t),u(t)) = \{h_i(x(t),u(t)) : i \in I\}$

We shall say that the function $h : E^{n+m} \rightarrow E^{q}$ is <u>regular</u> at a point $(\hat{x}(.),\hat{u}(.)) \in E^{n+m}$, if and only if for every $(\xi(.),\eta(.)) \in E^{n+m}$, $(\xi(.),\eta(.)) \neq 0$, such that equality $x = \hat{x} + \xi$, $u = \hat{u} + \eta$ implies the inequality

$$\bar{h}(x(t),u(t)) = \bar{h}((\hat{x}+\xi)(t),(\hat{u}+\eta)(t))$$

$$= \bar{h}(\hat{x}(t),\hat{u}(t)) + D_x\bar{h}(\hat{x}(t),\hat{u}(t)).\xi(t)$$

$$+ D_u\bar{h}(\hat{x}(t),\hat{u}(t)).\eta(t) \lneq 0$$

there exists a function $\omega : [0,1] \rightarrow E^{n+m}$ with the following properties,

i) $D\omega(\epsilon).\tau$ exists for $0 \leq \varepsilon \leq 1$

ii) $(\hat{x}(.),\hat{u}(.)) = \omega(o)$

iii) $\bar{H}[\omega(\varepsilon)] \leq 0, \ 0 \leq \varepsilon \leq 1$

iv) $(\xi(.), \eta(.)) = D\omega(o).\tau$

We now assume,

a) The functions F, L, f, G and h are twice continuously differentiable with respect to x and u,

b) the functions L, f and h are strictly convex with respect to u, that is,

$$L(x(t),u(t),t) > L(x(t),\hat{u}(t),t) + D_u L(x(t),\hat{u}(t),t).(u(t) - \hat{u}(t))$$

and similar conditions for f and h,

c) the functions F, L, f, G and h are convex with respect to x, that is,

$$L(x(t),u(t),t) \geq L(\hat{x}(t),u(t),t) + D_x L(\hat{x}(t),u(t),t).(x(t) - \hat{x}(t))$$

etc.,

d) the matrix $D_x G(x(t_f),t_f)$ is non-singular.

A pair $(\hat{u}(.),\hat{x}(.))$ satisfying (4.2), 4.3) and $h(x,u) \leq 0$ is said to be optimal if $P(x(t_o),\hat{u}) < P(x(t_o),u)$ for all $(u(.),x(.))$ satisfying (4.2), 4.3) and $h(x,u) \leq 0$.

<u>Proposition 6. Let u(.) and x(.) be an admissible control and trajectory satisfying equations (4.2), (4.3) and the inequality constraint $h(x(t),u(t)) \leq 0$. Let there exist multipliers $\hat{\lambda}(.)$, $\hat{\mu}(.)$ and $\hat{\nu}$, such that</u>

$$\dot{\hat{\lambda}} = -D_x H(\hat{x}, \hat{u}, \hat{\lambda}, \mu, t); \quad \hat{\lambda}(t_f) = D_x F(\hat{x}(t_f), t_f) + (D_x G(\hat{x}(t_f), t_f)^T \hat{\nu} \tag{4.40}$$

$$D_u H(\hat{x}, \hat{u}, \hat{\lambda}, \hat{\mu}, t) = 0 \tag{4.41}$$

$$\langle \hat{\mu}, h(\hat{x}, \hat{u}) \rangle = 0 \tag{4.42}$$

$$\hat{\lambda}(.) \geq 0, \ \hat{\mu}(.) \geq 0, \ \hat{\nu} \geq 0 \tag{4.43}$$

where $H(x,u,\lambda,\mu,t) = L(x,u,t) + \langle \lambda, f(x,u,t) \rangle + \langle \mu, h(x,u) \rangle$ (4.44)

Then $\hat{u}(.)$, $\hat{x}(.)$ is optimal.

Proof: Due to our convexity assumptions,

$$P(x(t_o),u) - P(x(t_o),\hat{u}) = F(x(t_f),t_f) - F(\hat{x}(t_f),t_f) + \int_{t_o}^{t_f} [L(x,u,t) - L(\hat{x},\hat{u},t)]dt$$

$$> \langle D_x F(\hat{x}(t_f),t_f), x(t_f) - \hat{x}(t_f) \rangle + \int_{t_o}^{t_f} [\langle D_u L(\hat{x},\hat{u},t), u-\hat{u} \rangle + \langle D_x L(\hat{x},\hat{u},t), x-\hat{x} \rangle]dt$$

$$= \langle D_x F(\hat{x}(t_f),t_f), x(t_f) - \hat{x}(t_f) \rangle - \int_{t_o}^{t_f} [\langle (D_u f(\hat{x},\hat{u},t))^T.\hat{\lambda} + (D_u h(\hat{x},\hat{u}))^T.\hat{\mu}, u-\hat{u} \rangle$$
$$+ \langle \dot{\hat{\lambda}} + (D_x f(\hat{x},\hat{u},t)^T.\hat{\lambda} + (D_x h(\hat{x},\hat{u}))^T.\hat{\mu}, x-\hat{x} \rangle]dt \tag{4.45}$$

Integrating by parts,

$$- \int_{t_o}^{t_f} \langle \dot{\hat{\lambda}}, x-\hat{x} \rangle dt = -\langle \hat{\lambda}(t_f), x(t_f)-\hat{x}(t_f) \rangle + \langle \lambda(t_o), x(t_o)-\hat{x}(t_o) \rangle$$

$$+ \int_{t_o}^{t_f} \langle \hat{\lambda}, f(x,u.t) - f(\hat{x},\hat{u},t) \rangle dt$$

$$> -\left[\langle D_x F(\hat{x}(t_f),t_f) + (D_x G(\hat{x}(t_f),t_f))^T \hat{\nu} \, , \, x(t_f)-\hat{x}(t_f) \rangle\right]$$

$$+ \int_{t_o}^{t_f} \langle \hat{\lambda}, D_x f(\hat{x},\hat{u},t).(x-\hat{x}) + D_u f(\hat{x},\hat{u},t).(u-\hat{u}) \rangle dt, \quad (4.46)$$

since $\hat{\lambda}(.) \geq 0$ and f is convex in x and u.

From (4.45) and (4.46),

$$P(x(t_o),u) - P(x(t_o),\hat{u}) > - \langle \hat{\nu}, D_x G(\hat{x}(t_f),t_f).(x(t_f)-\hat{x}(t_f) \rangle$$

$$- \int_{t_o}^{t_f} \langle \hat{\mu}, D_x h(\hat{x},\hat{u}).(x-\hat{x}) + D_u h(\hat{x},\hat{u}).(u-\hat{u}) \rangle dt \quad (4.47)$$

But $G(x(t_f),t_f) \geq G(\hat{x}(t_f),t_f) + D_x G(\hat{x}(t_f),t_f).(x(t_f)-\hat{x}(t_f))$.

But $G(x(t_f),t_f) = G(\hat{x}(t_f),t_f) = 0$

Hence, $D_x G(\hat{x}(t_f),t_f).(x(t_f) - \hat{x}(t_f)) \leq 0$

Also, $h(x,u) > h(\hat{x},\hat{u}) + D_x h(\hat{x},\hat{u}).(x-\hat{x}) + D_u h(\hat{x},\hat{u}).(u-\hat{u})$

But $h(x,u) \leq 0, \quad h(\hat{x},\hat{u}) \leq 0$

and $\hat{\mu} \geq 0 \quad \langle \hat{\mu}, h(\hat{x},\hat{u}) \rangle = 0$.

Therefore (4.47) can be written as,

$P(x(t_o),u) - P(x(t_o),u) > 0$, which proves the proposition.

It can be shown

<u>Proposition 7: Let h be regular at (u(.),x(.)). Conditions (4.40)-(4.42) and $\mu(.) \geq 0$ are also necessary for (u(.),x(.)) to be optimal.</u>

A general result of this type is proved in Chapter V. As a corollary of propositions 6 and 7, we can isolate a class of problems for which the necessary conditions are also sufficient (provided our convexity assumptions hold).

This is the class of systems for which

$f(x,u,t) = A(t)x(t) + B(t)u(t)$

$G(x(t_f),t_f) = Mx(t_f)$, where M is a pxn non-singular matrix, $p \leq n$

It is precisely the properties of this class of systems that we make use of in the modified second variation successive approximation method.

In problems in which the inequality constraint is of the form $h(x(t)) \leq 0$ (the state constrained problem), we can reduce it to the type of constraint we have considered by writing,

$$h(x(t)) = h(x(t_o)) + \int_{t_o}^{t} \dot{h}(\tau)d\tau$$

$$= h(x(t_o)) + \int_{t_o}^{t} D_x h(x(\tau)).f(x,u,\tau)d\tau$$

$$= h(x(t_o)) + C(x,u,t)$$

Certainly the most stringent requirement in the sufficiency arguments is the requirement that $\hat{\lambda}(.)$ be positive everywhere in $[t_o, t_f]$. A <u>sufficient condition</u> for $\hat{\lambda}(.)$ to be positive everywhere is given by the following proposition.(51)

<u>Proposition 8</u>.

Sufficient conditions for $\hat{\lambda}(.)$ to be non-negative everywhere in $[t_o, t_f]$ are

i) $\hat{\lambda}(t_f) \geq 0$

ii) $D_x L(.) + (D_x h)^T(.)\hat{\mu} \geq 0$

iii) $\frac{\partial f^j}{\partial x_i} \geq 0, \quad i \neq j$

<u>Proof</u>: Let $g(.) = D_x L(.) + (D_x h)^T(.)\hat{\mu}$.

Rewriting equation (4.40), componentwise,

$$\dot{\hat{\lambda}}_i = - \sum_{j=1}^{n} \frac{\partial f^j}{dx^i} \hat{\lambda}_j - g_i(t); \quad \hat{\lambda}_i(t_f) = \frac{\partial F}{\partial x^i} + \sum_j \frac{\partial G^j}{\partial x^i} \hat{\vartheta}_j \qquad (4.48)$$

Introduce the transformation,

$$\lambda_i = \exp\left[\int_{t_o}^{t} \frac{\partial f^i}{\partial x^i} d\tau \right] y_i$$

This converts the set of equations (4.48) to

$$\dot{y}_i = - \sum_{j \neq i} \frac{\partial f^j}{\partial x^i} y_j - g_i(t) \exp\left[- \int_{t_o}^{t} \frac{\partial f^i}{\partial x^i} ds \right].$$

Hence the sufficiency of the conditions is obvious.

CHAPTER 5

SECOND ORDER COMPUTATIONAL METHODS FOR THE SOLUTION OF OPTIMAL CONTROL PROBLEMS

5.1. Introduction

In the previous chapter we have considered second order necessary conditions and sufficient conditions for a class of optimal control problems. We have also indicated that second order conditions are important for some computational methods. In this chapter we present some second order computational methods for a general class of optimal control problems. The class of problems considered is known as the Bolza Problem in the Calculus of Variations. The algorithms considered are extensions of the gradient methods due to Kelley(68) and Bryson(69) and similar to the methods proposed by Merriam(70). Merriam however does not consider problems with terminal constraints. The algorithm presented is formally equivalent to Newton's Method in Function Space(71),(72) and indeed in some problems it would be better to use Newton's Method.

The development in this paper is formal and indicates how we solve these problems on a digital computer.

However, under the asumptions we have made a rigorous treatment of these successive approximation methods can be given.

The ~~paper~~ chapter may be divided into 6 sections. In Section 5.2 we formulate the problem and state the assumptions we have made. In Section 5.3 we state the first-order necessary conditions of optimality. These are the Euler-Lagrange equations and the transversality condition.

Section 5.4 is devoted to Second Variation Successive Approximation Methods and certain modifications to it.

In Section 5.5 we show how the second variation method is formally equivalent to Newton's Method and also indicate how the linear two point boundary value problem arising in Newton's Method can be solved in essentially the same way as in the Second Variation Method.

In Section 5.6 we point out certain advantages and disadvantages of the Second Variation Method.

5.2 Problem Statement

We consider the following Bolza problem. Find the optimal control function $\hat{u}$ and the corresponding optimal trajectory $\hat{x}$ so that the performance functional

$$P(x(t_o),u) = F(x(t_f),t_f) + \int_{t_o}^{t_f} L(x(t),u(t),t)dt \quad (5.1)$$

is minimised subject to the constraints

$$\frac{dx}{dt} = f(x(t),u(t),t) \ ; \ x(t_o) = c \text{ given} \ ; \quad (5.2)$$

$$G(x(t_f),t_f) = 0 \quad (5.3)$$

Here $x(t) \in E^n$, $u(t) \in E^m$, f is a function mapping E^{n+m+1} to E^n and G is a function mapping E^{n+1} to E^p, $p \leq n$. The time t_f may be explicitly or implicitly specified.

Assumptions. The computational method we present is an iterative method. At each iteration stage, we have to solve a minimization problem similar to the Accessory Minimization problem we were investigating in the previous chapter. We have to make suitable assumptions so that the auxiliary minimization problem occurring at each iteration stage is well defined and has a proper minimum. Hence the following assumptions:

i) The original problem has a unique minimum and this minimum is attained by some admissible control function u and corresponding trajectory x. Let Ω be a bounded open set of E^m. If $u \in \Omega$, u is called admissible.

ii) The system is locally completely controllable in

$(t_o, t_f]$ along any trajectory $\bar{x}$ corresponding to an admissible control $\bar{u}$.

iii) All functions are assumed to have continuous second derivatives with respect to x, u and t.

At every iteration stage .

iv) $D_x G(x(t_f), t_f)$ is non-singular.

v) The matrix of partial derivatives $D_u^2 H$ is positive definite and the matrix of partial derivatives $D_x^2 \Psi$ and $D_x^2 H - D_{xu}^2 (D_u^2 H)^{-1} D_{ux}^2$ are positive semidefinite where

$$\Psi(x(t_f), t_f, \nu) = F(x(t_f), t_f) + \langle \nu, G(x(t_f), t_f) \rangle \quad (5.4)$$

and $H(x(t), u(t), \lambda(t), t) = L(x(t), u(t), t) + \langle \lambda(t), f(x(t), u(t), t) \rangle$

(5.5)

5.3 First Order Necessary Conditions

Let $\hat{u}(.)$ and $\hat{x}(.)$ be the optimal control and the optimal trajectory and let $\hat{\lambda}(.) \in E^n$ and $\hat{\nu} \in E^p$ be the multiplier functions and the terminal constraint multipliers. Let Ψ and H be defined by (5.4) and (5.5) respectively. Then $\hat{u}(.)$, $\hat{x}(.)$, $\hat{\lambda}(.)$, $\hat{\nu}$ satisfy

$$\dot{x}(t) = f(x(t), u(t), t) \; ; \quad x(t_o) = c \quad (5.6)$$

$$\dot{\lambda}(t) = -D_x H(x(t), u(t), \lambda(t), t) \; ; \; \lambda(t_f) = D_x \Psi(x(t_f), t_f, \nu)$$

(5.7)

$$D_u H(x(t), u(t), \lambda(t), t) = 0 \quad (5.8)$$

$$G(x(t_f), t_f) = 0 \quad (5.9)$$

$$W(x(t_f),t_f,\nu) = H(x(t_f),u(t_f),\lambda(t_f),t_f) + D_t\Psi(x(t_f),t_f,\nu) = 0 \tag{5.10}$$

The last equation is the transversality condition.

5.4 A Second Order Successive Approximation Method

Let $P = P(x(t_o),u)$. The successive approximation method consists of constructing a sequence of functions $u_o(.)$, $u_1(.)$, ...$u_n(.)$ and $x_o(.)$, $x_1(.)$, ...$x_n(.)$, such that $P(x(t_o),u_{n+1}) \leq P(x(t_o),u_n)$, and $\underset{n\to\infty}{\text{Lt}}\ P(x(t_o),u_n) = \hat{P}$ $\underset{n\to\infty}{\text{Lt}}\ u_n = \hat{u}$, $\underset{n\to\infty}{\text{Lt}}\ x_n = \hat{x}$, where u and x satisfy (5.6) and (5.9). We shall (formally) construct sequences such that at each iteration stage equation (5.6) and (5.7) are satisfied and $P(x(t_o), u_{n+1}) \leq P(x(t_o),u_n)$. Also after a finite number of iterations

$$\| D_u H \| = \max_{1\leq i\leq m}\ \max_{t\in[t_o,t_f]} \left|\frac{\partial H}{\partial u^i}\right| \leq \epsilon_1$$

and

$$\| G(x(t_f),t_f \| = \max_{1\leq i\leq p} | G^i(x(t_f),t_f)| \leq \epsilon_2 \text{ and}$$

$$| W(x(t_f),t_f,\nu)| \leq \epsilon_3.$$

ϵ_1, ϵ_2 and ϵ_3 are suitably small positive numbers selected from numerical considerations. This means that we shall satisfy equations (5.8), (5.9) and (5.10) arbitrarily closely. It is a second order method since the nature of the convergence is quadratic (the number of correct digits at each iteration is doubled). In principle, for a linear system and a quadratratic performance criterion, we obtain one-step convergence.

We shall consider two different cases of the problem.

Case i) Final time t_f is given explicitly

Let us assume that we have chosen a nominal control u(.) in the time interval $[t_o, t_f]$ and the terminal Lagrange multiplier ν. The nominal trajectory is then obtained by integrating the system equations

$$\dot{x} = f(x,u,t) \ ; \quad x(t_o) = c$$

in the forward direction*. The equation,

$$\dot{\lambda} = -D_x H(x,u,\lambda,t)$$

is then integrated backwards with the boundary condition $\lambda(t_f) = D_x \Psi(x(t_f), t_f, \nu)$. All derivatives are evaluated at the nominal control and trajectory. The performance functional may now be written as

$$P(x(t_o),u) = \Psi(x(t_f),t_f) + \int_{t_o}^{t_f} [H(x(t),u(t),\lambda(t),t) - \langle\lambda(t),\dot{x}(t)\rangle] \, dt \qquad (5.11)$$

Let us expand P in a generalized Taylor's series,

---

*To simplify the notation we shall often drop arguments of a function.

$$P(x(t_o),u+\xi)=P(x(t_o),u)+\int_{t_o}^{t_f}\langle D_uH,\xi\rangle\,dt$$

$$+\frac{1}{2}\int_{t_o}^{t_f}[\langle\bar{D}_u^2H.\xi,\xi\rangle+\langle\bar{D}_u^2H.\eta,\eta\rangle+2\langle\bar{D}_{ux}^2.\eta,\xi\rangle]\,dt$$

$$+\frac{1}{2}\langle\bar{D}_x^2\psi(x(t_f),t_f,\nu).\eta(t_f),\eta(t_f)\rangle$$

Where the bar notation indicates that the derivatives are calculated on the line segments joining u and $u+\xi$, x and $x+\eta$.

Let us now assume that ξ and η are sufficiently small so that it is sufficient to retain terms up to second order. Call $\xi = \delta u$ and $\eta = \delta x$. Let the new control be $u_{new} = u_{old} + \delta u$. The improvement in control δu is obtained by minimising

$$\int_{t_o}^{t_f}[\varepsilon_1\langle D_uH,\delta u\rangle+\frac{1}{2}\langle D_u^2H.\delta u,\delta u\rangle+\frac{1}{2}\langle D_x^2H.\delta x,\delta x\rangle$$
$$+\langle D_{ux}^2H.\delta x,\delta u\rangle]\,dt$$
$$+\frac{1}{2}\langle D_x^2\psi(x(t_f),t_f,\nu).\delta x(t_f),\delta x(t_f)\rangle \tag{5.12}$$

Subject to the constraints,

$$\delta\dot{x} = D_xf.\delta x + D_uf.\delta u\ ;\quad \delta x(t_o) = 0 \tag{5.13}$$

$$\varepsilon_2G(x(t_f),t_f) + D_xG(x(t_f),t_f).\delta x(t_f) = 0 \tag{5.14}$$

where ε_1 and $_2$ are suitably chosen small positive numbers $0 < \varepsilon_1, \varepsilon_2 \leq 1$. We notice that to obtain the improvement in control we have to solve a new variational problem. It is however a linear problem with a quadratic performance criterion and linear terminal constraints, and can be solved. The complete solution of this variational problem is given in Appendix C. A discussion of why the parameters ε_1 and ε_2 are introduced in the particular manner is also given in Appendix C.

The solution of this new variational problem is given by

$$\delta\dot{x} = D_x f.\delta x + D_u f.\delta u \;; \quad \delta x(t_o) = 0 \tag{5.15}$$

$$\varepsilon_1 \Delta\dot{\lambda} = -D_x^2 H.\delta x - D_{xu}^2 H.\delta u - \varepsilon_1 (D_x f)^T.\Delta\lambda \tag{5.16}$$

$$\varepsilon_1 \Delta\lambda(t_f) = D_x^2 \psi(x(t_f), t_f, \nu).\delta x(t_f) + \varepsilon_2 (D_x G)^T.\Delta\nu \tag{5.17}$$

$$\delta u = -(D_u^2 H)^{-1} [\varepsilon_1 D_u H + D_{ux}^2 H.\delta x + \varepsilon_1 (D_u f)^T.\Delta\lambda] \tag{5.18}$$

where $\varepsilon_1 \Delta\lambda$ and $\varepsilon_2 \Delta\nu$ are the multipliers for the auxiliary minimization problem.

Let us introduce the matrices and vectors,

$$\begin{aligned} A &= D_x f - D_u f \; (D_u^2 H)^{-1} D_{ux}^2 H \\ B &= -D_u f.(D_u^2 H)^{-1} (D_u f)^T \\ C &= D_x^2 H - D_{xu}^2 H (D_u^2 H)^{-1} D_{ux}^2 H \\ v &= -D_u f \; (D_u^2 H)^{-1} D_u H \\ w &= -D_{xu}^2 H (D_u^2 H)^{-1} D_u H \end{aligned} \tag{5.19}$$

Using (5.19) and substituting (5.18) into (5.15) and (5.16), we obtain the following linear two-point boundary value problem,

$$\delta\dot{x} = A(t)\delta x + \varepsilon_1 B(t)\Delta\lambda + \varepsilon_1 v \; ; \quad \delta x(t_o) = 0 \tag{5.20}$$

$$\varepsilon_1 \Delta\dot{\lambda} = -C(t)\delta x - \varepsilon_1 A^T(t)\Delta\lambda - \varepsilon_1 w \tag{5.21}$$

$$\varepsilon_1 \Delta\lambda(t_f) = D_x^2 \psi(x(t_f),t_f).\delta x(t_f) + \varepsilon_2 (D_x G)^T.\Delta\nu \tag{5.22}$$

$$\varepsilon_2 G(x(t_f),t_f) + D_x G.\delta x(t_f) = 0 \tag{5.23}$$

Before proceeding further it is necessary to show that the particular choice of δu given by (5.18) indeed reduces the value of the performance functional. In Appendix C we show that for this choice of δu the sum of the first and second variations is indeed negative.

The linear two-point boundary value problem is perhaps best solved by introducing the linear transformations,

$$\varepsilon_1 \Delta\lambda = \varepsilon_1 l + K(t)\delta x + \varepsilon_2 N(t)\Delta\nu \tag{5.24}$$

$$\delta G = \varepsilon_1 m + N^T(t)\delta x + \varepsilon_2 P(t)\Delta\nu \tag{5.25}$$

where $\delta G = -\varepsilon_2 G(x(t_f),t_f)$.

(That such linear transformations exist may be easily shown by writing the solutions of the differential equations (5.20) - (5.21) explicitly.)

Differentiating (5.24) and (5.25), using equations (5.20) - (5.23) and equating coefficients of various terms to zero (see Appendix C), we obtain the following equations for ℓ, m, K, N and P :

$$\dot{\ell} + (KB+A^T)\ell + Kv+w = 0 \;; \quad \ell(t_f) = 0 \tag{5.26}$$

$$\dot{K}+KA+A^TK+KBK+C = 0 \;; \quad K(t_f) = D_x^2\Psi(x(t_f),t_f,\nu) \tag{5.27}$$

$$\dot{N} + (KB+A^T)N = 0 \;; \quad N(t_f) = (D_xG(x(t_f),t_f))^T \tag{5.28}$$

$$\dot{m} + N^T(B\ell + v) = 0 \;; \quad m(t_f) = 0 \tag{5.29}$$

$$\dot{P} + N^TBN = 0 \;; \quad P(t_f) = 0 \tag{5.30}$$

In order to be able to compute δu, we have to determine $\Delta\nu$. Equations (5.26) - (5.30) can be integrated backwards. Having done the integration backwards, we may calculate $\varepsilon_2\Delta\nu$ at time t_o from the relation,

$$\varepsilon_2\Delta\nu = P^{-1}(t_o)[\delta G - \varepsilon_1 m(t_o) - N^T(t_o)\delta x(t_o)] \tag{5.31}$$

Due to the assumptions we have made $P^{-1}(t_o)$ exists.

