RESEARCHES IN OPTIMAL AND SUBOPTIMAL CONTROL THEORY

G. Kang

March 1968
Report No. 68-10

Report No. 68-10
March 1968

RESEARCHES IN OPTIMAL AND SUBOPTIMAL CONTROL THEORY

G. KANG

DEPARTMENT OF ENGINEERING
UNIVERSITY OF CALIFORNIA
LOS ANGELES

FOREWORD

The research described in this report, "Researches in Optimal and Suboptimal Control Theory," Number 68-10, by Garfield Kang, was carried out under the direction of C. T. Leondes, Principal Investigator, in the Department of Engineering, University of California, Los Angeles.

This project is supported in part by the National Aeronautics and Space Administration Contract NsG 237-62 to the Institute of Geophysics and Interplanetary Physics of the University.

This report was the basis for a Doctor of Philosophy dissertation submitted by the author.

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont.)

TABLE OF CONTENTS (Cont.)

LIST OF SYMBOLS

t	Time
x	State vector in Euclidean n-space, in general.
u	Control vector in Euclidean r-space, in general.
y	Response vector in Euclidean m-space, in general.
$\mathbb{R}$	The real line. $\mathbb{R}^q$ = Euclidean q-space.
(t, x)	Time-state pair $\big($an $(n+1)$-tuple$\big)$; also called a phase.
G	Set of permissible phases (t, x).
G_o	Set of feasible phases $(t, x) \epsilon G$.
S	Set of target phases $(t, x) \epsilon G$.
$\mathcal{U}$	Class of admissible control functions $u(\cdot)$.
U	Permissible Control set $\subset \mathbb{R}^r$.
J	Performance functional or function.
L', L	Integrand of performance functional
L_∞	Space of essentially bounded, measurable r-vector functions.
$[t_o, t_1]$	Closed, control interval, commencing at time t_o and terminating at time t_1.
x_o, x_1	Initial and terminal state vectors.
θ_o, θ_1	Initial and terminal phases, (t_o, x_o) and (t_1, x_1).
δ	Control law or procedure.
u_δ	Control function over $[t_o, t_1]$ corresponding to δ.
P	Co-state vector $(P_1, P_2, \ldots, P_n)$; a row vector.
P_o	Co-state variable.
$\hat{P}$	(P_o, P), an $(n+1)$ row vector.
$\dfrac{\partial a}{\partial b}$	$(m\text{x}n)$ matrix having elements $(\partial a/\partial b)_{ij} \triangleq (\partial a_i/\partial b_j)$, if $a \epsilon \mathbb{R}^m$ and $b \epsilon \mathbb{R}^n$.
$\perp$	"Orthogonal"
wrt	"With respect to"
δ_*, J_*	Optimal control law and performance function.

u_*	Optimal control function over $[t_o, t_1]$.
$R(\theta_o; u_1, u_2)$	Relative regret or performance loss of $u_1(\cdot)$ with respect to $u_2(\cdot)$ for the initial phase θ_o.
$R(\theta_o; u)$	Regret function for $u(\cdot)$ with respect to the optimal control function u_* for θ_o.
$R_\delta(\theta_o)$	Regret function for δ with respect to the optimal control procedure δ_*.
$P[G_o]$	Class of piecewise continuously differentiable control laws defined on G_o.
$\mathbb{O}$	The null matrix
ϕ	The null set
ϵ, ϵ_p	Arbitrarily small, positive constants.
$o(\epsilon), O(\epsilon)$	Order notation.

CHAPTER 1

ELEMENTS OF CONTROL PROCESSES

This chapter is introductory in nature, serving to introduce terminology, notions, and assumptions. It may be skimmed through by readers intimately familiar with control theory.

1.0 THE PROCESS EQUATIONS

In this thesis we are concerned with control of processes which are characterized by differential equations of the form

$$\frac{dx_i}{dt} = f_i(t, x_1, x_2, \ldots, x_n, u_1, u_2, \ldots, u_r), \quad (i=1, 2, \ldots, n)$$

or, in vector notation,

$$\frac{dx}{dt} = f(t, x, u) \tag{1.1}$$

in which $x \in \mathbb{R}^n$ is an n-vector in Euclidean space $\mathbb{R}^n$, and $u \in \mathbb{R}^n$ is an r-vector which we shall call the control vector. The vector x is called the state vector. Equation (1.1) is an example of an important class of processes which are known as state-determined processes.[1] That is, at any instant of time t, an observed response $y \in \mathbb{R}^n$ of the process to a control history $u(\tau)$, $t_o \le \tau \le t$, may be expressed as

$$y(t) = g\Big(t;\ x(t_o),\ u(t_o \le \tau \le t)\Big) \tag{1.2}$$

where g has the property

$$g\Big(t;\ x(t_o),\ u(t_o \le \tau \le t)\Big) = g\Big(t;\ x(t),\ u(t)\Big) \tag{1.3}$$

Equations (1.2) and (1.3) indicate that the response depends on the past, but that this dependence can be projected into a state quantity x at some time and the control history since that time.

The process equations (1.1) is an important element in the modern statement of a control problem. There are four further elements.

1. A permissible set G of phases (t, x) to which the process is restricted;
2. A target set $S \subset G$ which must be attained;
3. A class $\mathcal{U}$ of admissible control time functions;
4. A performance functional J which maps a control function $u(\cdot)$ and its corresponding state trajectory $x(\cdot)$ into $\mathbb{R}$.

These are discussed below.

1.1 THE SET OF PERMISSIBLE PHASES, G

For convenience, we shall adopt Kalman's term of <u>phase</u> for the time-state pair (t, x) [which is actually an (n+1)-tuple].

One can expect that certain restrictions will exist on the totality of phases which apply to a given physical process. Thus, an important element of the control problem is a set $G \subset \mathbb{R} \times \mathbb{R}^n$ on which the process equation (1.1) is defined (for a given u) and to which the process is restricted. A form of G which is general enough for most problems is the following:

$$G = \left\{ (t, x) \in \mathbb{R} \times \mathbb{R}^n \;\middle|\; t \in T \text{ and } x \in G_t \right\}, \tag{1.4}$$

where T = an interval (finite or infinite) in $\mathbb{R}$

G_t = a connected subset of $\mathbb{R}^n$.

There are many reasons why G may be a proper subset of phase space $\mathbb{R} \times \mathbb{R}^n$. In some cases they arise quite naturally from physical constraints as demonstrated by the following two aerospace examples.

Example 1

An earth satellite is to be tracked by a ground radar antenna. The tracking period is limited to an interval $T = [t_1, t_2]$ due to tracking visibility. Over this period the antenna's tracking axis must be controlled so as to minimize the antenna pointing error in some sense, assuming radar track acquisition commences at t_1.

In this example the response vector y may be taken to be the antenna's elevation and azimuth tracking angles, E and A. The state vector would consist of these angles and their derivatives up to a sufficient order to enable the relation of antenna servocontrol torques (u_E, u_A) to state vector to have the form of (1.1).

The phase cross-section $G_t \subset \mathbb{R}^n$ would consist of states satisfying constraints on E, A, and their time derivatives. The limitations on the time derivatives may be due, for example, to constraints derived from disk and mount structural factors. Normally, G_t is invariant with time.

Example 2

The trajectory or orbit of a space probe about a planet is to be controlled over an interval T without incurring impact with the planet. If we choose x as the 6-vector of the probe's position and velocity components in some suitable reference frame, then

$$G = \left\{ (t, x) \in \mathbb{R} \times \mathbb{R}^6 \;\middle|\; t \in T \text{ and } \| x_p - x_{\oplus}(t) \| > a \right\}$$

in which x_p and $x_\oplus$ are the position vectors of the probe and planet, and a is the planet's radius.

(End of Example 2)

In addition to restrictions arising directly from problem constraints, there are those which can be levied by the control designer. In general, not all phases are feasible phases. That is, not all phases may lie on a trajectory which terminates on an assigned target set S. Therefore admissible solutions to a control process may exist for only a proper subset of phase space. Naturally, if such a subset can be determined ahead of time, the control designer will insist that motion in phase space be restricted to it.

Example 3

Consider the scalar process equation

$$\frac{dx}{dt} = ax + u \quad , \qquad a = \text{constant} > 0; \quad t \geq 0$$

We wish to select the control for an initial phase $(0, x_o)$ which causes the origin to be reached in minimum time. The control must be selected from the class

$$\mathcal{U} = \left\{ u(\cdot) \in L_\infty(T) \,\middle|\, |u(t)| \leq 1 \right\}$$

Note that the set of states which are reachable from x_o under the class $\mathcal{U}$ are those which satisfy the following equation.

$$x(t) = e^{at}\left[x_o + \int_o^t e^{-a\tau} u(\tau)d\tau\right] \quad , \ t \geq 0$$

Since $\mathcal{U}$ represents the unit sphere of the space of bounded, measurable functions on T, and hence is convex, the set of responses commencing with x_o is convex. Thus the reachable states at any time $t \geq 0$ satisfy the inequality

$$e^{at}\left[x_o - \frac{1}{a}(1-e^{-at})\right] \leq x(t) \leq e^{at}\left[x_o + \frac{1}{a}(1-e^{-at})\right]$$

This condition indicates that the origin can be reached for some $t\epsilon[0,\infty)$ if and only if $|x_o| < \frac{1}{a}$.

Thus, the set G of permissible phases for this example is given by

$$G = \left[(t,x)\epsilon \mathbb{R} \times \mathbb{R} \,\middle|\, t\epsilon\,[0,\infty) \text{ and } |x| < \frac{1}{a}\right]$$

1.2 THE SET OF TARGET PHASES, S

We may assign a set S of permissible phases $(t,x)\epsilon\, G$ which constitutes the terminal objectives or "right-end" boundary conditions of the control process.

It is possible to define S in a manner similar to G, i.e.,

$$S = \left\{(t,x)\epsilon\, \mathbb{R} \times \mathbb{R}^n \,\middle|\, t\epsilon T_1 \text{ and } x\epsilon S_t \subset G_t\right\}$$

where

$$T_1 = \text{ an interval (finite, degenerate, or infinite)} \subset T.$$

However, the assignment of a target set in concrete problems usually arises from rather specific statements such as the set of phases satisfying a system of equalities

$$F_i(t,x) = 0 \qquad i = 1,\ldots,S \leq (n+1)$$

or possibly a system of inequalities

$$F_i(t,x) \leq 0 \qquad i = 1,\ldots,S \leq 2(n+1)$$

We shall consider that S is specified by either of these or by a possible mixture, and that the functions $F_1,\ldots,F_S$ are continuously differentiable with respect to (t, x) at least over neighborhoods of those phases where one or more of the equalities obtain.

The definition of S is thus sharpened as follows:

$$S = \left\{ (t, x) \in G \;\middle|\; F_i(t, x) \leqq 0, \; i = 1, \ldots, S \right\} \tag{1.5}$$

where $\leqq$ indicates the possibility of inequality as well for some or all values of the index i.

<u>Example 4</u>

The <u>time-free, fixed right end</u> target set of Example 3 can be put into the above form.

<u>Example 5</u>

<u>Fixed time, free right end</u> problems are characterized by

$$S = \left\{ (t, x) \in G \;\middle|\; t - t_1 = 0 \right\}$$

where t_1 is a constant such that $[t, t_1] \subset T$.

<u>Example 6</u>

If, in the space probe problem of Example 2, we wish instead to impact with the planet, then we may set $G = T \times \mathbb{R}^6$ and specify

$$S = \left\{ (t, x) \in G \;\middle|\; \| x_p - x_\oplus(t) \| - a \leq 0 \right\}$$

1.3 THE CLASS OF ADMISSIBLE CONTROLS, $\mathcal{U}$

The control designer is usually confronted with control constraints due to design limitations or conditions corresponding to physical realizability. Thus, we may expect that the value of the control vector u at any instant will be restricted to a subset $U \subset \mathbb{R}^r$

in order to reflect these conditions and limitations. We shall assume that U satisfies the following conditions:

1. U is invariant with respect to the instantaneous phase (t, x)
2. U is a closed, convex set in $\mathbb{R}^r$ containing an open r-dimensional sphere.

In addition to restricting the range of the control functions to the set U, we shall impose the relatively weak requirement that the control functions be measurable and essentially bounded over finite time intervals. This requirement is not severe, since in actual practice we would certainly deal with control functions which are at least piecewise continuous and finite.

Given an initial phase $(t_o, x_o) \epsilon G$, an admissible control is an element of the function class $\mathcal{U}_{t_o}$ defined as follows:

$$\mathcal{U}_{t_o} = \left\{ u(\cdot) \epsilon L_\infty[t_o, t_1] \,\middle|\, [t_o, t_1] \subset T \text{ and } u \epsilon U, \forall t \epsilon [t_o, t_1] \right\} \tag{1.6}$$

Finally, the class of all admissible controls is defined as

$$\mathcal{U} = \bigcup_{t_o \epsilon T} \mathcal{U}_{t_o}$$

We shall sometimes write $\mathcal{U}_{t_o}[U]$ and $\mathcal{U}[U]$ in order to indicate a particular set U upon which the admissible controls depend.

<u>Example 7</u> <u>Unbounded Controls</u>

With $U = \mathbb{R}^r$, we obtain the (maximal) class of controls which have finite L_∞-norms, and hence finite L_1- norms, over finite intervals. Since these norms are usually related to energy or total available resources in concrete problems, the admissible class still retains physical meaning even though we choose to approximate our control set U by $\mathbb{R}^n$.

Example 8 Bounded Controls

If U is a bounded set, then the admissible controls are uniformly bounded, measurable functions. In concrete problems the bounds might correspond to saturation limits on torques and voltages or to mechanical restraints on actuator deflections.

1.4 THE PERFORMANCE FUNCTIONAL, J

The final element of our optimum control problem is the criterion by which we measure the relative merit of admissible control functions. Various measures of performance are possible such as the time or control effort required to attain the target set, the deviation of the state, at an assigned final time, from a desired state, or the integral square error of the state trajectory wrt a desired trajectory. These performance measures, as well as others in common use, can be characterized by functionals which map the function triple $(\bullet, x(\cdot), u(\cdot))$ over $[t_o, t_1] \subset T$ into the real line.

Strictly speaking, the performance functional is defined only for function triples which satisfy the following properties:

1. The control function is admissible, i.e., $u(\cdot) \in \mathcal{U}_{t_o}$,
2. The phase trajectory produced by $u(\cdot)$ remains within the set G of permissible phases and reaches the set S of target phases.

We shall use the term feasible to describe admissible controls and trajectories which cause condition 2 to obtain.

In this thesis we shall consider performance functions of the following form:

$$J(t_o, x_o; u_\delta) = \lambda\Big(x(t_1)\Big) + \int_{t_o}^{t_1} L'\Big(t, x(t), u_\delta(t)\Big)\, dt, \tag{1.7}$$

where (t_o, x_o) = initial phase in G,

$\Big(t_1, x(t_1)\Big)$ = terminal phase in S,

u_δ = a feasible control function,

and λ is twice differentiable in x and $L'(t, x, u)$ is continuously differentiable in (t, x, u).

Owing to the differentiability properties of λ and L', we may replace (1.7) by a criterion of the form:

$$J(t_o, x_o; u_\delta) = \int_{t_o}^{t_1} L\Big(t, x(t), u_\delta(t)\Big)\, dt, \tag{1.8}$$

where

$$L(t, x, u) \triangleq \frac{\partial\lambda}{\partial x} \cdot f(t, x, u) + L'(t, x, u),$$

and

$$\frac{\partial\lambda}{\partial x} \triangleq \left(\frac{\partial\lambda}{\partial x_1}, \frac{\partial\lambda}{\partial x_2}, \ldots, \frac{\partial\lambda}{\partial x_n}\right), \text{ a row vector.}$$

This form differs from (1.7) by an additive term $\lambda(x_o)$ which does not enter into relative comparisons of different control functions.

Note that we have indicated that J is dependent on the initial phase (t_o, x_o) and the feasible control function u_δ, even though the forms of (1.7) and (1.8) make it clear that J is a functional of the triple $\Big(\bullet, x(\cdot), u_\delta(\cdot)\Big)$. The state-determined property of the process [Equation (1.3)] allows this identification.

For an initial phase (t_o, x_o), the performance functional induces a linear ordering on the set $\mathcal{U}_{t_o}^{o}$ of feasible control functions.

We shall assume that λ, L', or L are defined so that a control u_α is considered as good for (t_o, x_o) as a control u_β if

$$J(t_o, x_o; u_\alpha) = J(t_o, x_o; u_\beta)$$

and better for (t_o, x_o) than u_β if

$$J(t_o, x_o; u_\alpha) < J(t_o, x_o; u_\beta)$$

If the control problem is formulated correctly and meaningfully, then we can expect that

(1) $$\inf_{u_\delta \in \mathcal{U}^o_{t_o}} J(t_o, x_o; u_\delta) > -\infty$$

(2) There will be control $u_* \in \mathcal{U}^o_{t_o}$ such that

$$J(t_o, x_o; u_*) = \inf_{u_\delta \in \mathcal{U}^o_{t_o}} J(t_o, x_o; u_\delta)$$

A feasible control u_* satisfying (2.) is said to be <u>optimal for</u> (t_o, x_o) (or merely <u>optimal</u> if it is clear that we have a particular initial phase in mind).

If condition 1 holds without 2 being necessarily true, then a sequence $< u_\alpha >$ of feasible controls will exist such that

$$\lim_{\alpha \to \infty} J(t_o, x_o; u_\alpha) = \inf_{u_\delta \in \mathcal{U}^o_{t_o}} J(t_o, x_o; u_\delta)$$

A control u_α satisfying

$$J(t_o, x_o; u_\alpha) < \inf_{u_\delta \in \mathcal{U}^o_{t_o}} J(t_o, x_o; u_\delta) + \epsilon$$

for $\epsilon > 0$ will be called <u>ϵ-optimal for (t_o, x_o)</u>.

Example 9 Final Value Loss

Given the target $S = \{(t, x) \epsilon G \mid t - t_1 = 0\}$ corresponding to a fixed time, free right end problem, we may wish to minimize the deviation of the terminal state $x(t_1)$ from a desired state x_f. Thus

$$J(t_o, x_o; u_\alpha) = \lambda\left(x(t_1)\right) = h^2\left(\| x(t_1) - x_f \|\right)$$

where $h(\cdot)$ is a differentiable, monotone increasing function of its argument. In this case, J assumes the role of a final value loss function. Such a performance measure might be applicable to control of the final position and velocity of a rocket at an assigned thrust termination time.

