PROJECT ON EFFICIENCY OF DECISION MAKING

IN

ECONOMIC SYSTEMS
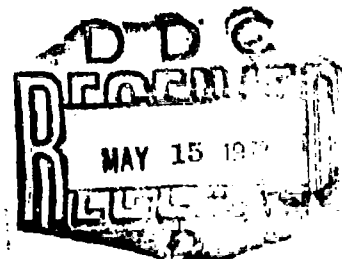
COMPARATIVE DYNAMICS

(SENSITIVITY ANALYSIS)

IN OPTIMAL CONTROL THEORY

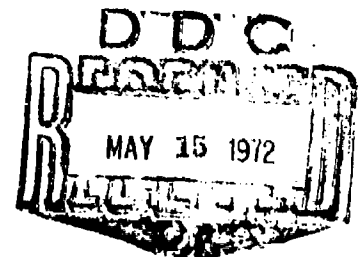Hajime Oniki

HARVARD UNIVERSITY

COMPARATIVE DYNAMICS

(SENSITIVITY ANALYSIS)

IN OPTIMAL CONTROL THEORY

Hajime Oniki

Technical Report No. 10

Harvard University
1737 Cambridge Street, Room 401
Cambridge, Massachusetts 02138

April, 1972

DOCUMENT CONTROL DATA - R & D

Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified

1. ORIGINATING ACTIVITY (Corporate author)	2a. REPORT SECURITY CLASSIFICATION
Project on Efficiency of Decision Making in Economic Systems, 1737 Cambridge St., #401, Harvard University, Cambridge, Mass. 02138	Unclassified
	2b. GROUP

3. REPORT TITLE

Comparative Dynamics (Sensitivity Analysis) in Optimal Control Theory

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Technical Report No. 9

5. AUTHOR(S) (First name, middle initial, last name)

Hajime Oniki

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
March, 1972	39	10

8a. CONTRACT OR GRANT NO	9a. ORIGINATOR'S REPORT NUMBER(S)
N00014-67-A-0298-0019	Technical Report No. 10
b. PROJECT NO NR 047-004	
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its publication is unlimited. Reproduction in whole or in part is permitted for any purpose of the United States Government.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Logistics & Mathematical Statistics Branch, Dept. of the Navy, Office of Naval Research, Washington, D.C. 20360

13. ABSTRACT

The paper considers an optimal control problem with a parameter and develops a systematic method for comparative dynamics. A sufficient condition for the optimum solution to be differentiable with respect to the parameter is provided. Formulas for computing the derivative are given in the form of initial-value problems of linear differential equations. Possibility of discontinuous optimal controls is fully taken care of. The derivative of the maximized objective function with respect to the parameter is expressed in a simple form by using the auxiliary variables. An example of the comparative dynamics is given in terms of a model of optimal capital accumulation.

COMPARATIVE DYNAMICS (SENSITIVITY ANALYSIS)

IN OPTIMAL CONTROL THEORY*

Hajime Oniki

I. Introduction

Optimal control theory was developed by Pontryagin and his associates [9] as a renovation of the classical theory of calculus of variations. It provides a convenient method for analyzing a wide class of economic problems such as planning the optimal capital accumulation of an economy and investigating the process of investment by a firm or by an individual. Although many applications of Pontryagin's theory to economic problems have been published, only a few of them have paid attention to the problem of comparative dynamics (sensitivity analysis), i.e., that of analyzing the effect of a parameter on the optimum solution.[1] In planning problems comparative dynamics may be useful for investigating the dependence of a plan on exogenous factors. In dynamic behavioral models it could serve as a tool for deriving the demand or the supply functions of an economic agent in question.

The present paper develops a systematic method dealing with comparative dynamics in optimal control problems. To facilitate the basic idea, let us consider a simple problem of maximizing the function $f(x,\theta)$ in x, where θ is a parameter to be regarded as fixed while the maximization is carried out. What we usually do is to obtain the first-order condition:

(1) $f_x(x,\theta) = 0,$

and to solve this for x. Assume that the solution x thus obtained
is a unique optimum. The dependence of this optimum on the parame-
ter θ may be studied by calculating the derivative of x with
respect to θ:

(2) $f_{xx} \cdot x'(\theta) + f_{x\theta} = 0$, or

   $x'(\theta) = - (f_{xx})^{-1} \cdot f_{x\theta}$.

This method can be extended to constrained maximum problems. In
the theory of household behavior, for example, the solution is
known in such terms as Slutzky equations, Hicksians, etc., where
the role of θ is played by commodity price or income.

It is noted that Pontryagin's optimum condition for a control
problem (composed of the maximum principle, auxiliary differential
equations, and transversality conditions) is essentially the first-
order condition; in this sense it is an extension of (1).[2/] What
I intend to do in this paper under the name of comparative dynamics
is to extend (2) to optimal control problems. Under certain as-
sumptions we shall provide a set of formulas by means of which
the derivatives of the optimum solution of a control problem with
respect to a parameter may be calculated.