From (5.31) and (5.24),

$$\begin{aligned}\varepsilon_1\Delta\lambda(t) = {} & \varepsilon_1[\ell(t)-N(t)P^{-1}(t_o)m(t_o)] - N(t)P^{-1}(t_o)\delta G \\ & - N(t)P^{-1}(t_o)N^T(t_o)\delta x(t_o) + K(t)\delta x(t)\end{aligned} \tag{5.32}$$

Hence from (5.32) and (5.18), we get

$$\delta u = -\,\varepsilon_1 (D_u^2 H)^{-1}[D_u H + (D_u f)^T \ell(t)]$$
$$+ (D_u^2 H)^{-1}(D_u f)^T N(t) P^{-1}(t_o)\ \delta G$$
$$+ (D_u^2 H)^{-1}(D_u f)^T N(t)\ P^{-1}(t_o) N^T(t_o)\delta x(t_o)$$
$$- (D_u^2 H)^{-1}[D_{ux}^2 H + (D_u f)^T K]\ \delta x \qquad (5.33)$$

Let
$$\varepsilon_1 (D_u^2 H)^{-1}[D_u H + (D_u f)^T \ell(t)] = r_1$$
$$-\ \varepsilon_1 (D_u^2 H)^{-1}(D_u f)^T N(t) P^{-1}(t_o) m(t_o) = r_2$$
$$-\ (D_u^2 H)^{-1}(D_u f)^T N(t) P^{-1}(t_o) = K_1 \qquad (5.34)$$
$$(D_u^2 H)^{-1}[D_{ux}^2 H + (D_u f)^T K] = K_2$$

Since $\delta x(t_o) = 0$, using (5.34) equation (5.33) may be written as

$$\delta u(t) = -r_1(t) - r_2(t) - K_1(t)\delta G - K_2(t)\delta x \qquad (5.35)$$

Equation (5.35) is a linear feedback equation. It is therefore clear that the process of improvement in control function is a linear feedback process. This is perhaps brought out more clearly in Fig. 5.1.

We have also solved the neighbouring optimal feedback control problem. In this case the nominal trajectory is optimal and hence $D_u H = 0$. This implies that $\ell(t) = m(t) = 0$ everywhere in $[t_o, t_f]$ and hence the feedforward terms $r_1(t) = r_2(t) = 0$ everywhere in $[t_o, t_f]$. But $\delta x(t_o)$, which is the deviation from the nominal

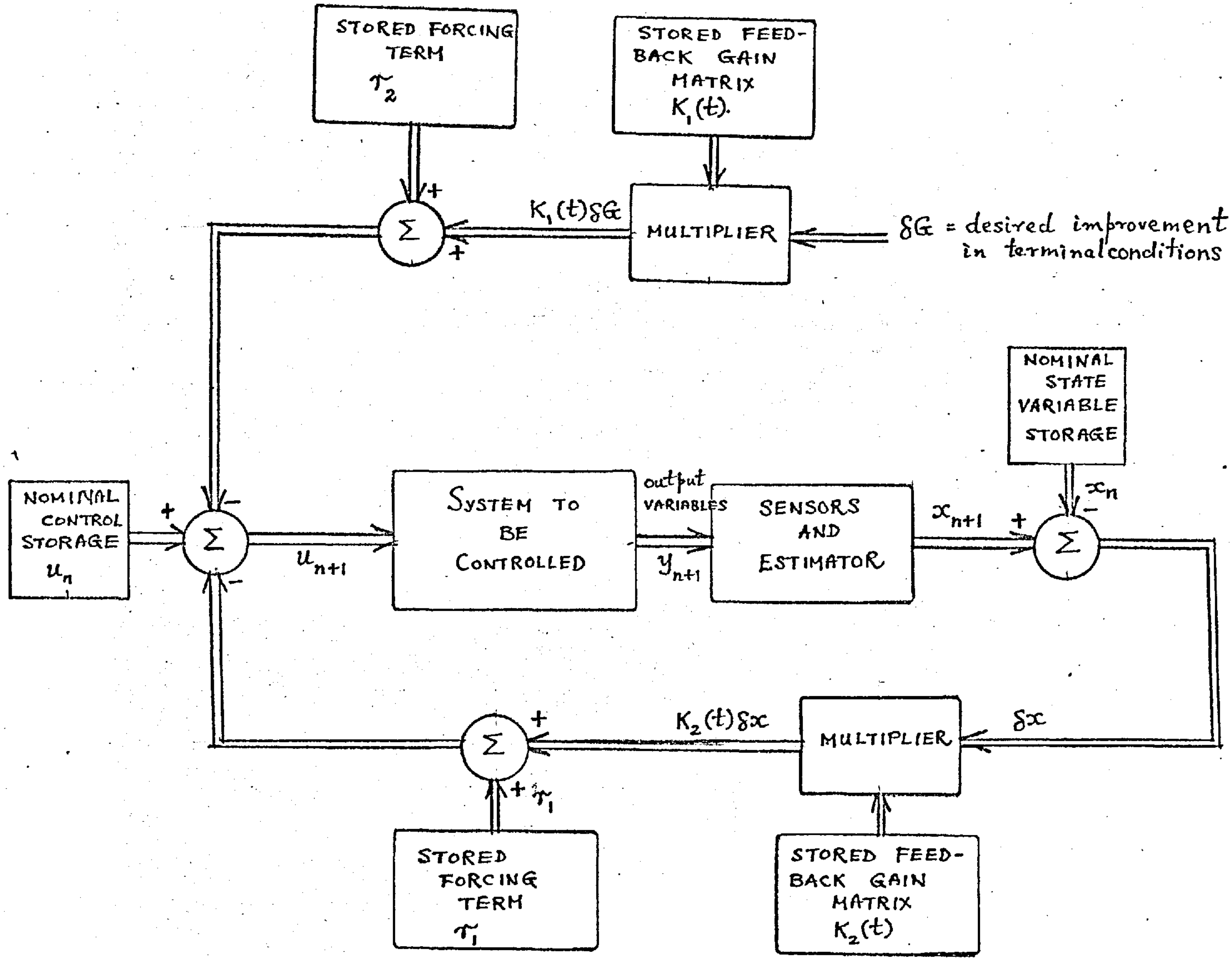


FIG. 5.1 LINEAR FEEDBACK CONTROL SCHEME ; ALSO I ITERATION STAGE (NEGLECTNG SENSORS AND ESTIMATOR)

trajectory due to a disturbance is not zero. Supposing that the state of the system is continuously measured, from (5.33) and (5.34), the neighbouring optimal time-varying feedback control is given by

$$\delta u(t) = -K_3(t)\delta x(t) + K_2(t)\delta G \qquad (5.36)$$

where $K_3(t) = K_1(t) - (D_u^2 H)^{-1}(D_u f)^T N(t) P^{-1}(t) N^T(t)$

What is interesting in expression (5.35) and (5.36) is that we can control independently the contributions to δu due to desired change in terminal conditions, initial condition change or desired change in $D_u H$ (or desired change in cost function). This is important in numerical computation, since it is very difficult to satisfy terminal conditions.

Fig. 5.1 could thus be considered as the design of a linear feedback control scheme for a non-linear system. The state variables may not be directly observable and may have to be estimated. If there is additive noise present, the problem can also be handled. In this case we have to minimise the expected value of the second variation. But the important point is, that for a linear system and a quadratic performance functional, the control and filtering (estimation) problems separate(73),(74).

Before considering the case when the time t_f is given only implicitly, let us consider an example which we solved on the computer to verify that for a linear system and a quadratic performance functional, the procedure yields one-step convergence.

<u>Example</u>: Consider the linear system,

$$\dot{x} = u \; ; \quad x(0) = 0$$

The performance criterion is $P(u,0) = \frac{1}{2} \int_0^T (x^2+u^2)dt$

$$T = 5$$

The minimum value of P is obviously 0 and the optimal control law is u = 0.

For this problem

$$H = \frac{1}{2}(x^2+u^2) + \lambda u$$

$$\frac{\partial H}{\partial u} = u+\lambda \; ; \quad \frac{\partial H}{\partial x} = x \; ; \quad \frac{\partial^2 H}{\partial u^2} = 1 \; ; \quad \frac{\partial^2 H}{\partial x^2} = 1 \; ; \quad \frac{\partial^2 H}{\partial u \partial x} = \frac{\partial^2 H}{\partial x \partial u} = 0$$

$$\frac{\partial f}{\partial u} = 1 \; ; \quad \frac{\partial f}{\partial x} = 0$$

From (5.19),

$$A = 0$$

$$B = -1$$

$$C = 1$$

$$v = -1$$

$$w = 0$$

From (5.26) and (5.27),

$$\dot{\ell} - K\ell - K = 0 \; ; \quad (5) = 0$$

$$\dot{K} - K^2 + 1 = 0 \; ; \quad K(5) = 0$$

$$\delta u = -(u+\lambda+\ell) - K\delta x$$

The initial guess was taken to be $u(t) = 1$ everywhere in $[t_o, t_f]$ and the corresponding value of the cost function was 22. After one iteration the cost reduced to 0.14×10^{-3}. A crude finite difference method was used for the integration.

<u>Case ii) Final time t_f not given explicitly</u>

The basic procedure for this case is basically the same as for Case (i). The expressions for first and second variations in the Taylor Series expansion are now more complicated since we now have to consider the variation in the final time t_f.

The control function u and the terminal time t_f is guessed and the system equations,

$$\dot{x} = f(x,u,t) \; ; \quad x(t_o) = c \text{ (given)}$$

is integrated in the forward direction. A value for ν is guessed and the Euler-Lagrange equation

$$\dot{\lambda} = -D_x H(x,u,\lambda,t)$$

is then integrated backwards with the boundary condition $\lambda(t_f) = D_x \psi(x(t_f), t_f, \nu)$. All derivatives are calculated at the nominal trajectory.

$P(x(t_o),u)$ is then expanded in a generalized Taylor series as in case (i). Let δP and $\delta^2 P$ be the first and second variations. Then (see Appendix C),

$$\delta P = (H+D_t\psi)_{t=t_f} dt_f + \int_{t_o}^{t_f} \langle D_u H, \delta u\rangle \, dt \tag{5.37}$$

$$\begin{aligned}
\tfrac{1}{2}\delta^2 P = \tfrac{1}{2}[&\langle D_x^2 \psi(x(t_f),t_f,\nu)\Delta x(t_f),\Delta x(t_f)\rangle + \\
&+ \langle D_x \psi(x(t_f),t_f,\nu), \Delta^2 x(t_f)\rangle \\
&+ D_t^2 \psi(x(t_f),t_f,\nu)dt_f^2 + 2\langle D_{xt}^2 \psi(x(t_f),t_f,\nu),\Delta x(t_f)dt_f\rangle] \\
&+ \langle D_u H(x(t_f),u(t_f),\lambda(t_f),t_f),\delta u(t_f)dt_f\rangle \\
&+ \tfrac{1}{2}\tfrac{d}{dt}[H(x(t_f),u(t_f),\lambda(t_f),t_f) - \langle\lambda(t_f),\dot{x}(t_f)\rangle]dt_f^2 \\
&+ \tfrac{1}{2}\int_{t_o}^{t_f} [\langle D_u^2 H.\delta u,\delta u\rangle + \langle D_x^2 H.\delta x,\delta x\rangle + 2\langle D_{ux}^2 H.\delta x,\delta u\rangle]dt
\end{aligned} \tag{5.38}$$

where $\Delta x(t_f) = \delta x(t_f) + \dot{x}(t_f)dt_f$

$$\Delta^2 x(t_f) = \ddot{x}(t_f)dt_f^2 + 2\delta x(t_f)dt_f$$

Neglecting third order terms (see Appendix C), δu is determined by minimising,

$$\begin{aligned}
&\varepsilon_2 (H+D_t\psi)_{t=t_f} dt_f + \varepsilon_1 \int_{t_o}^{t_f} \langle D_u H,\delta u\rangle dt \\
&+ \tfrac{1}{2}[\langle D_x^2\psi(\delta x+\dot{x}dt_f),\delta x+\dot{x}dt_f\rangle + \langle D_x\psi, \ddot{x}dt_f^2 + 2\delta\dot{x}dt_f\rangle \\
&\quad + D_t^2\psi\, dt_f^2 + 2\langle D_{xt}^2\psi, \delta x+\dot{x}dt_f\rangle dt_f] \\
&+ \tfrac{1}{2}\tfrac{d}{dt}[L(x(t),u(t),t]_{t=t_f} dt_f^2
\end{aligned}$$

$$+ \frac{1}{2} \int_{t_o}^{t_f} [\langle D_u^2 H.\delta u, \delta u\rangle + \langle D_x^2 H.\delta x, \delta x\rangle + 2\langle D_{ux}^2 H.\delta x, \delta u\rangle] dt \quad (5.39)$$

subject to the constraints,

$$\delta\dot{x} = D_x f.\delta x + D_u f.\delta u \; ; \quad \delta x(t_o) = 0 \quad (5.40)$$

$$\varepsilon_3 G(x(t_f),t_f) + D_x G(x(t_f),t_f).\delta x(t_f) + D_x G(x(t_f),t_f).\dot{x}(t_f) dt_f$$
$$+ D_t G(x(t_f),t_f) dt_f = 0 \quad (5.41)$$

The solution of this auxiliary minimization problem is given by, (see Appendix C),

$$\delta\dot{x} = A(t)\delta x + \varepsilon_1 B(t)\Delta\lambda + \varepsilon_1 v \; ; \quad \delta x(t_o) = 0 \quad (5.42)$$

$$\varepsilon_1 \Delta\dot{\lambda} = -C(t)\delta x - \varepsilon_1 A^T(t)\Delta\lambda - \varepsilon_1 w \quad (5.43)$$

$$\varepsilon_1 \Delta\lambda(t_f) = D_x^2 \psi(x(t_f),t_f,\nu).\delta x(t_f) + \varepsilon_2 (D_x G(x(t_f),t_f))^T.\Delta\nu$$
$$+ [D_x^2 \psi(x(t_f),t_f,\nu).\dot{x}(t_f) + D_{xt}^2 \psi(x(t_f),t_f,\nu) + (D_x f)^T.D_x\psi(x(t_f),t_f)]$$
$$dtf \quad (5.44)$$

$$\delta G = D_x G(x(t_f),t_f).\delta x(t_f) + [D_x G(x(t_f),t_f).\dot{x}(t_f) + D_t G(x(t_f),t_f)]$$
$$dt_f \quad (5.45)$$

$$\delta W = \langle D_x^2 \psi(x(t_f),t_f,\nu).\dot{x}(t_f) + D_{xt}^2 \psi(x(t_f),t_f,\nu) + (D_x f)^T.$$
$$D_x \psi(x(t_f),t_f,\nu), dx(t_f)\rangle$$
$$+ \varepsilon_2 \langle D_x G(x(t_f),t_f).\dot{x}(t_f) + D_t G(x(t_f),t_f), \Delta\nu\rangle + s(t_f) dt_f$$
$$(5.46)$$

where

$$\delta G = - \varepsilon_2 G(x(t_f),t_f)$$

$$\delta W = - \varepsilon_3 W = - \varepsilon_3 \, (H+D_t\psi)_{t=t_f}$$

$$s(t_f)=[\langle D_x^2\psi \, .\dot{x},\dot{x}\rangle+\langle D_x\psi \, ,D_x f.\dot{x}+D_u f.\dot{u}+D_t\rangle+\langle D_x L,\dot{x}\rangle$$

$$- \langle D_u L,\dot{u}\rangle + D_t L + D_t^2\psi + 2 \, \langle D_{xt}^2\psi \, ,\dot{x}\rangle]_{t=t_f}$$

The linear two-point boundary value problem (5.42) - (5.46) is solved in exactly the same way as in case (i) by introducing

$$\varepsilon_1 \Delta\lambda(t) = \varepsilon_1 \ell(t)+K(t)\delta x(t)+ \varepsilon_2 N(t)\Delta\nu+ p(t)dt_f \qquad (5.47)$$

$$\delta G = \varepsilon_1 m(t)+N^T(t)\delta x(t)+ \varepsilon_2 P(t)\Delta\nu+ q(t)dt_f \qquad (5.48)$$

$$\delta W = \varepsilon_1 n(t)+\langle p(t),\delta x(t)\rangle +\varepsilon_2\langle q(t),\Delta\nu\rangle+s(t)dt_f \qquad (5.49)$$

In a manner similar to that of case (i) we obtain equations (5.26) - (5.30) and the following set of differential equations for p,q,n and s:

$$\dot{p} + (A^T+KB)p = 0 \qquad (5.50)$$

$$\dot{q} + N^T Bp = 0 \qquad (5.51)$$

$$\dot{n} + \langle p, \, B\ell + v\rangle = 0 \qquad (5.52)$$

$$\dot{s} + \langle p,Bp\rangle = 0 \qquad (5.53)$$

$\Delta\nu$ and dt_f are now determined by solving equations (5.26) - (5.30) and (5.50) - (5.53) backwards and solving the equations

$$dG = \varepsilon_1 m(t_o)+N^T(t_o)\delta x(t_o)+ \varepsilon_2 P(t_o)\Delta\nu + q(t_o)dt_f \quad (5.54)$$

$$dW = \varepsilon_1 n(t_o)+\langle p(t_o),\delta x(t_o)\rangle+ \varepsilon_2\langle q(t_o),\Delta\nu\rangle+s(t_o)dt_f \quad (5.55)$$

Computing Procedure

We now indicate the computing procedure for the case when t_f is given explicitly.

i) Guess the control function u and integrate the system equation $\dot{x} = f(x,u,t)$ forwards with $x(t_o) = c$. Store u and the corresponding trajectory x.

ii) Guess a value for the multiplier ν and integrate the equation $\dot{\lambda} =-D_x H(x,u,\lambda,t)$ backwards with the boundary condition $\lambda(t_f) = D_x\Psi(x(t_f),t_f,\nu)$. Along the trajectory calculate the partial derivatives necessary to evaluate A,B,C, v and w (eqns. 5.19). Simultaneously integrate the differential equations for L,K,N, m and P backwards. Compute r_1,r_2,K_1 and K_2 and store them.

iii) Repeat step i) using,

$$u_{new} = u_{old} - r_1 - r_2 - K_1\delta G - K_2\delta X$$

iv) Repeat step ii)

v) Stop computation when

$$\max_{1\leq i\leq m} \max_{t\in[t_o,t_f]} \left|\frac{\delta H_1}{\delta u}\right| \leq \varepsilon_3 \text{ and } \max_{1\leq i\leq p} \left|G^i(x(t_f),t_f)\right| \leq \varepsilon_4$$

where ε_3 and ε_4 are small positive numbers.

Note: It is necessary to include an adjustment procedure for ε_1 and ε_2. In a suitable neighbourhood of the optimum they can be set equal to 1.

In obtaining the second order algorithm, we had to assume

i) $D_u^2 H$ positive definite

ii) $D_x^2 H - D_{xu}^2 H (D_u^2 H)^{-1} D_{ux}^2 H$ positive semi-definite

iii) $D_x^2 \psi(x(t_f), t_f, \nu)$ positive definite

at every iteration. For many problems if the nominal control function is not sufficiently near to the optimum, and if the terminal conditions are missed by a large amount, these assumptions may not be satisfied. In that case better estimates for u and ν are necessary. These improved estimates may be obtained by a gradient method. Alternatively the following successive approximation scheme may be used till the above assumptions are satisfied. For the subsequent development, we shall assume

$D_x^2 F(x(t_f), t_f)$ is positive semi-definite

$D_u^2 L$ is positive definite

$D_x^2 L - D_{xu}^2 L (D_u^2 L)^{-1} D_{ux}^2 L$ is positive semi-definite.

Note that these assumptions are assumptions on the performance criterion (which to some extent is at our choice) and hence much weaker assumptions than the previous ones.

We consider the case when t_f is fixed. We again choose a nominal control function u and integrate the system equations in the forward direction. The function F and the integrand L is now expanded in a Taylor's series round the nominal control u and the nominal trajectory x and terms up to the second order retained, thus obtaining

$$\delta P = \epsilon_1 \langle D_x F(x(t_f),t_f), \delta x(t_f)\rangle + \epsilon_1 \int_{t_o}^{t_f} [\langle D_x L, \delta x\rangle + \langle D_u L, \delta u\rangle] dt$$

$$\frac{1}{2}\delta^2 P = \frac{1}{2}\langle D_x^2 F(x(t_f),t_f).\delta x(t_f), \delta x(t_f)\rangle$$

$$+ \frac{1}{2} \int_{t_o}^{t_f} [\langle D_u^2 L.\delta u, \delta u\rangle + \langle D_x^2 L.\delta x, \delta x\rangle + 2\langle D_{ux}^2 L.\delta x, \delta u\rangle] dt$$

The control improvement δu is obtained by minimising $\delta P + \frac{1}{2}\delta^2 P$ subject to,

$$\delta \dot{x} = D_x f.\delta x + D_u f.\delta u \; ; \quad \delta x(t_o) = 0 \tag{5.56}$$

$$\epsilon_2 G(x(t_f),t_f) + D_x G(x(t_f),t_f).\delta x(t_f) = 0 \tag{5.57}$$

where $0 < {}_1, {}_2 \leq 1$. In view of our assumptions on F and L, this problem has a proper minimum, and the necessary conditions of optimality are also sufficient.

The Euler-Lagrange equations of this problem are,

$$\delta \dot{x} = D_x f.\delta x + D_u f.\delta u \; ; \quad \delta x(t_o) = 0 \tag{5.58}$$

$$\epsilon_1 \Delta\dot{\lambda} = - \epsilon_1 D_x L - D_x^2 L.\delta x - D_{xu}^2 L.\delta u - \epsilon_1 (D_x f)^T . \Delta\lambda \tag{5.59}$$

$$\delta u = -(D_u^2 L)^{-1}[\epsilon_1 D_u L + D_{ux}^2 L.\delta x + \epsilon_1 (D_u f)^T \Delta\lambda] \tag{5.60}$$

$$\epsilon_1 \Delta\lambda(t_f) = \epsilon_1 D_x F(x(t_f),t_f) + (D_x G(x(t_f),t_f))^T . \delta\nu + D_x^2 F(x(t_f),t_f) . \delta x(t_f) \quad (5.61)$$

Substituting (5.60) into (5.58) and (5.59), we get

$$\delta\dot{x} = \bar{A}\delta x + \bar{B}\Delta\lambda + \epsilon_1 \bar{v} \quad (5.62)$$

$$\lambda = -\bar{c}\delta x - \bar{A}^T \lambda - \bar{w} \quad (5.63)$$

where $\bar{A} = D_x f - D_u f (D_u^2 L)^{-1} D_{ux}^2 L$

$\bar{B} = -D_u f (D_u^2 L)^{-1} (D_u f)^T$

$\bar{c} = D_x^2 L - D_{xu}^2 L (D_u^2 L)^{-1} D_{ux}^2 L$

$\bar{v} = -D_u f (D_u^2 L)^{-1} D_u L$

$\bar{w} = D_x L - D_{xu}^2 L (D_u^2 L)^{-1} D_u L$

The way this linear two-point boundary value problem is solved is precisely the same as our previous method and we shall omit the details here. The differential equations for ℓ, k etc. we obtain are also of the same type. It can also be shown that this choice of δu reduces the value of the performance functional. The proof is exactly similar to the proof we present in Appendix C for the original second order method.

5.5 Relationships with Newton's Method

For simplicity we consider the case when there are no terminal constraints present. The method and conclusions are valid for the general Bolza problem. Solving the variational problem by Newton's Method means solving the Euler-Lagrange Equations by an iterative method. The method consists in guessing a nominal control function, a nominal trajectory and a nominal multiplier function and then linearizing the Euler-Lagrange equations round the guessed functions. A linear two-point boundary value problem is then solved which yields corrections to the guessed functions. The linear two-point boundary value problem to be solved is

$$\dot{x} + \delta\dot{x} = f + D_x f.\delta x + D_u f.\delta u \; ; \quad \delta x(t_o) = 0$$

$$\dot{\lambda} + \delta\dot{\lambda} = -D_x H - D_x^2 H.\delta x - D_{xu}^2 H.\delta u - D_{x\lambda}^2 H.\delta\lambda \; ; \quad \delta\lambda(t_f) = 0$$

$$D_u H + D_u^2 H.\delta u + D_{ux}^2 H.\delta x + D_{u\lambda}^2 H.\delta\lambda = 0$$

But for the fact that the system equations and the Euler-Lagrange equations are not satisfied by the initially guessed functions, these equations are precisely the same as equations (5.15) - (5.18). Thus the methods we have used in solving equations (5.15) - (5.18) may be used in solving the linear two-point boundary value problem in Newton's Method. As we have indicated previously from the viewpoint of numerical stability it is advantageous to solve the two-point boundary value problem in the way we have indicated. In problems where there is a constraint of the form $x(t_f) = a$ it may be better to use Newton's Method since we can guess the nominal trajectory to satisfy the boundary condition.