Example 10 Servomechanism Loss

We may wish to control the process so that its state trajectory over $[t_o, t_1]$ approximates a reference trajectory $\phi(t)$ in some optimum sense. For example,

$$J(t_o, x_o; u_\alpha) = \int_{t_o}^{t_1} L'\left(t, x(t)\right) dt = \int_{t_o}^{t_1} h\left(\| x(t) - \phi(t) \|\right) dt$$

where $h(\cdot)$ has the same properties as in the previous example.

This type of performance measure might be used in the antenna servo problem of Example 1.

Example 11 Control Cost

The effort expended to reach a target set S is an important performance measure which is frequently called upon. Thus, if

$$J(t_o, x_o; u_\alpha) = \int_{t_o}^{t_1} C\left(t, u_\alpha(t)\right) dt$$

then we shall call J a control cost functional. For example $C(t, u)$

may represent instantaneous power flow in an electrical network with u as a control voltage or current; J would then correspond to expended electrical energy. Or, $C(t, u)$ may be the instantaneous magnitude of thrust in a throttable rocket engine, and hence J would correspond to the total impulse requirement.

Example 12 Performance Loss with Cost Constraints

We may wish to combine final value loss λ with control cost to meet the requirements of Example 9 when control effort is limited to a level which is low enough to be significant. Thus

$$J_\rho(t_o, x_o; u_\alpha) = h^2\left(\| x(t_1) - x_f \|\right) + \rho \int_{t_o}^{t_1} C\left(t, u_\alpha(t)\right) dt$$

would be used, where ρ is a Lagrange multiplier which is adjusted to meet the constraint on control cost.

Similarly, the servomechanism loss of Example 10 may be combined in this way with the control cost.

1.5 FEEDBACK CONTROL LAWS

The control functions mentioned thus far are time functions over some interval. In many practical problems it would be desirable if the control vector u could be generated by some function δ which would map the instantaneous value of a response vector $y \in \mathbb{R}^m$ into U. Such a function would constitute a feedback control law, or briefly, control law.

A control law $\delta : y \in \mathbb{R}^m \rightarrow u \in \mathbb{R}^r$ will be called admissible if it generates an admissible control function $u_\delta \in \mathcal{U}$ for every response function $y(\cdot)$. Similarly, the terms feasible or optimal are applied to it if it generates feasible or optimal control functions for every response function.

In dealing with a control law δ we shall denote its performance measure by $J(t_o, x_o; u_\delta)$. Sometimes it will be necessary to abuse the notation by denoting it as $J(t_o, x_o; \delta)$.

One important example of a control law is one in which the response y is the instantaneous phase (t, x) itself.

CHAPTER 2

OPTIMUM CONTROL THEORY

2.0 STATEMENT OF THE CONTROL PROBLEM

With the assumptions and definitions introduced in Chapter 1, we may state a restricted version of the modern control problem as follows:

"A control process is defined by a given quintuple ($\mathcal{P}$, G, S, $\mathcal{U}$, J), where $\mathcal{P}$ denotes a state-determined process defined by

$$\frac{dx}{dt} = f(t, x, u) \tag{2.1}$$

in which $(t, x) \in G \subset \mathbb{R} \times \mathbb{R}^n$, $u \in U \subset \mathbb{R}^r$, and f is a mapping from $G \times U$ into $\mathbb{R}^n$ which is continuously differentiable in an open region of $\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^r$ containing $G \times U$.

We wish to determine controls $u(\cdot) \in \mathcal{U}$ which cause the phase (t, x) to reach S so that J is minimized."

This is essentially the statement given by Pontryagin et al.,[2] except that we have required differentiability with respect to (t, x, u) rather than merely requiring continuity of f with respect to (t, x, u) and differentiability with respect to (t, x) for each $u \in U$. We shall summarize Pontryagin's method of solution in Section 2.2.

If we further specify that the control functions $u(\cdot)$ be generated by feedback control laws δ_u, then under certain conditions Bellman's method of solution is applicable. This method is summarized in Section 2.3.

2.1 EXISTENCE OF OPTIMAL CONTROLS

The methods of Bellman and Pontryagin presume existence of optimum solutions so that the necessary conditions on which the two methods are based become sufficient if it turns out that only one control can satisfy the respective necessity conditions. The question of existence of optimal controls was taken up by Fillipov,[3] Markus and Lee,[4] Roxin,[5] and others.[6,7,8] Fillipov provided existence conditions for the time optimal problem even when the control set U is dependent on (t, x). Markus and Lee gave existence conditions in processes $\mathcal{P}$ in which the control vector u appears linearly in the state equations (2.1) and performance functional J. Roxin's conditions are more general than those of Markus and Lee. We shall state a theorem based on his results.

For this theorem we assume that assumptions given in Chapter 1 hold (except, of course, the assumption of existence of an optimum control), and that $G = T \times \mathbb{R}^n$.

<u>Theorem 1</u> (Roxin)

Suppose we reexpress the performance functional J given in Equation (1.7) in the form of (1.8).

$$J(t_o, x_o; u_\alpha) = \int_{t_o}^{t_1} L\Big(t, x(t), u(t)\Big)\, dt \tag{2.2}$$

where the state $x \in \mathbb{R}^n$ is governed by

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

so that $\hat{f} \triangleq (f, L) \in \mathbb{R}^{n+1}$ is a mapping from $T \times \mathbb{R}^n \times U$ into $\mathbb{R}^{n+1}$, where $U \subset \mathbb{R}^r$ is <u>compact.</u> Let there be a feasible solution from (t_o, x_o) to a closed target set S.

Then an optimal control $u_* \in \mathcal{U}_{t_o}[U]$ exists if $\hat{f}$ satisfies the following:

(i) $\hat{f}$ is continuous in $(x,u) \in \mathbb{R}^n \times U$ for each $t \in T$, and integrable over T for fixed $(x,u) \in \mathbb{R}^n \times U$.

(ii) For $(t,u) \in T \times U$, there is a constant K so that
$\|\hat{f}(t,x_2,u) - \hat{f}(t,x_1,u)\| \leq K\|x_2 - x_1\|$.

(iii) For all $u \in U$ (uniformly),
$\|\hat{f}(t,x,u)\| \leq \mu(t)\, g(\|x\|)$

where $\mu(t) \in L_1$ over every finite interval, and g is finite for finite arguments but non-integrable over intervals of the form $[c,\infty]$, $C \geq 0$.

(iv) For each $(t,x) \in T \times \mathbb{R}^n$, the range of $\hat{f}(t,x,u)$ as u describes the set U is convex.

Proof

Roxin showed that the set $R(t_o, x_o) \subset T \times \mathbb{R}^n \times \mathbb{R}$ of all reachable points of the form $(t, x(t), J(t))$, $t \geq t_o$, using the class $\mathcal{U}_{t_o}[U]$, where U is compact, is closed under conditions (i) through (iv).

Since S is closed, the product set $S \times \mathbb{R} \subset T \times \mathbb{R}^n \times \mathbb{R}$ is closed. Thus, $(S \times \mathbb{R}) \cap R(t_o, x_o)$ is closed and, further, nonempty by the assumption of existence of a feasible solution. This set consists of all points (t,x,J) such that $(t,x) \in S$. In other words, the J-components of this set form the totality of performance values J for feasible controls. It is closed and nonempty.

We now use the assumption of Chapter 1 that J is bounded from below by a real number to complete our proof.

(End of Proof)

Conditions (i) through (iii) of the theorem are essentially Caratheodory's condition for existence, uniqueness, and boundedness for an absolutely continuous solution $\hat{x}(t) = (x(t), J(t))$ which satisfies

$\hat{x}(t_o) = (x_o, 0)$ for every $u(\cdot) \epsilon \mathcal{U}_{t_o}$. If in Section 2.0 we set $G = T \times \mathbb{R}^n$ and require boundedness of the partial derivatives of $f(t,x,u)$ and $L(t,x,u)$ with respect to x, then conditions (i), (ii), (iii) easily obtain. Condition (iv) and the requirement that U be compact, however, are still essential [see Neustadt[6] on removal of (iv)].

We cannot apply the theorem directly to the process as defined in Section 2.0 if G_t, $t \epsilon T$, is a proper subset of $\mathbb{R}^n$. If we add the condition that all phase trajectories produced by $\mathcal{U}_{t_o}[U]$ and commencing from $(t_o, x_o) \epsilon G$ do not leave G, then existence of an optimum control follows.

2.2 PONTRYAGIN'S METHOD OF SOLUTION

In presenting Pontryagin's method we shall assume that an optimal control u_* exists, and that there is a neighborhood of admissible controls about u_* which produce feasible trajectories. We first state his necessary conditions, then discuss how one might solve the two-point boundary value problem which arises in general.

Notation

In the following

$$P \triangleq (P_1, \ldots, P_n), \text{ a row-vector}$$

$$\hat{P} \triangleq (P_o, P), \text{ an } (n+1) \text{ row-vector}$$

In general for an m-tuple $a = (a_1, a_2, \ldots, a_m)$ and an n-tuple $b = (b_1, b_2, \ldots, b_n)$

$$\frac{\partial a}{\partial b} \triangleq (m \times n) \text{ matrix of elements } \left(\frac{\partial a}{\partial b}\right)_{ij} = \left(\frac{\partial a_i}{\partial b_j}\right)$$

2.2.1 Necessary Conditions for Optimality[2, 25]

Theorem 2 (Pontryagin)

A feasible control $u(t)$ and its corresponding trajectory $\big(t, x(t)\big)$, commencing from (t_o, x_o) and terminating at $(t_1, x_1) \epsilon S$ cannot be optimal unless the following conditions hold:

(i) (The Minimum Principle)

There exist an absolutely continuous solution to the co-state system

$$\frac{dP_o}{dt} = - P \frac{\partial f}{\partial t}\big(t, x(t), u(t)\big) - \frac{\partial L}{\partial t}\big(t, x(t), u(t)\big)$$

$$\frac{dP}{dt} = - P \frac{\partial f}{\partial x}\big(t, x(t), u(t)\big) - \frac{\partial L}{\partial x}\big(t, x(t), u(t)\big)$$

so that $H(t, x, u, \hat{P}) \triangleq L(t, x, u) + Pf(t, x, u) + P_o$

satisfies

$$H\big(t, x(t), u(t), \hat{P}(t)\big) = \inf_{u \epsilon U} H\big(t, x(t), u, \hat{P}(t)\big)$$

almost everywhere on $[t_o, t_1]$, and $H\big(t_1, x(t_1), u(t_1), \hat{P}(t_1)\big) = 0.$

(ii) (Transversality)

At the terminal phase $(t_1, x_1) \epsilon S$, the vector $\big(P_o(t_1), P(t_1)\big)$ is orthogonal to S.

(End of Assertion)

Remarks

If L, f, and the set S_t of target states do not depend on time, then P_o may be dropped from all considerations.

If a solution can be found so that transversality holds and $H\big(t, x(t), u(t), \hat{P}(t)\big) = \inf_{u \epsilon U} H\big(t, x(t), u, \hat{P}(t)\big)$ with $H = 0$ at t_1, then H will automatically be zero a.e. $[t_o, t_1]$. This is not true in reverse, however.

It is easy to show that Pontryagin's principle leads to a two-point boundary value problem. Let $u_2 = k(t, x, P)$ be the (hopefully) unique function which minimizes $H(t, x, u, \hat{P}) = L(t, x, u) + Pf(t, x, u) + P_o$ over the control set U. Then substituting $k(t, x, P)$ into the state and co-state equations, we have

$$\frac{dx}{dt} = f\Big(t, x, k(t, x, P)\Big)$$

$$\frac{dP}{dt} = -P\, \frac{\partial f}{\partial x}\Big(t, x, k(t, x, P)\Big) - \frac{\partial L}{\partial x}\Big(t, x, k(t, x, P)\Big)$$

subject to boundary conditions:

$$(t_o, x_o),\ (t_1, x_1) \in S,\ \Big(P_o(t_1),\ P(t_1)\Big) \perp S,$$

and

$$H(t_1, x_1, k_1, \hat{P}_1) = 0$$

If a solution to this, generally nonlinear, two-point boundary value problem is found, then $u_*(t) = k\Big(t, x(t), P(t)\Big)$, $[t_o, t_1]$.

2.2.2 Sucessive Approximation Techniques

Approximation in Co-state Space

With suitable assurance for uniqueness and existence of a solution, the equations just presented may be solved by numerical iteration techniques. The solution $\Big(x(\cdot),\ P(\cdot)\Big)$, such that the right-end boundary conditions are satisfied, depends on finding the correct value for $P(t_o)$. This can be approximated by an appropriate iteration method. The idea here is to integrate from a set of trial initial values $\Big\{\Big(x_o,\ P_{(\alpha)}(t_o)\Big)\Big\}$, so that tabular functions relating the behavior of the right-end constraint expressions to $P(t_o)$ can be obtained. A functional approximation is then made to these tabular functions and a value for $P(t_o)$ is solved for which yields the apparent solution for $P(t_o)$ on this approximate surface. The entire procedure is then repeated for a set of perturbations about this apparent solution. Thus, the method would utilize an iteration procedure such as Newton's method, Secant-methods, or Muller's method.

The difficulty with this method is that the co-state equations are generally unstable for the forward-integration steps involved. Thus, the solution will be very sensitive to small changes in $P(t_o)$ and numerical integration errors may hamper convergence.

Approximation in Control Space

This method involves a sequence of trial control functions $<\mu_\alpha(t)>$ which are successively generated by the following technique.[9,10]

Given a trial function $u_\alpha(t)$, the state equations are integrated forward from (t_o, x_o) to yield the state-trajectory $x_\alpha(t)$. Using (x_α, u_α) the linear co-state equations are integrated backward from $\left(t_1, P(t_1)\right)$, so that the co-state trajectory $P_\alpha(t)$ is obtained. Then using the minimum principle we generate $u_{\alpha+1}(t)$ by

$$H\left(t, x_\alpha(t), u_{\alpha+1}(t), \hat{P}_\alpha(t)\right) = \inf_{u_{\alpha+1} \in U} H\left(t, x_\alpha(t), u_{\alpha+1}, \hat{P}_\alpha(t)\right)$$

The process is repeated until the sequence $< H\left(t, x_\alpha(t), u_{\alpha+1}(t), \hat{P}_\alpha(t)\right) >$ converges to a function which has an acceptably small deviation from zero over $[t_o, t_1]$. Why this method works at all and what constitutes an acceptably small deviation function will be the subject of Chapter 5, Suboptimal Control Sequences.

There is one difficulty associated with the method as outlined. We have no assurance that the trial control functions lead to the target set, nor do we necessarily know the value $P(t_1)$ to use in the backward integration of the co-state equations. In the case of a fixed-time, free right-end problem, however, this difficulty vanishes, since $S = \{t_1\} \times \mathbb{R}^n$ and $P(t_1) = 0$ by transversality.

Since in actual practice we do not expect perfect knowledge of the initial state nor perfect execution of a desired control function, the attainment of a precise target set may be relaxed. Thus, an optimal solution to the free right-end problem wherein the performance functional is augmented by a final value loss with respect to a target set S_{t_1}, for fixed final time t_1, would normally be close enough for practical purposes. If the target set S is not restricted to a single final time, then optimal solutions over an appropriate range of final times must be found so that the optimal final time may be selected. This is the so-called <u>penalty function</u> method due originally to Courant[11] and applied by Kelly,[12] Ostrovskii,[13] and Okamura[14] to control problems. The latter has provided proofs that the modified problem converges to the original problem as one assigns greater weights to the added final value loss.

2.3 BELLMAN'S METHOD OF SOLUTION

Bellman's method, based on his dynamic programming concepts,[15,16,9] is aimed at the derivation of optimal control <u>laws</u> which generate optimum control functions as a function of the instantaneous phase (t, x). As one might expect of a method which solves an entire class of problems [i.e., for all permissible initial phases] at once, certain conditions must hold.

<u>Basic Assumptions</u>

1. Attention is restricted to an open, connected subset $G_o \subset G$ for which an optimum control law δ_* exists.

2. The optimum performance function $J(t,x;\delta_*)$, $(t,x)\epsilon G_o$, has continuous partial derivatives with respect to time and state components, i.e., $J(t,x;\delta_*)$ is everywhere differentiable in G_o.

2.3.1 Necessary Conditions for Optimality[16]

Theorem 3 (Bellman)

In order that a feasible control law δ, defined on G_o, be optimal, its performance function $J = J(t,x;\delta)$ must satisfy the following condition.

$$\inf_{u\epsilon U}\left[L(t,x,u) + \frac{\partial J}{\partial x} f(t,x,u) + \frac{\partial J}{\partial t}\right] = 0 \qquad \text{for all } (t,x)\epsilon G_o$$

and

$$J(t,x;\delta) = 0 \qquad \text{for } (t,x)\epsilon S$$

(End of Assertion)

Remarks

Kalman[17] and Bridgeland[18,19] have provided theorems wherein the above conditions also become sufficient if further conditions are hypothesized. However, these amount to existence and uniqueness arguments for the optimal control law. We shall not pursue this matter since it would take us far from our objectives.

The conditions of the theorem may be expressed as

$$\frac{\partial J}{\partial t} + \inf_{u\,\epsilon\,U}\left[\frac{\partial J}{\partial x} f(t,x,u) + L(t,x,u)\right] = 0 \tag{2.3}$$

with boundary condition $J(t,x;\delta) = 0$ for $(t,x)\epsilon S$. This leads to two conditions which one must satisfy:

1. $\delta_* = k\left(t, x, \frac{\partial J_*}{\partial x}\right)$, where k minimizes

$\frac{\partial J_*}{\partial x} f(t, x, k) + L(t, x, k)$ for all $(t, x) \in G_o$.

2. The optimal performance function $J_* = J(t, x; \delta_*)$ must satisfy the partial differential equation

$$\frac{\partial J_*}{\partial t} + \frac{\partial J_*}{\partial x} \cdot f\left(t, x, k\left(t, x, \frac{\partial J_*}{\partial x}\right)\right) + L\left(t, x, k\left(t, x, \frac{\partial J_*}{\partial x}\right)\right) = 0$$

with $J_* = 0$ on S.

If the (generally nonlinear) partial differential equation can be solved, then condition 1 defines the optimal control law.