It is well known that under certain conditions a solution of
a system of differential equations is differentiable with respect
to a parameter appearing in the system (the theorem of variational
differential equations).[3/] Since Pontryagin's condition contains
differential equations, it is suggested that one might make use
of this theorem for comparative dynamics. However, an immediate
application is difficult, since the theorem presupposes that the

differential equations are continuous in state variables, while
those appearing in Pontryagin's condition frequently exhibit dis-
continuities in state variables (e.g., bang-bang controls).[4/]
To resolve this, we shall extend the theorem of variational equa-
tions to the discontinuous case (Section III). Once this is at-
tained, we can readily derive formulas for comparative dynamics
from Pontryagin's condition (Section IV).

The method of comparative dynamics to be presented in this
paper is general enough to allow a parameter to appear almost
anywhere in the original control problem; e.g., the objective func-
tion may contain it, the initial state may be a function of it, or
the constraint on the control may be affected by it (or any com-
bination of these). The paper deals with the effect of a parameter
on the optimal control, on the state variables, and on the objec-
tive function. Both cases of finite and infinite horizons are
considered.

The following section (Section II) formulates our problem.
Section III is devoted to extending the theorem of variational
differential equations to the case where discontinuities arise.
The main results will be presented in Section IV. It will state
a set of conditions sufficient for the optimum solution to be
differentiable with respect to a given parameter, together with
formulas for computing the derivative. Section V deals with the
effect of a parameter on the objective function. The last section
(Section VI) is devoted to an example in which the method of com-
parative dynamics is applied to a model of optimal capital accumu-
lation.

II. The Problem

We shall be concerned with an optimal control problem with a parameter. It is of fixed time, variable end-points, autonomous differential equations, and fixed constraints on controls:[5,6] the objective function

$$(3) \qquad v = \int_{t_0}^{t_1} f^0(x(t),u(t),\theta)dt,$$

is to be maximized in $x(t)$ and $u(t)$, subject to

$$(4) \qquad \dot{x}(t) = f(x(t),u(t),\theta),$$

$$(5) \qquad g(u(t),\theta) \geq 0 \qquad (t_0 \leq t \leq t_1),$$

$$(6) \qquad \emptyset^i(x(t_i),\theta) = 0 \qquad (i = 0,1),$$

where $(t_0,t_1)(-\infty \leq t_0 < t_1 \leq +\infty)$ is a fixed interval of time, x is an n-vector (the state variables), u is an m-vector (the controls), θ is a number (the parameter), and the functions f^0: $R^{n+m+1} \to R^1$, f: $R^{n+m+1} \to R^n$, $\emptyset^i$: $R^{n+1} \to R^{r_i}(0 \leq r_i \leq n)$, g: $R^{m+1} \to R^k(k \geq 0)$ are all assumed continuously differentiable with respect to the arguments, R^j being a j-dimensional space. Let us put

$$\dot{v}(t) = f^0(x(t),u(t),\theta) \qquad (t_0 \leq t \leq t_1),$$

$$(7) \qquad v(0) = 0,$$

so that

$$(3a) \qquad v = v(t_1).$$

If the functions $u(t,\theta)$, $x(t,\theta)$, and $v(t,\theta)$ maximize (3a) subject to (4)-(7) for a given parameter θ, they are called optimum. In addition, we call such $u(t,\theta)$ optimal control.

The appearance of θ in the above functions reflects the fact that, in general, an optimum depends on it. Comparative dynamics deals with how a change in θ affects an optimum. First of all, it is easily established that for each t the set of optimum solutions $u(t,\theta)$ and $x(t,\theta)$ is upper semicontinuous in θ (with respect to the inclusion relation), since everything is continuous in θ. Furthermore, the objective function

(3b) $\qquad v = v(t_1) = v(t_1,\theta)$

is continuous in θ. In the present paper, we shall focus our attention on the differentiability of an optimum solution with respect to θ, assuming its uniqueness.

It is seen that without losing generality the differentiability of an optimum solution may be examined with the assumption $\theta = 0$. For simplicity, when θ is set equal to zero, we may suppress the number 0 for θ in the argument of a function. Thus, $f(x,u) = f(x,u,0)$, $x(t) = x(t,0)$, etc.

Next, we state Pontryagin's optimum condition for the control problem.[7,8] First of all, define the Hamiltonian

(8) $\qquad H(p,x,u,\theta) \equiv f^0(x,u,\theta) + p \cdot f(x,u,\theta),$

and its maximum in u subject to (5)

(9) $\qquad M(p,x,\theta) \equiv \underset{g(u,\theta)\geqq 0}{\text{Max}} \quad H(p,x,u,\theta)$

where p is an n-vector (<u>the auxiliary variables</u>) and the dot denotes the inner product. If $u(t)$ and $x(t)$ is an optimum for a problem in which $\theta = 0$, then there exists a nontrivial function $p(t)$ on (t_0,t_1) such that

(the maximum principle)

(10) $\qquad H(p(t),x(t),u(t)) = M(p(t),x(t)),$

(the auxiliary differential equations)