5.6 A Discussion of Various Methods of Solving Optimal Control Problems

A number of methods have been proposed for the solution of two-point boundary value problems arising in optimal control problems. These may be subdivided into three main classes:

i) Boundary Condition Iteration Method

ii) Control function Iteration Method

iii) Newton type Iteration Methods

The choice of the method to be adopted depends on the problem and on the nature of the application. Each problem will have a certain structure and exhibit certain stability properties, although in a non-linear problem it might be very difficult to isolate either. Further the nature of the control application may impose various constraints. For example, if on-line control is envisaged, rapidity of convergence may over-ride other factors. For some problems it may be necessary to obtain extremely accurate trajectories, while in others convergence of the performance functional to within a pre-assigned tolerance may be sufficient. In spite of this, certain advantages and disadvantages of each of these methods may be pointed out and certain recommendations made.

ii) Boundary Condition Iteration:

In this method, typically the control function u is eliminated from the first two Euler-Lagrange equations by solving $H_u = 0$ and the resulting first two Euler-Lagrange equations are solved by iteration on one of the unknown boundary values say, $\lambda(t_o)$. A suitable scalar terminal error function $V[x(t_f,\lambda(t_o)),\lambda(t_f,\lambda(t_o))]$ is then constructed. The boundary value $\lambda(t_o)$ is then adjusted till the error function goes to zero. The adjustment requires the computation of the gradient V Systematic methods for doing this are available[76]. These methods have certain computer programming advantages. Computer logic is simple and fast storage requirements are small. In problems where the method is successful accurate trajectories are obtained. The main disadvantage is the inherent instability of one of the Euler-Lagrange equations. To determine whether the method is applicable a preliminary analysis of the problem may possibly be carried out in the following way: let the unforced system equation be linearized round the given initial condition. An eigen-value analysis of the linearized system matrix could now be made. If the matrix turns out to be essentially self-adjoint boundary iteration methods are quite suitable. If not and if $t_f - t_o$ is substantially greater than the dominant system time-constant, severe instabilities may be encountered.

ii) Control Function Iteration

Control function iteration methods using both gradient techniques and steepest descent technique have been proposed in the literature. In these methods the control function is successively improved till $\|D_u H\| \leq \varepsilon$ where $\| \ \|$ is some suitable norm of the $D_u H$ function and ε is a small positive number. The primary advantage of this method is that computations are always performed in the stable direction. However convergence tends to be intolerably slow in a certain neighbourhood of the optimum. To improve convergence the size-step cannot be increased since this leads to instability. The iteration methods we have presented in this paper may be considered to be direct extensions of gradient or steepest descent techniques. We have stated previously that the second variation method is formally equivalent to Newton's method in function space. In a suitable neighbourhood of the optimum convergence is therefore quadratic. Computations here are also always performed in the stable direction. In fact in a suitable neighbourhood of the optimum, the inherent stability properties of linear feedback control systems inhibits the propagation of numerical errors. As a by-product we obtain linear time-varying feedback gains for neighbouring optimum feedback control.

On the other hand the conditions that various matrices be positive definite or semi-definite may be too strong. In such cases it may be necessary to get better estimates of the control function by using gradient methods or use the alternative successive approximation method we have indicated in conjunction with the second variation method. Numerical difficulties may also be encountered in integrating the matrix Ricatti equations, specially if the dynamic system is unstable. It is also to be noted that the matrix D_u^2H is to be inverted. Computer storage requirements are also greater since the feedback gain matrices have to be stored.

Some computational effort may be saved. For example, it is not necessary to compute $(D_u^2H)^{-1}$ at every iteration. In fact in practice this may be held constant after two or three iterations. Convergence will necessarily be slower.

For ordinary minimization problems some very efficient computational algorithms have recently been proposed[77]. These algorithms may be considered to lie somewhere between gradient and Newton's method. A distinctive feature of these methods is that use is made of information generated in previous iterations. Generalisations of these methods to function spaces should be possible.

In this paper we have not considered inequality constraints. The assumption was made that these could be approximated by means of penalty functions. Extensions of the techniques presented here to problems with inequality constaints on control and state variables appear to be possible. The auxiliary minimization problem then has additional linear inequality constraints. In this case the corresponding dual maximization problem could be solved to obtain the improvement in control function.

iii) Newton's Method

Newton's method was first proposed by Hestenes[78] to solve fixed and point problems of the Calculus of Variations. A complete analysis of the method for this class of problems was given by Stein[79]. In the context of function space, the method dates back to Kantorovich[80].

Kalaba[81] has also used this method for a special class of problems and called it 'quasi-linearisation'. Recently the method has been applied to some optimal control problems by Kopp and McGill[82]. They eliminated the control function u from the first two Euler-Lagrange equations by using the equation $D_u H = 0$. The linearised Euler-Lagrange equations are then integrated for n-linearly independent boundary conditions. The unknown boundary value $\delta\lambda(t_o)$ is found by using linear interpolation and a matrix inversion. Improvements $\delta x(t)$ and $\delta\lambda(t)$ are then obtained by one more integration.

If the linear two-point boundary value problem is solved in this way, the method suffers from the instability disadvantages of boundary interation methods.

In our view, the methods advocated in this paper could be used to solve the linear two-point boundary value problem arising within Newton's Method.

CHAPTER 6

Computer Solutions of Optimal Control Problems

6.1 Introduction

In this chapter we present results of computer solutions of two optimal control problems using the second order computational method we developed in the previous chapter. Rather than solve many problems, we have preferred to solve only two problems but with great care.

The first problem is one arising in the control of boilers. The solutions given might have potential applications in boiler control. The second problem is an idealized rocket problem. These two problems were chosen since some numerical results[87,88] are available, thereby providing a basis for comparison.

The programming was very ably done by David Stone of Imperial College, who has made an independent contribution in the numerical integration procedures used.

6.2 Overall Organization of the Computer Programme.

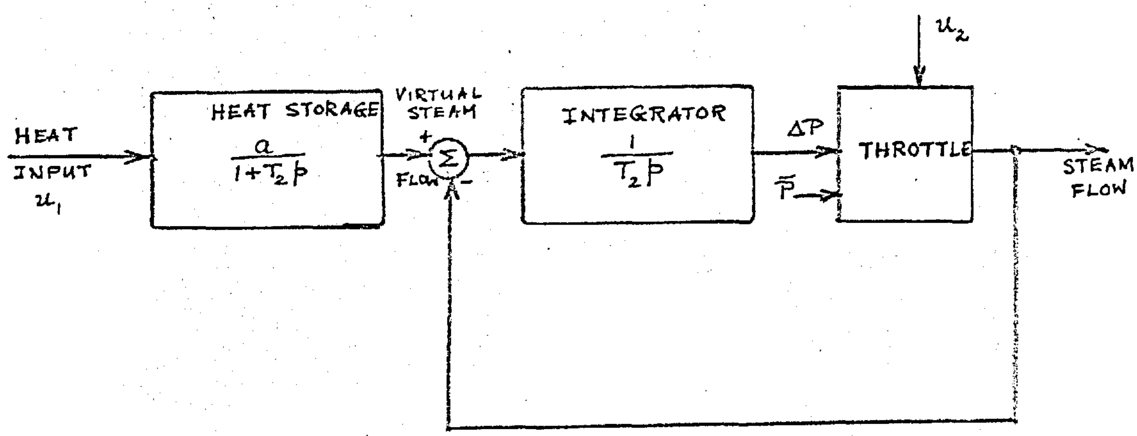
The overall flow chart of the computer programme is shown in Fig.6.1. The programme was organized as a hybrid gradient and second variation method. The least necessary condition for the second variation procedure to work is $D_u^2 H > 0$ at each point of the trajectory. If this assumption is not satisfied the programme automatically switched to a gradient method. It is very difficult to write a programme for the solution of optimal control problems which will be general enough to handle any non-linear system. Nevertheless the two programmes were written entirely in sub-routine form and made as general as possible. Thus if sufficient experience could be gained by solving a variety of problems of this kind, one could attempt to write a general programme for the solution of non-linear optimal control problems.

The rocket problem was solved first. The integration routine used here was a Runge-Kutta routine. One of the disadvantages of using a Runge-Kutta routine is that it is necessary to use a Lagrangian Interpolation routine to obtain the values of λ, ι and K at intermediate points within the basic integration interval. This is so because we need the values of these variables at the intermediate points in order to carry out the forward integration of the dynamic equations again. This fact enormously complicates the computer programme. A more satisfactory method is to use a Predictor Corrector method, together with a suitable starter. A Hamming Predictor Corrector method, together with an Euler-Newton starter was coded by David Stone of Imperial College. This integration procedure was used in the Boiler problem.

The programmes were written in Fortran IV and the computations were performed on the IBM 7090 computer of Imperial College.

6.3 <u>The Boiler Problem</u>.

The basic model is a "virtual steam flow" model of a boiler due to Profos.[88] This is shown in Fig.6.2.



<u>Fig.6.2. Virtual Steam Flow Model of Boiler.</u>

The system equations written in state variable form are:

$$\frac{dx_1}{dt} = \frac{1}{T_1}(au_1 - x_1). \qquad (6.1)$$

$$\frac{dx_2}{dt} = \frac{1}{T_2}[x_1 - u_2(x_2 + \bar{P})] \qquad (6.2)$$

where x_1 = virtual steam flow (including heat storage)-lbs./sec.

x_2 = pressure deviation from steady state $\bar{P}$ - lbs./in^2

T_2 = steam storage integration constant

a, T_1 = constants derived from a larger model

u_1 = fuel input rate control

u_2 = throttle control.

The interest(89) of the problem is in developing a control system which will allow larger swings of the load of a power boiler than is at present possible. Existing process-type controllers on a boiler are designed for their regulatory function and their piecemeal implementation gives rise to severe restrictions on load changing. A suitable performance criterion for this problem is

$$\text{Minimise} \quad \int_{t_0}^{t_f} (\text{steam flow-demanded flow})^2 dt \qquad (6.3)$$

However, during the load change it is necessary to constrain temperatures within certain limits. This can be linked to a pressure state constraint. Also there is a desire to end the period of load change in the steady state. There is thus a constraint on the end-point pressure. These two factors are met by a time-weighted penalty term in the performance criterion. The final performance criterion is,

$$\text{Minimise} \quad \int_{t_o}^{t_f} [(u_2(x_2+\bar{F})-F)^2 + C.e^{\alpha t}x_2^2]dt \qquad (6.4)$$

The simplification is made here of considering a step change in fuel input, u_1, at the beginning of the control interval and hence only one control u_2 is free.

The numerical values of the various constants are,

$a = 1.6 \times 10^{-3}$

$T_1 = 200$

$T_2 = 4.7$

$t_f - t_o = 400$ secs.

$c = 0.1$

$\alpha = 0.001$

Initial Conditions are

$x_1(t_o) = 241$

$x_2(t_o) = 0$

<u>Steady State Values are</u>

$$\bar{F} = 225$$

$$x_2 = 0$$

$$u_2 = 0.14423077$$

u_1 held constant at 140,625.00

The equations to be integrated backwards are,

$$\dot{\lambda}_1 = \frac{\lambda_1}{T_1} - \frac{\lambda_2}{T_2} \; ; \quad \lambda_1(t_f) = 0 \tag{6.5}$$

$$\dot{\lambda}_2 = \frac{\lambda_2 u_2}{T_2} - 2[U_2(x_2+\bar{F})-F]u_2 - 2ce^{\alpha t}x_2 \; ; \quad \lambda_2(t_f) = 0 \tag{6.6}$$

$$\dot{l}_1 - [\frac{a^2}{2T_1^2(x_2+\bar{F})^2} K_{11} + \frac{1}{T_1}]l_1 - [\frac{1}{2T_2^2} K_{12} - \frac{1}{T_2}]l_2$$

$$- \frac{1}{2} \frac{a^2}{T_1^2(x_2+\bar{F})^2} K_{11} + \frac{1}{2T_2} [2u_2(x_2+\bar{F})-2F-\frac{\lambda_2}{T_2}]K_{12}=0 \tag{6.7}$$

$$\dot{l}_2 - \frac{a^2}{2T_1^2(x_2+\bar{F})^2} K_{12} l_1 + [\frac{u_2}{T_2} - \frac{\lambda_2}{2T_2^2(x_2+\bar{F})} - \frac{F}{T_2(x_2+\bar{F})}$$

$$- \frac{1}{2T_2^2} K_{22}]l_2$$

$$- \frac{1}{2} \frac{a^2}{T_1^2(x_2+\bar{F})^2} K_{12} + \frac{1}{2T_2} [2u_2(x_2+\bar{F}) - 2F - \frac{\lambda_2}{T_2}]K_{22}$$

$$- \frac{1}{2(x_2+\bar{P})} [2u_2(x_2+\bar{P}) - 2F - \frac{\lambda_2}{T_2}][4u_2(x_2+\bar{P}) - 2F - \frac{\lambda_2}{T_2}]$$

$$= 0 \qquad (6.8)$$

$$\dot{K}_{11} - \frac{a^2}{2T_1^2(x_2+\bar{P})^2} K_{11}^2 - \frac{1}{2T_2^2} K_{12}^2 - \frac{2K_{11}}{T_1} + \frac{2K_{12}}{T_2} = 0 \qquad (6.9)$$

$$\dot{K}_{12} - \frac{a^2}{2T_1^2(x_2+\bar{P})^2} K_{11}K_{12} - \frac{1}{2T_2^2} K_{12}K_{22} + [\frac{u_2}{T_2} - \frac{\lambda_2}{2T_2^2(x_2+\bar{P})} -$$

$$\frac{F.}{T_2(x_2 + \bar{P})} - \frac{1}{T_1}]K_{12} + \frac{K_{22}}{T_2} = 0 \qquad (6.10)$$

$$\dot{K}_{22} - \frac{a^2}{2T_1^2(x_2+\bar{P})^2} K_{12}^2 - \frac{1}{2T_2^2} K_{22}^2 + 2[\frac{u_2}{T_2} - \frac{\lambda_2}{2T_2^2(x_2 + \bar{P})} -$$

$$\frac{F.}{T_2(x_2 + \bar{P})}]K_{22} + 2u_2^2 + 2c.e^{\alpha t}$$

$$- \frac{1}{2(x_2 + \bar{P})^2} [4u_2(x_2+\bar{P}) - \frac{\lambda_2}{T_2} - 2F]^2 = 0 \qquad (6.11)$$

The boundary conditions for the l and K equations are

$$l_1(t_f) = l_2(t_f) = K_{11}(t_f) = K_{12}(t_f) = K_{22}(t_f) = 0$$

The improvement in control is given by

$$\delta u_2 = r + M_{21}[(x_1)_{new}-(x_1)_{old}] + M_{22}[(x_2)_{new}-(x_2)_{old}] \qquad (6.12)$$

$$r = -\frac{1}{2(x_2+\overline{P})}\left\{2u_2(x_2+\overline{P}) - 2F - \frac{\lambda_2}{T_2} - \frac{l_2}{T_2}\right\}$$

$$M_{21} = +\varepsilon\frac{a}{2T_2(x_2+\overline{P})}K_{12}$$

$$M_{22} = -\frac{2u_2}{(x_2+\overline{P})} + \frac{F}{(x_2+\overline{P})^2} + \frac{\lambda_2}{2T_2(x_2+\overline{P})^2} + \varepsilon\frac{1}{2T_2(x_2+\overline{P})}K_{22}$$

It is noted that in this problem

$$D_u^2 H > 0$$

and hence there is no need to include the gradient subroutine in the programme.

$M_{21}(t)$ and $M_{22}(t)$ are optimal linear time-varying feedback gains. These are obtained directly in the last iteration.

6.3.1 Discussion of Results.

The results of this problem are shown in Figs.6.1-6.7. From Fig.6.1, it is seen that as far as reduction in cost is concerned we get almost one-step convergence. The control essentially converges to the optimal control in 2-steps. The difference in the control function between the second and third iterations is in the sixth decimal place. The difference in the state variables between the second and third iterations is also very small. However, as seen from Fig.6.2 max $\frac{\partial H^{\cdot}}{\partial u_2}$ has a value of 3×10^4 at the second iteration and 7×10^{-2} at the fourth iteration, even though the control is virtually the same at these two iterations.

The optimal linear time-varying feedback gains are shown in Figs.6.6 and 6.7. These can be used for optimal linear feedback control against small disturbances.

For this particular problem, the second variation method appears to be vastly superior to gradient methods, since approximately 60 iterations were needed for convergence to the optimum using a gradient method.

In this problem ϵ was set equal to one and no halving operations were necessary.

400 steps were used for the various integrations.

6.4 The Rocket Problem.[87]

It is desired to launch a rocket in fixed time to a given attitude with a given final vertical velocity component with maximum horizontal velocity component. The problem is simplified by making the following assumptions and approximations:

i) thrust varies with mass so as to produce constant acceleration

ii) the earth is flat

iii) gravitational acceleration is constant.

The following relevant quantities are defined:

x_1 = attitude

x_2 = velocity component in the horizontal direction

x_3 = velocity component in the vertical direction

a = constant acceleration due to thrust

g = constant gravitational acceleration

u = inclination of the thrust vector to the horizontal.

The motion of the rocket is governed by the differential equations

$$\dot{x}_1 = x_3$$

$$\dot{x}_2 = a \cos u$$

$$\dot{x}_3 = a \sin u - g$$

Given $x_1(t_o) = x_2(t_o) = x_3(t_o) = 0$

$t_o = 0, \quad t_f = 100$

$a = 64$ ft./sec.^{2}

$g = 32$ ft./sec.^{2}

find u(.) such that,

$x_2(t_f$ is a maximum

and $x_1(t_f) = 100{,}000$

$x_3(t_f) = 0$

It was decided to approximate the terminal constraints by means of penalty functions and thus the following modified cost function was formed.

Minimise $-x_2(t_f) + \alpha_1(x_1(t_f) - 100{,}000)^2 + \alpha_2(x_3(t_f))^2$

The equations to be solved backwards are,

$$\dot{\lambda}_1 = 0 \;; \quad \lambda_1(t_f) = 2\alpha_1(x_1(t_f) - 100{,}000)$$

$$\dot{\lambda}_2 = 0 \;; \quad \lambda_2(t_f) = -1$$

$$\dot{\lambda}_3 = -\lambda_1 \;; \quad \lambda_3(t_f) = 2\alpha_2 x_3(t_f)$$

$$\dot{l}_3 - \frac{aK_{33}\cos u}{\lambda_3 \sin u - \cos u}[\sin u - (l_3 - \lambda_3)\cos u] = 0; \quad l_3(t_f) = 0$$

$$-\dot{K}_{11} = \frac{a\cos^2 u}{\lambda_3 \sin u - \cos u} K_{13}^2 \;; \quad K_{11}(t_f) = 2\alpha_1$$

$$-\dot{K}_{13} = K_{11} + \frac{a\cos^2 u}{\lambda_3 \sin u - \cos u} K_{33}^2 \;; \quad K_{13}(t_f) = 0$$

$$-\dot{K}_{33} = 2K_{13} + \frac{a\cos^2 u}{\lambda_3 \sin u - \cos u} K_{33}^2 \;; \quad K_{33}(t_f) = 2\alpha_2$$

The improvement in control is given by

$$\delta u = \frac{1}{(\lambda_3 \sin u - \cos u)}[(\lambda_3 \cos u + \sin u + l_3 \cos u) + \cos u(K_{13}\delta x_1 + K_{33}\delta x_3)]$$

6.4.1 Discussion of Results

Extensive numerical experimentation has been done on this problem. The general conclusion is that the second variation method is not very suitable for this problem, unless a sufficiently good nominal trajectory is found. This is mainly because terminal constraints have to be satisfied and we have had considerable difficulty with approximating terminal constraints by means of penalty functions. The two terminal constraints in this problem are point constraints and depending on whether we 'overshoot' or 'undershoot' the constraints the function $\frac{\partial^2 H}{\partial u^2}$ changes radically.

Computing was done with two different nominal trajectories. The first nominal trajectory used was a linear approximation to the control. The same nominal control was used by Dreyfus. The control functions are shown in Fig. 6.10. The general shape of the trajectory is the same as Dreyfus's, excepting the 'transient' which occurs towards the end of the time interval. It is thought that this happens because the penalty function coefficients are too tight. On the other hand, when the value of these coefficients is made smaller, the terminal constraints are not satisfied within the specified tolerances. One possibility is to start with a low value for these coefficients and converge to a solution. These coefficients can then be updated and we can re-converge to the solution. We have had some difficulties in reconvergence too. Nevertheless engineering solutions for this problem have been obtained, although they are not as accurate as that obtained by Dreyfus. The behaviour of the cost as a function of the no. of iterations has been plotted in Fig. 3.11.

Results obtained with a different choice of penalty function coefficients is shown in Fig. 3.12.

Let us also mention that for this case the hybrid programme generally worked on the first variation.

The total computing time including input output was 5 mins.

A different set of results with a parabolic approximation to the optimal control as the nominal is shown in Fig. 3.12 and Fig. 3.13.

In this case the programme generally worked on the second variation.

It should also be mentioned that we have sometimes had some difficulty in integrating the Riccati equations in the Second Variation Algorithm. Generally speaking, these were cured by taking a smaller step size.

On the basis of the experience we have obtained, it is probably fair to say that satisfying terminal constraints is considerably more difficult than free end-point problems. More computing by the direct methods of satisfying terminal constraints we have presented in Chapter 5 needs to be done. Perhaps the most appropriate way of solving this problem is by the Newton-Raphson method.