One method of solution would consist of finding J_* in terms of a polynomial series in (t, x). Alternatively, we may resort to a purely numerical approach and solve (2.3) over a numerical grid in G_o using Bellman's flooding procedure.[15] Finally, there is a method of successive approximations which is described in the next section.

2.3.2 Method of Successive Approximation

In this method we generate a sequence $< \delta_\alpha >$ of control laws using the following algorithm.

1. Given a feasible control law δ_α we solve the linear partial differential equation

$$\frac{\partial J_\alpha}{\partial t} + \frac{\partial J_\alpha}{\partial x} f(t, x, \delta_\alpha) + L(t, x, \delta_\alpha) = 0$$

with $J_\alpha = 0$ on S. J_α will actually be δ_α's performance function. Thus it can also be found by direct calculations on

$$J(t_o x_o; \delta_\alpha) = \int_{t_o}^{t_1} L(t, x, \delta_\alpha)\, dt,$$

where $dx/dt = f(t, x, \delta_\alpha)$, and (t_o, x_o) is allowed to range over G_o.

2. Having $J_\alpha(t, x)$ we generate $\delta_{\alpha+1}$ by

$$\delta_{\alpha+1} = k\left(t, x, J_\alpha(t, x)\right)$$

3. We terminate the sequence when

$$\frac{\partial J_\alpha}{\partial t} + \inf_{u \in U} \left[\frac{\partial J_\alpha}{\partial x} f(t, x, u) + L(t, x, u)\right],$$

which is always nonpositive, is sufficiently close to zero over G_o.

(End of Algorithm)

This method, called approximation in policy space, was first suggested by Bellman[16] who also showed that the sequence $< J_\alpha >$ was monotone decreasing, and hence convergent for performance functions J which are bounded from below. Recently, Leake and Liu[10] provided convergence theorems with sufficient mathematical rigor.

Again the important question is one concerned with termination of the algorithm. What constitutes a residual function which is sufficiently close to zero? We shall consider this criterion in Chapter 5.

CHAPTER 3
SUBOPTIMAL CONTROL THEORY

3.1 THE REGRET FUNCTION

Given an initial phase $\theta_o = (t_o, x_o)$ and two feasible control functions u_1 and u_2 the relative performance loss of u_1 with respect to u_2 is defined as

$$R(\theta_o\, ;\, u_1, u_2) \triangleq J(\theta_o\, ;\, u_1) - J(\theta_o\, ;\, u_2) \tag{3.1}$$

The performance of a feasible control u relative to the optimal control u_* is of interest. For this purpose we define the regret function for u

$$R(\theta_o ; u) \triangleq \sup_{u_2 \in \mathcal{U}_{t_o}} R(\theta_o ; u, u_2) = J(\theta_o ; u) - \inf_{u_2 \in \mathcal{U}_{t_o}} J(\theta_o ; u_2) \tag{3.2}$$

Assuming the optimal control u_* exists, Equation (3.2) becomes

$$R(\theta_o ; u) = J(\theta_o ; u) - J(\theta_o ; u_*)$$

Let us now consider the fact that the methods of Pontryagin and Bellman yield control procedures δ which are uniformly optimal with respect to initial phases. That is, for every phase $\theta = (t, x)$ in a set G_o of feasible phases their procedure yields $J_*(\theta) = \inf_{u \in \mathcal{U}_t} J(\theta ; u)$.

With respect to these uniformly optimal procedures we say that a procedure δ is ϵ-optimal for $\theta_o \in G_o$ if its control function u_δ for the initial phase θ_o yields:

$$R_\delta(\theta_o) \triangleq J(\theta_o ; u_\delta) - J_*(\theta_o) \leq \epsilon \tag{3.3}$$

We say that the procedure δ is <u>(uniformly) ϵ-optimal</u> if

$$\sup_{\theta_o \in G_o} R_\delta(\theta_o) = \sup_{\theta_o \in G_o} \Big(J(\theta_o; u_\delta) - J_*(\theta_o)\Big) \le \epsilon$$

Finally, if we have a probability measure $\mu(\theta_o)$ defined on G_o, we say that a procedure δ is <u>ϵ-Bayes with respect to μ</u> if

$$\int_{G_o} R_\delta(\theta_o) d\mu(\theta_o) = \int_{G_o} \Big[J(\theta_o; u_\delta) - J_*(\theta_o)\Big] d\mu(\theta_o) \le \epsilon$$

3.2 INTEGRAL REPRESENTATION OF THE REGRET FUNCTION

In order to create a structure which allows us to relate to the (indirect) variational methods of Pontryagin and Bellman we must seek an integral representation of the regret function $R_\delta(\theta_o)$. Thus Theorem 4 below will be key in all our further developments.

<u>Theorem 4</u>

Let $J_*(t,x)$ be continuously differentiable on an open, connected subset $G_1 \subset G_o$ and let $\bar{G}_1 \cap S$ be nonempty. Let the procedure δ have a control u_δ for $\theta_o = (t_o, x_o) \in G_1$ which produces the feasible trajectory $\big(t, x(t)\big) \in G_1$ terminating at $(t_1, x_1) \in S$. Then the regret for θ_o has the form

$$R_\delta(\theta_o) = \int_{t_o}^{t_1} \Big[L\big(t, x(t), u_\delta(t)\big) + \frac{\partial J_*}{\partial x}\big(t, x(t)\big) \bullet f\big(t, x(t), u_\delta(t)\big) + \frac{\partial J_*}{\partial t}\big(t, x(t)\big)\Big] dt$$

<u>Proof</u>

By definition,

$$R_\delta(t_o, x_o) = J(t_o, x_o; u_\delta) - J_*(t_o, x_o)$$
$$= \int_{t_o}^{t_1} L\Big(t, x(t), u_\delta(t)\Big) dt - J_*(t_o, x_o)$$

Since $J_*(t,x)$ is continuously differentiable in G_1, $J_*\big(t, x(t)\big)$ is absolutely continuous over $[t_o, t_1]$. Thus, it may be represented as an indefinite integral

$$J_*\big(t, x(t)\big) = J_*(t_o, x_o) + \int_{t_o}^{t} \dot{J}_*\big(t, x(t)\big)\, dt$$

where $\dot{J}_*\big(t, x(t)\big) \in L_1[t_o, t]$ for all $t \epsilon [t_o, t_1]$.

In fact, we have

$$\dot{J}_*\big(t, x(t)\big) = \frac{\partial J_*}{\partial t}\big(t, x(t)\big) + \frac{\partial J_*}{\partial x}\big(t, x(t)\big) \bullet \frac{dx(t)}{dt}$$

Using $dx(t)/dt = f\big(t, x(t), u_\delta(t)\big)$ a.e. $[t_o, t_1]$, we obtain

$$J_*(t_1, x_1) = J_*(t_o, x_o) + \int_{t_o}^{t_1} \left[\frac{\partial J_*}{\partial t}\big(t, x(t)\big) + \frac{\partial J_*}{\partial x}\big(t, x(t)\big) \right.$$

$$\left. \bullet\, f\big(t, x(t_1) u_\delta(t)\big) \right] dt$$

Noting that on the target set all performance measures are equal to zero, we obtain $J_*(t_1, x_1) = 0$, and the result of our theorem follows easily.

Remark

The theorem has been proved only for a subset G_1 [of the set G_o of feasible phases] over which $J_*(\theta)$ is continuously differentiable. If $J_*(\theta)$ is only piecewise continuously differentiable over G_o, then we need further conditions in order for the representation to hold. However, for problems consistent with Bellman's assumptions, we have $G_1 = G_o$, and we may state the following corollary to Theorem 4.

Corollary 1

Let $\mathcal{D}_o$ be a nonempty class of feasible control laws δ defined over a set G_o of feasible phases such that $J_*(\theta) = \inf_{\delta \in \mathcal{D}_o} J(\theta; \delta)$ is continuously differentiable on G_o[an optimal control law $\delta_* \in \mathcal{D}_o$ need not exist]. A feasible control law $\delta \in \mathcal{D}_o$ is ϵ-optimal for $\theta_o \epsilon G_o$ if:

$$\int_{t_o}^{t_1} \left[L\left(t, x(t), u_\delta(t)\right) + \frac{\partial J_*}{\partial x}\left(t, x(t)\right) \bullet f\left(t, x(t), u_\delta(t)\right) + \frac{\partial J_*}{\partial t}\left(t, x(t)\right) \right] dt \leq \epsilon$$

for the pair $\left(x(\cdot),\ u_\delta(\cdot)\right)$ which δ generates commencing with θ_o. It is ϵ-optimal if this property holds for all $\theta_o \epsilon G_o$.

Remark

We may state another corollary dealing with properties of the optimum control law itself. This is an alternate form of Bellman's conditions given in Theorem 3.

Corollary 2 [Alternate form of Bellman's conditions]

In order that a feasible control law δ be optimal, its performance function $J(t, x)$, presumably differentiable on $(t, x) \epsilon G_o$, must satisfy

$$\inf_{u \epsilon U} \left[L\left(t, x(t), u\right) + \frac{\partial J}{\partial x}\left(t, x(t)\right) \bullet f\left(t, x(t), u\right) + \frac{\partial J}{\partial t}\left(t, x(t)\right) \right] = 0$$

almost everywhere $[t_o, t_1]$ for every feasible trajectory $\left(t, x(t)\right)$ generated by δ.

Proof

If δ is optimal, then it is ϵ-optimal for every $\epsilon > 0$. Thus for any $\theta_o \epsilon G_o$

$$\int_{t_o}^{t_1}\left[L\left(t, x(t), u_\delta(t)\right) + \frac{\partial J_*}{\partial x}\left(t, x(t)\right) \bullet f\left(t, x(t), u_\delta(t)\right) + \frac{\partial J_*}{\partial t}\left(t, x(t)\right)\right] dt = 0$$

No admissible variation of u_δ can cause the integral to be less than zero since the regret function is always non-negative. Thus

$$\inf_{u \in U}\left[L\left(t, x(t)u\right) + \frac{\partial J_*}{\partial x}\left(t, x(t)\right) \bullet f\left(t, x(t), u\right) + \frac{\partial J_*}{\partial t}\left(t, x(t)\right)\right] = 0$$

a.e. $[t_o, t_1]$ for every trajectory $\left(t, x(t)\right)$ of δ. Finally, if δ is optimal, then its performance function $J(t,x)$ must equal $J_*(t,x)$ everywhere on G_o, and hence we may substitute $J(t,x)$ for $J_*(t,x)$ in our last equation.

(End of Proof)

Corollary 2 is a more precise statement of Bellman's necessity condition (Theorem 3), and it brings the alternate formulations of Pontryagin and Bellman into a somewhat closer relation.

3.3 INTEGRAL REPRESENTATION OF REGRET FOR A SPECIFIC CLASS OF CONTROL LAWS

The representation derived in the previous section is restricted to optimum performance functions $J_*(t,x)$ which are continuously differentiable in a set $G_1 (\subset G_o)$ having limit points on S. Optimum control laws which are continuously differentiable on such a set will have J_*'s with this property [see Appendix A]. However, there is a large class of problems in which $J_*(t,x)$ is non-differentiable along a locus of phases through which we may wish to pass a feasible, suboptimal trajectory. Pontryagin et al.,[2] have given several examples of time-optimal problems where the optimum control law is discontinuous in G_o along certain hypersurfaces or switching boundaries. In these examples, $J_*(t,x)$ turned out to be continuous but non-differentiable on such boundaries. In many cases these boundaries were shaped in such a manner that a choice of an appropriate region G_1 would be severely restricted. Thus, we wish to obtain an integral representation of the regret function for laws which are not limited to points at which J_* is continuously differentiable.

It will turn out that the integral representation will still hold under certain conditions, and that, if the trajectory $\big(\cdot,\ x(\cdot)\big)$ moves along a locus where J_* is non-differentiable, the co-state variables can be used in place of the partials. We shall derive these results for J_*'s which correspond to optimal controls of a certain class.

3.3.1 The Class of Control Laws, $P[G_o]$

A control law δ is said to be of class P on a domain G_o of feasible phases (or briefly, of class $P[G_o]$), if it has the following properties:

- It is feasible, i.e., produces feasible trajectories contained in G_o and generates admissible control functions.
- There exists a finite partition $\{G_1, G_2, \ldots, G_K\}$ of G_o into regions such that

  (i) $G_o = \bigcup_{k=1}^{K} G_k$; $G_k \cap G_\ell = \phi$; G_k = connected set.

  (ii) $\delta(t, x)$ is continuously differentiable on each such region.

  (iii) The boundary between any two regions is of the form $b(t, x) = 0$, where $b(t, x)$ is continuously differentiable at its solution points in $G_o - S$.

  (iv) The boundary points in $G_o - S$ will form a collection of connected sets after what has been said. Each such connected set must belong entirely to one region or another.

Example 13

For the simple time-optimal problem

$$\frac{dx}{dt} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

where

$$x = (x_1, x_2) \in \mathbb{R}^2$$

$$u = \text{scalar control from } U = [-1, 1]$$

$$S = \left\{ (t, x) \in \mathbb{R} \times \mathbb{R}^2 \mid x = (0, 0) \right\}$$

The optimum control law[2] is

$$\delta(t, x) = \begin{cases} 1 & \text{if } (t, x) \in G_1 \\ -1 & \text{if } (t, x) \in G_{-1} \end{cases}$$

in which

$$G_1 = \left\{ (t, x) \mid x_2 + \sqrt{2x_1} \leq 0 \text{ for } x_1 \geq 0, \text{ or } x_2 + \sqrt{-2x_1} < 0 \text{ for } x_1 \leq 0 \right\}$$

$$G_{-1} = \left\{ (t, x) \mid x_2 + \sqrt{2x_1} > 0 \text{ for } x_1 \geq 0, \text{ or } x_2 + \sqrt{-2x_1} \geq 0 \text{ for } x_1 \leq 0 \right\}$$

This partition satisfies our requirements. The boundary is given by

$$b(t, x) = x_2 + \sqrt{2|x_1|} \ \text{sgn}\, x_1 = 0$$

which is continuously differentiable at all solution points except at the target state (0, 0). The two branches of the boundary [one for $x_1 > 0$ and one for $x_1 < 0$] satisfy condition (iv).

(End of Example)

We now give an assertion which will clarify the basis for a lemma to follow.

Assertion 1

For the control problem of Section 2.0, let G_o be an open, connected set of feasible phases and δ be a control law of class $P[G_o]$. Let $x(t; t_o, x_o)$ be a trajectory solution starting from the initial phase (t_o, x_o) in the interior of one of the regions. Then $x(t; t_o, x_o)$ is continuously differentiable with respect to (t_o, x_o).

Justification

Even though the governing differential equation

$$\frac{dx}{dt} = g(t, x) \triangleq f\big(t, x, \delta(t, x)\big)$$

may have discontinuous right-hand sides, we nonetheless have a succession of regions G_k, along the given trajectory, within which $g(t, x)$ is continuously differentiable. Since the boundaries between the regions have differentiable forms, then the phase (t_β, x_β) at which the trajectory first meets a boundary will be continuously differentiable with respect to the initial phase (t_o, x_o). But the function $g(t, x)$ will be continuously differentiable with respect to boundary phases for the next region, and hence motion within this next region will be differentiable with respect to (t_o, x_o). Thus, by a chain-rule of derivatives we can infer continuous differentiability throughout the entire motion in $G_o - S$.

For details of this reasoning, the reader is referred to problem 6, pp. 39-40, of Coddington and Levinson.[20]

(End of Justification)

Assertion 2

Under the same conditions of Assertion 1, the performance function $J(t_o, x_o; \delta)$ is continuously differentiable with respect to initial phases (t_o, x_o) in the interior of one of the regions.

Justification

The reader is referred to Appendix A for an explicit demonstration of the conclusion. However, one can readily appreciate that if the conclusion of Assertion 1 holds, then it should hold for the system of differential equations with

$$\frac{dJ}{dt} = L\left(t, x, \delta(t, x)\right)$$

adjoined to it.

(End of Justification)

We are now in a position to give the following lemma.

Lemma 1

Let $\delta_*(t, x)$ be an optimum control law belonging to $P[G_o]$. Suppose $\delta(t, x)$ is a feasible control law which causes the trajectory which it produces to move <u>along</u> one of δ_*'s switching boundaries <u>only</u> if it moves optimally. Then the regret function for δ has the form

$$R_\delta(t_o, x_o) = \int_{[t_o, t_1]-\omega_*} \left[L\left(t, x(t), u_\delta(t)\right) + \frac{\partial J_*}{\partial x}\left(t, x(t)\right) \cdot f\left(t, x(t), u_\delta(t)\right) + \frac{\partial J_*}{\partial t}\left(t, x(t)\right)\right] dt$$

where

ω_* = union of time intervals of <u>positive</u> measure during which $x(t)$ moves along a <u>switching</u> boundary of δ_*.

Proof

The proof of Theorem 4 was based on the fact that $J_*(t,x)$ was continuously differentiable, and hence $J_*\ t, x(t)$ was absolutely continuous on $[t_o, t_1]$. The same idea applies here as to its absolute continuity except for some modifications. We consider two types of subarcs of $x(\cdot)$, those which lie entirely within the interior of some region G_k of δ_*, and those which pass along a boundary for a positive interval of time.

Let us consider $[t_o, t_1]$ partitioned into consecutive open intervals $\{(\tau_i, \tau_{i+1});\ i=0,1,2,\ldots,I\}$ such that $x(\cdot)$ is either entirely within some region G_k or passing along a switching boundary of δ_*. Then,

$$J_*(t_1, x_1) - J_*(t_o, x_o) = \sum_{i=0}^{I}\left[J_*\left(\tau_{i+1}, x(\tau_{i+1})\right) - J_*\left(\tau_i, x(\tau_i)\right)\right]$$

where

$$\left(\tau_o, x(\tau_o)\right) \triangleq (t_o, x_o) \text{ and } \left(\tau_{I+1}, x(\tau_{I+1})\right) \triangleq (t_1, x_1)$$

For an interval (τ_i, τ_{i+1}) corresponding to passage of $x(\cdot)$ through a region G_k, we have as before

$$J_*\left(\tau_{i+1}, x(\tau_{i+1})\right) - J_*\left(\tau_i, x(\tau_i)\right) = \int_{\tau_i}^{\tau_{i+1}}\left(\frac{\partial J_*}{\partial x} f(t,x,u_\delta) + \frac{\partial J_*}{\partial t}\right)dt.$$

Otherwise, we have

$$\begin{aligned} J_*\left(\tau_{i+1}, x(\tau_{i+1})\right) - J_*\left(\tau_i, x(\tau_i)\right) &= -\int_{\tau_i}^{\tau_{i+1}} L\left(t, x_*(t), u_*(t)\right) dt \\ &= -\int_{\tau_i}^{\tau_{i+1}} L\left(t, x(t), u_\delta(t)\right) dt, \end{aligned}$$

since along the boundaries $\left(x(\cdot),\ u_\delta(\cdot)\right)$ must coincide with some optimal pair $\left(x_*(\cdot),\ u_*(\cdot)\right)$ by hypothesis.