(11) $\qquad \dot{p}(t) = -H_x(p(t),x(t),u(t)) \qquad (t_0 \leqq t \leqq t_1),$

(the transversality conditions)

(12) $\qquad p(t_i) \in L^i \equiv L^i(x(t_i)) \qquad (i = 0,1),$

provided that

(13) $\qquad \mathrm{rank}(\emptyset_x^i(x(t_i)) = r_i \qquad (i = 0,1),$

where H_x, $\emptyset_x^i$ are, respectively, the partial derivatives of H, $\emptyset^i$ with respect to x; H_x will be treated as an n-vector and $\emptyset_x^i$ as an (r_i,n)-matrix. Further, L^i is the subspace of R^n spanned by the row-vectors of $\emptyset_x^i(x(t_i))$.

The optimum solution characterized by the above condition ranges over a very wide variety. In order to isolate a class of optimum solutions which can be studied by means of comparative dynamics, we define a regular optimum solution in the following way:

DEFINITION:

(a) A pair (p,x) is _regular for_ θ, if the control u satisfying the maximum principle

(14) $\qquad H(p,x,u,\theta) = M(p,x,\theta)$

is unique and is a continuously differentiable function

(15) $\qquad u = u(p,x,\theta)$

in a neighborhood of (p,x,θ).

(b) An optimum control $u(t,\theta)$ is _regular at_ t _for_ θ, if $(p(t,\theta), x(t,\theta))$ is regular for θ and $u(t,\theta) = u(p(t,\theta), x(t,\theta),\theta)$.

(c) An optimum control $u(t,\theta)$ is <u>regular on</u> (t_0,t_1) <u>for</u> θ, if the optimum control is regular for θ at t_0, at t_1, and at each t of (t_0,t_1) except for a <u>finite</u> number of points, say s_j (called <u>switching time-points</u>) $(j=1,\ldots,q;\ q \geqq 0)$, where $t_0 < s_1 < s_2 < \ldots < s_q < t_1$.

It is easily seen that not all optimum solutions are regular. In control problems encountered in applications, however, optimum solutions usually turn out to be regular. In the sequel, we shall be concerned only with regular optimum solutions.

The concept of regularity defined above has a close relation to (in fact, is originated from) the method of phase-diagrams, used commonly in applications for solving control problems. The following is a typical regular optimum for a problem with $\theta = 0$ described in terms of phase-diagrams: The whole interval (t_0,t_1) is divided into $(q+1)$ subintervals by the switching time-points $s_1,\ldots,s_q$. We consider an optimal control $u(t)$ which is regular (hence smooth) at the interior of any subinterval. It may be discontinuous at a switching time-point. On the other hand, the (p,x)-space is divided into regions of regular points, each of which is an open set (from (a) of the definition of regularity). The path of the optimum solution $(p(t),\ x(t))$ starts at an interior point of a region at $t=t_0$, crosses its boundary at $t=s_1$, stays in the interior of another region for $s_1 < t < s_2$, crosses its boundary at $t=s_2$, and so on. It terminates at an interior point of some region at $t=t_1$. The path is smooth in the interior of any region. At a boundary point it is continuous but may not be smooth. If we characterize the boundary of regions which the path crosses at $t=s_j$ by the equation $h^j(p,x) = 0$, then

(16) $h^j(p(s_j),\ x(s_j)) = 0$ $(j=1,\ldots g)$.

We shall call $(p(s_j), x(s_j))$ <u>switching point</u>. Figures 1a and 1b illustrate a regular optimum solution for a case of n=m=1 and q=2.

There are several reasons that an optimum control "switches" at a time-point, say t = s. This might arise from the fact that the set of effective constraints of (5) is changed from one to another at t = s. Also, it might arise from the fact that the control satisfying (14) jumps from a local maximum to another local maximum at t = s. We shall study the properties of the optimum control at a switching time-point more in detail in Section IV

We are now able to state our problem. First, to simplify the notation, let us introduce

$$z \equiv (p\ x),$$

(17) $F(z,\theta) \equiv (-H_x(z,u(z,\theta),\theta)\quad f(x,u(z,\theta),\theta))$

where z is a (2n)-vector and F: $R^{2n+1} \to R^{2n}$ is a continuously differentiable function for regular z = (p,x). It may be discontinuous at a point satisfying $h^j(z) = 0$. Then, the constraints (4)-(6), the optimum condition (10)-(12), and the switchings (16) for a problem in which $\theta = 0$ can be reduced to:

> <u>(the original and the auxiliary differential equations together with the maximum principle)</u>