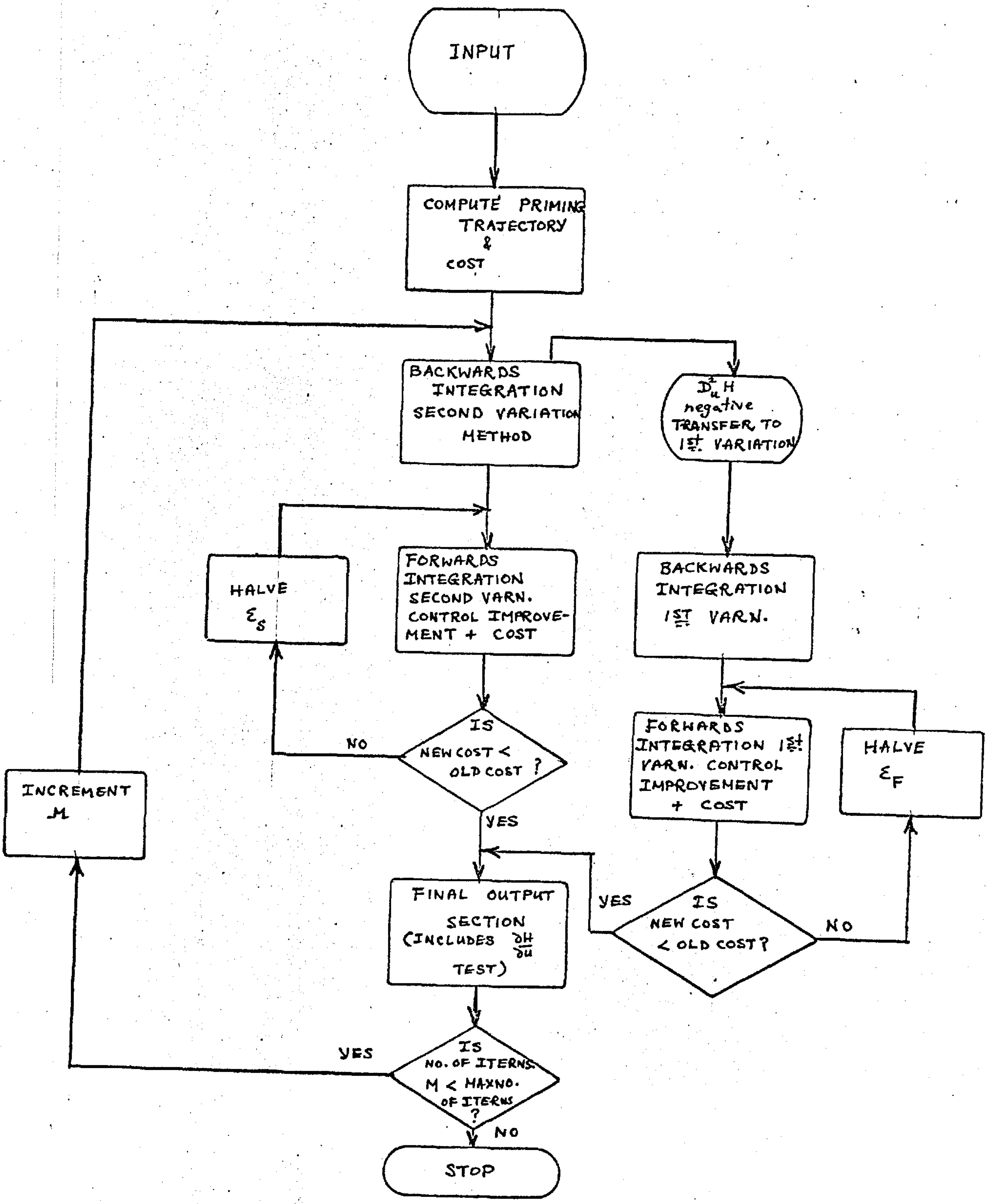


FIG. 6.1 OVERALL ORGANIZATION OF COMPUTER PROGRAMME

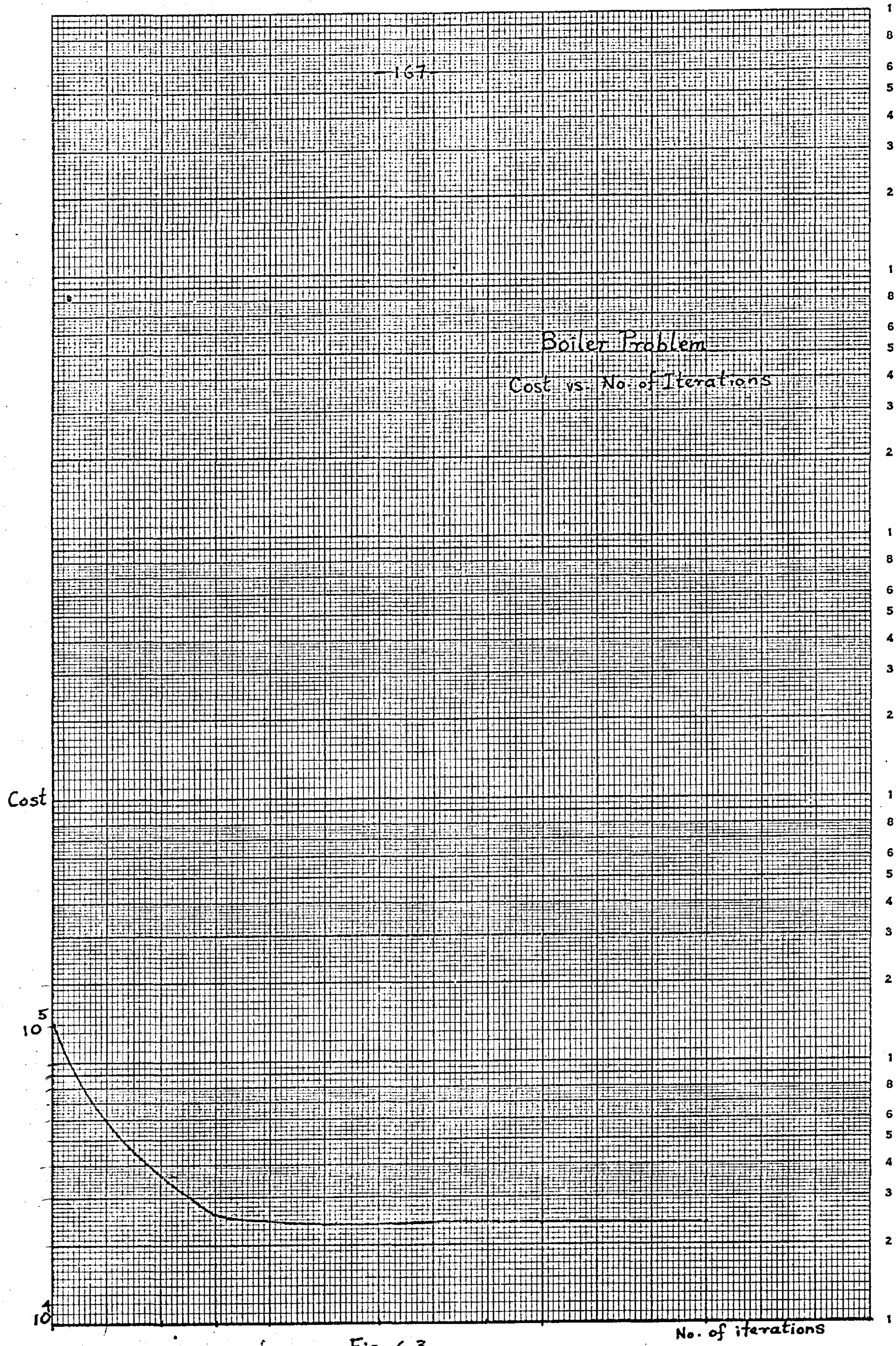

Boiler Problem
Cost vs. No. of Iterations
Cost
10^5
10^4
No. of iterations

Fig. 6.3

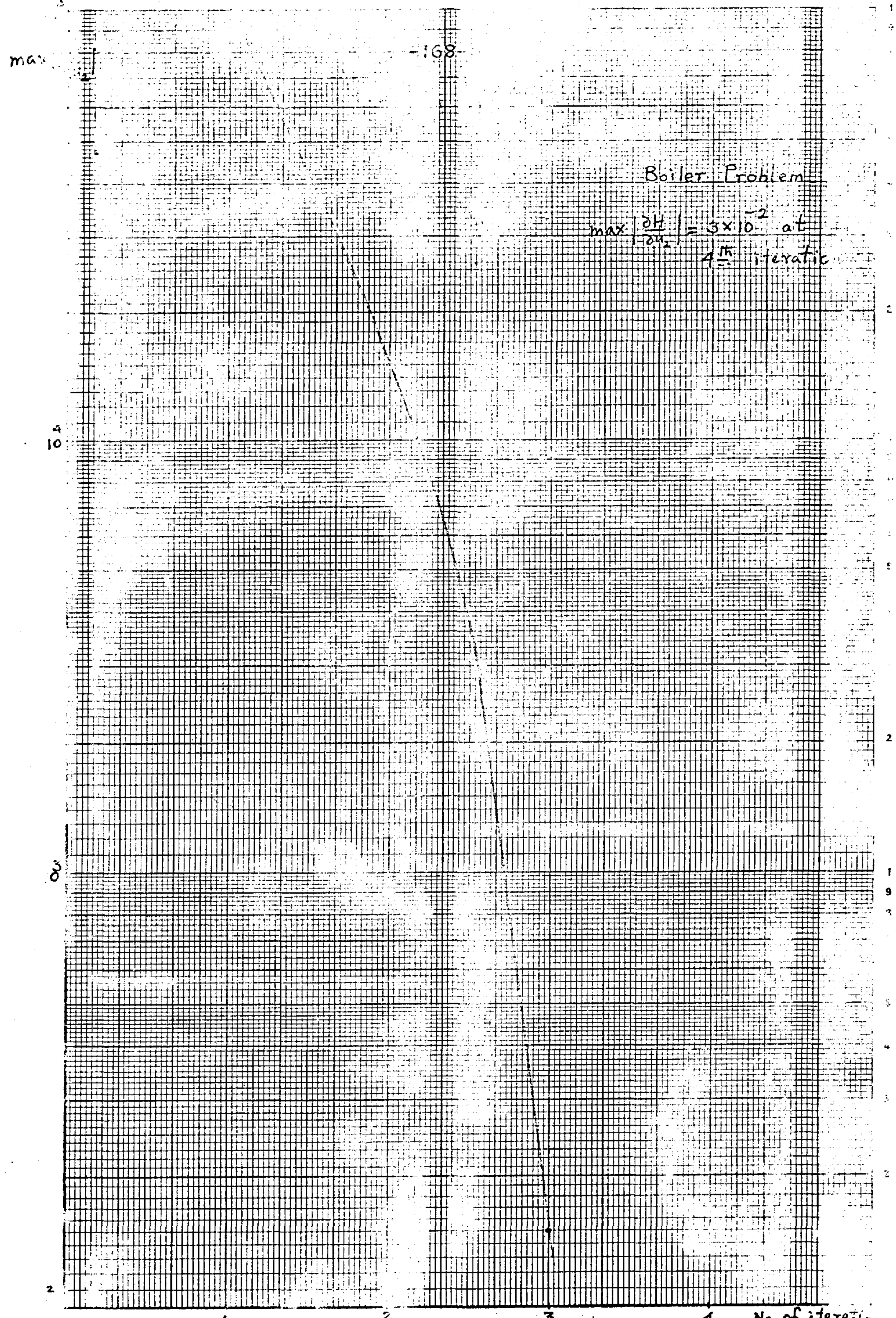
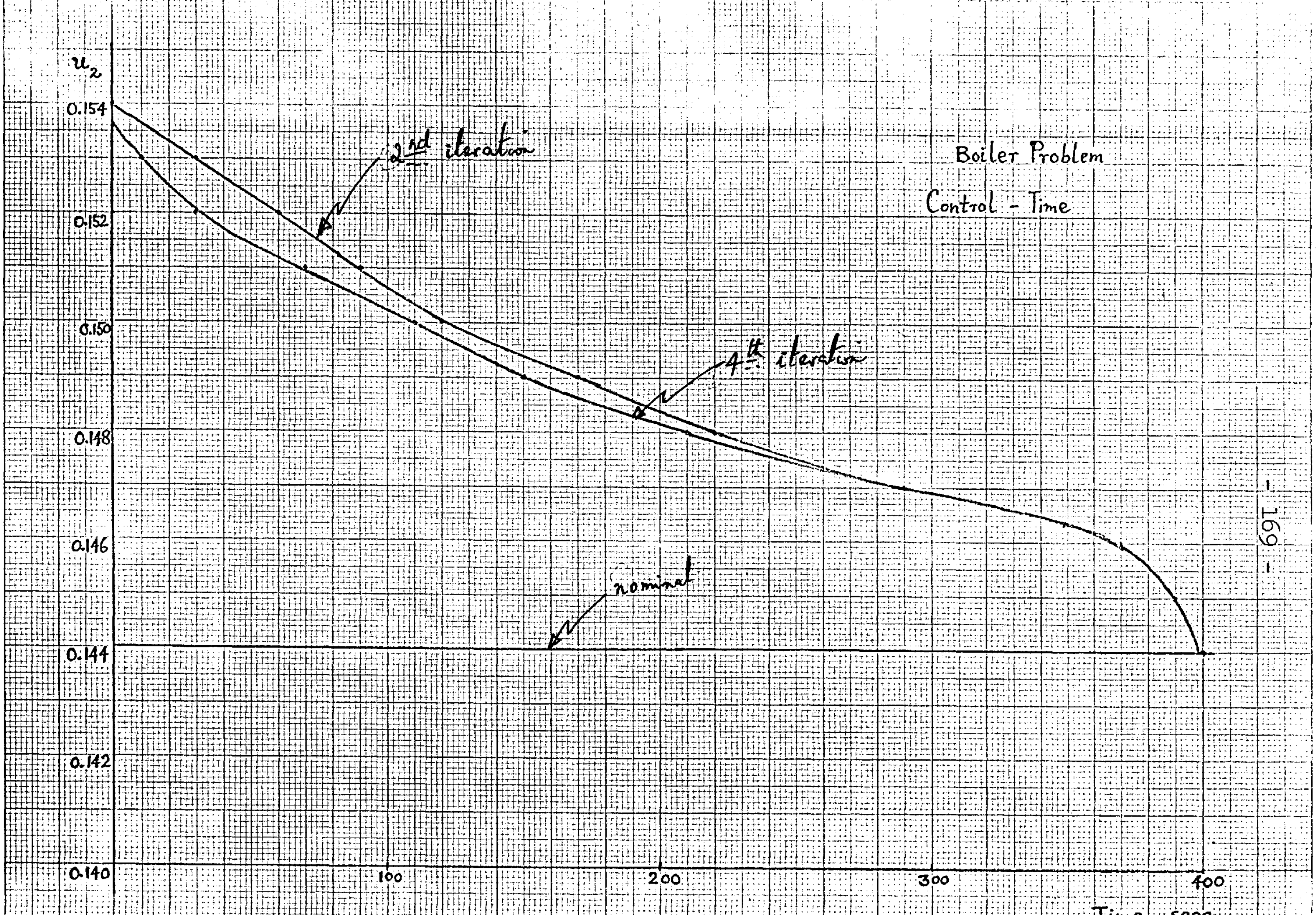
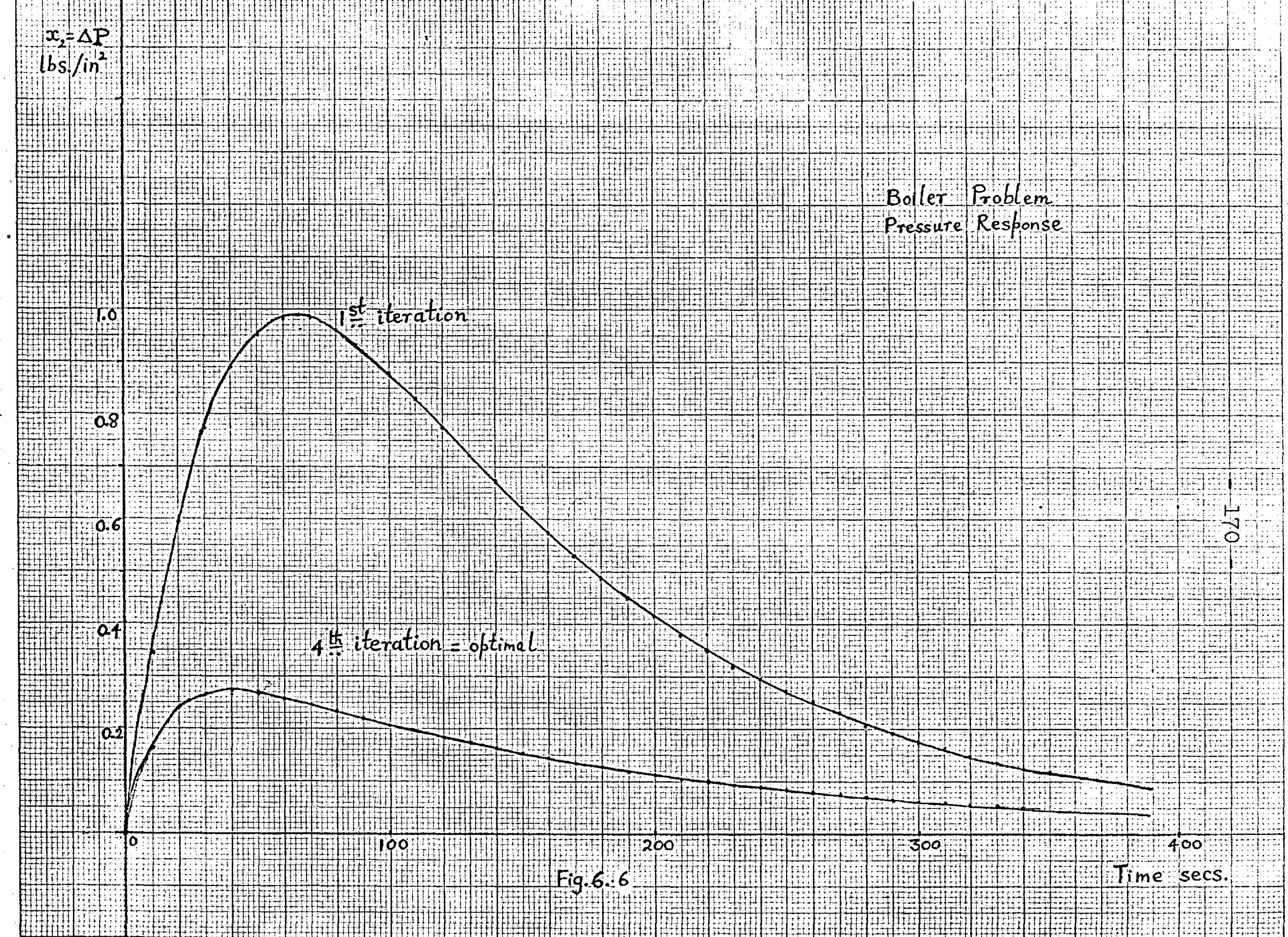

Boiler Problem
max $\left|\frac{\partial H}{\partial u_2}\right| = 3\times 10^{-2}$ at
4^{th} iteratic
max
10^4
10^3
2
1
2
3
4
No. of iterati.

Fig. 6.4



u2
0.154
0.152
0.150
0.148
0.146
0.144
0.142
0.140
100
200
300
400
Time - secs.
2nd iteration
4th iteration
nominal
Boiler Problem
Control - Time

Fig. 6.5


$x_2 = \Delta P$
lbs./in²
1.0
0.8
0.6
0.4
0.2
0
100
200
300
400
Time secs.
Boiler Problem
Pressure Response
1st iteration
4th iteration = optimal
Fig.6.6

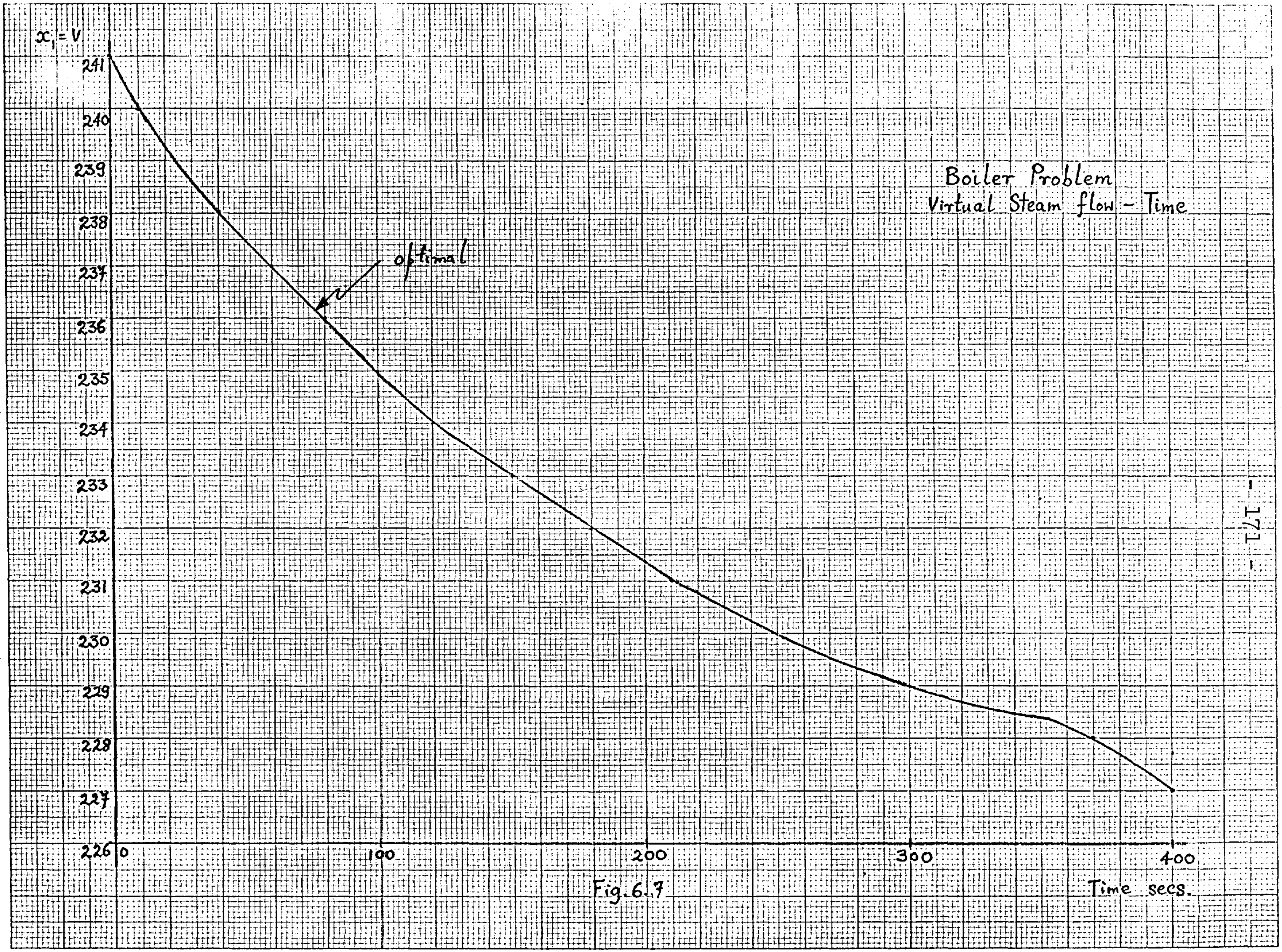

Boiler Problem
Virtual Steam flow - Time
optimal
$x_1 = V$
241
240
239
238
237
236
235
234
233
232
231
230
229
228
227
226
0
100
200
300
400
Time secs.

Fig. 6.7

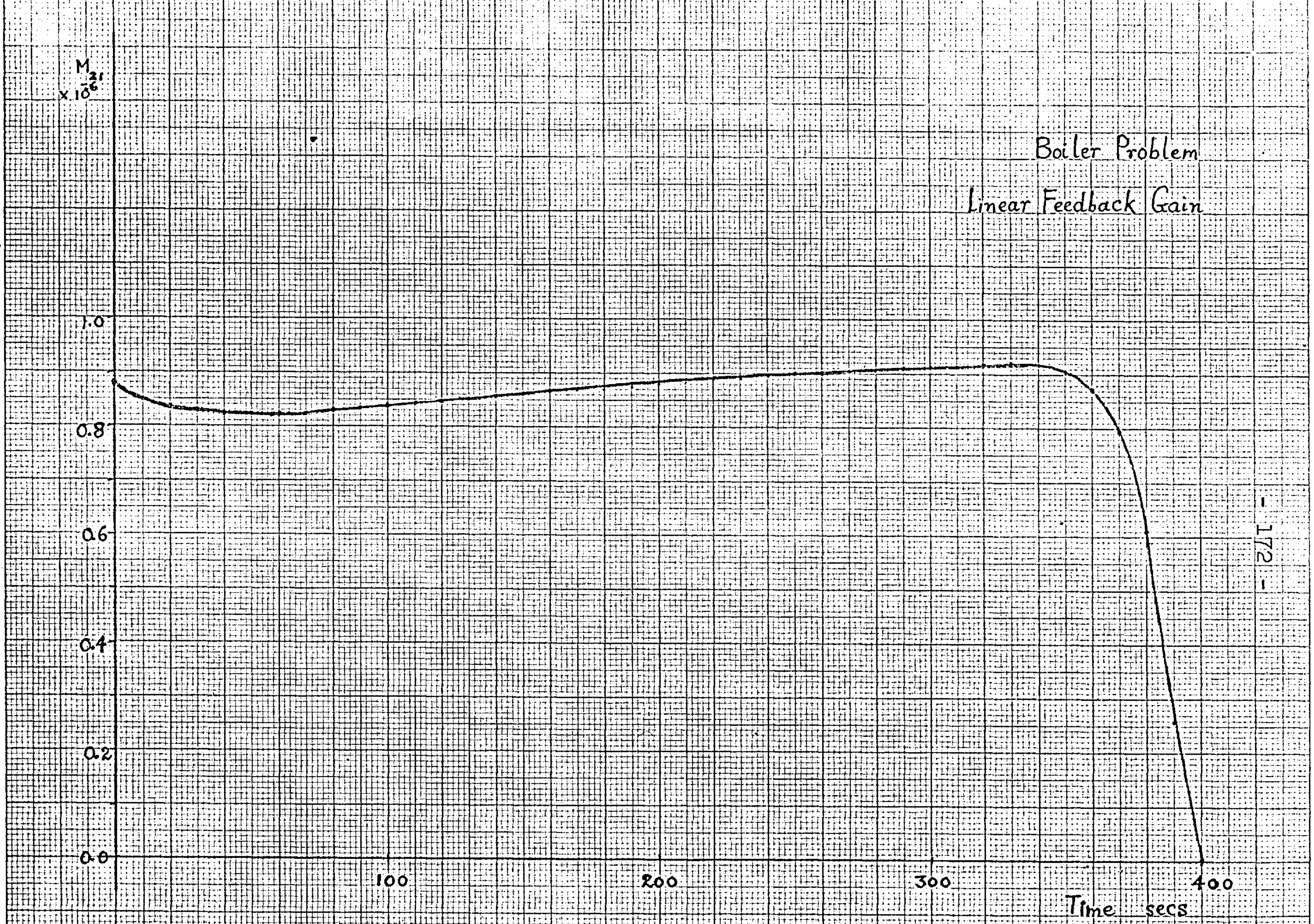

M21 x 10^-6
Boiler Problem
Linear Feedback Gain
1.0
0.8
0.6
0.4
0.2
0.0
100
200
300
400
Time secs