Thus,

$$J_*(t_1, x_1) = J_*(t_o, x_o) + \int_{[t_o, t_1] - \omega_*} \left(L(t, x, u_\delta) + \frac{\partial J_*}{\partial x} f(t, x, u_\delta) + \frac{\partial J_*}{\partial t} \right) dt$$

$$+ \int_{\omega_*} L(t, x, u_\delta) dt$$

where ω_* is the set defined in the lemma.

Since, $J_*(t_1, x_1) = 0$ and, by definition,

$$R_\delta(t_o, x_o) = \int_{[t_o, t_1]} L(t, x, u_\delta) dt - J_*(t_o, x_o)$$

the desired result is readily obtained.

(End of Proof)

This result may be applied to some of the two-dimensional processes pointed out by Pontryagin et al.,[2] in which the boundaries coincide with optimal trajectories. In order to move along these boundaries, motion must therefore be optimal.

3.3.2 Integral Representation in Terms of Co-state Variables

Control procedures based on Pontryagin's method are control laws in the following sense. Suppose the optimum control problem is solved for all initial phases $(t_o, x_o) \epsilon\ G_o$. Then the initial value of the co-state vector $P(t_o)$ is available for each initial phase in G_o. Naturally it will depend on the initial phase, so we shall denote it by $P(t_o, x_o)$. Thus, the optimum control vector $u_*(t_o)$ at the phase (t_o, x_o) would be one which minimizes $L(t_o, x_o, u) + P(t_o, x_o) f(t_o, x_o, u)$, in accordance with the minimum principle. The optimum control law would then have the form

$$\delta_*(t,x) = Q\Big(t, x, P(t,x)\Big)$$

for every $(t,x) \epsilon G_o$.

We shall now give a lemma, whose proof is given in Appendix B, and finally a theorem for integral representation in terms of co-state variables.

Lemma 2

For optimum control laws of class $P[G_o]$,

$$P(t,x) = \frac{\partial J_*}{\partial x}(t,x)$$

$$P_o(t,x) = \frac{\partial J_*}{\partial t}(t,x)$$

wherever $J_*(t,x)$ is continuously differentiable, and wherever $\partial J_*/\partial x\Big(t, x(t)\Big)$ and $\partial J_*/\partial t\Big(t, x(t)\Big)$ are absolutely continuous over a time interval containing t.

(End of Statement)

This result deserves some comment even though it is not itself an end objective of this work. It has been shown by Pontryagin,[2] Kalman,[17] Rozonoer,[21] and others,[9,22] that whenever $J_*(t,x)$ is twice differentiable in G_o, the co-state variables may be everywhere equated to the partials of J_*. Lemma 2 goes a little further. J_* need not be twice differentiable throughout G_o in order to equate the two sets of variables at given points in G_o. In other words only local conditions need be satisfied, and these conditions do not involve second partials for J_*. Naturally this holds only for J_*'s arising from optimal controls in the class $P[G_o]$. However, this class is a rather broad one, encompassing the majority of optimal control laws which are synthesized in practice.

Theorem 5

Let the optimal control law δ_* be of class $P[G_o]$ and $\big(P(t,x), P_o(t,x)\big)$ be the optimum co-state variables for initial phases $(t,x)\in G_o$. Then a feasible control law δ, which permits its trajectory $x(\cdot)$ to move along one of δ_*'s switching boundaries only if it moves optimally, has the following regret function:

$$R_\delta(t_o, x_o) = \int_{t_o}^{t_1}\Big[L\big(t, x(t), u_\delta(t)\big) + P\big(t, x(t)\big) f\big(t, x(t), u_\delta(t)\big) + P_o\big(t, x(t)\big)\Big] dt$$

Proof

By Lemma 1 and Lemma 2

$$R_\delta(t_o, x_o) = \int_{[t_o, t_1] - \omega_*} \Big[L(t, x, u_\delta) + P(t,x)\, f(t, x, u_\delta) + P_o(t,x)\Big] dt$$

For any time interval ω_i comprising ω_*, motion is optimal by hypothesis. Over this interval, the integrand becomes

$$L\big(t, x_*(t), u_\delta(t)\big) + P(t)\, f\big(t, x_*(t), u_*(t)\big) + P_o(t) = 0$$

almost everywhere on ω_i, according to Pontryagin's principle.[2] Hence

$$\int_{\omega^*} [L + Pf + P_o]\, dt = 0.$$

(End of Proof)

3.4 SUFFICIENT CONDITIONS FOR ϵ-OPTIMALITY

The following theorem summarizes the results. It is felt that far more general results apply. However, they have proven to be rather elusive. A summary of conjectures is given in Chapter 6.

Theorem 6

Let $J_*(t,x)$ correspond to a procedure δ_* which can be represented as a control law in $P[G_o]$. Let the procedure δ have a control function u_δ for the initial phase $(t_o, x_o) \in G_o$ which produces a feasible trajectory $\big(\cdot,\ x(\cdot)\big)$ such that J_* is a.e. continuously differentiable along it. Then δ is ϵ-optimal for (t_o, x_o) if:

$$\int_{t_o}^{t_1} \left[L(t,x,u_\delta) + \frac{\partial J_*}{\partial x}(t,x) \cdot f(t,x,u_\delta) + \frac{\partial J_*}{\partial t}(t,x) \right] dt \leq \epsilon$$

(time arguments omitted for simplicity)

Furthermore, if the partials $\partial J_*/\partial x\big(t, x(t)\big)$ and $\partial J_*/\partial t\big(t, x(t)\big)$ are absolutely continuous on $[t_o, t_1]$, Pontryagin's co-state variables $P(t,x)$ and $P_o(t,x)$ may be used in place of the partials. If the trajectory moves optimally through phases where J_* is non-differentiable, then the above condition still applies in terms of co-state variables.

(End of Statement)

The next two chapters will deal with the application of these results to two different types of problems:

- Approximations to known optimal controls.
- Termination conditions for successive approximations to unknown optimal solutions.

CHAPTER 4

APPROXIMATIONS TO LINEAR OPTIMAL CONTROLS

4.0 UTILIZATION OF THE REGRET CRITERION

The criteria of Corollary 1 and Theorem 6 involve the optimum performance function $J_*(t, x)$, and hence could not be used directly for design purposes if J_* were unavailable. One application of the criteria, however, would be in rational approximation of known optimal control laws. It often happens that a known general solution to an optimal control problem is not mechanized by designers, because its sophistication may present some problems of implementation. The criteria provided may be useful in these situations.

Approximations to linear time-varying control laws are considered in this chapter.

4.1 APPROXIMATION CRITERION FOR A CLASS OF PROBLEMS

Let the process $\mathcal{P}$ be given by

$$\frac{dx}{dt} = g(t, x) + B(t)u \qquad (4.1)$$

where $x \in \mathbb{R}^n$, $u \in U = \mathbb{R}^r$, and $g: \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$. Let the performance functional have the form

$$J(t_o, x_o; u) = \int_{t_o}^{t_1} \left[q(t, x) + \ell(t, x)\, u + u^T N(t) u \right] dt \qquad (4.2)$$

where q, ℓ, and N are (1x1), (1xr), and (rxr) matrices, respectively.

In the following it is assumed that δ_* exists in $P[G_o]$ so that J_* has properties which enable the representations of Corollary 1 and Theorem 6.

Theorem 7 [Fixed Time, Quadratic Control Cost]

Let $\delta_*(t,x)$ be the optimum control law for the above problem for fixed time t_1 and $N(t)$ a positive definite (symmetric) (rxr) matrix for each $t \in [t_o, t_1]$. Then a feasible control law δ and its trajectory $\big(t, x(t)\big)$ are ϵ-optimal for (t_o, x_o) if:

$$\int_{t_o}^{t_1} \left\| \delta\big(t, x(t)\big) - \delta_*\big(t, x(t)\big) \right\|_N^2 dt \leq \epsilon$$

[Note: $\|b\|_P^2 \triangleq b^T P b$, where b, P are $(q \times 1)$, $(q \times q)$ matrices respectively.]

Proof

From Theorem 6 we have

$$\int_{t_o}^{t_1} \left[L(t,x,\delta) + \frac{\partial J_*}{\partial x} \bullet f(t,x,\delta) + \frac{\partial J_*}{\partial t} \right] dt \leq \epsilon$$

By adding and subtracting $L(t,x,\delta_*)$ and $f(t,x,\delta_*)$ we can arrange the following:

$$\int_{t_o}^{t_1} \left[L(t,x,\delta_*) + \frac{\partial J_*}{\partial x} \bullet f(t,x,\delta_*) + \frac{\partial J_*}{\partial t} \right] dt + \int_{t_o}^{t_1} \left\{ L(t,x,\delta) - L(t,x,\delta_*) + \frac{\partial J_*}{\partial x} \bullet \Big[f(t,x,\delta) - f(t,x,\delta_*) \Big] \right\} dt \leq \epsilon$$

Since δ_* is optimal and of class $P[G_o]$, the integrand of the first integral is zero along $\big(t, x(t)\big)$ wherever $\partial J_*/\partial x$ and $\partial J_*/\partial t$ are defined [see Corollary 2 and the remarks of Appendix B]. Since they exist a.e. $[t_o, t_1]$ the first integral is zero.

Substituting the definitions of L and f from Equations (4.1) and (4.2) into the remaining integral we obtain:

$$\int_{t_o}^{t_1}\left[\ell(t,x)(\delta-\delta_*) + \delta^T N\delta - \delta_*^T N\delta_* + \frac{\partial J_*}{\partial x} B(\delta-\delta_*)\right] dt \le \epsilon$$

Since δ_* is optimal we also have

$$\frac{\partial}{\partial u}\left[L(t,x,u) + \frac{\partial J_*}{\partial x} f(t,x,u) + \frac{\partial J_*}{\partial t}\right]\Bigg|_{u=\delta_*} = 0$$

or

$$\ell(t,x) + 2\delta_*^T N + \frac{\partial J_*}{\partial x} \dot{B} = 0$$

Substituting $\frac{\partial J_*}{\partial x}$ B from this expression, we readily obtain

$$\int_{t_o}^{t_1}\left[\delta^T N\delta - \delta_*^T N\delta_* - 2\delta_*^T N(\delta-\delta_*)\right] dt \le \epsilon ,$$

and the result follows easily.

(End of Proof)

<u>Remark</u>

The criterion involves only the trajectory $\left(t, x(t)\right)$ produced by δ. Thus, it involves the control function $\delta\left(t, x(t)\right)$ and the time function $\delta_*\left(t, x(t)\right)$. The latter should not be confused with $\delta_*\left(t, x_*(t)\right) = u_*(t)$ which is the optimum control function from the initial phase (t_o, x_o).

4.2 APPLICATION TO LINEAR TIME-VARYING CONTROLS

In the event that the process equations take the form

$$\frac{dx}{dt} = A(t)x + B(t)u \tag{4.3}$$

and the performance functional is of the form

$$J(t_o, x_o; u) = \int_{t_o}^{t_1} \left[x^T Q(t) x + 2x^T L(t) u + u^T N(t) u \right] dt \qquad (4.4)$$

it is known that δ_* has the form:[9,17,22]

$$\delta_*(t, x) = -N^{-1}(t) \left[B^T(t) M(t) + L^T(t) \right] x \qquad (4.5)$$

in which M(t) satisfies the matrix Riccati differential equation

$$\frac{dM}{dt} = -(Q - LN^{-1}L^T) - M(A - BN^{-1}L^T) - (A - BN^{-1}L^T)^T M^T + MBN^{-1}B^T M^T \qquad (4.6)$$

with boundary condition $M(t_1) = \Phi$. Since Q and N may be taken as symmetric, M may also be regarded as symmetric.

With the control law of (4.5) the optimum performance function assumes the form

$$J_*(t, x) = x^T M(t) x \qquad (4.7)$$

Instead of determining M(t) from (4.6) one may regard M(t) as the transformation matrix between the state and co-state vectors.

$$P^T(t) = M(t)\, x(t) \qquad (4.8)$$

This relation may be obtained by a process which involves the solution of the system of 2n first-order differential equations

$$\frac{dz}{dt} = \begin{bmatrix} A - BN^{-1}L^T, & -BN^{-1}B \\ -Q + LN^{-1}L^T, & -A^T + LN^{-1}B^T \end{bmatrix} z \qquad (4.9)$$

where z is the combined state — co-state vector

$$z \triangleq \begin{bmatrix} x \\ p^T \end{bmatrix} \tag{4.10}$$

subject to boundary conditions of Pontryagin's formulation.

In any case, Equation (4.5) shows that, aside from $N^{-1}(t)$ which is subject to the designer's definition, the feedback gain matrix may require approximation due to its general dependence on time. How should this be done without sacrificing proximity to optimality?

Suboptimum Designs

Suppose we choose to use a suboptimum form

$$\delta(t, x) = -N^{-1}(t)\, K(t)x \tag{4.11}$$

where $K(t)$ is an $(r \times n)$ matrix serving to approximate $B^T(t)M(t) + L^T(t)$. Then by Theorem 7

$$\int_{t_o}^{t_1} \left\| N^{-1}(t) \left[B^T(t)\, M(t) + L^T(t) - K(t) \right] x(t) \right\|_N^2 dt \le \epsilon$$

or

$$\int_{t_o}^{t_1} \left\| \left[B^T(t)\, M(t) + L^T(t) - K(t) \right] x(t) \right\|_{N^{-1}}^2 dt \le \epsilon \tag{4.12}$$

For specific problems we may wish to proceed in different ways from this juncture.

Example 14

Suppose that $B(t)$ and $L(t)$ are constant matrices or else simple enough in form so that in Equations (4.11) and (4.12) we choose the approximating form

$$K(t) = B^T(t)\, \tilde{M}(t) + L^T(t)$$

Then for ϵ-optimality we require

$$\int_{t_o}^{t_1} \| [M(t) - \tilde{M}(t)]\, x(t) \|^2_{BN^{-1}B^T} \, dt \leq \epsilon$$

If we define the (weighted) norm of any (n x m) matrix A with respect to the (n x n) weight matrix W by

$$\| A \|^2_W \triangleq \text{trace}\, [A^T W A]$$

then we obtain the inequality

$$\| [M(t) - \tilde{M}(t)] x(t) \|^2_{BN^{-1}B^T} \leq \| M(t) - \tilde{M}(t) \|^2_{BN^{-1}B^T} \cdot \| x(t) \|^2$$

Thus a sufficient condition for ϵ-optimality is

$$\int_{t_o}^{t_1} \| M(t) - \tilde{M}(t) \|^2_{BN^{-1}B^T} \, dt \leq \frac{\epsilon}{\sup\limits_{t_o \leq t \leq t_1} \| x(t) \|^2}$$

This criterion may be used to approximate M(t) if it is known. Alternatively, if M(t) were to be mechanized but must be found by numerical integration of (4.6), then this criterion could be used to determine the tolerable error in the numerical procedure.

It should be pointed out that in many cases we have that $\sup\limits_t \| x(t) \|^2 = \| x_o \|^2$ and that ϵ is frequently acceptable as some small proportion of $J_*(t_o, x_o)$. That is, from (4.7)

$$\epsilon \triangleq x_o^T M(t_o) x_o \; \epsilon_p$$

where ϵ_p is a small ($0 < \epsilon_p << 1$) proportionality constant. Then we would have

$$\int_{t_o}^{t_1} \| M(t) - \widetilde{M}(t) \|^2_{BN^{-1}B^T} dt \le \epsilon_p \left(\frac{x_o^T M(t_o) x_o}{x_o^T x_o} \right)$$

as the corresponding criterion.

<u>Example 15</u> <u>(Merriam,[9] pp. 97-99)</u>

This example is chosen to illustrate some of the ideas above. The optimum control law turns out to be simple enough since the problem is rather simple. However, it would be instructive to see how one would approximate it. Consider the scalar process

$$\frac{dy}{dt} = u(t) \qquad 0 \le t \le T$$

where it is desired that $y(t)$ be regulated wrt a reference value Y so that the following is minimized

$$J(o, x_o; u) = \int_o^T \left[\left(y(t) - Y \right)^2 + \frac{1}{\omega^2} u^2(t) \right] dt$$

In this last equation Y and ω are constants set forth by the designer. By defining $x(t) \triangleq y(t) - Y$, we see that $Q = 1$, $N = 1/\omega^2$, $B = 1$, $A = 0$, and $L = 0$ in Equations (4.3) and (4.4) with $t_o = 0$ and $t_1 = T$. Thus, Equations (4.5) and (4.6) yield

$$\delta_* = -\omega^2 m(t) x, \text{ where}$$

$$\frac{dm}{dt} = -1 + \omega^2 m^2(t) \, ; \quad m(T) = 0$$

The differential equation is readily integrated[23] to yield

$$m(t) = \frac{1}{\omega} \tanh \omega(T-t)$$

Thus,

$$\delta_*(t, x) = -x\omega \tanh \omega(T-t)$$

and from (4.7)

$$J_*(t,x) = \frac{x^2}{\omega} \tanh \omega(T-t)$$

We note that the feedback gain is time-varying even though the process and performance functionals were time-invariant. This is due to the finite control interval.

Let us now consider an ϵ-optimal law

$$\delta(t,x) = -x\omega\, g(t)$$

where we specify ϵ as

$$\epsilon = \epsilon_p\, J_*(o, x_o) = \epsilon_p \frac{x_o^2}{\omega} \tanh \omega T$$

From Theorem 7 or Equation (4.12)

$$\int_o^T \left(g(t) - \tanh \omega(T-t)\right)^2 x^2(t)dt \leq \epsilon_p \frac{x_o^2 \tanh \omega T}{\omega}$$

With the control law $\delta(t,x)$ it is easily verified that

$$x(t) = e^{-\int_o^t g(t)\omega\, dt} x_o$$

Thus, for ϵ-optimality we must have

$$\int_o^T e^{-2\omega \int_o^t g(t)dt} \cdot \left(g(t) - \tanh \omega(T-t)\right)^2 dt \leq \epsilon_p \frac{\tanh \omega T}{\omega}$$

This is assured for $g(t) \geq 0$ if

$$\int_o^T \left(g(t) - \tanh \omega(T-t)\right)^2 dt \leq \epsilon_p \frac{\tanh \omega T}{\omega}$$

With a change of variables this becomes

$$\int_0^{\omega T}\Big(h(x) - \tanh x\Big)^2 dx \le \epsilon_p \tanh \omega t$$

where $h\Big(\omega(T-t)\Big) = g(t)$.