(18) $\dot{z}(t) = F(z(t))$

for regular t in (t_0,t_1) (i.e., for all t except $s_1,\ldots,s_q$),

> <u>(the switchings)</u>
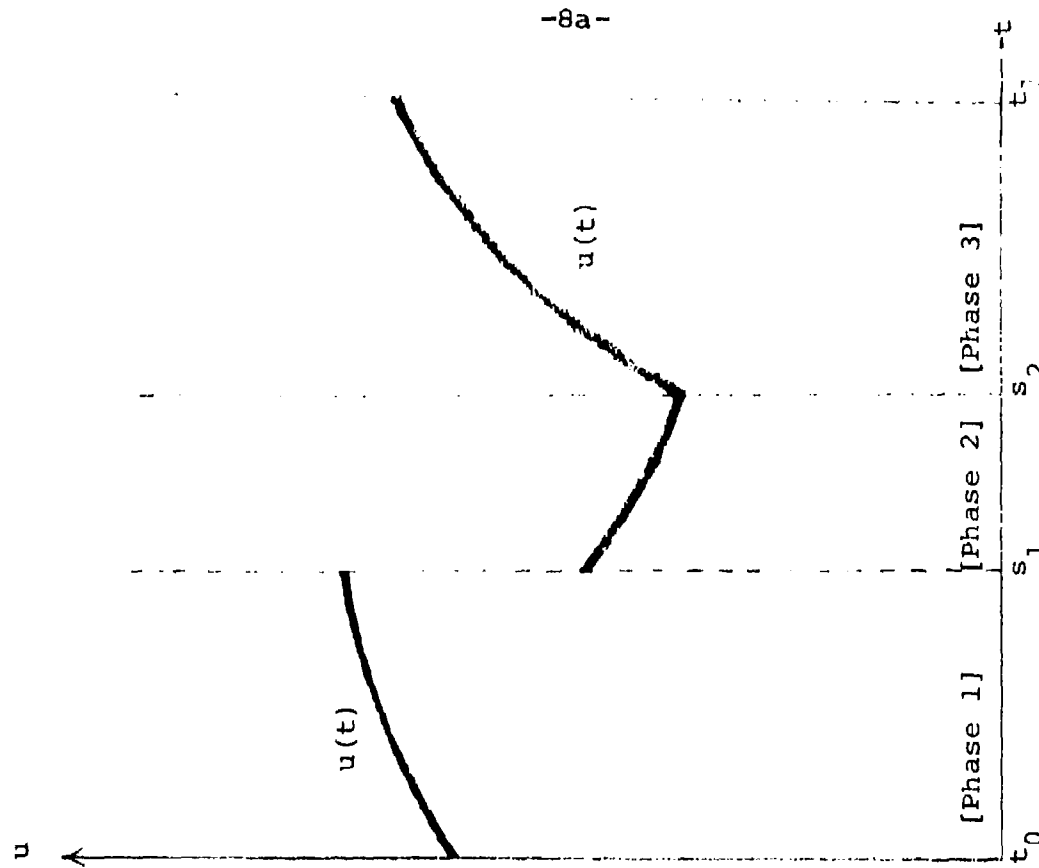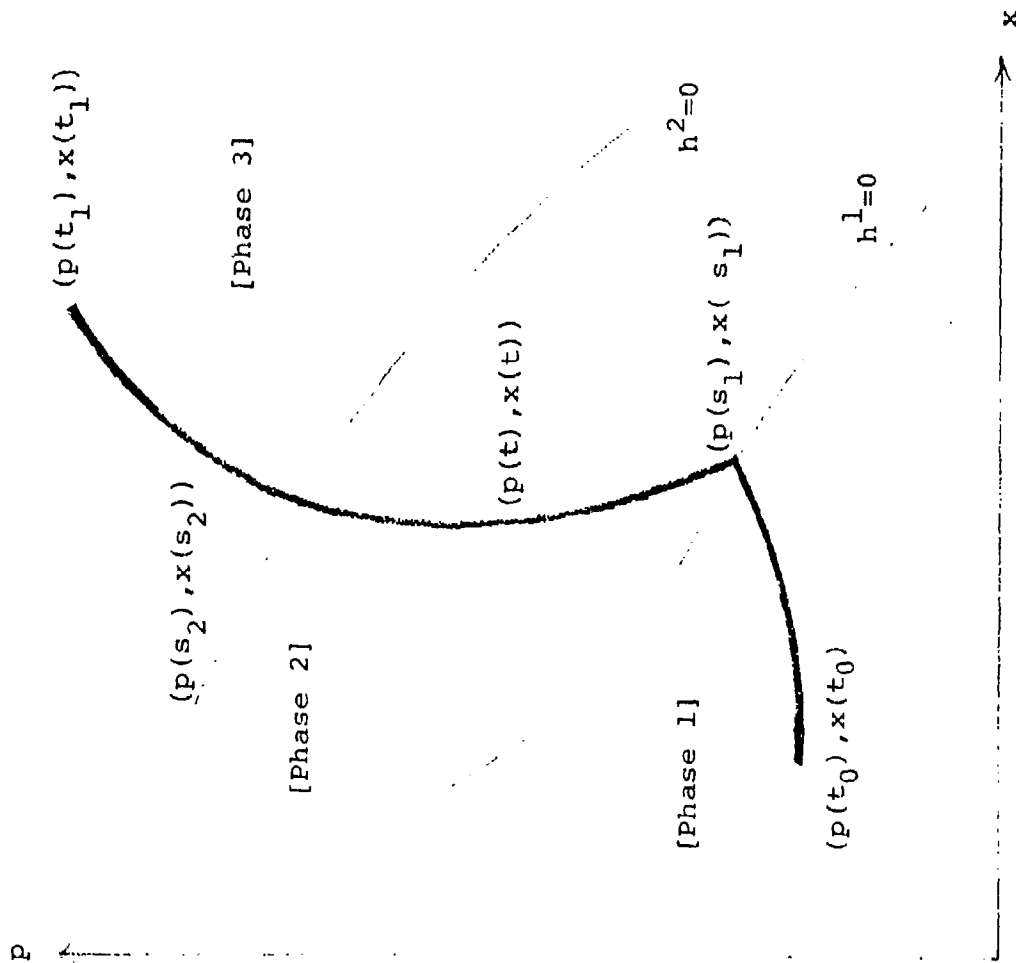
(19) $h^j(z(s_j)) = 0 \qquad (j=1,\ldots,q),$