Fig. 6.8

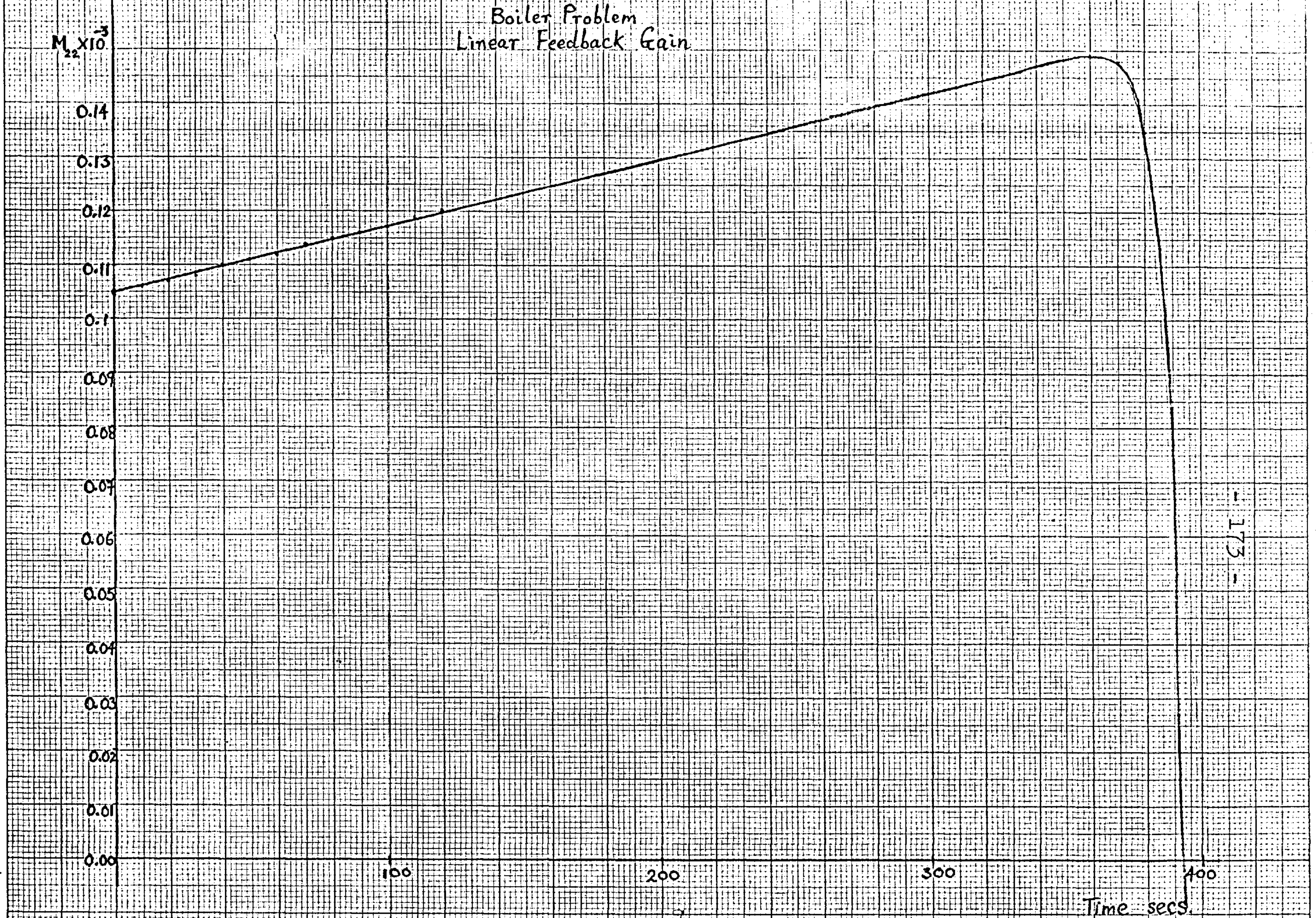
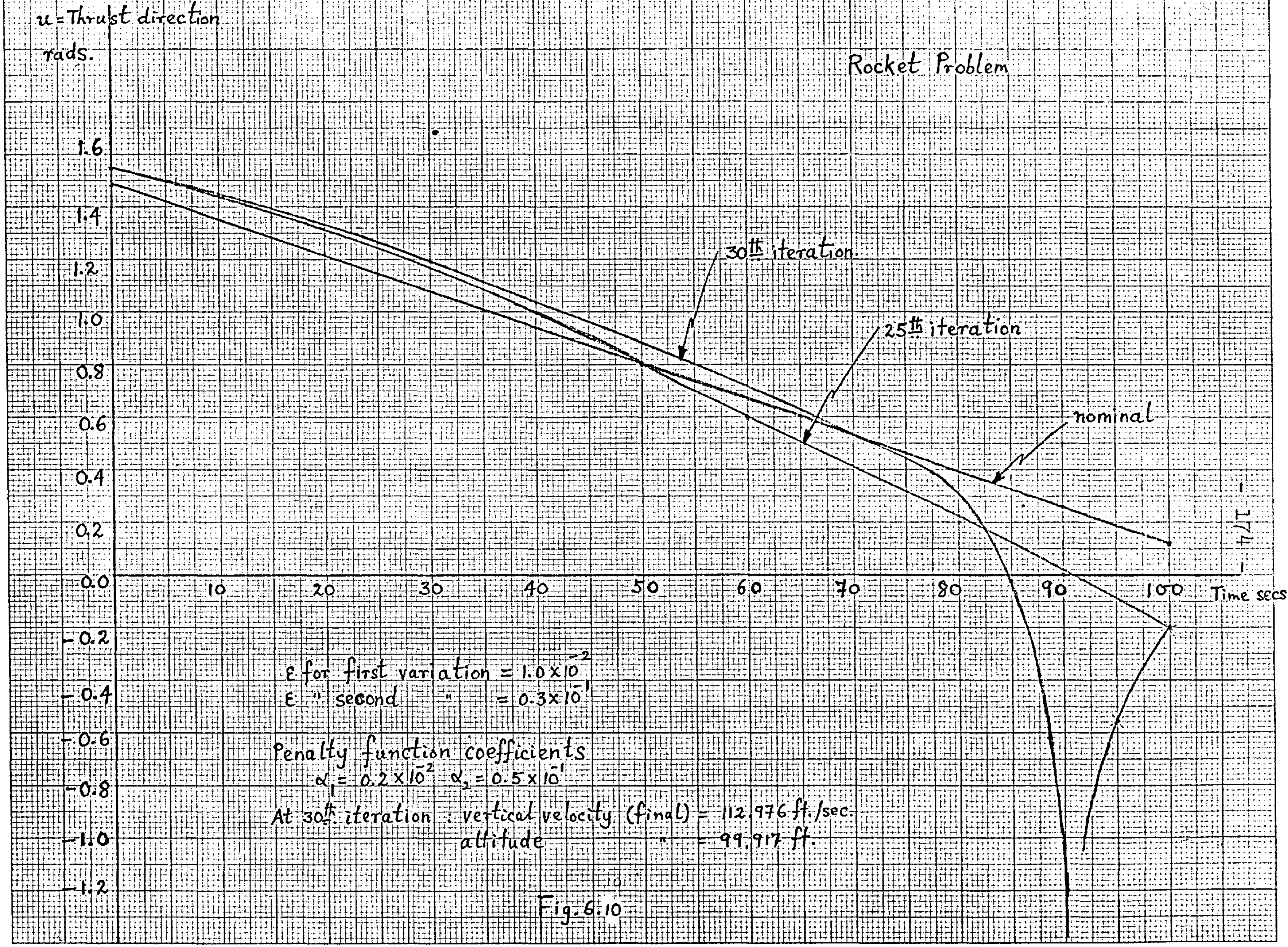

Boiler Problem
Linear Feedback Gain
M22 X10^-3
0.14
0.13
0.12
0.11
0.1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
100
200
300
400
Time secs.

Fig 6.9


u = Thrust direction
rads.
Rocket Problem
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
-1.2
10
20
30
40
50
60
70
80
90
100
Time secs
30th iteration.
25th iteration
nominal
ε for first variation = 1.0 x 10⁻²
ε " second " = 0.3 x 10⁻¹
Penalty function coefficients
α₁ = 0.2 x 10⁻² α₂ = 0.5 x 10⁻¹
At 30th iteration : vertical velocity (final) = 112.976 ft./sec.
altitude " = 99,917 ft.
Fig. 6.10

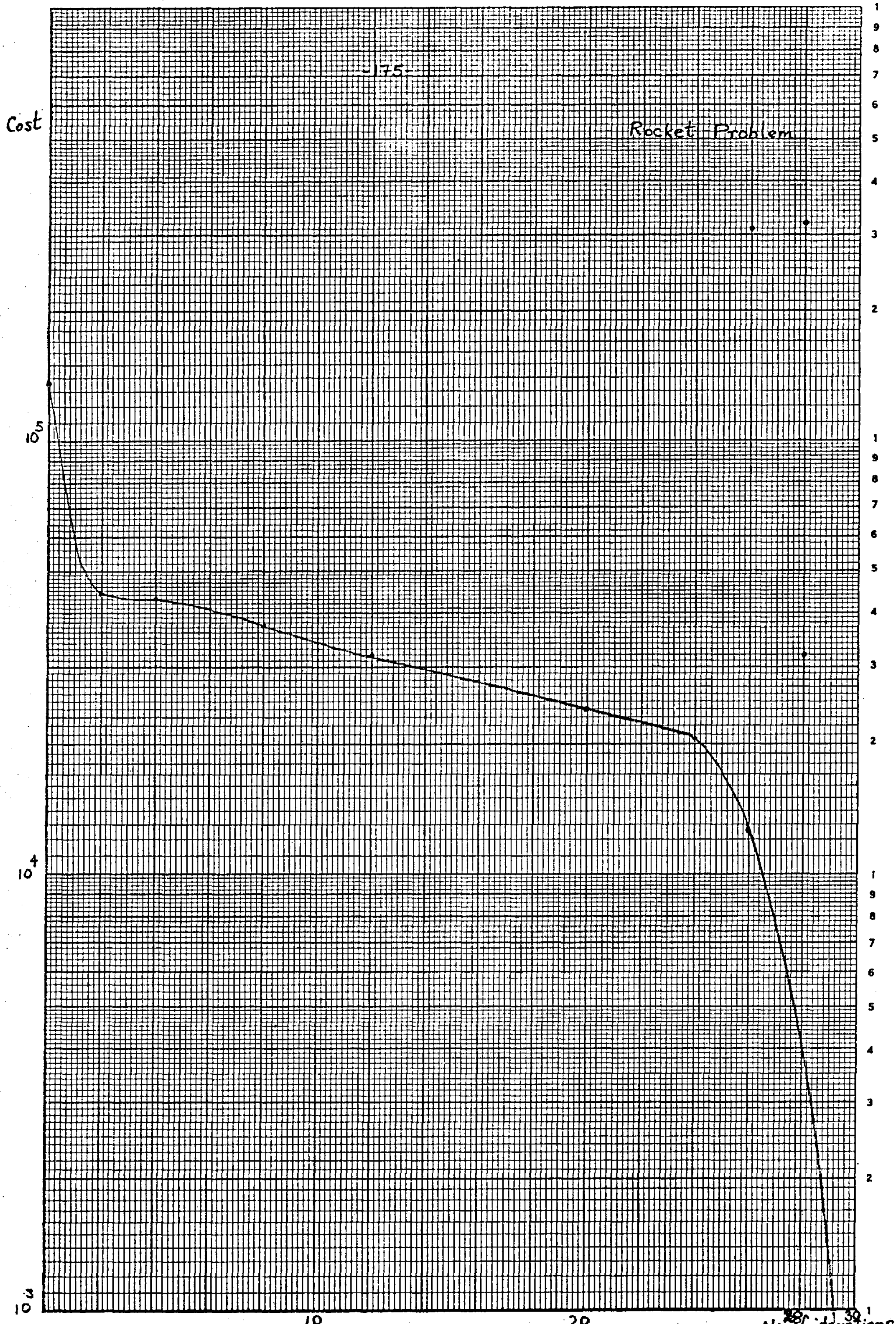


Fig. 6.11.

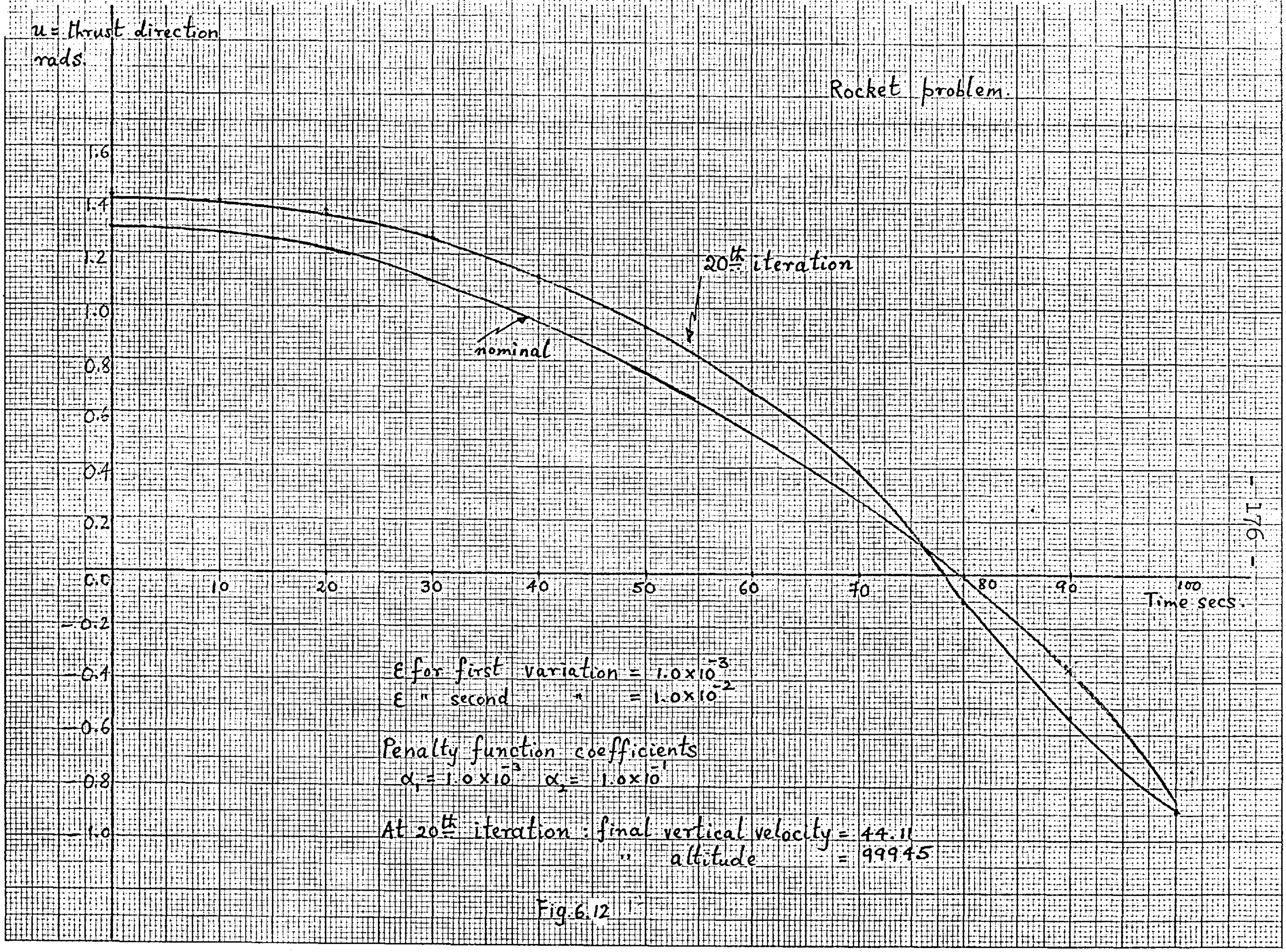

u = thrust direction
rads.
Rocket problem.
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
10
20
30
40
50
60
70
80
90
100
Time secs.
20th iteration
nominal
ε for first variation = 1.0 × 10⁻³
ε " second " = 1.0 × 10⁻²
Penalty function coefficients
α₁ = 1.0 × 10⁻³ α₂ = 1.0 × 10⁻¹
At 20th iteration : final vertical velocity = 44.11
" altitude = 99945

Fig. 6.12

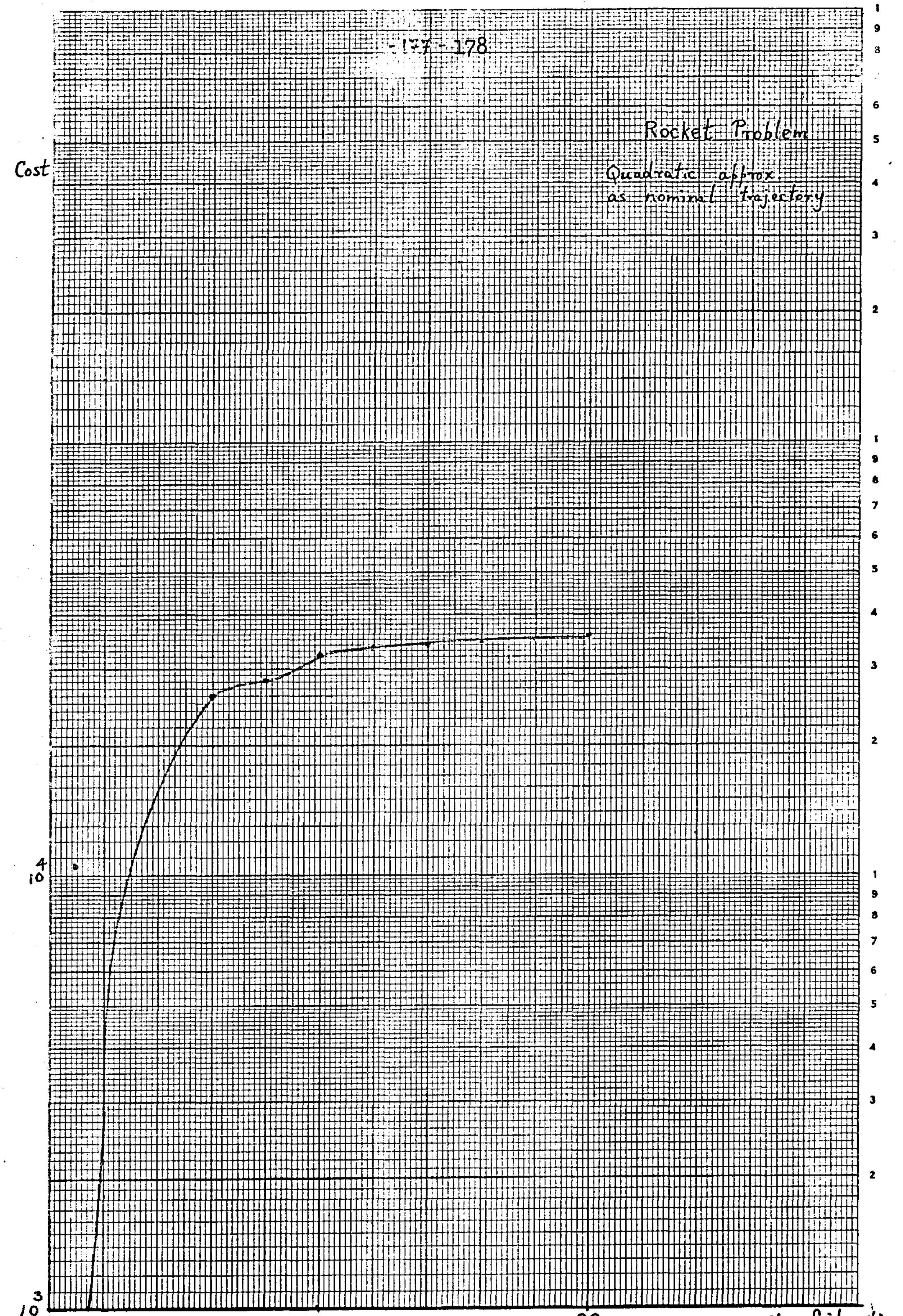

Rocket Problem
Quadratic approx. as nominal trajectory
Cost
10^4
10^3
10
20
No. of iteration

Fig. 3.13

CHAPTER 7

PROGRAMMING IN FUNCTION SPACE

7.1. Introduction

So far we have been concerned with systems having a finite dimensional state space. There are many systems for which the state space is infinite dimensional. Examples of such systems are provided by distributed parameter systems, systems with pure time delay and certain stochastic systems. For the theory of state space representation of such systems, we may refer to the recent work of Balakrishnan[90].

Our main objective in this chapter is to develop a suitable theory to handle optimal control problems having an infinite dimensional state space. Of course, the theory is equally applicable to systems having a finite-dimensional state space. We thus generalise and unify the results of previous chapters.

The theory that we develop is an extension of ideas from mathematical programming to solve programming problems in function spaces. Essentially we obtain generalisations of the Lagrange Multiplier Rule and the Kuhn-Tucker theorem of non-linear programming.

Before developing the theory, let us consider some examples of optimal control problems which may be solved by using the theory we develop. Let the distributed parameter system be described by the partial differential equation,

$$\frac{\partial f}{\partial t}(x,t) = Af(x,t) + u(x,t) \text{ ,...} \qquad (7.1)$$

where the system state at any instant of time is the function, $f(x,t)$, of time and a space variable x belonging to a suitable function space and the control $u(x,t)$ is distributed in the sense that it is a function of time and a space variable x. A is a linear operator usually unbounded. It is assumed that boundary conditions are taken care of by suitably restricting the domain of A.

Consider, for example, the problem of minimum energy control. This problem is to transfer an arbitrary initial state to the null state such that the control energy

$$\int_{t_o}^{t_f}\int_X u^2(x,t)dt \qquad (7.2)$$

is minimised.

Sometimes it may be more convenient to describe the system in terms of input-output behaviour. This might be the case if a model has been built from experimental data. For non-linear systems this could take the form of a Volterra Series

$$t(t) = \int_0^{t_f} K_1(t,\tau)u(\tau)d\tau + \int_0^{t_f}\int_0^{t_f} K_2(t,\tau_1,\tau_2)u(\tau_1) u(\tau_2)d\tau_1 d\tau_2 \qquad (7.3)$$

where u(.) is the input variable and y(.) the output variable. The above equation may be conveniently represented as,

$$y = N(u) \qquad (7.4)$$

where u and y are elements of suitable function spaces U and Y, and N is a non-linear mapping from U to Y. Assuming these function spaces to be normed linear spaces, we may require that the optimal input u(.) be chosen so that

$$\| N(u) - y_d \|^2_y \qquad (7.5)$$

be minimised, where y_d is the desired output and $\quad_y$ is the norm in Y-space, subject to the constraints

$$\| u \|^2_U \leq \alpha \qquad (7.6)$$

This chapter may be divided into two basic sections. In Section 7.2 we develop a theory of mathematical programming in Banach Spaces. In Section 7.3 we apply this theory to certain problems of optimal control, mainly to illustrate the theory.

7.2. Minimization Problems in Banach Spaces

7.2.1. Problem Statement

Let X and V be real Banach Spaces and let R be the space of reals. Let B be an open subset of X, and let Ω be a given closed set in X. Let the mappings f of X into R, g of X into V and h of X into R be given and we assume that f, g and h are twice continuously differentiable in the sense of Fréchet.

An element $x \in (B \cup \Omega)$ is said to be <u>admissible</u>. The problem to be solved is to find an admissible element $\hat{x}$ which satisfies the equation,

$$g(x) = 0_V \tag{7.7}$$

such that $f(\hat{x}) \leq f(x)$, for all $x \in (B \cup \Omega)$ satisfying equation (7.7).

The closed set Ω is defined by,

$$= [x : \ell(x) \leq 0] \tag{7.8}$$

We shall solve the problem by considering two cases:

i) $\hat{x} \in \text{Int}(\Omega)$, where $\text{Int}(\Omega)$ is the interior of

ii) $\hat{x} \in \partial\Omega$, where $\partial\Omega$ is the boundary of Ω.

7.2.2. Case (i) $x \in \text{Int}\,\Omega$

Let $N = [x : g(x) = 0_V,\ x \text{ admissible}]$

<u>Def:1</u>. The function g is said to be <u>regular</u> at the admissible point x if Dg(x) is surjective.*

*Throughout this chapter we shall generally follow the notation and terminology of Dieudonne[47].

Let g be <u>regular</u> at x and let

$$T = [h : Dg(x).h = 0_V].$$ T is a subspace of X

Def:2. Let $T_x = [x+h : h \in T]$. T_x is called the <u>linear tangent manifold</u> T_x of N at x

<u>Proposition 3.</u> (Liusternik and Sobolev)(85)

$(Dg(x))^{-1}$ exists and is linear.

<u>Assumption</u>: Range of Dg(x) is a closed subspace of V.*

<u>Proposition 4.</u> If g is not regular at x, there exists a $\hat{\lambda} \in \mathcal{L}(V;R)$, such that

$$\hat{\lambda}[Dg(\hat{x}).h] = 0 \tag{7.9}$$

for all $h \in X$.

<u>Proof</u>: Since range of Dg(x) has been assumed to be a closed subspace of V, from the Han-Banach theorem the above proposition follows.

<u>Proposition 5.</u> Let $\hat{x}$ be admissible and let g be regular at $\hat{x}$. Then $Df(\hat{x}).h = 0$ for all $h \in T$, if and only if there exists a $\hat{\lambda} \in \mathcal{L}(V;R)$ such that

$$Df(\hat{x}).h = \hat{\lambda}[Dg(\hat{x}).h]$$

<u>Proof</u>: The if part is trivial.

Let X/T be the factor space of cosets with respect to T. We note Dg(x) is a mapping from X/T onto V and

---

*We need to make this assumption, since from the fact that g is not regular at x we can only conclude that range of Dg(x) is not of IInd category.

from proposition 3 $[Dg(\hat{x})]^{-1}$ exists. Hence $[(Dg(\hat{x}))^*]^{-1}$ also exists and is linear. That is, the equation

$(Dg(\hat{x}))^* \hat{\lambda} = k$ has a unique solution for all $k \in [X/T]^*$.

But $Df(\hat{x}) \in [X/T]^*$. Hence the proposition.

We now look for conditions which guarantee the existence of a unique $\hat{\lambda} \in \mathcal{L}(V;R)$.

Assumption 1 (Constraint Qualification)

We shall say that the function $g(x)$ satisfies the constraint qualification at an admissible point $\hat{x}$ satisfying $g(x) = 0$ if and only if for every $h \in X$, $h \neq 0_X$, such that the equality $x = \hat{x} + h$ implies the equality

$g(x) = g(\hat{x}) + Dg(\hat{x}).h = 0$, there exists a function $\psi : [0,1] \to X$ with the following properties

i) $D\psi(\varepsilon).\tau$ exists for $0 \leq \varepsilon \leq 1$

ii) $\hat{x} = \psi(0)$

iii) $g[\psi(\varepsilon)] = 0$, $0 \leq \varepsilon \leq 1$

iv) $h = D\psi(0).\tau$, $\tau > 0$

Assumption 2. $X = X_1 \otimes X_2$ and $D_{x_2} g$ has an inverse.

Proposition 6. Assumption 2 implies Assumption 1.

Proof: Let $\hat{x} = (\hat{x}_1, \hat{x}_2)$. By the Implicit Function Theorem, there is an open neighbourhood U_1 of $\hat{x}_1$ in X_1 such that for every open connected neighbourhood U of $\hat{x}_1$, contained in U_1, there is a unique continuous mapping u of U into X_2 such that $u(\hat{x}_1) = \hat{x}_2$ and $g(x_1, u(x_1)) = 0$ for any

$x_1 \in U$. We also have, $u(x_1)$ is continuously differentiable in U and $Du(x_1) = -[D_{x_2} g(x_1,u(x_1))]^{-1}. [D_{x_1} g(x_1,u(x_1))]$

The result now clearly follows by considering $x_1 = \hat{x}_1 + \varepsilon h$, where ε is so small that $x_1 \in U$.