We may use orthogonal polynomials to carry out our approximation from this point on. However, if we have either of the asymptotic cases

(i) $\omega T >> 1$, or

(ii) $\omega T << 1$,

then it is easy to see that the following simple gains are possible

(i) $g(t) = \text{constant} = \dfrac{1}{\omega T}\displaystyle\int_0^{\omega T} \tanh x\, dx$, if

$$\epsilon_p \ge \int_0^{\omega T} \tanh^2 x\, dx - (\omega T) g^2$$

or,

(ii) $g(t) = \text{linear} = \omega(T-t)$, if $\epsilon_p \ge \dfrac{(\omega T)^6}{9 \times 7}$

Thus, over long control periods a simple constant can be used as a feedback gain. For relatively short control periods, a linear time-varying gain would suffice. For moderate control intervals a composite of the two would be appropriate. These approximations are of course, the type of approximations which a designer would have intuitively employed. The approximation criterion, however, provides the rational basis for such ideas.

CHAPTER 5

SUBOPTIMAL CONTROL SEQUENCES

5.0 SUCCESSIVE APPROXIMATIONS

In this chapter we shall consider suboptimal control sequences generated by successive approximations. Monotone approximation techniques are given for both the Bellman and Pontryagin conditions. Finally, we shall consider termination criteria for these methods. The following assumptions are made:

- A1. Fixed final time, free right-end problems. (Terminal constraints approximated by final value loss considerations.)
- A2. Existence of optimal solutions.
- A3. Unique solutions to the necessary conditions of Bellman and Pontryagin.

These considerations have already been discussed in Chapter 2.

5.1 CONTROL LAW SEQUENCES

The following technique is an extension of the work of Leake and Liu[10] who carried through Bellman's idea of approximation in policy space. The extension consists of allowing the successive control laws to be of class $P[G_o]$ rather than requiring them to be continuously differentiable throughout G_o.

Construction of the Sequence

Let $\delta_\alpha(t,x)$ be of class $P[G_o]$ and $J_\alpha(t,x) \triangleq J(t,x;\delta_\alpha)$ be its corresponding performance function obtained whether by direct calculation or by solution of the linear partial differential equation

$$\frac{\partial J_\alpha}{\partial t} + \frac{\partial J_\alpha}{\partial x} \bullet f(t, x, \delta_\alpha) + L(t, x, \delta_\alpha) = 0 \tag{5.1}$$

with boundary condition $J_\alpha(t_1, x) = 0$ (all $x \in \mathbb{R}^n$).

Having found $J_\alpha(t, x)$, another control law $\delta_{\alpha+1}(t, x)$ is generated by

$$\delta_{\alpha+1}(t, x) = k\left(t, x, \frac{\partial J_\alpha}{\partial x}(t, x)\right) \tag{5.2}$$

where

$$L(t, x, k) + \frac{\partial J_\alpha}{\partial x} \bullet f(t, x, k) = \inf_{u \in U}\left[L(t, x, u) + \frac{\partial J_\alpha}{\partial x} \bullet f(t, x, u)\right] \tag{5.3}$$

Since $\delta_\alpha \in P[G_o]$, $\partial J_\alpha/\partial x$ may not be defined at points (t, x) along a boundary of a region G_k for δ_α. Thus $\delta_{\alpha+1}$ in (5.2) may not be defined everywhere in G_o. We may extend its domain of definition as follows. We consider the statement of (5.1) as an identity defined everywhere in G_o. Hence $\partial J_\alpha/\partial t$ and $\partial J_\alpha/\partial x$ may be regarded as limits as (t, x) approach the partition boundary $\bar{G}_k \cap \bar{G}_\ell$ from either G_k^o or G_ℓ^o. We shall assume that finite limits exist everywhere along the boundary using interior points of either G_k or G_ℓ. Thus $\partial J_\alpha/\partial t$ and $\partial J_\alpha/\partial x$ are assumed to be regularized so that they are well defined mappings on G_o. With this extension, $\delta_{\alpha+1}$ is likewise defined everywhere in G_o.

Lemma 3

The control law $\delta_{\alpha+1}$ is uniformly as good as δ_α. Further if δ_α is not optimal, then $\delta_{\alpha+1}$ is better than δ_α for some $(t_o, x_o) \in G_o$.

Proof

Consider $\delta_{\alpha+1}$'s regret relative to δ_α

$$R(t_o, x_o; \delta_{\alpha+1}, \delta_\alpha) = \int_{t_o}^{t_1} \left[L(t, x_{\alpha+1}, \delta_{\alpha+1}) + \frac{\partial J_\alpha}{\partial x} \bullet f(t, x_{\alpha+1}, \delta_{\alpha+1}) + \frac{\partial J_\alpha}{\partial t} \right] dt$$

$$\leq \int_{t_o}^{t_1} \left[L(t, x_{\alpha+1}, \delta_\alpha) + \frac{\partial J_\alpha}{\partial x} \bullet f(t, x_{\alpha+1}, \delta_\alpha) + \frac{\partial J_\alpha}{\partial t} \right] dt,$$

where the inequality arises from (5.2) and (5.3). However, the integrand in this last integral is zero by virtue of (5.1). Thus, $\delta_{\alpha+1}$ is uniformly as good as δ_α since its relative regret is nonpositive for all $(t_o, x_o) \in G_o$.

We prove the second half of the lemma by a contrapositive argument. Suppose there were no phase $(t_o, x_o) \in G_o$ for which strict inequality holds above. Then $R(t_o, x_o; \delta_{\alpha+1}, \delta_\alpha) \equiv 0$ for all $(t_o, x_o) \in G_o$. This implies $J_{\alpha+1} \equiv J_\alpha$, which in turn implies

$$\left[x_{\alpha+1}(t),\ \delta_{\alpha+1}\left(t, x_{\alpha+1}(t)\right) \right] = \left[x_\alpha(t),\ \delta_\alpha\left(t, x_\alpha(t)\right) \right] \text{ a.e. } [t_o, t_1]$$

Thus, we have

$$\inf_{u \in U} \left\{ L(t, x_\alpha, u) + \frac{\partial J_\alpha}{\partial x} \bullet f(t, x_\alpha, u) + \frac{\partial J_\alpha}{\partial t} \right\} = L(t, x_\alpha, \delta_\alpha) + \frac{\partial J_\alpha}{\partial x} \bullet f(t, x_\alpha, \delta_\alpha) + \frac{\partial J_\alpha}{\partial t}$$

almost everywhere along every trajectory $(t, x_\alpha(t))$ produced by δ_α. Since the right side of this equation is zero by virtue of (5.1), then J_α satisfies the necessary condition for optimality given in Corollary 2, Chapter 3 (see the remark following the proof of Theorem A in Appendix B as to why Corollary 2 holds also for optimum control laws of class $P[G_o]$). Because of assumptions (A.2) and (A.3) of

Section 5.0, J_α and δ_α must correspond to the optimal solution J_* and δ_*.

(End of Proof)

Definition

A control law δ is said to be better than a control law δ' if its regret function $R_\delta(\theta)$ satisfies

$$R_\delta(\theta) \le R_{\delta'}(\theta) \quad \text{for all } \theta \epsilon G_o$$

and there is one $\theta \epsilon G_o$ for which strict inequality holds.

Theorem 8

Suppose each successive control law $\delta_{\alpha+1}$ is of class $P[G_o]$. Then the sequence $< \delta_\alpha >$ either converges to the optimum law in a finite number of steps or is monotonically better.

Proof

From lemma 3

$$R(\theta; \delta_{\alpha+1}, \delta_\alpha) \le 0 \quad \text{for all } \theta \epsilon G_o$$

and if $\delta_\alpha \neq \delta_*$ then strict inequality holds for some $\theta \epsilon G_o$.

$$\begin{aligned} R_{\delta_{\alpha+1}}(\theta) &= J(\theta; \delta_{\alpha+1}) - J_*(\theta) = J(\theta; \delta_{\alpha+1}) - J(\theta; \delta_\alpha) \\ &\quad + J(\theta; \delta_\alpha) - J_*(\theta) = R(\theta; \delta_{\alpha+1}, \delta_\delta) + R_{\delta_\alpha}(\theta) \end{aligned}$$

Thus

$$R_{\delta_{\alpha+1}}(\theta) \le R_{\delta_\alpha}(\theta) \quad \text{for all } \theta \epsilon G_o$$

and strict inequality holds for some $\theta \epsilon G_o$ if $\delta_\alpha \neq \delta_*$. Therefore, if there is no finite α for which $\delta_\alpha = \delta_*$, then $< \delta_\alpha >$ is a sequence of monotonically better control laws.

Remarks

The important fact about the sequence is that it is monotonically better whether it converges or not.

We have not given conditions which would guarantee that each successive law is of class $P[G_o]$. In many practical problems this will turn out to be the case a posteriori. Sufficient conditions which come immediately to mind, however, are

1. Use of control laws which are piecewise analytic and having partition boundaries defined by analytic functions, and

2. The bounded control set is defined by an analytic function on $\mathbb{R}^r$.

These conditions are perhaps much too strong for many problems of interest, however.

5.2 CONTROL FUNCTION SEQUENCES

In many instances in practice, an optimal solution for a specific initial phase (t_o, x_o) is desired rather than for an entire set of initial phases. In other words an optimal or near-optimal control function $u(\cdot)$ is sought over an interval $[t_o, t_1]$ for a specific initial state x_o. Whether or not the designer implements this solution as an open-loop time function $u(\cdot)$ or a control law δ_u which generates $u(\cdot)$ depends on the problem's external considerations.

A method is given below which yields a sequence of monotonically better control solutions for a given initial phase (t_o, x_o). The problem of synthesis of the feedback control laws corresponding to these control functions is solved simultaneously, since control laws are inherent in the method.

Construction of the Sequence

Let $u_\alpha(t)$ be any admissible control function for the initial phase (t_o, x_o). From this we can obtain some control law $\delta_{\alpha+1}$ as follows:

Consider the artifice of regarding $u_\alpha(t)$ as a control law:

$$\delta_\alpha(t,x) \triangleq u_\alpha(t) \,, \quad t \in [t_o, t_1] \tag{5.4}$$

The performance function J_α for this control law is

$$J_\alpha(t,x) \triangleq J(t,x;u_\alpha) = \int_t^{t_1} L\big(\tau, x_\alpha(\tau), u_\alpha(\tau)\big)\, d\tau \tag{5.5}$$

where $\big(\tau, x(\tau)\big)$ is the trajectory produced by u_α starting from the initial phase (t,x) [Note: (t,x) need not be the initial phase (t_o, x_o) of the problem].

As in Section 5.1 $\delta_{\alpha+1}$ is chosen so that

$$\delta_{\alpha+1}(t,x) = k\left(t, x, \frac{\partial J_\alpha}{\partial x}\right)$$

where

$$L(t,x,k) + \frac{\partial J_\alpha}{\partial x} \bullet f(t,x,k) = \inf_{u \in U} \left[L(t,x,u) + \frac{\partial J_\alpha}{\partial x} \cdot f(t,x,u)\right]$$

Finally $u_{\alpha+1}(t)$ is generated by integrating the process equations from (t_o, x_o), using the control law $\delta_{\alpha+1}$, and setting

$$u_{\alpha+1}(t) \triangleq \delta_{\alpha+1}\big(t, x_{\alpha+1}(t)\big) \tag{5.6}$$

Remark

Since only $\partial J_\alpha / \partial x$ is involved in the choice of $\delta_{\alpha+1}$, it is possible to obtain $\delta_{\alpha+1}$ without first solving for J_α either directly from (5.5) or as a solution to the partial differential equation (5.1).

From (5.5) we have

$$\frac{\partial J_\alpha}{\partial x}(t,x) = \int_t^{t_1} \frac{\partial L}{\partial x}\Big(\tau, x_\alpha(\tau), u_\alpha(\tau)\Big) \bullet \frac{\partial x_\alpha(\tau)}{\partial x}\, d\tau \tag{5.7}$$

where $\partial x_\alpha(\tau)/\partial x$ is the fundamental matrix for the process

$$\frac{dx(\tau)}{d\tau} = f\Big(\tau, x(\tau),\ u_\alpha(\tau)\Big) \quad \text{a.e. } [t, t_1]$$

with initial condition $x(t) = x$.

Relation (5.7) would be particularly useful if the process equations were linear, since the fundamental matrix would be independent of x and u_α, and $\partial J_\alpha/\partial x$ would depend on x only through $\partial L/\partial x$.

<u>Theorem 9</u>

Let $u_o(\cdot)$ be an admissible control function for (t_o, x_o) and let each successive control function obtained by continuing the above process be admissible. Then the sequence of control functions $< u_\alpha >$ generated by $< \delta_\alpha >$ either converges a.e. $[t_o, t_1]$ to the optimal control u_* for (t_o, x_o) in a finite number of steps or is monotonically better for (t_o, x_o).

<u>Proof</u>

Because each u_α is admissible, $J_\alpha(t,x)$ defined by (5.5) is seen to be continuously differentiable wrt $(t,x) \in G_o$. Thus, Theorem 4, Chapter 3, allows us to express the relative regret of $\delta_{\alpha+1}$ wrt $\delta_\alpha \triangleq u_\alpha(t)$ as

$$R(t_o, x_o; \delta_{\alpha+1}, \delta_\alpha) = \int_{t_o}^{t_1}\left[L(t, x, \delta_{\alpha+1}) + \frac{\partial J_\alpha}{\partial x} \bullet f(t, x, \delta_{\alpha+1}) + \frac{\partial J_\alpha}{\partial t}\right]dt$$

$$\leq \int_{t_o}^{t_1}\left[L(t, x, \delta_\alpha) + \frac{\partial J_\alpha}{\partial x} \bullet f(t, x, \delta_\alpha) + \frac{\partial J_\alpha}{\partial t}\right]dt$$

But this last integral is zero for the same reason as in the proof of lemma 3, Section 5.1. In fact we have

$$L(t, x, \delta_{\alpha+1}) + \frac{\partial J_\alpha}{\partial x} \bullet f(t, x, \delta_{\alpha+1}) + \frac{\partial J_\alpha}{\partial t} \leq 0, \text{ all } (t, x) \in G_o$$

If equality holds for the relative regret expressions above, then

$$L(t, x_{\alpha+1}, \delta_{\alpha+1}) + \frac{\partial J_\alpha}{\partial x} \bullet f(t, x_{\alpha+1}, \delta_{\alpha+1}) + \frac{\partial J_\alpha}{\partial t} = 0$$

a.e. along $\left(t, x_{\alpha+1}(t)\right)$.

But it is also true that

$$L(t, x_{\alpha+1}, \delta_\alpha) + \frac{\partial J_\alpha}{\partial x} \cdot f(t, x_{\alpha+1}, \delta_\alpha) + \frac{\partial J_\alpha}{\partial t} = 0$$

everywhere along $\left(t, x_{\alpha+1}, (t)\right)$ since this is an identity for every $(t, x) \in G_o$.

This implies [see remark following proof for an expansion of this point] that $\delta_{\alpha+1}\left(t, x_\alpha(t)\right) = \delta_\alpha \triangleq u_\alpha(t)$ a.e. $[t_o, t_1]$. In other words $x_{\alpha+1}(t) = x_\alpha(t)$, and hence

$$\inf_{u \in U}\left[L(t, x_\alpha, u) + \frac{\partial J_\alpha}{\partial x} \bullet f(t, x_\alpha, u) + \frac{\partial J_\alpha}{\partial t}\right] = 0 \quad \text{a.e. } [t_o, t_1].$$

If it can be established that $\partial J_\alpha / \partial t$ and $\partial J_\alpha / \partial x$ are the co-state variables along $\left(t, x_\alpha(t)\right)$, then this last equation implies that $u_\alpha = u_*$ a.e. $[t_o, t_1]$ due our assumption of uniqueness of solution to

Pontryagin's necessary conditions. Lemma A of Appendix B establishes the identification of the partials with co-state variables.

We have thus shown that if the relative regret of $\delta_{\alpha+1}$ wrt δ_α is zero for (t_o, x_o), then $\delta_\alpha \triangleq u_\alpha(t)$ is equal a.e. $[t_o, t_1]$ to the optimal control function for (t_o, x_o). A contrapositive argument establishes that $u_{\alpha+1}(t) \triangleq \delta_{\alpha+1}\left(t, x_{\delta+1}(t)\right)$ is better for (t_o, x_o) if u_α is not optimal. Thus, our sequence $< u_\alpha >$ is monotonically better for (t_o, x_o), and if for some finite α this is not so, then the sequence has converged a.e. $[t_o, t_1]$ to u_*.

(End of Proof)

<u>Remark</u>

The crucial step in the above proof was the equating of $\delta_{\alpha+1}$ to δ_α a.e. along $\left(t, x_{\alpha+1}(t)\right)$ because $\delta_{\alpha+1}$ and δ_α both satisfy

$$L(t, x_{\alpha+1}, \delta) + \frac{\partial J_\alpha}{\partial x} \bullet f(t, x_{\alpha+1}, \delta) + \frac{\partial J_\alpha}{\partial t} = 0$$

almost everywhere along $\left(t, x_{\alpha+1}(t)\right)$. It is claimed that if this is not so then the control problem has been ill-posed. Consider the control law δ_λ defined by

$$\delta_\lambda(t, x) = \lambda \delta_{\alpha+1}(t, x) + (1-\lambda)\, \delta_\alpha(t, x), \quad 0 \le \lambda \le 1$$

This will be an admissible control law with the property

$$L(t, x, \delta_{\alpha+1}) + \frac{\partial J_\alpha}{\partial x} \bullet f(t, x, \delta_{\alpha+1}) + \frac{\partial J_\alpha}{\partial t} \le L(t, x, \delta_\lambda) + \frac{\partial J_\alpha}{\partial x} \bullet f(t, x, \delta_\lambda) + \frac{\partial J_\alpha}{\partial t} \le 0$$

for all $(t, x) \in G_o$. Thus, along $\left(t, x_{\alpha+1}(t)\right)$

$$L(t, x_{\alpha+1}, \delta_\lambda) + \frac{\partial J_\alpha}{\partial x} \bullet f(t, x_{\alpha+1} \delta_\lambda) + \frac{\partial J_\alpha}{\partial t} = 0$$

almost everywhere. If $\delta_{\alpha+1}$ and δ_α were different along the trajectory over a nonzero measure of time, then a nondenumerable set of phases $(t, x_{\alpha+1}(t))$ would exist for which $\delta_{\alpha+1} \triangleq k(t, x, \partial J_\alpha/\partial x)$ cannot be uniquely defined. This is because δ_λ would be equally as effective in the minimization process involved in deriving the function k. Moreover this would be true for all $\lambda \in [0, 1]$. A situation such as this will arise if $L(t, x, u)$ and/or $f(t, x, u)$ are ill-defined or if one or more components of the control vector have no influence on the process behavior. We, of course, assume that the problem has been posed properly so that k is well defined, except possibly on a set of phases which is at most denumerable.