Figure 1b: Example of a regular optimum control.

Figure 1a: Example of a regular optimum solution.

(the end and the transversality conditions)

(20) $\psi^i(z(t_i)) = 0$ (i=0,1),

where ψ^i: $R^{2n+1} \rightarrow R^n$ is a function of (z,θ) and (20) is equivalent to (5) and (12). To see that this is possible, it suffices to observe from (13) and the definition of L_i that (12) is equivalent to $(n-r_i)$ linear constraints on $z(t_i)$.

Our task will then be to investigate how the function $z(t) \equiv z(t,0)$ is shifted when the parameter θ is changed near $\theta = 0$. In Section IV, we shall obtain a set of conditions sufficient for $z(t,\theta)$ to be differentiable with respect to θ at $\theta = 0$ and formulas to compute the derivative.

It is noted that the condition (18)-(20) is only necessary for optimum. Hence it is possible that $z(t) = z(t,0)$ satisfying (18)-(20) is an optimum but $z(t,\theta)$ satisfying equations like (18)-(20) for a small $\theta \neq 0$ is not an optimum. If this is the case, the derivative of $z(t,\theta)$ with respect to θ at $\theta = 0$, though it exists and is computable, does not describe the shift of the optimum solution. Such a case may arise in applications, even if the optimum for $\theta = 0$ is regular. (If the optimum solution is not unique at $\theta = 0$, then usually this will be the case.) It seems that there is no systematic way of dealing with this kind of complexity; it can be analyzed case by case only.[9] Yet it is true that in many applications the function $z(t,\theta)$ satisfying (18)-(20) remains as the unique optimum for all θ near $\theta = 0$. In such an "interior-maximum" case, our method will be useful directly for comparative dynamics.

III. Variational Differential Equations

In this section, we shall state two lemmas on variational differential equations. The first is well known in the theory of ordinary differential equations. The second is an extension of the first to the case where the equations are allowed to be discontinuous in state variables. The notation in this section is independent of that in the previous and the following sections.

Consider the following system of differential equations together with an initial condition, both containing a parameter θ:

(21) $\dot{x} = F(x,\theta)$,

(22) $x(\tau(\theta)) = \xi(\theta)$,

where x and ξ are vectors, F a vector-valued function, τ denotes the initial time, and ξ the initial point of the state variable x. The functions F, τ, ξ are all continuously differentiable. Further, F_x, F_θ, τ_θ, etc., are (partial) derivatives, and $F(x,0) \equiv F(x)$, $\xi(0) \equiv \xi$, etc. It is assumed that a fixed interval of time $T \equiv (t_0, t_1)$ is given and that

(23) $t_o < \tau(\theta) < t_1$.

LEMMA 1 (Peano): Suppose that a solution $x(t)$ of (21) and (22) in which $\theta = 0$ exists on the entire interval T:

(21a) $\dot{x}(t) = F(x(t))$, $(t \in T)$,

(22a) $x(\tau) = \xi$.

Then, there exists a positive number, say $\hat{\theta}$, for which the following is true:

(i)   For each $\theta\,(\,|\theta\,|\,<\,\hat{\theta})$, a unique solution $x(t,\theta)$ of (21), (22) exists on T;

(ii)   For each $\theta\,(\,|\theta\,|\,<\,\hat{\theta})$ and each $t\,(t\,\epsilon\,T)$, the following expressions exist and are continuous in (t,θ):

(24) $\dot{x}(t,\theta) \equiv \dfrac{\partial x(t,\theta)}{\partial t}$

$x_\theta(t,\theta) \equiv \dfrac{\partial x(t,\theta)}{\partial \theta}$,

$\dot{x}_\theta(t,\theta) \equiv \dfrac{\partial^2 x(t,\theta)}{\partial t \partial \theta} = \dfrac{\partial^2 x(t,\theta)}{\partial \theta \partial t}$.