Combining propositions 5 and 6, we get,

Theorem 1 (Necessary Conditions)

Let $\hat{x}$ be a regular admissible point satisfying $g(x) = 0_V$. Then if assumption 2 holds and if $f(x)$ has a minimum at $x = \hat{x}$, then the point $\hat{x}$ must satisfy

$$\lambda_o Df(\hat{x}).h - \hat{\lambda}[Dg(\hat{x}).h] = 0 \qquad (7.10)$$

for all $h \in X$, where $\hat{\lambda} \in \mathcal{L}(V;R)$ is unique and $\lambda_o = 1$

Remark: From (7.10) we see that the problem of minimising $f(x)$ subject to minimising $g(x) = 0_V$ is equivalent to minimising the unconstrainted problem

$$F(x,\lambda) = f(x) - \lambda[g(x)] \qquad (7.11)$$

Case (ii) $\hat{x} \in \partial\Omega$

In view of our previous remark we may consider the problem of minimising,

$$f(x) - \lambda[g(x)] \qquad (7.12)$$

subject to $\ell(x) \leq 0$ (7.13)

The inequality constraint ℓ is required to satisfy the following constraint qualification:

The function ℓ satisfies the constraint qualification at an admissible point $x = \hat{x}$, satisfying $g(x) = 0_V$, if and only if for every $h \in X$, $h \neq 0_X$, such that the equality $x = \hat{x} + h$ implies the inequality

$\ell(x) = \ell(\hat{x}) + D\ell(\hat{x}).h \leq 0$, there exists a function $\psi : [0,1] \rightarrow X$ with the following properties

i) $D\psi(\varepsilon).\tau$ exists for $0 \leq \varepsilon \leq 1$

ii) $\hat{x} = \psi(0)$

iii) $\ell[\psi(\varepsilon)] \leq 0, \quad 0 \leq \varepsilon \leq 1$

iv) $h = D\psi(0).\tau\ , \ \tau > 0.$

Proposition 7. Let $\hat{x}$ minimise $f(x) - \hat{\lambda}[g(x)]$ subject to $\ell(x) \leq 0$ and suppose ℓ satisfies the constraint qualification at x. It then follows

$\ell(\hat{x}) + D\ell(\hat{x}).h \leq 0$ implies $-[Df(\hat{x}).h - \hat{\lambda}[Dg(\hat{x}).h] \leq 0$

Proof: Since ℓ satisfies the constraint qualification at $\hat{x}$, there exists a function $\psi(\varepsilon)$ with the properties shown above. Consider the function

$$\phi(\varepsilon) = f[\psi(\varepsilon)] - \hat{\lambda}[g(\psi(\varepsilon))]$$

Since $\phi(\varepsilon)$ has a minimum at $\varepsilon = 0$,

$$D\phi(0).\tau = Df(\psi(0)).D\psi(0).\tau - \hat{\lambda}[Dg(\psi(0)).D\psi(0).\tau]$$

$$= Df(\hat{x}).h - \hat{\lambda}[Dg(\hat{x}).h] \geq 0.$$

Proposition 7 is a generalisation of a result in Kuhn and Tucker's paper which is stated without any proof.

Proposition 8. (Farkas Lemma) (Ky Fan)

In order that $g(x) \geq 0$ be a consequence of $\ell(x) \geq 0$, it is necessary and sufficient that

$g(x) = \hat{\mu}f(x)$ where μ is a number ≥ 0.

Theorem 2. Let ℓ satisfy the constraint qualification at $\hat{x}$ and let $f(x) - \hat{\lambda}[g(x)]$ have a minimum at x subject to $\ell(x) < 0$. Then the point $\hat{x}$ must satisfy

$$Df(\hat{x}).h - \hat{\lambda}[Dg(\hat{x}).h] + \mu D\ell(\hat{x}).h = 0 \quad (7.14)$$

for all $h \in X$

$$\hat{\mu}\,\ell(\hat{x}) = 0\ ,\ \hat{\mu} \geq 0 \quad (7.15)$$

Proof: From propositions 7 and 8, we get (7.14).

Since x lies on the boundary of Ω,

$$\hat{\mu}\,\ell(\hat{x}) = 0.$$

Summarising,

Theorem 3. Let $X = X_1 \otimes X_2$ and let $\hat{x} = (\hat{x}_1, \hat{x}_2)$ be a regular admissible point, such that $[D_{x_2} g(\hat{x}_1, \hat{x}_2)]^{-1}$ exists. Let the inequality constraint ℓ satisfy the constraint qualification at $\hat{x}$. If f has a minimum at $\hat{x}$ subject to $g(x) = 0_v$, $\ell(x) \leq 0$, it is necessary that x satisfy

$$Df(\hat{x}).h - \hat{\lambda}[Dg(\hat{x}).h] + \hat{\mu}D\ell(\hat{x}).h = 0 \text{ for all } h \in X \quad (7.16)$$

$$\hat{\mu} \geq 0,\ \hat{\mu}\,\ell(\hat{x}) = 0 \quad (7.17)$$

We shall now put further restrictions on the functions f, g and ℓ. In particular we shall assume

i) $g(x) = A_1x_1 + A_2x_2 - c = 0$. (7.18)

where A_1 and A_2 are linear operators mapping X_1 to V and X_2 to V respectively such that A_2^{-1} exists, and c is a given element of V.

ii) $f(x)$ and $\ell(x)$ are convex functions of x. That is,

$$f(x) \geq f(\hat{x}) + Df(\hat{x}).(x-\hat{x})$$

and $$\ell(x) \geq \ell(\hat{x}) + D\ell(\hat{x}).(x-\hat{x}).$$

<u>Theorem 4</u>. (Necessary and Sufficient Conditions)

<u>If the above assumptions hold and if the assumptions of theorem 3 hold, then $\hat{x}$ minimises f(x) subject to</u> $A_1x_1 + A_2x_2 - c = 0$, $\ell(x) \leq 0$, <u>if and only if</u>

$$Df(\hat{x}).h - \hat{\lambda}[A_1h_1 + A_2h_2] + \hat{\mu}\, D\ell\,(\hat{x}).h = 0 \text{ for all } h \in X$$

$$\hat{\mu}\;\ell\;(\hat{x}) = 0\;,\; \hat{\mu} \geq 0.$$

<u>Proof</u>: The only if part follows from Theorem 3.

To prove the if part,

$$\begin{aligned} f(x) - f(\hat{x}) &\geq Df(\hat{x}).(x-\hat{x}), \text{ from convexity of } f \\ &= \hat{\lambda}[A_1(x_1-\hat{x}_1) + A_2(x_2-\hat{x}_2)] - \hat{\mu}D\,\ell\,(\hat{x}).(x-\hat{x}) \\ &\geq -\hat{\mu}\;(\;\ell(x) - \ell\,(\hat{x}))\;,\text{ from convexity of} \\ &= -\hat{\mu}\;\ell(x) \\ &\geq 0. \end{aligned}$$

We now want to extend the duality theory of non-linear programming to Banach Spaces. We shall assume that the various assumptions (including convexity) we have made hold for the discussion on duality. We first define the two dual problems:

I <u>Primal Problem</u>

Find $\hat{x} = (\hat{x}_1, \hat{x}_2)$ such that x satisfies

$$A_1x_1 + A_2x_2 - c = 0 \qquad (7.19)$$

$$\ell(x) \leq 0 \qquad (7.20)$$

and f(x) is minimised

II <u>Dual Problem</u>

Find $\hat{x}$, $\hat{\lambda}$, $\hat{\mu}$ satisfying

$$Df(x).h - \lambda[A_1h_1 + A_2h_2] + \mu D\ell(x).h = 0 \qquad (7.21)$$

for all $h = (h_1, h_2) \in X$,

$$\mu \geq 0 \qquad (7.22)$$

such that

$$L(x,\lambda,\mu) = f(x) - \lambda[A_1x_1 + A_2x_2' - c] + \mu\ell(x) \qquad (7.23)$$

is maximised.

<u>Theorem 5 (Duality Theorem)</u>

<u>If $\hat{x} = (\hat{x}_1, \hat{x}_2)$ minimises f(x) subject to (7.19) and (7.20), then $(\hat{x}, \hat{\lambda}, \hat{\mu})$ maximises $L(x,\lambda,\mu)$ subject to (7.21) and (7.22) and further $f(\hat{x}) = L(\hat{x}, \hat{\lambda}, \hat{\mu})$.</u>

Proof: From Theorem 4, we know that there exists $(\hat{x},\hat{\lambda},\hat{\mu})$ such that (7.21) and (7.22) are satisfied. Let us assume the existence of (x,λ,μ) such that

$$L(x,\lambda,\mu) > L(\hat{x},\hat{\lambda},\hat{\mu})$$

$$\begin{aligned} L(\hat{x},\hat{\lambda},\hat{\mu}) &= f(\hat{x}) - \hat{\lambda}[A_1\hat{x}_1+A_2\hat{x}_2-c] + \hat{\mu}\, \ell(\hat{x}) \\ &\geq f(\hat{x}) - \lambda[A_1\hat{x}_1+A_2\hat{x}_2-c] + \mu\, \ell(\hat{x}) \\ &= L(\hat{x},\hat{\lambda},\mu). \end{aligned}$$

Hence $L(x,\lambda,\mu) > L(\hat{x},\hat{\lambda},\hat{\mu}) \geq L(\hat{x},\hat{\lambda},\mu)$

$$L(\hat{x},\hat{\lambda},\mu) \geq L(x,\lambda,\mu) + Df(x).(\hat{x}-x) - \lambda[A_1(\hat{x}_1-x_1) + A_2(\hat{x}_2-x_2)] + \mu\, D\ell(x).(\hat{x}-x) \qquad (7.24)$$

But since $L(x,\lambda,\mu) > L(\hat{x},\hat{\lambda},\mu)$, we must have from (7.24)

$$Df(x).(\hat{x}-x)-\lambda[A_1(\hat{x}_1-x) + A_2(\hat{x}_2-x_2)] + \mu\, D\ell(x).(\hat{x}-x) < 0$$

which is a contradiction.

Hence $\hat{x},\hat{\lambda},\hat{\mu}$ maximises L. From the construction of the dual problems, it is clear

$$L(\hat{x},\hat{\lambda},\hat{\mu}) = f(\hat{x}).$$

Hence the theorem.

7.3. Applications to Optimal Control Problems

7.3.1. Linear System with Quadratic Performance Functional

Let H_1 and H_2 be Hilbert Spaces.

Consider the linear dynamical system, whose state space

H_1 is infinite dimensional, and whose evolution of state is governed by the linear differential equation

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) \tag{7.25}$$

and the initial state $x(t_o)$ is given. For each $t \in [t_o, t_f]$, $x(t) \in H_1$ and $u(t) \in H_2$ and $A(t)$ and $B(t)$ are linear continuous mappings mapping $H_1 \to H_1$ and $H_2 \to H_1$ respectively. It is assumed that the differential equation has a unique solution given $u(.)$ and $x(t_o)$. For suitable conditions, see Dieudonne[47].

We consider the problem of minimising

$$P(x(t_o).u) = \frac{1}{2} \int_{t_o}^{t_f} [\langle x(t),P(t)x(t)\rangle_1 + \langle u(t),R(t)u(t)\rangle_2]dt \tag{7.26}$$

where $R(t)$ is a symmetrix, positive definite operator, bounded away from zero mapping $H_1 \to H_1$ and $P(t)$ is a symmetrix positive definite operator mapping $H_2 \to H_1$ and $\langle .,.\rangle_1$ and $\langle .,.\rangle_2$ represent inner products in H_1 and H_2 respectively.

The solution of equation (7.25) is given by

$$x(t) = C(t,t_o).x(t_o) + \int_{t_o}^{t} C(t,\tau)B(\tau)u(\tau)d\tau \tag{7.27}$$

where $C(t,t_o)$ is the solution of the homogeneous equation,

$\dot{C}(t,t_o) = A(t).C(t,t_o)$; $C(t_o,t_o) = I.$

We shall assume that the system is completely controllable in $[t_o,t_f]$ where

Def:(91) A phase $(x(t_o),t_o)$ is said to be controllable if there exists some finite $t_1 > t_o$ and some admissible control $u(t)$ which transfers $(x(t_o),t_o)$ to $(0,t_1)$. If every phase is controllable, then the system is said to be completely controllable.

Let us define the linear operator,

$$L(t_o,t_f) = \int_{t_o}^{t_f} C(t_f,\tau)B(\tau)u(\tau)d\tau \qquad (7.28)$$

and let $L^*(t_o,t_f)$ be the operator adjoint to L.

Let $S(t_o,t_f) = L(t_o,t_f)\, L^*(t_o,t_f)$.

By analogy from the finite dimensional case(91), it is clear that a necessary and sufficient condition for the system to be completely controllable is that the self-adjoint operator $S(t_o,t_f)$ have a bounded inverse.

Introduce the transformation

$$x(t) = x(t_o) + \int_{t_o}^{t} \dot{x}(\tau)d\tau \qquad (7.29)$$

Then the problem reduces to minimising,

$$P(x(t_o),u) = \frac{1}{2} \int_{t_o}^{t_f} (\langle x(t_o)+ \int_{t_o}^{t} \dot{x}(\tau)d\tau \, , \, P(t)[x(t_o)+$$

$$+ \int_{t_o}^{t} \dot{x}(\tau)d\tau \,]\rangle_1 + \langle u(t),R(t)u(t)\rangle)dt \qquad (7.30)$$

subject to $\dot{x}(t) = A(t)[x(t_o)+ \int_{t_o}^{t} \dot{x}(\tau)d\tau\,]+B(t)u(t)$ (7.31)

Applying Theorem 1 and simplifying, if $\hat{u}(.)$ and $\hat{x}(.)$ are optimal they satisfy,

$$\lambda_o R(t)u(t) + B^*(t)\hat{\lambda}(t) = 0 \qquad (7.32)$$

where $\lambda_o \geq 0$ and $B^*(t)$ is the adjoint operator to $B(t)$*,

and

$$\hat{\lambda}(t) = \int_{t}^{t_f} [\lambda_o P(\tau)x(\tau) + A^*(\tau)\hat{\lambda}(t)]d\tau$$

If $\lambda_o = 0$, the controllability assumption is violated. Hence $\lambda_o > 0$ and can be set equal to 1 which determines $\hat{\lambda}(.)$ uniquely.

We now have to investigate, whether the equations

---

*The assumption of regularity corresponds to $\lambda_o > 0$.

$$x(t) = x(t_o) + \int_{t_o}^{t} [A(\tau)x(\tau) - B(\tau)R^{-1}(\tau)B^*(\tau)\hat{\lambda}(\tau)]d\tau \tag{7.33}$$

$$\hat{\lambda}(t) = \int_{t}^{t_f} [P(\tau)x(\tau) + A^*(\tau)\hat{\lambda}(\tau)]d\tau \tag{7.34}$$

have a unique solution.

Let $K_{11}(t,\tau) = A(\tau) \quad , 0 \leq \tau \leq t \leq t_f$

$= 0 \quad , 0 \leq t < \tau \leq t_f$

$K_{12}(t,\tau) = -B(\tau)R^{-1}(\tau)B(\tau) \quad , 0 \leq \tau \leq t \leq t_f$

$= 0 \quad , 0 \leq t < \tau \leq t_f$

$K_{21}(t,\tau) = 0 \quad , 0 \leq \tau \leq t \leq t_f$

$= P(\tau) \quad , 0 \leq t < \tau \leq t_f$

$K_{22}(t,\tau) = 0 \quad , 0 \leq \tau \leq t \leq t_f$

$= A^*(\tau) \quad , 0 \leq t < \tau \leq t_f$

The pair of equations (7.33) and (7.34) can be written as

$$z(.) = z(t_o) + K(t_o,t_f)\, z(.) \tag{7.35}$$

where $z(.) = \begin{pmatrix} x(.) \\ \hat{\lambda}(.) \end{pmatrix}$, $z(t_o) = \begin{pmatrix} x(t_o) \\ 0 \end{pmatrix}$

and $K(t,t_f)$ is the integral operator

$$\int_{t_o}^{t_f} \begin{pmatrix} K_{11}(t,\tau) & K_{12}(t,\tau) \\ K_{21}(t,\tau) & K_{22}(t,\tau) \end{pmatrix} (.)\, d\tau$$

If H_1 and H_2 are finite-dimensional and if the kernel of the integral operator is square summable, then it is easily shown that the Fredholm Alternative holds and hence the solution exists and is unique. In this case the solution is given by,

$$z(.) = z(t_o) + C(t,t_f)\, z(.) \tag{7.36}$$

where $C(t_o,t_f)$ is the resolvent transformation, and the optimal control is given by,

$$u = -R^{-1}(t)B\ (t)C_{21}(t,t_f)\ [I + C_{11}(t,t_f)]^{-1}x,$$

provided $[I + C_{11}(t,t_f)]^{-1}$ exists.
The control thus is in linear feedback form.

For the general case application of a fixed point principle will yield a sufficient condition for the existence and uniqueness of a solution.

Consider now the same finite dimensional problem, but with the performance functional,

$$\int_{t_o}^{t_f} \|x_d(t) - x(t)\|^2_{E^n} dt$$

and an energy type constraint

$$\int_{t_o}^{t_f} \|u\|^2 dt \leq M$$

Assuming that $H_1 = L_n^2[t_o, t_f]$ and $H_2 = L_m^2[t_o, t_f]$ with inner products

$$[x,y]_1 = \int_{t_o}^{t_f} \langle x(t), y(t)\rangle_{E^n} dt$$

and $$[x,y]_2 = \int_{t_o}^{t_f} \langle x(t), y(t)\rangle_{E^m} dt$$

and introducing the linear transformation,

$$Lu = \int_{t_o}^{t} \Phi(t,\tau)\, B(\tau)\, u(\tau) d\tau$$, we have

the following convex-programming problem in H_2:

Minimise $[x_d - Lu,\ x_d - Lu]_1$

subject to $[u,u]_2 \leq M$

We can now apply theorem 4.

We form the Lagrangian,

$$\phi(u,\mu) = [x_d - Lu,\ x_d - Lu] + \hat{\mu}([u,u] - M)$$

$$D_u\phi, h] = 2[-Lh,\ x_d - Lu] + 2\hat{\mu}([h,u]) = 0$$

from which the optimal $u = \hat{u}$ satisfies

$$-L^{*}(x_d - Lu) + \hat{\mu}u = 0.$$

and also the conditions $\hat{\mu} \geq 0$ and $\hat{\mu}([\hat{u},\hat{u}] - M) = 0$.

If $\hat{\mu} > 0$,

$$\hat{u} = (L^{*}L + \hat{\mu}I)^{-1} L^{*}x_d$$

Essentially the same result has been obtained by Balakrishnan[30] from a slightly different point of view.

It is interesting to formulate the Dual problem. The dual problem is,

Maximise $[x_d - Lu, x_d - Lu]_1 + \mu([u,u]_2 - M)$

subject to $(L^{*}L + \mu I)u = L^{*}x_d$

$\mu \geq 0$

The dualproblem appears to be closely connected with the theory of filtering and prediction.

7.3.2. A Quadratic Programming Problem

Consider the linear system

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) \; ; \; x(t_o) = 0 \qquad (7.37)$$

where for each t, $x(t) \in E^n$ and $u(t) \in E^m$ and A and B are $n \times n$ and $n \times m$ matrices. A(t) and B(t) are assumed to be bounded for each t.

It is required to find u(.) and x(.) such that,

$$\int_{t_o}^{t_f} [\langle x(t),x(t)\rangle + \langle u(t),u(t)\rangle]dt$$

is minimised and the constraint

$$a \leq x(t_f) \leq b \qquad (7.38)$$

is satisfied.

In exactly the same way as in the previous case the problem can be reduced to

Minimise $[Lu,Lu]_1 + [u,u]_2$

subject to $a \leq \int_{t_o}^{t_f} K(t_f,\tau)u(\tau)d\tau \leq b$, where

$K(t_f,\tau) = \Phi(t_f,\tau)B(\tau)$.

Denote the rows of the matrix $K(t_f,\tau)$ by $k_1,k_2,\ldots k_n$. Hence the constraints can be written as

$$[k_i,u] \leq b$$

$$-[k_i,u] \leq a \qquad i = 1,2,\ldots..n.$$

If the system is now assumed to be completely controllable in $[t_o,t_f]$ then L^*L is positive definite.

We can now introduce the new inner product and norm as,

$$(u,v) = [u,v] + [u,L^*Lv]$$

and the norm $\|u\|^2 = [u,L\ Lu] + [u,u]$

$$[u,k_i] = [u,(I+L^*L)^{-1}(I+L^*L)k_i]$$

$$= [(I+L^*L)^{-1}u,(I+L^*L)k_i]$$

$$= ((I+L^*L)^{-1}u,\ k_i)$$

Hence we finally have the problem,

Minimise $\|u\|^2$

$$\text{subject to } ((I+L^*L)^{-1}u,\ k_i) \leq b \qquad i = 1,2,\ldots.n$$
$$-(I+L^*L)^{-1}u,\ k_i) \leq a$$

This problem can now be solved using certain results of Ky Fan[52].

7.3.3. Distributed Parameter Systems

We shall only briefly indicate how we could solve distributed parameter systems using this theory. It is possible to use the theory to derive formally Euler-Lagrange equations for a distributed parameter problem. As an example consider the linear system,

$$\frac{\partial f}{\partial t}(x,t) = A\ f(x,t) + Bu(x,t)\ , \text{ where}$$

A is the infinitesimal generator of a strongly continuous semi-group, f and u are elements of suitable function spaces (which we shall take to be a Hilbert Space) and B is a linear operator mapping u into the domain of A.

Consider the problem of minimising

$$\frac{1}{2} \int_{\Omega}\int_{t_o}^{t_f} (f^2(x,t) + u^2(x,t))dt$$

Then a formal application of our results yield the

Euler-equations,

$$\frac{\partial f}{\partial t}(x,t) = Af(x,t) - B\lambda(x,t)$$

$$\frac{\partial \lambda}{\partial t}(x,t) = -A^{*}\lambda(x,t) - f(x,t)$$

A more rigorous approach is possible by using the theory of semi-groups.

7.4. Iterative Procedures

In principle, gradient procedures or Newton's method could be applied to solve problems of distributed parameter systems. In practice, the problems would have to be approximated by lumped parameter models and then solved by some iterative procedure.

Methods for solving non-linear programming problems like the gradient projection method and the method of feasible directions can be generalised to Hilbert Spaces. These however are not too constructive; for example, to use the gradient projection method, explicit knowledge of the projection operator is needed.

We would also like to mention that sufficient conditions for convergence of the iterative procedures presented in Chapter 5 have been obtained using a function space approach. These have been omitted from this thesis since it is of only mathematical interest.

7.5. Multi-level Control and Programming

We would briefly like to mention that the results we have obtained enable us, in principle, to develop a theory of multi-level programming and control for dynamical systems. The details of this are beyond the scope of the present thesis.

CHAPTER 8

CONCLUSIONS AND FURTHER RESEARCH

In this thesis we have demonstrated how a wide variety of optimal control problems could be solved using ideas of variational calculus and mathematical programming. We have also extended results of mathematical programming to deal with problems in function space. Indirectly, we have tried to show that dynamic problems of mathematical economics and operations research are essentially the same as problems of optimal control.

In Chapter 2 we have solved a general class of non-linear discrete time optimal control problems using methods of mathematical programming. Many erroneous results on discrete time control problems have been published in the literature. Mathematical programming methods allow us to derive very simply but vigorously various results for control problems.

In Chapter 3 we have shown how a dynamic allocation problem could be reformulated as an optimal control problem with control and state variable inequality constraints and solved using mathematical programming methods.

Chapter 4 was devoted to second order necessary conditions and sufficient conditions for a class of continuous time optimal control problems. Second order conditions have so far been neglected in the literature of optimal control.

In Chapter 5 we consider second order iterative algorithms for solving optimal control problems. The second variation method we have presented can be used to obtain feedback solutions in a suitably small region of the state space.

In Chapter 6 we present computer results for the solution of two optimal control problems. In particular, for the boiler problem the open loop programme and the feedback gains could probably be implemented in practice.

Finally, in Chapter 7 we extend existing results of mathematical programming to function spaces and show how they can be applied for the solution of infinite dimensional control problems.

Various areas of further research suggest themselves. We would like to highlight three areas where results could be obtained.

I. For discrete-time optimal control problems two areas of research using mathematical programming methods are

i) Use of stochastic programming to solve stochastic control problems

ii) application of the theory of games to control problems. Much of the groundwork for this is contained in the book of Blackwell and Girschik.[93]

II. Much more computational experience is needed for the second variation methods we have presented. In particular it would be interesting to obtain some idea of the region of state space where the methods work without any modification.

Two main extensions are

a) The method should be extended to handle inequality constraints. It is felt that ideas of duality could be used to advantage here.

b) Assume a model for the second variation and up-date the parameters on the basis of information obtained from previous iterations. This would then be analogous to methods of Fletcher and Powell for ordinary minimization problems.

III. Pearson's results on Duality could be considered as a special case of our Duality result in Chapter 7. It would be interesting to develop a duality theory for a class of stochastic control problems. The relationship of the duality between optimal control and optimal filtering and our duality results should be investigated. Application of the theory presented in Chapter 7 to obtain more detailed results for stochastic and distributed parameter systems should be attempted.

Appendix A

Some Results of Non-Linear Programming

In this Appendix se shall briefly summarise some results of non-linear programming.

Let $x \in E^n$. The problem we are trying to solve is:

Minimise the real-valued function $f(x)$ subject to the constraints

$$g_i(x) = 0 \ , \quad i = 1,2,\ldots,m \tag{A.1}$$

$$h_i(x) \leq 0 \ , \quad i = 1,2,\ldots,p \tag{A.2}$$

Let $A = \{x : h.(x) \leq 0\}$. Let $\hat{x}$ be the minimising x. Let us first assume that $\hat{x}$ is in the interior of A, that is, the constraints $h_i(x) \leq 0$ are not operative. We can then solve the problem by the Lagrange Multiplier rule. In order to ensure the existence and uniqueness of the Lagrange multiplier $\hat{\lambda} \in E^m$, we have to make some assumptions.