(End of Remark)

A very important by-product of Theorem 9 is the fact that the procedure of this section allows feedback control synthesis of any optimal control function. It is not always true that the feedback configuration will be simpler to implement than merely storing $u_*(t)$ in a suitable memory unit. However this may be the case in certain problems. The following corollary would be of value in such problems.

Corollary 3 [Synthesis of Optimal Controls]

Let $u_*(t)$ be optimal for (t_o, x_o), then

$$\delta(t, x) = k\left(t, x, \frac{\partial J}{\partial x}\right)$$

where

$$J(t, x; u_*) \triangleq \int_t^{t_1} L\big(\tau, x(\tau), u_*(\tau)\big)\, d\tau$$

is a feedback realization which is optimal for (t_o, x_o).

5.3 TERMINATION CRITERIA

The final theorem of this dissertation is concerned with a criterion which may be used to terminate the sequences. This criterion guarantees ϵ-optimality.

Theorem 10

Let $g(t,x)$ be a non-negative function from G_o into $\mathbb{R}$ such that

$$\int_{t_o}^{t_1} g\big(t, x(t)\big)\, dt \le \epsilon$$

for all feasible trajectories. Then the sequences $< \delta_\alpha >$ of Theorems 8 and 9 consist of ϵ-optimal laws for all $\alpha \ge N$ if

$$-g(t,x) \le \inf_{u \in U} \left[L(t,x,u) + \frac{\partial J_N}{\partial x} \bullet f(t,x,u) + \frac{\partial J_N}{\partial t} \right]$$

for all $(t,x) \in G_o$.

Proof

Since for all $(t,x) \in G_o$

$$\inf_{u \in U} \left[L(t,x,u) + \frac{\partial J_N}{\partial x} \bullet f(t,x,u) + \frac{\partial J_N}{\partial t} \right] \le L(t,x,\delta_*) + \frac{\partial J_N}{\partial x} \bullet f(t,x,\delta_*) + \frac{\partial J_N}{\partial t}$$

then

$$-g(t,x) \le L(t,x,\delta_*) + \frac{\partial J_N}{\partial x} \bullet f(t,x,\delta_*) + \frac{\partial J_N}{\partial t}$$

Therefore, along the optimal trajectory $\big(t, x_*(t)\big)$

$$-\int_{t_o}^{t_1} g(t, x_*)dt \le \int_{t_o}^{t_1} \left[L(t, x_*, \delta_*) + \frac{\partial J_N}{\partial x} \bullet f(t, x_*, \delta_*) + \frac{\partial J_N}{\partial t} \right] dt$$

The integral on the left is greater than $(-\epsilon)$ by hypothesis. The integral on the right is the regret for δ_* with respect to δ_N or, alternatively, the negative of δ_N's regret function $R_{\delta_N}(t_o, x_o)$. Thus,

$$R_{\delta_N}(t_o, x_o) \leq \epsilon \quad \text{for all } (t_o, x_o) \epsilon G_o$$

By the monotonicity of $< \delta_\alpha >$, all successors to δ_N are also ϵ-optimal.

(End of Proof)

The simplest choice of g for uniform ϵ-optimality is

$$g(t, x) = \frac{\epsilon}{t_1 - t_o}$$

However, if the designer has some rough idea of how the optimal trajectories will behave, then other non-trivial choices for g may be appropriate for reducing the number of iterations.

If $L(t, x, u)$ is non-negative for all $(t, x) \epsilon G_o$ and $u \epsilon U$, and it is desired that

$$R_{\delta_N}(t_o, x_o) \leq \epsilon_o + \epsilon_p J_*(t_o, x_o) \tag{5.8}$$

where ϵ_o and ϵ_p are small positive constants, then g can be selected as

$$g(t, x) = \frac{\epsilon_o}{t_1 - t_o} + \epsilon_p \inf_{u \epsilon U} L(t, x, u) \tag{5.9}$$

This follows directly from the fact that

$$\int_{t_o}^{t_1} \inf_{u \epsilon U} L\left(t, x_*(t), u\right) dt \leq \int_{t_o}^{t_1} L\left[t, x_*(t), \delta_*\left(t, x_*(t)\right)\right] dt$$

CHAPTER 6

CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE STUDIES

6.0 CONCLUSIONS

The major contribution of this dissertation is the introduction of the regret function and its integral representation based on Weierstrass ideas.[26,27] The most direct application of this representation is the rational approximation of known optimal control policies. This was illustrated in Chapter 4 for linear, time-varying optimal controls (Theorem 7).

In Chapter 5 we utilized the ideas of Chapter 3 to solve the problem of suboptimal designs when optimal solutions to fixed time, free right end problems were not known a priori, as in the case of Chapter 4. Two methods of successive approximations were given, corresponding to the control function and control law approaches of Pontryagin and Bellman, respectively. Theorem 8 represents an extension of the work of Leake and Liu for continuously differentiable control laws. The extension to piecewise continuously differentiable laws is considered significant, since problems with bounded control sets are likely to result in laws of class $P[G_o]$. The iterative method for control functions (Section 5.2) is new insofar as this investigator knows. The feedback synthesis given in Corollary 3 is apparently also new.

Finally, the termination criterion given in Theorem 10 is felt to be an important contribution to the design of ϵ-optimal controls. We may summarize this in terms of the following corollary to Theorem 10.

<u>Corollary 4</u> [Fixed time, free right end problem]

Let $g_\epsilon(t, x)$ be a non-negative function defined on G_o such that

$$\int_{t_o}^{t_1} g_\epsilon(t, x)\, dt \le \epsilon$$

for all feasible trajectories $\big(t, x(t)\big)$ in G_o. A feasible control law $\delta \in P[G_o]$, with performance function $J(t, x)$, is ϵ-optimal if

$$\inf_{u \in U} \left[L(t, x, u) + \frac{\partial J}{\partial x}(t, x) \cdot f(t, x, u) + \frac{\partial J}{\partial t}(t, x) \right] \ge - g_\epsilon(t, x)$$

(End of Statement)

This condition does not involve knowledge of the optimal solution (δ_*, J_*), although it is based very much on properties which the optimal solution must have. Corollary 4, based on Theorem 10, is the ϵ-optimal extension of Corollary 2, based on Theorem 4. In the case of Corollary 2, the condition given is sufficient for <u>optimality</u> if an optimal solution exists and is a unique solution to the condition given there. In the case of Corollary 4, a sufficient condition for <u>ϵ-optimality</u> is given under the same assumptions of existence and uniqueness of solution to the necessary conditions for optimality.

The above corollary is useful in design problems in the following sense. At the onset of a control problem the designer will usually have a control scheme in mind, which he knows will work, and which, if not optimal, will be close to being so. If physical reasoning leads to the conclusion that an optimal solution exists and is the unique extremum, then Corollary 4 allows him to verify the merits of his control scheme. The methods of Chapter 5 may be employed to iterate on his initial choice, if its performance is felt to be in need of improvement.

6.1 RECOMMENDATIONS FOR FUTURE STUDIES

In Chapter 3 we sought an integral representation for regret in terms of partials of the optimal performance function $J_*(t, x)$. In order to proceed from the class of continuously differentiable control laws, a class $P[G_o]$ of control laws was hypothesized and the optimal law was assumed to be in such a class. Is the representation valid for a more general class? The problem here is twofold. First, the regret can be given an integral representation only if $J_*(t, x)$ is absolutely continuous for all trajectories $\big(t, x(t)\big)$ produced by the feasible control in question. Second, even if the regret is expressible as an integral over $[t_o, t_1]$, does the integrand have the form

$$L\big(t, x(t),\ u_\delta(t)\big) + \frac{\partial J_*}{\partial x} \cdot f\big(t, x(t),\ u_\delta(t)\big) + \frac{\partial J_*}{\partial t} \quad ?$$

It would appear that a form such as this would be desirable, since it relates directly to the Hamiltonian conditions of Bellman and Pontryagin. We have succeeded in showing that, under certain conditions, the form above holds in terms of the co-state variables even if $\partial J_*/\partial x$ and $\partial J_*/\partial t$ were undefined over a positive measure of time along a trajectory. The conditions were that $\delta_* \in P[G_o]$ and that δ cause the state to move optimally if it moves at all, for a positive measure of time, through states for which the partials are undefined. Is this latter condition necessary? In other words, will it hold in terms of co-state variables regardless of how the state moves along boundaries of the regions $G_k \subset G_o$? (In studying the pertinent examples of Reference 2 this investigator found that in many cases a feasible trajectory could not move <u>along</u> a boundary unless the boundary were a manifold of optimal trajectories.)

A question which has not been settled by this dissertation, but whose answer has been long suspected by researchers is the following. Are the co-state variables $P(t_o, x_o)$ and $P_o(t_o, x_o)$ limits of $\partial J_* / \partial t$ as $(t, x) \rightarrow (t_o, x_o)$ in some appropriate way?

Extension of the Research to Statistical Systems

We have not touched on the matter of control schemes for statistical processes or processes which are not perfectly observable. That is, what are sufficient conditions for ϵ-optimality of a control law using an estimate of the true phase (t, x) which is corrupted by observation noise? In addition, how would the solution be affected if the formulation also included random control execution errors which depend on the control decisions?

For the case of linear systems with normally distributed observation and control errors, optimal solutions are known. These turn out to be rather complex if the error processes have non-trivial covariance matrices. The approximation criterion of Theorem 7 may be useful if it is appropriately extended to the statistical case.

BIBLIOGRAPHY

1. Zadeh, L.A. and C.A. Desoer, Linear System Theory, McGraw-Hill Book Co., New York (1963).

2. Pontryagin, L.S., V.G. Boltyanskii, R.V. Gamkrelidze, and E.F. Mishchenko, The Mathematical Theory of Optimal Processes, Interscience, New York (1962).

3. Flippov, A.F., "On Certain Questions in the Theory of Optimal Control," J. SIAM Control (A), 1 (1962), 76-84.

4. Lee, E.B. and L. Markus, "Optimal Control of Nonlinear Processes," Arch. Rational Mech. Anal., 8 (1961), 36-58.

5. Roxin, E., "The Existence of Optimal Control," Michigan Math. J., 9 (1962), 109-119.

6. Neustadt, L.W., "The Existence of Optimal Controls in the Absence of Convexity Conditions," J. Math. Anal. Appl. 7 (1963), 110-117.

7. Cesari, L., "Existence Theorems for Optimal Solutions in Pontryagin and Legrange Problems," J. SIAM Control (A), 3 (1966), 475-498.

8. Stoddart, A.W.J., "Existence of Optimal Controls," Pacific J. Math., 20 (1967), 167-177.

9. Merriam, C.W., III, Optimization Theory and the Design of Feedback Control Systems, McGraw-Hill Book Co., New York (1964).

10. Leake, R.J. and R.W. Liu, "Construction of Suboptimal Control Sequences," J. SIAM Control (A), 5 (1967), 54-63.

11. Courant, R., Calculus of Variations and Supplementary Notes and Exercises (mimeographed lecture notes), New York University Institute of Mathematical Sciences, New York (1956).

12. Kelley, H.J., "Method of Gradients," Optimization Techniques (ed., G. Leitmann), Academic Press, New York (1962), Chapter 6.

BIBLIOGRAPHY (Cont.)

13. Ostrovskii, G.M., "On a Method of Solving Variational Problems," Automation Remote Control, 23 (1962), 1284-1289.

14. Okamura, K., "Some Mathematical Theory of the Penalty Method for Solving Optimum Control Problems," J. SIAM Control (A), 2 (1965), 317-331.

15. Bellman, R., Dynamic Programming, Princeton University Press, Princeton, New Jersey (1957).

16. Bellman, R., Adaptive Control Processes: A Guided Tour, Princeton University Press, Princeton, New Jersey (1961).

17. Kalman, R.E., "The Theory of Optimal Control and the Calculus of Variations," Mathematical Optimization Techniques (ed., R. Bellman), University of California Press, Berkeley and Los Angeles (1963), Chapter 16.

18. Bridgeland, T.F., Jr., "On the Existence of Optimal Feedback Controls," J. SIAM Control (A), 1 (1963), 261-274.

19. Bridgeland, T.F., Jr., "On the Existence of Optimal Feedback Controls, II," J. SIAM Control (A), 2 (1965), 137-150.

20. Coddington, E.A. and N. Levinson, Theory of Ordinary Differential Equations, McGraw-Hill Book Co., New York (1955).

21. Rozonoer, L.I., "The Maximum Principle of L.S. Pontryagin in Optimal System Theory – Part III," Automation Remote Control, 21 (1960), 1517-1532.

22. Kopp, R.E., "Pontryagin Maximum Principle," Optimization Techniques (ed., G. Leitmann), Academic Press, New York (1962), Chapter 7.

23. Lee, E.B., "Design of Optimum Multivariable Control Systems," Trans. ASME (D), 83 (1961), 85-90.

24. Kamke, E., Differentialgleichungen (3rd edition), Chelsea Publishing Company, New York (1959).

BIBLIOGRAPHY (Cont.)

25. Hestenes, M.R., Calculus of Variations and Optimal Control Theory, John Wiley & Sons, New York (1966).

26. Bolza, O., Lectures on the Calculus of Variations, Dover Publishing Co., New York (1961), 84-114.

27. Hilbert, D., "Mathematische Probleme," Archiv der Mathematik und Physik, Third Series, 1, 213-237, B.G. Teubner, Leipzig and Berlin (1901).

APPENDIX A

EXISTENCE AND CONTINUITY OF PERFORMANCE PARTIALS

We wish to give an explicit demonstration of Assertion 2 of Chapter 3. These results will then be used in Appendix B to prove Lemma 2 of the same chapter.

Let the control law $\delta \epsilon P[G_o]$ have the solution $x(t; t_o, x_o)$ over $[t_o, t_1(t_o, x_o)]$ where (t_o, x_o) is an interior point of one of δ's regions $\{G_1, G_2, \ldots, G_K\}$ of continuous differentiability. Let us denote the sequence of regions G_k which $\left(t, x(t; t_o, x_o)\right)$ passes through by $< G_{k_i} >$, where $i=1,2,\ldots,I$ (finite).

By definition, δ's performance for (t_o, x_o) is given by

$$J(t_o, x_o; \delta) = \int_{t_o}^{t_1(t_o, x_o)} L\left[t, x(t; t_o, x_o), \ \delta\left(t, x(t; t_o, x_o)\right)\right] dt$$

For a neighboring phase $(t_\lambda, x_\lambda) \triangleq (t_o + \lambda_o, x_o + \lambda)$, where $(\lambda_o, \lambda) \epsilon \mathbb{R} \times \mathbb{R}^n$ such that $|\lambda_o|$ and $\|\lambda\|$ are less than a small number ε, the solution $\left(t, x(t; t_\lambda, x_n)\right)$ will pass successively through the same sequence $< G_{k_i} >$ of regions (this due essentially to condition (iv) of the definition for $P[G_o]$).

To shorten the length of the expressions to follow we introduce the following notation:

$$\theta_o = (t_o, x_o) \ ; \quad \theta_\lambda = (t_\lambda, x_\lambda)$$

$$x(t; \theta_o) = x(t; t_o, x_o) \ ; \quad x(t; \theta_\lambda) = x(t; t_\lambda, x_\lambda)$$

$$\theta_o(t) = \left(t, x(t; \theta_o)\right) \ ; \quad \theta_\lambda(t) = \left(t, x(t; \theta_\lambda)\right)$$

Thus, for the solution $\theta_\lambda(t)$ we have

$$J(\theta_\lambda; \delta) = \int_{t_\lambda}^{t_1(\theta\ )} L\left[\theta_\lambda(t), \delta\left(\theta_\lambda(t)\right)\right] dt$$

For control laws of class $P[G_o]$ the solutions $x(t;\theta_o)$ are continuously differentiable with respect to $\theta_o \epsilon \overset{o}{G}_{k_1}$, where $\overset{o}{G}_{k_1}$ denotes the interior of G_{k_1}. Thus, we have

$$\theta_\lambda(t) = \theta_o(t) + \frac{\partial\theta_o(t)}{\partial\theta_o} \bullet \begin{bmatrix} \lambda_o \\ \lambda \end{bmatrix} + o(\varepsilon) \tag{A.1}$$

throughout $[t_o, t_1(\theta_o)]$, and

$$\delta\left(\theta_\lambda(t)\right) = \delta\left(\theta_o(t)\right) + \frac{\partial\delta}{\partial\theta}\left(\theta_o(t)\right) \bullet \frac{\partial\theta_o(t)}{\partial\theta_o} \bullet \begin{bmatrix} \lambda_o \\ \lambda \end{bmatrix} + o(\varepsilon) \tag{A.2}$$

whenever $\theta_\lambda(t)$ and $\theta_o(t)$ are contained in a <u>single region</u> G_{k_i}. Whenever $\theta_\lambda(t)$ and $\theta_o(t)$ are in two different regions G_{k_i} and $G_{k_{i+1}}$, then one solution or the other must have reached the boundary first at a boundary phase $\theta_o(t_B)$ or $\theta_\lambda(t_B')$. Let us assume that the unperturbed solution does so first. In this case for, $t \epsilon [t_B, t_B + \varepsilon]$

$$\delta\left(\theta_o(t)\right) = \delta_+\left(\theta_o(t_B)\right) + \frac{\partial\delta}{\partial\theta_+}\left(\theta_o(t_B)\right) \bullet \frac{d\theta_o(t)}{dt_+}(t-t_B) + o(\varepsilon)$$

and (A.3)

$$\delta\left(\theta_\lambda(t)\right) = \delta_-\left(\theta_o(t_B)\right) + \frac{\partial\delta}{\partial\theta_-}\left(\theta_o(t_B)\right) \bullet \frac{d\theta_o(t)}{dt_-}(t-t_B) + o(\varepsilon)$$

where subscripts $(\pm)$ indicate right or left limits as we approach $\theta_o(t_B)$ from region $G_{k_{i+1}}$ or G_{k_i}, respectively. In the case where $\theta_\lambda(t)$ meets the boundary first, the same expressions will apply except for interchanging $\left(\theta_o(t),\ \theta_o(t_B),\ t_B\right)$ with $\left(\theta_\lambda(t),\ \theta\ (t_B'),\ t_B'\right)$ wherever they appear.