(iii)   The derivative $x_\theta(t) \equiv x_\theta(t,0)$ satisfies the system of variational equations

(25) $\dot{x}_\theta(t) = F_x(x(t))x_\theta(t) + F_\theta(x(t))$, on T,

and the initial condition

(26) $x_\theta(\tau) = \xi_\theta - \dot{x}(\tau)\tau_\theta$.[10/]

Next, let $h(x,\theta)$ be a scalar-valued function. We shall deal with a system of differential equations with discontinuities in the state variables specified in the following way:

(27) $\dot{x} = F(x,\theta)$,   if $h(x,\theta) \geqq 0$,

          $G(x,\theta)$,   if $h(x,\theta) < 0$,

(28) $x(\tau(\theta)) = \xi(\theta)$,

where F and G are vector-valued functions. We assume that the functions F, G, h, τ, and ξ are all continuously differentiable with respect to (x,θ) or θ. We asume, for definiteness, that

(29) $h(\xi(0),0) = h(\xi) > 0.$

LEMMA 2:   Suppose that a solution $x(t)$ $(t \in T)$ of (27) and (28) for $\theta = 0$ exists. The function $x(t)$ satisfies:

(30) $x(t)$ is continuous on T, and

there exists a switching time s $(t_0 < s < t_1)$ such that

$\dot{x}(t) = F(x(t)),$

(31) $h(x(t)) > 0,$ if $t_0 \leq t < s,$

(32) $h(x(s)) = 0,$

$\dot{x}(t) = G(x(t)),$

(33) $h(x(t)) < 0,$ if $s < t \leq t_1;$

(34) $x(\tau) = \xi.$

Suppose, further, that

(35) $h_x(x(s)) \cdot F(x(s)) \neq 0,$

$h_x(x(s)) \cdot G(x(s)) \neq 0,$

where the dot denotes the inner product.

Then, there exists a positive number, say $\hat{\theta}$, for which the following is true:

(i)   For each θ $(|\theta| < \hat{\theta})$, there uniquely exists a switching time $s(\theta)$ and a solution $x(t,\theta)$ of (27), (28) on T. The functions $s(\theta)$, $x(t,\theta)$ satisfy:

(36) $\dot{x}(t,\theta) = F(x(t,\theta),\theta),$

$h(x(t,\theta),\theta) > 0,$ if $t_0 \leq t < s(\theta);$

(37) $h(x(s(\theta),\theta),\theta) = 0;$

(38) $x(t,\theta) = G(x(t,\theta),\theta),$

         $h(x(t,\theta),\theta) < 0,$ if $s(\theta) < t \leq t_1;$

(39) $x(\tau(\theta),\theta) = \zeta(\theta).$

(ii) For each θ ($|\theta| < \hat{\theta}$) and each t ($t \in T,$ $t \neq s(\theta)$), the following derivatives exist and are continuous:

(40) $s_\theta(\theta) \equiv \dfrac{ds(\theta)}{d\theta}$,

         $\dot{x}(t,\theta) \equiv \dfrac{\partial x(t,\theta)}{\partial t}$,

         $x_\theta(t,\theta) \equiv \dfrac{\partial x(t,\theta)}{\partial \theta}$,

         $\dot{x}_\theta(t,\theta) \equiv \dfrac{\partial^2 x(t,\theta)}{\partial t \partial \theta} = \dfrac{\partial^2 x(t,\theta)}{\partial \theta \partial t}$.

(iii) The derivatives $x_\theta(t) \equiv x_\theta(t,0)$ $(t \neq s)$ and $s_\theta \equiv s_\theta(0)$ satisfy:

(41) $\dot{x}_\theta(t) = F_x(x(t))x_\theta(t) + F_\theta(x(t)),$ if $t_0 \leq t < s;$

(42) $x_\theta(\tau) = \zeta_\theta - x(\tau)\tau_\theta;$

(43) $\dot{x}_\theta(t) = G_x(x(t))x_\theta(t) + G_\theta(x(t)),$ if $s < t \leq t_1;$

(44) $x_\theta(s+0) = x_\theta(s-0) - s_\theta \cdot [x(s+0) - \dot{x}(s-0)];$

(45) $s_\theta = -\dfrac{h_x(x(s)) \cdot x_\theta(s-0) + h_\theta(x(s))}{h_x(x(s)) \cdot \dot{x}(s-0)}$

         $= -\dfrac{h_x(x(s)) \cdot x_\theta(s+0) + h_\theta(x(s))}{h_x(x(s)) \cdot \dot{x}(s+0)}$.

(iv) The derivative

(46) $$\frac{dx(s)}{d\theta} \doteq \frac{dx(s(\theta),\theta)}{d\theta}\Bigg|_{\theta = 0}$$

exists and is given by

(47) $$\frac{dx(s)}{d\theta} = \dot{x}(s-0)\cdot s_\theta + x_\theta(s-0)$$

$$= \dot{x}(s+0)\cdot s_\theta + x_\theta(s+0).\underline{11/}$$

Figure 2 illustrates LEMMA 2. It is noted that the formulas (41) through (45) are two successive initial-value problems of linear differential equations; a program capable to solve an initial-value problem of linear differential equations can also solve (41)-(45). For, we may first solve (41) for $x_\theta(t)$ on (t_0,s) given the initial condition (42). From this we may compute $x_\theta(s-0)$, and hence s_θ by means of (45). Then, we may proceed to solve (43) for $x_\theta(t)$ on (s,t_1) given the initial condition $x_\theta(s+0)$ computed from (44). (Note that the original solution $x(t)$ is regarded as fixed, during the time we are solving (41)-(45).)