Let $G = (g_1 \ldots g_m f)^T$

$$DG(x) = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \cdots\cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & & \\ \frac{\partial g_m}{\partial x_1} & \cdots\cdots & \frac{\partial g_m}{\partial x_n} \\ \frac{\partial f}{\partial x_1} & \cdots\cdots & \frac{\partial f}{\partial x_n} \end{bmatrix}$$

$$Dg(x) = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \cdots\cdots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_m}{\partial x_1} & \cdots\cdots & \frac{\partial g_m}{\partial x_n} \end{bmatrix}$$

If Rank $(DG(\hat{x}))$ = Rank $(Dg(\hat{x})) = m$, there exists a unique multiplier vector $\hat{\lambda} \in E^m$, such that

$$Df(x) - Dg(x))^T\lambda = 0$$

Thus, if we define the function

$\phi(x,\lambda) = f(x) - \langle\lambda, g(x)\rangle$, where $g(x) = (g_1 \ldots g_m)^T$

the problem of minimising $f(x)$ subject to $g(x) = 0$ is equivalent to the problem of minimising the unconstrained function $\phi(x,\lambda)$.

Let us now assume that x lies on the boundary of the constraint set A.

Let $I = \{i : h_i(x) = 0\}$ and let $h'(x) = \{h_i(x) : i \in I\}$. That is, we are dividing the inequality constraint set

$h_i(x) = 0$, $i = 1,2,\ldots,p$, into two halves, those for which $h_i(x) = 0$ and those for which $h_i(x) < 0$.

The following results are due to Kuhn and Tucker.[86] Let the following constraint qualification be satisfied: For any δx satisfying $Dh'(x).\delta x \leq 0$, there corresponds a differentiable arc $x = a(\theta)$, $0 \leq \theta \leq 1$ contained in the constraint set with $x = a(o)$ and some positive scalar α such that

$$\left.\frac{da}{d\theta}\right|_{a = a(o)} = \alpha\, \delta x .$$

This assumption is designed to rule out singularities on the boundary of the constrained set such as an outward pointing cusp.

Kuhn Tucker Theorem.

Let $\hat{x}$ solve the minimum problem. Under the above assumptions there exists a unique $\lambda \in E^m$ and some $\mu \in E^p$ such that

$$Df(\hat{x}) + (Dg(\hat{x}))^T\hat{\lambda} + (Dh(\hat{x}))^T\hat{\mu} = 0 \qquad (A.3)$$

$$\hat{\mu} \geq 0 \qquad (A.4)$$

$$\langle \hat{\mu}, h(\hat{x})\rangle = 0 \qquad (A.5)$$

A more general theorem of this type is proved in Chapter 7.

Appendix B

Mathematical Background

The purpose of this chapter is to set forth various definitions and concepts from functional analysis which are used in the thesis. Part of this material is taken from the thesis of A. E. Pearson.[84] Otherwise the principal sources are the books of Dieudonné[47] and Liusternik and Sobolev.[85]

Mathematically, the thesis deals mainly with operators on normed linear spaces. Therefore, the definition of normed linear spaces, the convergence of sequences in these spaces, and certain aspects of the calculus of operators are among those topics of major concern here.

B.1 Normed Linear Spaces.

Before introducing the particular space of functions with which the thesis is mainly concerned, the general definition of a normed linear space will be stated for the sake of completeness.

A set X of elements or points x,y,z... is a real linear space if the following conditions are satisfied:

A. For any two elements $x, y \in X$ (read "x and y belong to the set X") there is a uniquely defined third element $z = x + y$, called their sum, such that

1. $x + y = y + x$
2. $x + (y+z) = (x+y) + z$
3. there exists an element 0 having the property that $x + 0 = x$ for all $x \quad X$
4. for every $x \quad X$ there exists an element $-x$ such that $x + (-x) = 0$.

B. For arbitrary real numbers α, β and any element $x \quad X$, there is defined an element $\alpha\, x$ such that

1. $\alpha(\beta x) = (\alpha\beta)x$
2. $1.x = x$

C. The operations of addition and multiplication are related in the manner that

1. $(\alpha+\beta)x = \alpha\, x + \beta\, x$
2. $\alpha(x+y) = \alpha\, x + \alpha\, y$

A linear space X is said to be <u>normed</u> if to each element $x \in X$ there is associated a non-negative real number $\|x\|$, called the <u>norm</u> of x, satisfying the conditions

$$1. \quad \|x\| = 0 \text{ if and only if } x = 0$$
$$2. \quad \|\alpha x\| = |\alpha| \cdot \|x\| \qquad \text{(B.1)}$$
$$3. \quad \|x+y\| \leq \|x\| + \|y\|$$

When applied to the difference between two elements $x, y \in X$, the norm $\|x-y\|$, has the geometric interpretation of being the distance between x and y in the space X ($\|x-y\|$ defines a metric for the space X).

Since as an analysis tool, functional analysis is a generalization of well known concepts from the fields of algebra, geometry, and calculus, similar interpretations to many of the concepts used in the thesis will present themselves during the course of the development.

Concerning a further interpretation to the norm $\|x-x_o\|$ where x_o is a particular point $\in X$, the set of all points $x \in X$ for which

$$\|x - x_o\| \leq r ,$$

defines a closed <u>sphere</u> of radius r centered about x_o in the space X. The sphere will be denoted by $S(x_o,r)$, that is

$$S(x_o,r) = \left\{ \begin{array}{l} \text{the set of all points } x \in X \\ \text{such that } \|x - x_o\| \leq r \end{array} \right\} \qquad \text{(B.2)}$$

Two definitions pertaining to sets and used in the thesis concern a subset of a set and boundedness of a set. A set X_1 is said to be a subset of a set X_2 if all the elements belonging to X_1 are also contained in X_2 (including the possibility that X_1 and X_2 are equal). A set X is said to be bounded if there exists a unique constant K such that $\| x \| < K$ for all $x \in X$, i.e. if the set lies in a sphere.

B1.1 $L^P[0,T]$ Space.

A convenient way to define a particular set or space X is to specify a particular norm on that set. Of the many types of normed linear spaces, the type with which the thesis is principally concerned is the class of function spaces $L^P[0,T]$ with norm defined by

$$\| x \| = \left[\int_0^T |x(t)|^p dt \right]^{\frac{1}{p}}, \; p \geq 1. \qquad (B.3)$$

For a chosen number p, the norm (B-3) defines a set X composed of all functions x(t), $t \in [0,T]$ for which the norm exists, i.e. $\| x \| < \infty$. The space $L^1(0,T)$ defines the set of all functions with bounded "area", the space $L^2[0,T]$ the set of all square integrable functions, etc. The space $L^{\infty}[0,T]$ defines the set of all functions bounded on the interval $[0,T]$ for which the corresponding norm becomes

$$\|x\|_{p=\infty} = \max_{t\in[0,T]} |x(t)|$$

The space $L^{\infty}[0,T]$ is of particular interest in the time optimal control problem, discussed briefly in Section 1.5.1, as a result of the equivalence between a closed sphere in $L^{\infty}[0,T]$ space and the amplitude constraint imposed upon the input signals.

Concerning the relationship between $L^p[0,T]$ spaces corresponding to different values of p, any function x(t), $t\in[0,T]$ which belongs to $L^{p_1}[0,T]$ also belongs to $L^{p_2}[0,T]$ for $p_2 < p_1$. That is to say, the space $L^{p_2}[0,T]$ is a subset of the space $L^{p_1}[0,T]$ for $p_2 \leq p_1$.

B.1.2 Convergence of Sequences

The concept of convergence of sequences in normed linear spaces is especially important in the thesis because in general to synthesize an optimal control scheme it is necessary to achieve optimization in a step by step manner through the construction of a sequence of inputs $\{u_n(t)\}$, $n = 1,2,\ldots$, t $[0,T]$, which converges to a solution of the criteria for optimal performance.

A definition of convergence (called "strong" convergence, or convergence in the norm) is the following: A sequence of elements $\{x_n\}$, $n = 1,2,\ldots$, of a normed linear space X converges to an element $x \in X$ if $\|x_n - x\| \to 0$ as $n \to \infty$, that is, if

$$\lim_{n \to \infty} \|x_n - x\| = 0 . \quad (B.4)$$

It should be remarked that, in general, a sequence $\{x_n\}$, $n = 1,2,\ldots$, may have a limiting element x which does not belong to the same space as the elements of the sequence. When this is the case, the sequence is not convergent according to Definition (B.4) because the norm $\|x_n - x\|$ is only defined for those elements x_n and x which are contained in the same space. However, for the type of spaces of principal interest in the thesis, i.e. $L^p(0,T)$ space, the limit element of a sequence, if it has one, will always belong to the same space as the elements of the sequence. The reason for this is a direct result of the fact that $L^p[0,T]$ space is a "complete" space.

Specifically, a normal linear space X is called complete (also a Banach space) if every fundamental sequence of the space has a limit in X. A fundamental sequence is defined as follows: A sequence of elements x_n of a metric space X is called a fundamental sequence or Cauchy sequence if for every number $\varepsilon > 0$ there exists an index number N such that $\|x_m - x_n\| < \varepsilon$ for all $m, n \geq N$.

The definition of strong convergence in the form of Equation (B.4) is not very practical for testing the convergence of a particular sequence for it requires knowledge of the limiting element. A more convenient form is the following: If the distance between successive pairs of elements of a sequence $\{x_n\}$ in a complete normed linear space X becomes progressively smaller with n, i.e. if

$$\|x_{n+1} - x_n\| < \|x_n - x_{n-1}\| \tag{B.5}$$

holds for all $n > 1$, then the sequence is <u>strongly convergent</u>.

If, less restrictively, inequality (B.5) holds for a sequence $\{x_n\}$ in a normed linear space, that is, without the completeness condition, then the sequence is <u>fundamental</u>.

To prove the equivalence of the above statements to the original definitions of strong convergence and a fundamental sequence respectively, it is first noted that if condition (B.5) holds, then it is certainly true that the inequality

$$\|x_{n+1} - x_n\| \leq \beta \|x_n - x_{n-1}\| , \quad \beta < 1 \tag{B.6}$$

holds for $n > 1$ and some number β, $0 \leq \beta < 1$, irrespective of whether the space X is complete or not. Iterating on (B.6),

$$\|x_3 - x_2\| \leq \beta \|x_2 - x_1\|$$

$$\|x_4 - x_3\| \leq \beta \|x_3 - x_2\| \leq \beta^2 \|x_2 - x_1\| \qquad (B.7)$$

$$\cdots\cdots\cdots\cdots\cdots$$

$$\|x_{n+1} - x_n\| \leq \beta^{n-1} \|x_2 - x_1\|, \; n > 1$$

Now consider the identity

$$\|x_{n+m}-x_n\| = \|(x_{n+m}-x_{n+m-1}) + (x_{n+m-1}-x_{n+m-2}) + (x_{n+m-2} - x_{n+m-3}) + \ldots. + (x_{n+1} - x_n)\| \qquad (B.8)$$

Applying Property 3 of Equation (B.1), the triangle inequality, successively to Equation (B.8),

$$\|x_{n+m}-x_n\| \leq \|x_{n+m}-x_{n+m-1}\| + \|x_{n+m-1}-x_{n+m-2}\| + \ldots + \|x_{n+1} - x_n\|$$

which on the basis of (B.7) becomes

$$\|x_{n+m} - x_n\| \leq (\beta^{n+m-2} + \beta^{n+m-3} + \ldots + \beta^{n-1}) \|x_2-x_1\| \qquad (B.9)$$

$$= \beta^{n-1}(\beta^{m-1} + \beta^{m-2} + \ldots + 1) \|x_2-x_1\|, \; n > 1.$$

Since $\frac{1}{1-\beta} = \sum_{i=0}^{m-1} \beta^i + \sum_{i=m}^{\infty} \beta^i \geq \sum_{i=0}^{m-1} \beta^i$ for $0 \leq \beta < 1$,

inequality (B.9) can be written as

$$\|x_{n+m}-x_n\| \leq \frac{\beta^{n-1}}{1-\beta} \|x_2-x_1\|, \quad n > 1 \qquad (B.10)$$

Then given any $\epsilon > 0$, it is evident from (B.10) that the norm $\|x_{n+m}-x_n\|$ can be made less than ϵ by choosing n sufficiently large; thus the sequence satisfying inequality (B.5) is a fundamental sequence.

If the space X is complete, then the limiting element x of the fundamental sequence x_n must belong to the space X, by definition. Letting $m \to \infty$ in (B.10), it follows that

$$\|x_n - x\| \leq \frac{\beta^{n-1}}{1-\beta} \|x_2-x_1\| , \quad n > 1 \qquad \text{(B.11)}$$

where x is the limit point of the sequence. Thus a sequence $\{x_n\}$ satisfying condition (B.5) in a complete normed linear space is strongly convergent according to Definition (B.4).

It is clear from inequalities (B.10) and (B.11) that strong convergence implies that the sequence is fundamental. However, the converse is not always true unless the space is complete.

There is yet another type of convergence for sequences in normed linear spaces called "weak" convergence which relates to the behaviour of a sequence under linear functional transformations in the space. However, the strong convergence of sequences (which implies convergence in the weak sense) is the type of convergence referred to in the thesis, and therefore weak convergence will not be considered here.

B.2 Operators on Normed Linear Spaces.

If there exists a correspondence between the elements of one space X and the elements of another space Y, then the mechanism by which the relationship is established is called an operator or mapping. The relationship between an element $x \in X$ and its image element $y \in Y$ may be denoted by

$$y = F(x)$$

or (B.12)

$$y = Fx$$

The space X is called the domain of the operator F and Y the range. It is assumed here that the spaces X and Y are normed linear spaces.

If in particular the image space Y is a subset of the real line, i.e. the image elements y are simply real numbers, then the operator is called a functional.

An operator F is said to be bounded if there exists a constant C such that

$$\|F(x)\| \leq C\|x\| \qquad (B.13)$$

for all $x \in X$.

An operator F is said to be continuous if for every number $\varepsilon > 0$ there exists a number $\delta > 0$ such that

$$\|F(x_1) - F(x_2)\| < \varepsilon$$

when (B.14)

$$\|x_1 - x_2\| < \delta$$

for all $x_1, x_2 \in X$.

Note that the norms in Definitions (B.13) and (B.14) must be interpreted according to whether the domain X is involved or the image space Y. That is, $\|F(x)\|$ refers to the norm in the image space Y while $\|x\|$ refers to the norm in the domain X.

All the operators with which the thesis is concerned are assumed to be bounded and continuous.

B.2.1 Linear Operators.

An operator F is said to be linear if it satisfies the condition

$$F(\alpha x_1 + \beta x_2) = \alpha F(x_1) + \beta F(x_2) \tag{B.15}$$

for any two elements $x_1, x_2 \in X$, and arbitrary real numbers α, β. If a linear operator is continuous, then it is also bounded. The reverse is true as well for linear operators.

The norm of a linear operator F, denoted by $\|F\|$, is defined as the greatest lower bound of the numbers C which satisfy the boundedness condition (B.13). The norm of a linear operator is given equivalently by the expressions

$$\| F \| = \sup_{\|x\| \neq 0} \frac{\| F(x) \|}{\| x \|}$$

$$= \sup_{\|x\| \leq 1} \| F(x) \| \qquad (B.16)$$

$$= \sup_{\|x\| = 1} \| F(x) \|$$

If in particular the operator is a linear functional, then its norm becomes

$$\| F \| = \sup_{\|x\| \leq 1} | F(x) | \qquad (B.17)$$

Definition (B.17) follows from (B.16) only if the norm in the image space Y (the real line for a functional) is taken as the normalized norm. That is to say, if $Y = L^p[0,T]$ space for example, then the norm $F(x)$ in (B.16) should be taken as

$$\| F(x) \| = \left[\frac{1}{T} \int_0^T | F(x) |^p dt \right]^{\frac{1}{p}}$$

rather than the unnormalized form (B.3), in order that Definition (B.16) reduce to (B.17) when F is a linear functional.

If F_1 and F_2 are two linear operators on a linear normed space X, then the inequality

$$\| F \| \leq \| F_1 \| + \| F_2 \| \qquad (B.18)$$

holds for their sum $F = F_1 + F_2$.

If F is a linear operator from the space X to the space Y, and G a linear operator mapping elements from Y into a third space Z, then the composite operator H = GF defined by

$$z = H(x) = G(F(x)), \quad x \in X, \quad z \in Z \tag{B.19}$$

is called the <u>product</u> of the operators F and G. The norm $\|H\|$ satisfies the inequality

$$\|H\| \leq \|G\| . \|F\| . \tag{B.20}$$

B.2.2 <u>Adjoint Operators</u>.

Before defining an adjoint operator, it is worthwhile to introduce the concepts of a conjugate space and an inner product on conjugate spaces.

It is known that the set of all linear functionals defined on a normed linear space X forms itself a normed linear space called the <u>conjugate space</u> of X and denoted by X^*. Considering in particular $L^p[0,T]$ space, the conjugate space is $L^q[0,T]$ space where p and q are related by

$$\frac{1}{p} + \frac{1}{q} = 1 . \tag{B.21}$$

That is to say, $L^q[0,T] = (L^p[0,T])^*$ in the above notation.

The fact that $L^p[0,T]$ and $L^q[0,T]$ are conjugate to one another is a result of the proof that the general form of a linear functional on $L^p[0,T]$ space is

$$F(x) = \int_0^T f(t)x(t)dt, \quad x \in L^p[0,T] \qquad (B.22)$$

The norm of functional (B.22), as determined by Definition (B.17) in conjunction with Hölder's inequality, is given by

$$\| F \| = \left[\int_0^T |f(t)|^q dt \right]^{\frac{1}{q}} \qquad (B.23)$$

Thus whatever function $f(t)$, $t \in [0,T]$ is used in specifying a linear functional on $L^p[0,T]$ space, that function must belong to the conjugate space $L^q[0,T]$.

Notwithstanding the fact that Equation (B.22) is the general form for a linear functional on $L^p[0,T]$ space, the expression is seen to be linear with respect to either x or f. Such a bilinear expression is called an inner product between the two elements.

In general, the <u>inner product</u> between two elements x and y belonging to spaces which are conjugate to one another, i.e. $x \in X$, $y \in X^*$, is the bilinear functional associated with those spaces and denoted by $\langle x,y \rangle$. The inner product possesses the following properties:

$$
\begin{aligned}
&1. \quad \langle x,y\rangle = \langle y,x\rangle \\
&2. \quad \langle x_1 + x_2, y\rangle = \langle x_1,y\rangle + \langle x_2,y\rangle \text{ for all} \\
&\quad\;\; x_1, x_2\ X; \quad y\ X^* \\
&3. \quad \langle x,y_1 + y_2\rangle = \langle x,y_1\rangle + \langle x,y_2\rangle \text{ for all} \\
&\quad\;\; x \in X; \quad y_1, y_2 \in X^* . \\
&4. \quad \langle \alpha x,y\rangle = \langle x,\alpha y\rangle = \alpha\langle x,y\rangle \text{ for } \alpha \text{ any} \\
&\quad\;\; \text{real number.}
\end{aligned}
\tag{B.24}
$$

In particular, the inner product on the conjugate spaces $L^p[0,T]$ and $L^q[0,T]$, $\frac{1}{p} + \frac{1}{q} = 1$, is

$$\langle x,y\rangle = \int_0^T x(t)y(t)dt \tag{B.25}$$

where $x\ L^p(0,T)$ and $y\ L^q(0,T)$ or vice versa.

When the conjugate spaces are the same, i.e. $X = X^*$, then the spaces are said to be <u>self conjugate</u>. A self conjugate space is a <u>Hilbert space</u> for which the norm is derived from the inner product according to

$$\| x \| = \sqrt{\langle x,x\rangle} \tag{B.26}$$

and the additional property for the inner product

$$\langle x,x\rangle \geq 0 \text{ , with equality if and only if } x = 0,$$

may be included with those of (B.24). It is seen that $L^2[0,T]$ space is a (real) Hilbert space.

Consider now a linear operator L which maps elements from the space X into the image space Y, i.e.

$$y = L(x) = Lx$$

where $x \epsilon X$, $y \epsilon Y$, X and Y normed linear spaces. The <u>adjoint operator</u> of the linear operator L, denoted by L^*, can be defined by the inner product relation,

$$\langle z, Lx \rangle = \langle x, L^* z \rangle \tag{B.27}$$

where z is an arbitrary element ϵ Y^* (the space conjugate to Y), and x is an arbitrary element ϵ X. Thus the adjoint operator L^* maps elements from the space Y^* (the conjugate to the range Y of the linear operator L) into the space X^* (the conjugate to the domain X of the linear operator L).

The adjoint operator L^* of a linear operator L is also a linear operator, and

$$\|L^*\| = \|L\| \tag{B.28}$$

Three basic properties of adjoint operators are the following: If L_1 and L_2 are two linear operators with domain X and range in Y, then

1. the adjoint of their sum is equal to the sum of their adjoints,

$$(L_1 + L_2)^* = L_1^* + L_2^* \tag{B.29}$$

2. the adjoint of their product is equal to the product of their adjoints in reverse order

$$(L_1 L_2)^* = L_2^* L_1^* \quad \text{(B.30}$$

3. the adjoint of the identity operator is equal to the identity operator

$$I^{*} = I \ . \tag{B.31}$$

As an example, consider the linear operator L_1 defined by

$$y(t) = L_1 x = \int_0^t k(t-\tau)x(\tau)d\tau \tag{B.32}$$

where $x \epsilon L^p[0,T]$, $y \epsilon L^r[0,T]$, and the kernel $k(t-\tau)$ vanishes for $\tau > t$. The domain of the linear operator (B.32) is $X = L^p[0,T]$ and its range $Y = L^r[0,T]$. The corresponding conjugate spaces are $X^{*} = L^q[0,T]$ and $Y^{*} = L^s[0,T]$ where

$$\frac{1}{p} + \frac{1}{q} = 1 \ ; \quad \frac{1}{r} + \frac{1}{s} = 1 \ . \tag{B.33}$$

Substituting Equation (B.32) into the inner product relation (B.27)

$$\langle z, L_1 x \rangle = \int_0^T z(t) \int_0^t k(t-\tau)x(\tau)d\tau dt. \tag{B.34}$$

Keeping in mind the fact that $k(t-\tau) = 0$ for $\tau > t$, the upper limit of integration for the inner integral in (B.34) may be changed from t to T such that

$$\langle z, L_1 x \rangle = \int_0^T z(t) \int_0^T k(t-\tau)x(\tau)d\tau dt. \tag{B.35}$$

Reversing the orders of integration in (B.35)

$$\langle z, L_1 x \rangle = \int_0^T x(\tau) \int_0^T k(t-\tau)z(t)dt d\tau$$

Interchanging the notation for the variables of integration

$$\langle z, L_1 x\rangle = \int_0^T x(t) \int_0^T k(\tau - t) z(\tau) d\tau dt. \qquad (B.36)$$

But since $k(\tau - t) = 0$ for $t > \tau$, Equation (B.36) may be written as

$$\langle z, L_1 x\rangle = \int_0^T x(t) \int_t^T k(\tau - t) z(\tau) d\tau dt = (x, L_1^* z).$$

That is, the adjoint operator L_1^* of the linear operator (B.32) is given by

$$L_1^* z = \int_t^T k(\tau - t) z(\tau) d\tau \qquad (B.37)$$

which maps elements $z \in L^s[0,T]$ (the conjugate to the image space of operator L_1) into the space $L^q[0,T]$ (the conjugate to the domain of operator L_1). Note the transposition of arguments for the kernel function appearing in the adjoint operator (B.37) as compared to the kernel in (B.32).

In a similar manner, the adjoint operator of the linear operator L_2 given by

$$L_2 x = \int_0^T k(T + t - \tau) x(\tau) d\tau \qquad (B.38)$$

may be determined with the result that

$$L_2^* x = \int_0^T k(T + \tau - t) z(\tau) d\tau , \qquad (B.39)$$

B.2.3 Differential Calculus in Banach Spaces.

Let X and Y be Banach spaces and let A be an open subset of X. Let f, g be two mappings of A into Y. We say that f and g are tangent at a point $x_o \in A$ if

$$\mathop{\mathcal{L}t}_{\substack{x \to x_o \\ x \neq x_o}} \frac{\|f(x) - g(x)\|}{\|x - x_o\|} = 0 \qquad \text{(B.40)}$$

This implies, of course, $f(x_o) = g(x_o)$.

Among all functions tangent at x_o to a function f, there is at most one mapping of the form $x \to f(x_o) + u(x-x_o)$ where u is linear.

We say that a continuous mapping f of A into F is Fréchet differentiable at the point $x_o \in A$ if there is a linear mapping u of X into Y such that $x \to f(x_o) + u(x-x_o)$ is tangent to f at x_o. We have just seen that this mapping is unique. It is called the Fréchet derivative of f at x_o and written $Df(x_o)$.