The demonstration can now begin with these preliminaries away. Let

$$\Delta_\lambda J = J(\theta_\lambda;\delta) - J(\theta_o;\delta) = \int_{t_\lambda}^{t_1(\theta_\lambda)} L\left[\theta_\lambda(t), \delta\left(\theta_\lambda(t)\right)\right]dt$$

$$- \int_{t_o}^{t_1(\theta_o)} L\left[\theta_o(t),\ \delta\left(\theta_o(t)\right)\right]dt$$

$$= \int_{t_\lambda}^{t_1(\theta_\lambda)} \left\{L\left[\theta_\lambda(t),\ \delta\left(\theta_\lambda(t)\right)\right] - L\left[\theta_o(t),\ \delta\left(\theta_o(t)\right)\right]\right\}dt$$

$$+ \int_{t_1(\theta_o)}^{t_1(\theta_\lambda)} L\left[\theta_o(t),\ \delta\left(\theta_o(t)\right)\right]dt - \int_{t_o}^{t_\lambda} L\left[\theta_o(t),\ \delta\left(\theta_o(t)\right)\right]dt$$

in which $L\left[\theta_o(t),\ \delta\left(\theta_o(t)\right)\right]$ is extended as is necessary for these integrals by holding it constant at $L\left[\theta_o(t_1),\ \delta\left(\theta_o(t_1)\right)\right]$ or $L[\theta_o, \delta(\theta_o)]$ outside of $[t_o, t_1(\theta_o)]$.

We treat each integral separately for convenience.

<u>Integral 1</u>

$$I_1 = -\int_{t_o}^{t_\lambda} L\left[\theta_o(t), \delta\left(\theta_o(t)\right)\right]dt = -\lambda_o\ L\left(\theta_o, \delta\left(\theta_o\right)\right) + o(\epsilon),$$

since $t_\lambda = t_o + \lambda_o$, where $\lambda_o = 0(\epsilon)$.

<u>Integral 2</u>

$$\int_{t_1(\theta_o)}^{t_1(\theta_\lambda)} L\left[\theta_o(t),\ \delta\left(\theta_o(t)\right)\right]dt = L\left(\theta_1^o, \delta(\theta_1^o)\right) \bullet [t_1(\theta_\lambda) - t_1(\theta_o)] + o(\epsilon)$$

where $\theta_1^o \triangleq (t_1^o,\ x_1^o) \in S$ for our unperturbed solution. Of course, if we are dealing with a fixed time problem, this integral is zero. In the general case

$$t_1(\theta_\lambda) - t_1(\theta_o) = \frac{\partial t_1}{\partial \theta}(\theta_o) \bullet \begin{bmatrix} \lambda_o \\ \lambda \end{bmatrix} + o(\xi)$$

since the solutions are continuously differentiable wrt initial phases and the boundary ∂S of the target S is continuously differentiable wrt terminal phases [see Chapter 1]. Thus,

$$I_2 = \int_{t_1(\theta_o)}^{t_1(\theta_\lambda)} L\left[\theta_o(t), \delta\left(\theta_o(t)\right)\right] dt = L\left(\theta_1^o, \delta(\theta_1^o)\right) \frac{\partial t_1}{\partial \theta}(\theta_o) \bullet \begin{bmatrix} \lambda_o \\ \lambda \end{bmatrix} + o(\xi)$$

Integral 3

$$I_3 = \int_{t_\lambda}^{t_1(\theta_\lambda)} \left\{ L\left[\theta_\lambda(t), \delta\left(\theta_\lambda(t)\right)\right] - L\left[\theta_o(t), \delta\left(\theta_o(t)\right)\right]\right\} dt$$

$$= \int_{t_\lambda}^{t_1(\theta_\lambda)} \left\{ L\left[\theta_\lambda(t), \delta\left(\theta_\lambda(t)\right)\right] - L\left[\theta_o(t), \delta\left(\theta_\lambda(t)\right)\right]\right\} dt$$

$$+ \int_{t_\lambda}^{t_1(\theta_\lambda)} \left\{ L\left[\theta_o(t), \delta\left(\theta_\lambda(t)\right)\right] - L\left[\theta_o(t), \delta\left(\theta_o(t)\right)\right]\right\} dt$$

This easily becomes

$$I_3 = \int_{t_\lambda}^{t_1(\theta_\lambda)} \frac{\partial L}{\partial \theta}\left[\theta_o(t),\ \delta\left(\theta_o(t)\right)\right] \frac{\partial \theta_o(t)}{\partial \theta_o} \begin{bmatrix} \lambda_o \\ \lambda \end{bmatrix} dt + o(\xi)$$

$$+ \int_{t_\lambda}^{t_1(\theta_\lambda)} \left\{ L\left[\theta_o(t),\ \delta\left(\theta_\lambda(t)\right)\right] - L\left[\theta_o(t),\ \delta\left(\theta_o(t)\right)\right]\right\} dt$$

since $L(\theta, u)$ is continuously differentiable wrt (θ, u). The second integral on the right, call it I_{32}, may be treated as follows:

$$I_{32} = \int_{[t_\lambda, t_1(\theta_\lambda)] - \omega_\varepsilon} \frac{\partial L}{\partial u}\left[\theta_o(t), \delta\left(\theta_o(t)\right)\right] \frac{\partial \delta}{\partial \theta}\left(\theta_o(t)\right) \frac{\partial \theta_o(t)}{\partial \theta_o}\begin{bmatrix}\lambda_o \\ \lambda\end{bmatrix} dt$$

$$+ \int_{\omega_\varepsilon}\left\{L\left[\theta_o(t), \delta\left(\theta_\lambda(t)\right)\right] - L\left[\theta_o(t), \delta\left(\theta_o(t)\right)\right]\right\} dt + o(\varepsilon)$$

where ω_ε is the union of all time intervals over which $\theta_\lambda(t)$ and $\theta_o(t)$ are in two different regions G_{k_i} and $G_{k_{i+1}}$. Let ω_{ε_i} be any one of these time intervals. In fact, let us assume $\theta_o(t)$ crosses the boundary first at a boundary phase $\theta_o(t_B)$ and $\theta_\lambda(t)$ reaches it later at $\theta_\lambda(t'_B)$ where $t'_B = t_B + \varepsilon_i$. Then for this case

$$\int_{t_B}^{t_B+\varepsilon_i} L\left[\theta_o(t), \delta\left(\theta_\lambda(t)\right)\right] dt = \int_{t_B}^{t_B+\varepsilon_i} L\left[\theta_o(t), \delta_-\left(\theta_o(t_B)\right)\right] dt$$

$$+ \int_{t_B}^{t_B+\varepsilon_i} \frac{\partial L}{\partial u}\left[\theta_o(t), \delta_-\left(\theta_o(t_B)\right)\right] \frac{\partial \delta}{\partial \theta_-}\left(\theta_o(t_B)\right)$$

$$\bullet \frac{d\theta_o}{dt_-}(t - t_B)\, dt + o(\varepsilon_i)$$

and

$$\int_{t_B}^{t_B+\varepsilon_i} L\left[\theta_o(t), \delta\left(\theta_o(t)\right)\right] dt = \int_{t_B}^{t_B+\varepsilon_i} L\left[\theta_o(t), \delta_+\left(\theta_o(t_B)\right)\right] dt$$

$$+ \int_{t_B}^{t_B+\varepsilon_i} \frac{\partial L}{\partial u}\left[\theta_o(t), \delta_+\left(\theta_o(t_B)\right)\right] \frac{\partial \delta}{\partial \theta_+}\left(\theta_o(t_B)\right)$$

$$\bullet \frac{d\theta_o(t)}{dt_+}(t - t_B)\, dt + o(\varepsilon_i)$$

Thus, we have for an interval ω_{ε_i},

$$\int_{t_B}^{t_B+\varepsilon_i} \left\{ L\left[\theta_o(t),\ \delta\left(\theta_\lambda(t)\right)\right] - L\left[\theta_o(t),\ \delta\left(\theta_o(t)\right)\right]\right\} dt$$

$$= \int_{t_B}^{t_B+\varepsilon_i} L\left[\theta_o(t),\ \delta_-\left(\theta_o(t_B)\right)\right] - L\left[\theta_o(t),\ \delta_+\left(\theta_o(t_B)\right)\right] dt \ + o(\varepsilon_i)$$

Since $\theta_o(t)$ is continuous and L differentiable wrt θ,

$$\int_{t_B}^{t_B+\varepsilon_i} \left\{ L\left[\theta_o(t), \delta\left(\theta_\lambda(t)\right)\right] - L\left[\theta_o(t),\ \delta\left(\theta_o(t)\right)\right]\right\} dt = \varepsilon_i \Delta L\left(\theta_o(t_B)\right) + o(\varepsilon_i)$$

where $\Delta L\left(\theta_o(t_B)\right) \triangleq L\left[\theta_o(t_B),\ \delta_-\left(\theta_o(t_B)\right)\right] - L\left[\theta_o(t_B),\ \delta_+\left(\theta_o(t_B)\right)\right]$.
From this example one sees that the integral over ω_ε will have the form

$$\int_{\omega_\varepsilon} \left\{ L\left[\theta_o(t), \delta\left(\theta_\lambda(t)\right)\right] - L\left[\theta_o(t), \delta\left(\theta_o(t)\right)\right]\right\} dt = \sum_{i=1}^{I} \Delta L\left(\theta_o(t_{B_i})\right) \varepsilon_i$$

However, as in the case of Integral I_2

$$\varepsilon_i = t'_{B_i} - t_{B_i} = \frac{\partial t_{B_i}}{\partial \theta} \begin{bmatrix} \lambda_o \\ \lambda \end{bmatrix} + o(\varepsilon)$$

since our boundaries are assumed continuously differentiable. Thus,

$$I_{32} = \int_{[t_\lambda, t_1(\theta_\lambda)] - \omega_\varepsilon} \frac{\partial L}{\partial u} \frac{\partial \delta}{\partial \theta} \frac{\partial \theta_o(t)}{\partial \theta_o} \begin{bmatrix} \lambda_o \\ \lambda \end{bmatrix} dt$$

$$+ \sum_{i=1}^{I} \Delta L\left(\theta_o(t_{B_i})\right) \frac{\partial t_{B_i}}{\partial \theta} \begin{bmatrix} \lambda_o \\ \lambda \end{bmatrix} + o(\varepsilon)$$

Conclusion

Collecting all integrals I_1, I_2, I_{31}, I_{32} and letting $\varepsilon \to 0$, we have to first order

$$\Delta J_\lambda = \int_{t_o}^{t_1(\theta_o)} \left\{ \frac{\partial L}{\partial \theta}\left[\theta_o(t), \delta\left(\theta_o(t)\right)\right] + \frac{\partial L}{\partial u}\left[\theta_o(t), \delta\left(\theta_o(t)\right)\right] \frac{\partial \delta}{\partial \theta}\left(\theta_o(t)\right) \right\} \frac{\partial \theta_o(t)}{\partial \theta_o} \begin{bmatrix} \lambda_o \\ \lambda \end{bmatrix} dt$$

$$+ \sum_{i=1}^{I} \Delta L\left(\theta_o(t_{B_i})\right) \frac{\partial t_{B_i}}{\partial \theta} \begin{bmatrix} \lambda_o \\ \lambda \end{bmatrix} + L\left(\theta_1^o, \delta(\theta_1^o)\right) \frac{\partial t_1}{\partial \theta} \begin{bmatrix} \lambda_o \\ \lambda \end{bmatrix}$$

$$- \lambda_o L\left(\theta_o, \delta(\theta_o)\right)$$

This shows that $J(\theta_o; \delta)$ is differentiable for $\theta_o \in G^o_{k_1}$, and moreover that it is continuously so.

The partials of $J(\theta_o; \delta)$ are given below.

$$\frac{\partial J}{\partial t_o}(\theta_o; \delta) = \int_{t_o}^{t_1} \left\{ \frac{\partial L}{\partial x}\left[\theta(t), \delta\left(\theta(t)\right)\right] + \frac{\partial L}{\partial u}\left[\theta(t), \delta\left(\theta(t)\right)\right] \cdot \frac{\partial \delta}{\partial x}\left(\theta(t)\right) \right\} \frac{\partial x(t)}{\partial t_o} dt$$

$$+ \sum_{i=1}^{I} \Delta L\left(\theta(t_{B_i})\right) \frac{\partial t_{B_i}}{\partial t_o} + L\left(\theta_1, \delta(\theta_1)\right) \frac{\partial t_1}{\partial t_o} - L\left(\theta_o, \delta(\theta_o)\right) \tag{A.4}$$

$$\frac{\partial J}{\partial x_o}(\theta_o; \delta) = \int_{t_o}^{t_1} \left\{ \frac{\partial L}{\partial x}\left[\theta(t), \delta\left(\theta(t)\right)\right] + \frac{\partial L}{\partial u}\left[\theta(t), \delta\left(\theta(t)\right)\right] \cdot \frac{\partial \delta}{\partial x}\left(\theta(t)\right) \right\} \frac{\partial x(t)}{\partial x_o} dt$$

$$+ \sum_{i=1}^{I} \Delta L\left(\theta(t_{B_i})\right) \frac{\partial t_{B_i}}{\partial x_o} + L\left(\theta_1, \delta(\theta_1)\right) \frac{\partial t_1}{\partial t_o} \tag{A.5}$$

In these last equations we have dropped all superfluous subscripts since we are concerned with the single trajectory:

$$\theta(t) = \left(t, x(t; t_o, x_o)\right)$$

APPENDIX B

IDENTIFICATION OF CO-STATE VARIABLES WITH PERFORMANCE PARTIALS

Theorem A

Let the optimal performance function $J_*(t_o, x_o)$ correspond to a control law $\delta_* \epsilon P[G_o]$. Then $\partial J_*/\partial t_o$ and $\partial J_*/\partial x_o$ satisfy the co-state equations and boundary condition $\Big[L\big(t_o, x_o, \delta_*(t_o, x_o)\big) + \partial J_*/\partial x_o\,(t_o, x_o) \bullet f\big(t_o, x_o, \delta_*(t_o, x_o)\big) + \partial J_*/\partial t_o\,(t_o, x_o) = 0\Big]$ of Pontryagin's method (Theorem 2, Chapter 2) if:

(i) $(t_o, x_o) \epsilon G_k^o$, and

(ii) $\dfrac{\partial J_*}{\partial t}\big(t, x(t)\big)$ and $\dfrac{\partial J_*}{\partial x}\big(t, x(t)\big)$ are absolutely continuous over some interval containing t_o.

Proof

Condition (i) allows us to conclude that the performance function partials $\partial J_*/\partial t$, $\partial J_*/\partial x$ exist and are continuous in a neighborhood about (t_o, x_o). Condition (ii) enables us to differentiate with respect to time almost everywhere in this neighborhood.

From Equation (A.5) of Appendix A we have for $\big(t, x(t)\big)$ in this neighborhood,

$$\frac{\partial J_*}{\partial x}\big(t, x(t)\big) = \int_t^{t_1}\left[\frac{\partial L}{\partial x}\big(\tau, x(\tau), \delta_*(\tau)\big) + \frac{\partial L}{\partial u}\big(\tau, x(\tau), \delta_*(\tau)\big)\cdot\frac{\partial \delta_*(\tau)}{\partial x}\right] \bullet \frac{\partial x(\tau)}{\partial x(t)}\,d\tau + \sum_{i=1}^{I}\Delta L(t_{B_i}, x_{B_i})\,\frac{\partial t_{B_i}}{\partial x(t)} + L\big(t_1, x_1, \delta_*(t_1)\big)\frac{\partial t_1}{\partial x(t)}$$

[Note: For simplicity we have denoted $\delta_*\big(\tau, x(\tau)\big)$ by $\delta_*(\tau)$.]