Let $x - x_o = h$. The Frechet Differential, $Df(x_o).(x-x_o) = Df(x_o).h$ may be calculated from the formula

$$Df(x_o).h = \mathop{\mathcal{L}t}_{\xi \to 0} \frac{f(x_o + \xi h) - f(x_o)}{\xi} , \qquad \text{(B.41)}$$

where ξ is a real number, or from the formula

$$Df(x_o).h = \frac{\partial}{\partial \xi} f(x_o + \xi h)\Big|_{\xi=0} \qquad \text{(B.42)}$$

However, the operations on the right hand side of equations (B.41) and (B.42) do not in themselves suffice to define the Frechet differential, because there exist mappings f for which the operations on the right hand side are meaningful, but which themselves fail to satisfy the definition of a Frechet differential. When this is the case the functions defined by the right hand side of equation (B.41) and (B.42) is called the Gateaux Differential of the mapping and denoted by $Df(x_o;h)$.

For example the mapping f defined by

$$f(x) = \int_0^t |x(\tau)| \, d\tau, \quad x \in L^1[0,T] ,$$

possesses the Gateaux differential

$$Df(x_o;\, h) = \int_0^t h(\tau) \mathrm{Sgn}(x(\tau)) d\tau \;,$$

where $$\mathrm{Sgn}(x) = \begin{cases} +1 & \text{for } x > 0 \\ -1 & \text{for } x < 0 \end{cases}$$

but is not Frechet differentiable.

It is evident that if the Frechet differential $Df(x_o).h$ exists, then the Gateaux differential $Df(x_o;h)$ exists and the two are equal. However, the converse is not necessarily true as shown by the above example.

Conditions under which the existence of the weak differential implies the existence of the Frechet differential relate to the condition that $Df(x_o;h)$ be uniformly continuous in x_o and continuous in h in a certain neighbourhood of the point x_o.

The Fréchet differential $Df(x_o).h$ is linear with respect to the variation h, i.e.

$$Df(x_o).(h_1+h_2) = Df(x_o).h_1 + Df(x_o).h_2 \qquad (B.43)$$

whereas the Gateaux differential need not necessarily satisfy linearity with respect to its variation.

Let us remark that the definition of a Gateaux Differential is the same as that of first variation in the Calculus of Variations.

The Frechet Derivative (when it exists) of a continuous mapping f of A into Y at a point x_o A is thus an element of the Banach space $\mathcal{L}(X;Y)$ (the space of Linear mappings from X to Y).

We shall not continue with any further details of Differential Calculus in Banach Spaces. This Differential Calculus is in fact very similar to ordinary Calculus and counterparts of mean value theorem, Taylor's theorem, etc., exist here too. The interested reader is referred to Dieudonne[47], Chapter VIII.

APPENDIX C

DERIVATION OF SOME RESULTS FOR THE SECOND ORDER COMPUTATIONAL METHOD

C.1. Solution of the Variational Problem Given by Equations (5.12) - (5.14)

In a suitably small neighbourhood of the optimum, D_uH will be small. In fact it will be a differential quantity and thus of the same order of magnitude as $D_u^2H.\delta u$ etc. arising in (5.12). However if the nominal control function chosen is not sufficiently near to the optimum, D_uH will be large. Thus the choice of δu by minimising $\delta P + \frac{1}{2}\delta^2 P$ subject to the constraint (5.12) may render the linearization (5.13) invalid. Hence ε_1, $0 < \varepsilon_1 \leq 1$ is introduced in the integrand (5.12). In a suitably small neighbourhood of the optimum ε_1 can be set equal to 1.

Similarly for the nominal control function chosen, we may miss the terminal constraints by a large amount. Thus $G(x(t_f),t_f)$ will be large. Hence the desired improvement in terminal condition might have to be corrected in small steps. Hence the parameter ε_2 is introduced.

The variational problem (5.12) - (5.14) is solved by writing the Hamiltonian for the problem in the form

$$\mathcal{H} = \varepsilon_1 \langle D_u H, \delta u\rangle + \frac{1}{2} \langle D_u^2 H.\delta u, \delta u\rangle + \frac{1}{2} \langle D_x^2 H.\delta x, \delta x\rangle$$

$$+ \langle D_{ux}^2 H.\delta x, \delta u\rangle + \varepsilon_1 \langle \Delta\lambda, D_x f.\delta x + D_u f.\delta u\rangle \qquad \text{(C.1)}$$

where $\varepsilon_1 \Delta\lambda$ is the Lagrange Multiplier function for this problem. Because of the presence of the linear term in the integrand of (5.12), the multiplier is written as $\varepsilon_1 \Delta\lambda$. Differentiation of $\mathcal{H}$ with respect to δu and δx yields equations (5.16) and (5.18) in the usual way.

Also the boundary condition for $\varepsilon_1 \Delta\lambda$ is found by forming,

$$\frac{1}{2}\langle D_x^2 \psi(x(t_f), t_f, \nu).\delta x(t_f), \delta x(t_f)\rangle$$

$$+ \varepsilon_2 \langle \Delta\nu, \varepsilon_2 G(x(t_f), t_f) + D_x G(x(t_f), t_f).\delta x(t_f)\rangle$$

and differentiating the above term with respect to $\delta x(t_f)$. This yields equation (5.17).

C.2. Proof that Sum of First and Second Variations is Negative for the Choice of δu Given by (5.18)

We first assume that there are no terminal constraints present. Let us evaluate the terms in the integrand of (5.12) separately. We have,

$$=\langle D_u H, \delta u\rangle = -\ \mathcal{E}_1^2 \langle D_u H, (D_u^2 H)^{-1} . D_u H\rangle - \ \mathcal{E}_1 \langle D_u H, (D_u^2 H)^{-1} D_{ux}^2 H . \delta x\rangle$$

$$-\ \mathcal{E}_1^2 \langle D_u H, (D_u^2 H)^{-1} (D_u f)^T . \Delta\lambda \qquad (C.2)$$

$$\frac{1}{2}\langle D_u^2 H.\delta u, \delta u\rangle + \langle D_{ux}^2 H.\delta x, \delta u\rangle + \frac{1}{2}\langle D_x^2 H.\delta x, \delta x\rangle$$

$$= \frac{1}{2}\,\varepsilon_1^2 \langle D_u H, (D_u^2 H)^{-1} D_u H\rangle + \frac{1}{2}\langle (D_x^2 H - D_{xu}^2 H (D_u^2 H)^{-1} D_{ux}^2 H)\delta x, \delta x\rangle$$

$$+ \frac{1}{2}\,\varepsilon_1^2 \langle \Delta\lambda, D_u f (D_u^2 H)^{-1} (D_u f)^T \Delta\lambda\rangle + \frac{1}{2}\,\varepsilon_1^2 \langle D_u H, (D_u^2 H)^{-1} (D_u f)^T \Delta\lambda\rangle \qquad (C.3)$$

Combining (C.2) and (C.3), we obtain,

Integrand of (5.12)

$$= -\ \frac{1}{2}\,\varepsilon_1^2 \langle D_u H, (D_u^2 H)^{-1} D_u H\rangle - \frac{1}{2}\,\varepsilon_1^2 \langle B\delta\lambda, \delta\lambda\rangle + \langle C\delta x + \mathcal{E}_1 w, \delta x\rangle$$

$$-\ \frac{1}{2}\langle C\delta x, \delta x\rangle .$$

Writing $C\delta x + \mathcal{E}_1 w = -\ \mathcal{E}_1 \Delta\dot{\lambda} - \mathcal{E}_1 A^T \Delta\lambda$, performing integration by parts, and using

$\mathcal{E}_1 \Delta\lambda(t_f) = D_x^2 F.\delta x(t_f)$, we obtain,

Sum of first and second variations

$$= -\ \frac{1}{2}\,\varepsilon_1^2 \int_{t_o}^{t_f} [\langle D_u H + (D_u f)^T \Delta\lambda, (D_u^2 H)^{-1} (D_u H + (D_u f)^T \Delta\lambda)\rangle] dt$$

$$-\ \int_{t_o}^{t_f} \langle C\delta x, \delta x\rangle dt - \frac{1}{2}\langle D_x^2 F.\delta x(t_f), \delta x(t_f)\rangle$$

which is negative in view of our assumptions.

For the case where terminal constraints are present, the development is the same. In this case,

Sum of first and second variations

$$= -\frac{1}{2}\varepsilon_1^2 \int_{t_o}^{t_f} [\langle D_uH+(D_uf)^T\Delta\lambda,\ (D_u^2H)^{-1}(D_uH+(D_uf)^T\Delta\lambda)]dt$$

$$- \int_{t_o}^{t_f} \langle C\delta x, \delta x\rangle dt - \frac{1}{2}\langle D_x^2\psi.\delta x(t_f), \delta x(t_f)\rangle$$

$$- \langle P^{-1}(t_o)[\ \varepsilon_2 G + \varepsilon_1 m(t_o)],\ \varepsilon_2 G\rangle.$$

$P^{-1}(t_o)$ is positive definite. We now have to make the assumption that the last term within < > is positive definite, to be strictly correct. In computation, this means making $\varepsilon_1 m(t_o)$ sufficiently small.

C.3 Derivation of Equations (5.26) - (5.30)

Differentiating equation (5.24), we obtain,

$$\varepsilon_1\Delta\dot\lambda = \varepsilon_1\dot\ell + \dot K\delta x + K(A\delta x+ \varepsilon_1 B\Delta\lambda+ \varepsilon_1 v) + \varepsilon_2\dot N\Delta\vartheta$$

$$= \varepsilon_1\dot\ell + \varepsilon_1 KB\ell + \varepsilon_1 Kv+(\dot K+KA+KBK)\delta x+(\ \varepsilon_2\dot N+ \varepsilon_2 KBN)\Delta\vartheta$$

$$= -C\delta x - A^T(\ \varepsilon_1\ell + K\delta x+ \varepsilon_2 N\Delta\vartheta) - \varepsilon_1 w \quad \text{from (5.21)}$$

Hence

$$\varepsilon_1(\dot\ell + (KB+A^T)\ell + Kv+w)+(\dot K+KA+A^TK+KBK+C)\delta x$$
$$+ \varepsilon_2(\dot N+KBN+A^TN)\Delta\vartheta = 0 \qquad (C.4)$$

Since the above equation has to vanish for arbitrary

δx and $\Delta\nu$, we obtain

$$\dot{\ell} + (KB+A^T)\ell + Kv+w = 0 \tag{C.5}$$

$$\dot{K} + KA + A^TK + KBK + C = 0 \tag{C.6}$$

$$\dot{N} + (KB+A^T)N = 0 \tag{C.7}$$

Equating coefficients at the end point,

$$K(t_f) = D_x^2\psi(x(t_f),t_f,\nu) \tag{C.8}$$

$$N(t_f) = (D_xG(x(t_f),t_f))^T \tag{C.9}$$

Differentiating (5.25),

$$0 = \varepsilon_1\dot{m}+\dot{N}^T\delta x+N^T(A\delta x+ \varepsilon_1 B\Delta\lambda+ \varepsilon_1 v) + \varepsilon_2\dot{P}\Delta\nu$$

$$= \varepsilon_1\dot{m}+\dot{N}^T\delta x+N^TA\delta x+N^TB(\varepsilon_1\ell +K\delta x+ \varepsilon_2 N\Delta\nu)+ \varepsilon_1 N^Tv+ \varepsilon_2\dot{P}\Delta\nu$$

$$= \varepsilon_1(\dot{m}+N^TB\ell +N^Tv)+(\dot{N}^T+N^TA+N^TBK)\delta x+ \varepsilon_2(\dot{P}+N^TBN)\Delta\nu$$

Again, since the above equation has to vanish for arbitrary δx and $\Delta\nu$, we obtain

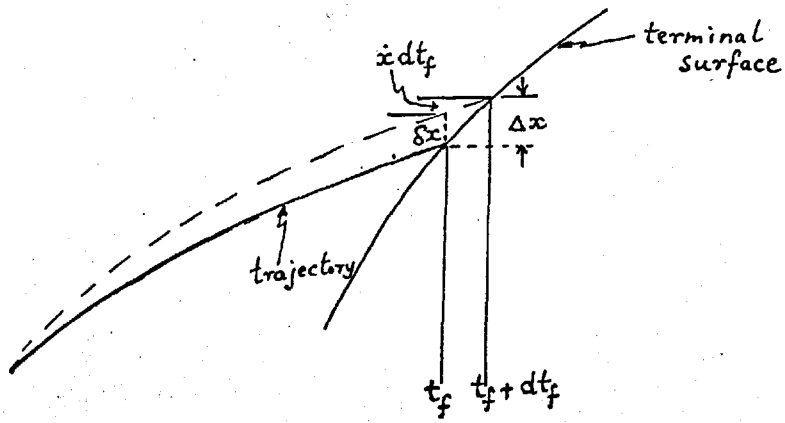
$$\dot{m} + N^T(B\ \ + v) = 0 \tag{C.10}$$

$$\dot{P} + N^TBN = 0 \tag{C.11}$$

The boundary conditions for m and P are clearly $m(t_f) = 0$, $P(t_f) = 0$.

C.4. Case When t_f is given Implicitly

The expressions for the first and second variations are now obtained according to the development in Bliss(18), pp 226-227. The main difference from the previous case is that at the end point we have to consider dependent and independent variations, $\Delta x(t_f) = \delta x(t_f) + \dot{x}(t_f)dt_f$.



Sketch C.1

In forming the auxiliary minimization problem, the term $\langle D_u H(x(t_f), u(t_f), \lambda(t_f), t_f), \delta u(t_f) dt_f \rangle$ is multiplied by ε_1 and then neglected as being of third order.

The solution of the auxiliary minimization problem is tedious but straightforward. Essentially we obtain equations corresponding to equations (5.6) - (5.10). Equation (5.46) is the transversality condition of the auxiliary minimization problem.

REFERENCES

1. M. A. Arbib: A General Framework for Automatic Theory and Control Theory: to be published S.I.A.M. J. on Control.
2. L. A. Zadeh and C. Desoer: Linear Systems Theory: State Space Approach: McGraw Hill, 1963.
3. L. D. Berkovitz: A Survey of Certain Aspects of the Mathematics of Control Problem: Memo RM-3309-PR, Rand Corporation, December 1962.
4. L. D. Berkovitz: Variational Methods in Problems of Control and Programming: J. Math. Annal. and Applns., Vol. 3, 1961, pp 145-169.
5. L. D. Berkovitz: On Control Problems with Bounded State Variables: The Rand Corporation, RM-3207-PR, July 1962.
6. D. W. Bushaw: Optimal Discontinuous Forcing Terms: Contributions to the Theory of Nonlinear Oscillations, Vol. 2, Princeton University Press, Princeton, 1958.
7. R. Bellman, I. Glicksberg and O. Gross: On the Bang Bang Control Problem: Quart. Appl. Math., Vol. 14, 1956, pp 11-18.
8. R. V. Gamkrelidze: The Theory of Time-Optimal Processes for Linear Systems, Izv. Akad. Nauk. SSSR, Ser. Mat., Vol. 22, 1958, pp 449-474.

9. J. P. LaSalle: The Time Optimal Control Problem, Contributions to the Theory of Nonlinear Oscillations, Vol. 5, Princeton University Press, Princeton, 1960, pp 1-24.

10. L. W. Neustadt, Synthesizing Time Optimal Control Systems, J. Math. Anal. Appl. Vol. 1, 1960, pp 484-493.

11. H. Halkin: A Generalization of La Salle's Bang Bang Principle: to be published S.I.A.M. J. on Control.

12. E. B. Lee and L. J. Markus, Optimal Control for Nonlinear Processes, Arch. Rational Mech. Anal. Vol. 8, 1961, pp 36-58.

13. E. Roxin, The Existence of Optimal Controls, Michigan Math. J. Vol. 9, 1962, pp 109-119.

14. A. F. Filippov, On Certain Questions in the Theory of Optimal Control, Vestnik Moskov. Univ. Ser. Mat. Meh. Astr. Fiz. Him., No. 2, 1959, pp 25-32. English Trans. S.I.A.M. J. Control Ser. A 1, 76-84 (1962).

15. L. W. Neustadt: On the Eixstence of Optimal Controls in the Absence of Convexity Conditions: J. Math. Annal. and Applns.

16. L. D. Berkovitz: loc. cit.

17. L. D. Berkovitz: loc. cit.

18. G. A. Bliss: Lectures on the Calculus of Variations, University of Chicago Press, Chicago, 1946.

19. E. J. McShane: On Multipliers for Lagrange Problems, Amer. J. Math. Vol. 6, 1939, pp 809-819.

20. M. R. Hestenes: A General Problem in the Calculus of Variations with Applications to Paths of Least Time: The Rand Corporation, Research Memorandum RM-100, Feb. 1949.

21. L. S. Pontryagin et al: The Mathematical Theory of Optimal Processes: Interscience 1962.

22. E. J. McShane: Necessary Conditions in Generalized-Curve Problems of the Calculus of Variations, Duke J. Math. Vol. 7, 1940, pp 1-27.

23. R. V. Gamkrelidze: The Theory of Time-Optimal Processes for Linear Systems, Izv. Akad. Nauk. SSSR, Vol. 24, 1960, pp 315-356.

24. C. Caratheodory: Variationsrechnung und Partielle Differentialgleichungen erster Ordnung, Teubner, Leipzig, 1935.

25. R. E. Kalman: The Theory of Optimal Control and the Calculus of Variations: RIAS Technical Report 61-3.

26. H. Halkin: On the Necessary Conditions of Optimal Control: J. Analyse Math., 1964, pp 1-82.

27. E. Roxin: A Geometric Approach to Pontryagin's Maximum Principle: Non-linear Differential Equations and Non-linear Mechanics, ed. La Salle and Lepschetz, Academic Press, 1963.

28. L. I. Rozonoer: On Sufficient Conditions for Optimality, Dokl. Akad. Nauk. SSSR, Vol. 127, 1959, pp 520-523.

29. E. B. Lee: A Sufficient Condition in the Theory of Optimal Control: S.I.A.M. J. on Control, Vol. 1, No. 3, pp 241-245 (1963).

30. A. V. Balakrishnan: An Operator Theoretic Formulation of a Class of Control Problems and a Steepest Descent Method of Solution: J.S.I.A.M. Control Ser. A, Vol. 1, No. 2, 1963.

31. R. E. Kalman: Contributions to the Theory of Optimal Control: Bol. Soc. Mat. Mexicana., 1960, p.102.

32. S. Katz: A Discrete Version of Pontryagin's Maximum Principle: J. Electron. and Control, 13, 179-184 (1962).

33. S. S. L. Chang: Digitized Maximum Principle: Proc. IRE pp 2030-31 (1960).

34. L. A. Zadeh and B. Whalen: On Optimal Control and Linear Programming: IRE Trans. on Automatic Control, Vol. AC7, No.4 July 1962.

35. J. B. Rosen: The Gradient Projection Method for Non-Linear Programming, Part 1, Linear Constraints, Journal of the S.I.A.M., 8 (1960); Part 2; Non-Linear Constraints: Journal of the S.I.A.M., 9 (1961).

36. S. K. Mitter: Discussion on paper by Kirchmayer and Ringlee: Optimal Control of Thermal Hydro System Operation, Proceedings of Second IFAC Congress, Basel, 1963.

37. R. Dorfman, P. A. Samuelson and R. M. Solow: Linear Programming and Economic Analysis: McGraw Hill, 1958.

38. I. M. Gelfand and S. V. Fomin: Calculus of Variations: English trans. Prentice Hall, 1963.

39. J. H. Kelley: Gradient Methods, chapter in Book 'Optimisation Techniques': ed. G. Leitmann, Academic Press, 1963.

40. A. Bryson and W. F. Denham: A Steepest Ascent Method of Solving Optimum Programming Problems: J. Appl. Mech., 29, 2, 1962.

41. C. W. Merriam III: Optimisation Theory and the Design of Feedback Control Systems. McGraw Hill 1964.

42. C. W. Merriam III: Direct Computational Methods for Feedback Control Optimization: Information and Control, April 1965.

43. H. J. Kelley, R. Kopp and G. Mayer: A Trajectory Optimization Technique Based upon the Theory of the Second Variation: AIAA Astrodynamics Conference, Yale University, August, 1963.

44. S. R. McReynolds and A. E. Bryson, Jr.: A Successive Sweep Method for Solving Optimal Programming Problems: Harvard Tech. Rept. No. 463, March 1965.

45. S. K. Mitter: Successive Approximation Methods for the Solution of Optimal Control Problems: to be published Automatica, 1965.

46. A. D. Michael: Le Calcul Differentiel dans les Espaces de Banach. Paris, Gauthier-Villars, 1958.

47. J. Dieudonne: Foundations of Modern Analysis, Chapter VIII, Academic Press, 1960.

48. L. Collatz: Funktionalanalysis und Numerische Mathematik: Springer-Verlag, 1964.

49. L. V. Kantorovich: On Newton's Method, Trudy Mat. Inst. Steklov, Vol. 28, pp 104-144, 1949.

50. J. Ragazzini and G. F. Franklin: Sampled Data Control Systems: McGraw Hill 1958.

51. R. Bellman: Introduction to Matrix Analysis: McGraw Hill, 1960.

52. H. W. Kuhn and A. W. Tucker (eds.): Linear Inequalities and Related Systems: Annals of Mathematics Studies No. 38, Princeton University Press, Princeton, N.J. 1956

53. O. L. Mangasarian and J. Ponstein: Minimax and Duality in Nonlinear Programming. Rept. P-1182, Shell Development Company, Emeryville, California.

54. J. D. Pearson: Ph.D. Thesis, University of London, 1963. Also: Reciprocity and Duality in Control and Programming: to be published J. Math. Annal. Applns. 1965.

55. J. D. Pearson: Duality and a Decomposition Technique: lecture given at Control and Programming Conference, Colorado Springs, April 1965.

56. G. P. 90 Programme, Share Library.

57. G. Zoutendijk: Method of Feasible Directions: Elsevier, 1962.

58. C. W. Merriam III: loc. cit. 41, footnote p.269.

59. I. Gelfand and S. Fomin: loc. cit.

60. G. A. Bliss: loc. cit., p.231

61. G. A. Bliss: loc. cit., p.228

62. J. R. Gantmekher: Matrix Theory: Chelsea, 1959, p.247.

63. I. Gelfand and S. Fomin: loc. cit. p.159.

64. S. Dreyfus: Dynamic Programming and the Jacobi Condition: Rand Report, 1964.

65. H. Osborn: On the Foundations of Dynamic Programming: J. Math. and Mech., Vol. 8, 1959, pp 867-872.

66. R. Bellman and S. Dreyfus: Applied Dynamic Programming: Princeton University Press, 1962.

67. H. J. Kelley: loc. cit.

68. H. J. Kelley: loc. cit.

69. A. E. Bryson: loc. cit.

70. C. W. Merriam: loc. cit.

71. M. O. Stein: loc. cit.

72. L. Cottatz: loc. cit.

73. J. J. Florentin: Optimal Control of Continuous time Markov Stochastic Systems: J. Electronics and Control, June 1961, pp 473-488.

74. W. M. Wonham: Stochastic Problems in Optimal Control RIAS Technical Report, 63-14 May 1963.

75. W. T. Reid: Riccati Matrix Diff. Eqns. and Non-Oscillation Criteria for Associated Linear Differential Systems: Pacific Journal of Mathematics, Vol. B, No. 2, Summer 1963.

76. M. Levine: A Steepest Descent Technique for Synthesizing Optimal Control Programmes: Paper 4,

Conference on Advances in Automatic Control. Nottingham, April 1965.

77. R. Fletcher and M. J. D. Powell: A Rapidly Convergent Descent Method for Minimization, The Computer Journal, Vol. 6, 1963.

78. M. R. Hestenes: Numerical Methods of Obtaining Solutions of Fixed End Point Problems in the Calculus Variations, RM-102, The Rand Corp., August, 1949.

79. M. L. Stein: Loc. cit.

80. L. V. Kantorivich: loc. cit.

81. R. Kalaba: On Nonlinear Differential Equations, the Maximum Operation and Monotone Convergence, Journal of Mathematics and Mechanics, Vol. 8, No. 4, pp 519-574, July 1959.

82. R. E. Kopp and R. McGill: Several Trajectory Optimization Techniques, in Computing Methods in Optimization Problems, ed. A. V. Balakrinshnan and L. W. Neustadt, Academic Press, 1964.

83. L. Lasdon: A Multi-level technique for Decomposition: Ph.D. Thesis, Case Inst. of Technology, 1964.

84. A. E. Pearson: Adaptive Optimal Control of Non-Linear Systems: Ph.D. Thesis, Columbia University, 1963.

85. L. Liusternik and V. J. Sobolev: Elements of Functional Analysis, Chapter VI, English Translation, Ungar, 1961.

86. H. W. Kuhn and A. W. Tucker: Nonlinear Programming. Proc. 2nd Berkeley Symposium on Math. Stat and Prob. 481-492, Univ. of Calif. Press, Berkeley, California.

87. S. Dreyfus, Variational Problems with State Variable Inequality Constraints, Rand Corp. Report No. -2605, July 1962.

88. T. Stafford: Variational Methods Applied to the Design of Controls for Boilers, Ph.D. thesis, University of London, 1965.

89. T. Stafford: loc. cit.

90. A. V. Balakrishnan: On the State Space Theory of Linear Systems: to be published Journal of Math. Analysis and Applns., 1965.

91. R. E. Kalman, Y. C. Ho and K. S. Narendra: Controllability of Linear Dynamical Systems: Contributions to Differential Equations, Vol. 1, No. 2, 1962, pp.189-213.

92. F. Riesz and Bela Sz-Nagy: Functional Analysis: Ungar, 1955.

93. D. Blackwell and M. A. Girschik: Theory of Games and Statistical Decisions: Wiley, 1954.