Differentiating with respect to t (denoting this by a dot), we have

$$\frac{\partial \dot{J}_*}{\partial x}\Big(t, x(t)\Big) = -\left[\frac{\partial L}{\partial x}\Big(t, x(t),\ \delta_*(t)\Big) + \frac{\partial L}{\partial u}\Big(t, x(t),\ \delta_*(t)\Big) \cdot \frac{\partial \delta_*(t)}{\partial x}\right]$$

$$+ \int_t^{t_1}\left[\frac{\partial L}{\partial x}\Big(\tau, x(\tau),\ \delta_*(\tau)\Big) + \frac{\partial L}{\partial u}\Big(\tau, x(\tau), \delta_*(\tau)\Big)\right.$$

$$\left. \bullet\ \frac{\partial \delta_*(\tau)}{\partial x}\right]\left(\frac{\partial \dot{x}(\tau)}{\partial x(t)}\right) d\tau + \sum_{i=1}^{I} \Delta L(t_{B_i}, x_{B_i})\ \frac{\partial t_{B_i}}{\partial x(t_{B_i})}\left(\frac{\partial \dot{x}(t_{B_i})}{\partial x(t)}\right)$$

$$+ L\Big(t_1, x_1, \delta_*(t_1)\Big)\ \frac{\partial t_1}{\partial x(t_1)}\ \left(\frac{\partial \dot{x}(t_1)}{\partial x(t)}\right)$$

But since $\partial x(\xi)/\partial x(t)$ is a state-transformation matrix, we have:[1]

$$\left(\frac{\partial \dot{x}(\xi)}{\partial x(t)}\right) = -\ \frac{\partial x(\xi)}{\partial x(t)} \bullet\ \frac{\partial g}{\partial x}\Big(t, x(t)\Big) \tag{B. 1}$$

corresponding to the process equations

$$\frac{dx}{dt} = g(t, x) \triangleq\ f\Big(t, x,\ \delta_*(t, x)\Big) \tag{B. 2}$$

Therefore using relation (B. 1) we have

$$\frac{\partial \dot{J}_*}{\partial x}\Big(t, x(t)\Big) = -\ \frac{\partial J_*}{\partial x}\Big(t, x(t)\Big) \bullet\ \frac{\partial g}{\partial x}\Big(t, x(t)\Big) - \left[\frac{\partial L}{\partial x}\Big(t, x(t), \delta_*(t)\Big)\right.$$

$$\left. +\ \frac{\partial L}{\partial u}\Big(t, x(t), \delta_*(t)\Big)\ \frac{\partial \delta_*(t)}{\partial x}\right]$$

But from relation (B. 2)

$$\frac{\partial g}{\partial x}(t, x) = \frac{\partial f}{\partial x}\Big(t, x, \delta_*(t, x)\Big) + \frac{\partial f}{\partial u}\Big(t, x, \delta_*(t, x)\Big) \bullet\ \frac{\partial \delta_*}{\partial x}(t, x)$$

Substituting this we obtain,

$$\frac{\partial \dot{J}_*}{\partial x}\big(t, x(t)\big) = -\frac{\partial J_*}{\partial x}\big(t, x(t)\big) \bullet \frac{\partial f}{\partial x}\big(t, x(t), \delta_*(t)\big) - \frac{\partial L}{\partial x}\big(t, x(t), \delta_*(t)\big)$$

$$-\left[\frac{\partial J_*}{\partial x}\big(t, x(t)\big) \bullet \frac{\partial f}{\partial u}\big(t, x(t), \delta_*(t)\big) + \frac{\partial L}{\partial u}\big(t, x(t), \delta_*(t)\big)\right]\frac{\partial \delta_*(t)}{\partial x}$$

Our next step is the crucial one. The expression in square brackets is equal to

$$\frac{\partial}{\partial u}\left[L(t, x, u) + \frac{\partial J_*}{\partial x}(t, x) \bullet f(t, x, u) + \frac{\partial J_*}{\partial t}(t, x)\right] \qquad \text{(B. 3)}$$

evaluated at $\big(t, x(t), \delta_*(t, x(t))\big)$. This quantity multiplied by the matrix $\partial \delta_*(t, x)/\partial x \ \big(t, x(t)\big)$ is indicative of the variation that one might obtain in $[L + \partial J_*/\partial x \bullet f + \partial J_*/\partial t\ ]$ by using values of u equal to $\delta_*(t, y)$, where y is a state vector drawn from a small neighborhood $N\big(x(t)\big)$ about $x(t)$. Two things are possible:

(i) $\delta_*\big(t, x(t)\big)$ is on a boundary of the (closed) control set U, or

(ii) $\delta_*\big(t, x(t)\big)$ is an interior point of U.

If (ii) occurs then expression (B. 3) must be a null vector (at least for almost every t in a neighborhood of t_o), since $\delta_*\big(t, x, (t)\big)$ is optimal and minimizes $L + \partial J_*/\partial x \ f + \partial J_*/\partial t$ [see Corollary 2, Chapter 3].

If (i) occurs and the closure $\overline{\delta_*\big(t, N(x(t))\big)}$ of the image of $N\big(x(t)\big)$ has $\delta_*\big(t, x(t)\big)$ as an interior point, then the product of expression (B. 3) with $\partial \delta_*/\partial x$ must be null for the same reason as above. The remaining possiblity is that $\delta_*\big(t, x(t)\big)$ is a boundary point of U and an extreme point of $\overline{\delta_*\big(t, N(x(t))\big)}$. In this case, since δ_* is continuously differentiable, we must have $\partial \delta_*/\partial x$ as a null matrix at $\big(t, x(t)\big)$.

Thus, we obtain

$$\frac{\partial \dot{J}}{\partial x}\left(t, x(t)\right) = -\frac{\partial J_*}{\partial x}\left(t, x(t)\right) \bullet \frac{\partial f}{\partial x}\left(t, x(t), \delta_*(t)\right) - \frac{\partial L}{\partial x}\left(t, x(t), \delta_*(t)\right) \quad \text{(B. 4)}$$

almost everywhere in a neighborhood of t_o.

Let us now treat $\partial J_*/\partial_t$. From Equation (A. 4) of Appendix A,

$$\frac{\partial J_*}{\partial t}\left(t, x(t)\right) = \int_t^{t_1}\left[\frac{\partial L}{\partial x}\left(\tau, x(\tau), \delta_*(\tau)\right) + \frac{\partial L}{\partial u}\left(\tau, x(\tau), \delta_*(\tau)\right)\frac{\partial \delta_*(\tau)}{\partial x}\right]\frac{\partial x(\tau)}{\partial t}\, d\tau$$

$$+ \sum_{i=1}^{I} \Delta L(t_{B_i}, x_{B_i})\ \frac{\partial t_{B_i}}{\partial t} + L\left(t_1, x_1, \delta_*(t_1)\right)\frac{\partial t_1}{\partial t}$$

$$- L\left[t, x(t),\ \delta_*\left(t, x(t)\right)\right] .$$

We note that the only difference from what we had for $\partial J_*/\partial x\left(t, x(t)\right)$ is that we have an additional term (-L), and that we shall be concerned with partials

$$\frac{\partial x(\xi)}{\partial t} \quad \text{instead of} \quad \frac{\partial x(\xi)}{\partial x(t)}$$

Using the following relation,

$$\frac{\partial x(\xi)}{\partial t} = -\frac{\partial x(\xi)}{\partial x(t)}\ g\left(t, x(t)\right) \qquad \text{(B. 5)}$$

and recalling (B. 2), (B. 4), as well as previous devices, we may derive

$$\frac{\partial \mathring{J}_*}{\partial t}\left(t, x(t)\right) = -\frac{\partial J_*}{\partial x}\left(t, x(t)\right) \bullet \frac{\partial f}{\partial t}\left(t, x(t),\ \delta_*(t)\right) - \frac{\partial L}{\partial t}\left(t, x(t), \delta_*(t)\right) \qquad \text{(B. 6)}$$

almost everywhere in a neighborhood of t_o. Thus, $\partial J_*/\partial t$ and $\partial J_*/\partial x$ satisfy the co-state equations of Theorem 2, Chapter 2.

The remaining step is to show that they satisfy the boundary condition

$$L\left(t_o, x_o, \delta_*(t_o)\right) + \frac{\partial J_*}{\partial x} \bullet f\left(t_o, x_o, \delta_*(t_o)\right) + \frac{\partial J_*}{\partial t} = 0$$

This follows from Corollary 2, Chapter 3, which states that this must hold a.e. along an optimal trajectory (where the partials exist and are continuous). In particular, since all quantities involved in the boundary condition above are defined and continuous everywhere in a neighborhood of t_o, the condition actually holds everywhere in the neighborhood.

(End of Proof)

Remark

The allusion to Corollary 2 in this proof deserves some expansion. Corollary 2 is based on the assumption that $J_*(t, x)$ is continuously differentiable in a region G_o containing S or having boundary points in common with the boundary of S. In the present situation we are dealing with a region G_{k_1} which in general is remote from S. We are alluding to the fact that δ_* must also be optimal for the problem in which the boundary ∂G_{k_1} is considered as a target, since each subarc of an optimal trajectory is optimal for its endpoints.

Thus, for this sub-problem

$$J_*^1(t, x) \triangleq J_*(t, x) - J_*(t_B^-, x_B^-) = \int_{t_o}^{t_B^-} L\left(t, x(t), \delta_*(t)\right) dt$$

where (t_B^-, x_B^-), corresponding to an initial (t, x) is the phase produced by δ_* at time $t_B^- = t_{B_1} - \epsilon$ (ϵ, small and positive). Since δ_* is optimal, J_*^1 according to Corollary 2 must satisfy

$$L\big(t, x(t), \delta_*(t)\big) + \frac{\partial J_*^1}{\partial x}\big(t, x(t)\big) \bullet f\big(t, x(t), \delta_*(t)\big) + \frac{\partial J_*^1}{\partial t}\big(t, x(t)\big) = 0$$

a.e. $[t_o, t_B^-]$. This can be expressed in terms of J_* as

$$\left[L_* + \frac{\partial J_*}{\partial x} \bullet f_* + \frac{\partial J_*}{\partial t}\right] - \frac{\partial J_*(t_B^-, x_B^-)}{\partial(t_B^-, x_B^-)} \bullet \frac{\partial(t_B^-, x_B^-)}{\partial(t, x)} \bullet \begin{bmatrix} 1 \\ f_* \end{bmatrix} = 0$$

However, the last term is zero because

$$\frac{\partial(t_B^-, x_B^-)}{\partial(t, x)} \bullet \begin{bmatrix} 1 \\ f_* \end{bmatrix} = \frac{d}{dt}(t_B^-, x_B^-) = 0$$

We do not change (t_B^-, x_B^-) by moving along the trajectory.

Application of Theorem A

Lemma 2, Chapter 3, can now be proved. The partials satisfy the (linear) co-state equations and the same inhomogeneous boundary condition as the co-state variables, whenever conditions (i) and (ii) of Theorem A hold. Thus, they must be equal to the co-state variables under these conditions.

Some Useful Relations

Certain relations are given below which the reader may find useful. These relations are straightforward, but perhaps not popularly recognized. Lemma A below is used in the proof of Theorem 9, Chapter 5.

Theorem B

Let $u(t)$ be an admissible control over $[t_o, t_1]$ which produces the solution $x(t)$. Consider the integral

$$J_u(t,x) \triangleq J(t,x;u) = \int_t^{t_1} L\big(\tau, x(\tau), u(\tau)\big)\, d\tau$$

for $t \in [t_o, t_1]$. The following relation holds a.e. $[t_o, t_1]$.

$$L\big(t, x(t), u(t)\big) + \frac{\partial J_u}{\partial x}\big(t, x(t)\big) \bullet f\big(t, x(t), u(t)\big) + \frac{\partial J_u}{\partial t}\big(t, x(t)\big) = 0$$

Proof

$$\frac{\partial J_u}{\partial x}\big(t, x(t)\big) = \int_t^{t_1} \frac{\partial L}{\partial x}\big(\tau, x(\tau), u(\tau)\big) \bullet \frac{\partial x(\tau)}{\partial x(t)}\, d\tau \qquad \text{(B. 7)}$$

and

$$\frac{\partial J_u}{\partial t}\big(t, x(t)\big) \overset{a.e.}{=} - L\big(t, x(t), u(t)\big) + \int_t^{t_1} \frac{\partial L}{\partial x}\big(\tau, x(\tau), u(\tau)\big) \bullet \frac{\partial x(\tau)}{\partial t}\, d\tau$$

But for almost all $t \in [t_o, t_1]$

$$\frac{\partial x(\tau)}{\partial t} = - \frac{\partial x(\tau)}{\partial x(t)} \bullet f\big(t, x(t), u(t)\big)$$

Substituting this relation into the expression for $\partial J_u / \partial t$ and using the identity for $\partial J_u / \partial x$, the theorem is proved.

(End of Proof)

Lemma A

Let $u(\cdot)$ be an admissible control over $[t_o, t_1]$ for a fixed time, free right end problem. Let its performance function $J_u\big(t, x(t)\big)$ have the property that

$$\inf_{u \in U} \left[L\big(t, x(t), u\big) + \frac{\partial J_u}{\partial x} \bullet f\big(t, x(t), u\big) + \frac{\partial J_u}{\partial t} \right] = 0 \qquad \text{(B. 8)}$$

along the trajectory $\left(t, x(t)\right)$ produced by $u(\cdot)$. Then the partials $\left[\frac{\partial J_u}{\partial x}\left(t, x(t)\right), \frac{\partial J_u}{\partial t}\left(t, x(t)\right)\right]$ are equal to the co-state variable $\left(P(t), P_o(t)\right)$ corresponding to $x(t)$ and $u(t)$.

Proof

We first prove that $\partial J_u/\partial x \left(t, x(t)\right) = P(t)$ even without (B.8). Differentiating with respect to t in (B.7) (denoting this by a dot), we have

$$\left[\frac{\partial J_u}{\partial x}^{\bullet}\left(t, x(t)\right)\right] = \int_t^{t_1} \frac{\partial L}{\partial x}\left(\tau, x(\tau), u(\tau)\right) \bullet \left(\frac{\partial \dot{x}(\tau)}{\partial x(t)}\right) d\tau - \frac{\partial L}{\partial x}\left(t, x(t), u(t)\right)$$

almost everywhere in $[t_o, t_1]$. Using the relation

$$\left(\frac{\partial \dot{x}(\tau)}{\partial x(t)}\right) = - \frac{\partial x(\tau)}{\partial x(t)} \bullet \frac{\partial f}{\partial x}\left(t, x(t), u(t)\right)$$

and (B.7), it is readily seen that $\partial J_u/\partial x$ is a solution to the co-state equations for P. Finally, $P(t_1) = 0$ for the free right end problem, and $\partial J_u/\partial x$ is seen to satisfy this boundary condition also. Thus, $\partial J_u/\partial x$ must be equal to P(t) over $[t_o, t_1]$.

The proof that $\partial J_u/\partial t$ is equal to $P_o(t)$ will require an allusion to a proof in Pontryagin's work.[2] First of all, by hypothesis and the first part of our proof, we have

$$\frac{\partial J_u}{\partial t}\left(t, x(t)\right) = \sup_{u \epsilon U} \left[-L\left(t, x(t), u\right) - P(t)\, f\left(t, x(t), u\right)\right] \quad \text{(B.9)}$$

The right member has been shown by Pontryagin et al., to be absolutely continuous on $[t_o, t_1]$ (see pp. 101-103, Chapter II of Reference 2). Thus, $\partial J_u/\partial t$ is differentiable with respect to t almost everywhere in $[t_o, t_1]$.

From Theorem B we have, a.e. $[t_o, t_1]$,

$$\frac{\partial J_u}{\partial t}\Big(t, x(t)\Big) = -L\Big(t, x(t), u(t)\Big) - P(t)f\Big(t, x(t), u(t)\Big) \tag{B.10}$$

Let $\tau \epsilon [t_o, t_1]$ be a regular point of $u(\cdot)$ at which $\partial J_u/\partial t$ is differentiable with respect to t. Then (B.10) holds for τ. Let t be any point in a small neighborhood $[\tau-\epsilon, \tau+\epsilon]$ of τ. By virtue of (B.9) and (B.10) we have

$$\frac{\partial J_u}{\partial t}\Big(t, x(t)\Big) \geq -L\Big(t, x(t), u(\tau)\Big) - P(t)f\Big(t, x(t), u(\tau)\Big)$$

For sufficiently small ϵ we have

$$L\Big(t, x(t), u(\tau)\Big) = L\Big(\tau, x(\tau), u(\tau)\Big) + \frac{\partial L}{\partial t}\Big(\tau, x(\tau), u(\tau)\Big) \bullet (t-\tau)$$
$$+ \frac{\partial L}{\partial x}\Big(\tau, x(\tau), u(\tau)\Big) \bullet \Big(x(t) - x(\tau)\Big) + o(\epsilon),$$

$$f\Big(t, x(t), u(\tau)\Big) = f\Big(\tau, x(\tau), u(\tau)\Big) + \frac{\partial f}{\partial t}\Big(\tau, x(\tau), u(\tau)\Big) \bullet (t-\tau)$$
$$+ \frac{\partial f}{\partial x}\Big(\tau, x(\tau), u(\tau)\Big) \bullet \Big(x(t) - x(\tau)\Big) + o(\epsilon),$$

and

$$P(t) = P(\tau) - \left[P(\tau)\frac{\partial f}{\partial x}\Big(\tau, x(\tau), u(\tau)\Big) - \frac{\partial L}{\partial x}\Big(\tau, x(\tau), u(\tau)\Big)\right]\Big(x(t) - x(\tau)\Big)$$
$$+ o(\epsilon)$$

Substituting these into the above inequality we have

$$\frac{\partial J_u}{\partial t}\Big(t, x(t)\Big) - \frac{\partial J_u}{\partial t}\Big(\tau, x(\tau)\Big) \geq (t-\tau)\left[-P(\tau)\frac{\partial f}{\partial t}\Big(\tau, x(\tau), u(\tau)\Big)\right.$$
$$\left. - \frac{\partial L}{\partial t}\Big(\tau, x(\tau), u(\tau)\Big)\right] + o(\epsilon)$$

For $(t-\tau) > 0$ we have

$$\lim_{t \to \tau^+} \frac{\frac{\partial J_u}{\partial t}\left(t, x(t)\right) - \frac{\partial J_u}{\partial t}\left(\tau, x(\tau)\right)}{t-\tau} \geq -P(\tau)\frac{\partial f}{\partial t}\left(\tau, x(\tau), u(\tau)\right) - \frac{\partial L}{\partial t}\left(\tau, x(\tau), u(\tau)\right)$$

For $(t-\tau) < 0$ we have

$$\lim_{t \to \tau^-} \frac{\frac{\partial J_u}{\partial t}\left(t, x(t)\right) - \frac{\partial J_u}{\partial t}\left(\tau, x(\tau)\right)}{t-\tau} \leq -P(\tau)\frac{\partial f}{\partial t}\left(\tau, x(\tau), u(\tau)\right) - \frac{\partial L}{\partial t}\left(\tau, x(\tau), u(\tau)\right)$$

Since $\partial J_u/\partial t$ is differentiable at τ, the right and left limits exist and are equal to $(\partial \dot{J}_u/\partial t)$ at $t = \tau$. Thus,

$$\frac{d}{d\tau}\left[\frac{\partial J_u}{\partial t}\left(\tau, x(\tau)\right)\right] = -P(\tau)\frac{\partial f}{\partial t}\left(\tau, x(\tau), u(\tau)\right) - \frac{\partial L}{\partial t}\left(\tau, x(\tau), u(\tau)\right)$$

at all regular points $\tau \in [t_o, t_1]$ for which $(\partial \dot{J}_u/\partial t)$ exists. This is the same equation which $P_o(\tau)$ must satisfy a.e. $[t_o, t_1]$. We also note that at the terminal time t_1,

$$L\left(t_1, x(t_1), u(t_1)\right) + \frac{\partial J_u}{\partial t}\left(t_1, x(t_1)\right) = 0$$

which is precisely the condition which $P_o(t_1)$ must satisfy for the free right end problem. Thus, $\partial J_u/\partial t$ must be equal to $P_o(t)$ on $[t_o, t_1]$.

(End of Proof)