Equation (47) provides a formula to compute the shift of the switching point $x(s(\theta),\theta)$; the first term of the right side expresses its shift arising from the change in the switching time-point $s(\theta)$ and the second that arising directly from the change in the function $x(t,\theta)$.

IV. Comparative Dynamics

In Section II we formulated Pontryagin's condition satisfied by a regular optimum solution $z(t)$ of a control problem into the system (18)-(20). This system is a two-point boundary-value
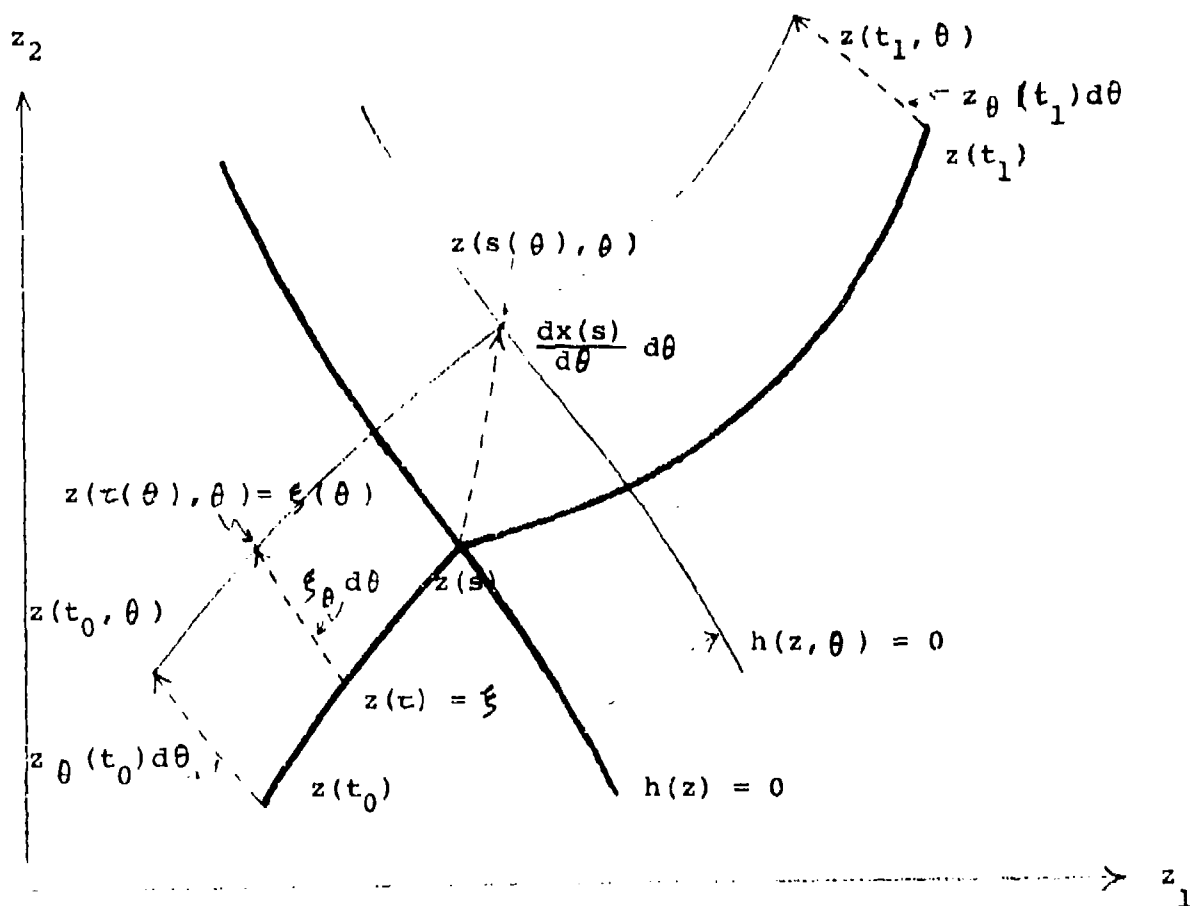
Figure 2: Shift of the solution of differential equations with discontinuities along a curve h = 0.

problem of ordinary differential equations possibly with discon-
tinuities in the state variables. In the present section we shall
make use of LEMMA 2 of the previous section to present a set of
conditions sufficient for the function $z(t) \equiv z(t,0)$ to be differ-
entiable with respect to θ near $\theta = 0$ and to obtain formulas to
compute the derivative:

THEOREM 1. Suppose that a unique (regular) solution $z(t)$ $(t_0 \leq$
$t \leq t_1)$ of (18)-(20) exists. The function $F(z,\theta)$ is assumed con-
tinuously differentiable for all (z,θ) such that θ is near 0 and
z is regular with respect to θ. The functions $h^j(z,\theta)$ and $\psi^i(z,\theta)$
are also continuously differentiable in neighborhoods of $(z(s_j),0)$
and $(z(t_i),0)$, respectively. Suppose further that the following
conditions are satisfied by $z(t)$:

(48) $\qquad h_z^j(z(s_j)) \cdot F(z(s_j \pm 0)) \neq 0 \qquad (j=1,\ldots,q)$

(49) $\qquad |A| \equiv \left| \begin{matrix} A^0 \\ \\ A^1 \end{matrix} \right| \equiv \left| \begin{matrix} \psi_z^0 \\ \\ \psi_z^1 \cdot y_\eta^1 \end{matrix} \right| \neq 0,$

where A is a $(2n,2n)$-matrix, $A^0 \equiv \psi_z^0$ and $A^1 \equiv \psi_z^1$. y_η^1 are $(n,2n)$-

matrices, $\psi_z^i \equiv \psi_z^i(z(t_i))$ is the derivative of $\psi^i(z,\theta)$ with respect

to z at $(z(t_i),0)$, $y_\eta^1 \equiv y_\eta(t_1)$ is a $(2n,2n)$-matrix to be defined

later by (66), and $|A|$ is the determinant of A.

Then, there exists a positive number, say $\hat{\theta}$, for which the
following is true:

(i) For each θ $(|\theta| < \hat{\theta})$, there uniquely exists switching

time-points $s_j(\theta)$ $(j=1,\ldots,q)$ and a unique regular solution $z(t,\theta)$ satisfying

(50) $\quad \dot{z}(t,\theta) = F(z(t,\theta),\theta) \quad$ for all $t \neq s_j(\theta)$,

(51) $\quad h^j(z(s_j(\theta),\theta),\theta) = 0 \quad (j=1,\ldots,q)$,

(52) $\quad \psi^i(z(t_i,\theta),\theta) = 0 \quad (i=0,1)$.

(ii) The derivatives

(53) $\quad s_{j\theta} \equiv s_{j\theta}(0) \equiv \dfrac{d}{d\theta} s_j(\theta)\Big|_{\theta=0}$

(54) $\quad z_\theta(t) \equiv z_\theta(t, \quad \equiv \dfrac{\partial}{\partial\theta} z(t,\theta)\Big|_{\theta=0} \quad (t \neq s_j)$,

(55) $\quad \dfrac{d}{d\theta} z(s_j) \equiv \dfrac{d}{d\theta} z(s_j(\theta),\theta)\Big|_{\theta=0}$

exist and satisfy

(56) $\quad \dot{z}_\theta(t) = F_z(z(t))\cdot z_\theta(t) + F_\theta(z(t)) \quad (t \neq s_j)$

(57) $\quad z_\theta(t_0) = -A^{-1}\cdot B$,

(58) $\quad z_\theta(s_j+0) = z_\theta(s_j-0) - s_{j\theta}\cdot[\dot{z}(s_j+0) - \dot{z}(s_j-0)]$,

(59) $\quad s_{j\theta} = - \dfrac{h_z^j(z(s_j))\cdot z_\theta(s_j{}^{\pm}0) + h_\theta^j(z(s_j))}{h_z^j(z(s_j))\cdot \dot{z}(s_j{}^{\pm}0)}$

(60) $\quad \dfrac{d}{d\theta} z(s_j) = \dot{z}(s_j{}^{\pm}0)\cdot s_{j\theta} + z_\theta(s_j{}^{\pm}0) \quad (j=1,\ldots,q)$,

where

(61) $\quad B \equiv \begin{bmatrix} B^0 \\ B^1 \end{bmatrix} \equiv \begin{bmatrix} \psi_z^0\cdot y_\theta^0 + \psi_\theta^0 \\ \psi_z^1\cdot y_\theta^1 + \psi_\theta^1 \end{bmatrix}$

is a $(2n)$-vector, $B^i = \psi^i_z \, y^i_\theta + \psi^i_\theta$ is an n-vector, $\psi^i_\theta = \psi^i_\theta(z(t_i))$

is the derivative of $\psi^i(z,\theta)$ with respect to θ at $(z(t_i),0)$ and

$y^i_\theta = y_\theta(t_i)$ is a $(2n)$-vector to be defined later by (66).

Some comments on this theorem follow. First, the assumption of continuous differentiability of $F(z,\theta)$ will be satisfied if the derivatives f_x, f_u, f_θ, f_{xx}, f_{xu}, $f_{x\theta}$, f^0_x, f^0_u, f^0_θ, f^0_{xx}, f^0_{xu}, and $f^0_{x\theta}$ exist and are continuous (see (17), (8), (15)); we need the second-order derivatives to obtain s_θ and $z_\theta(t)$. This resembles the case of simple maximization shown in (2).

Second, assumptions similar to (48) appeared in LEMMA 2 previously (see (35)). We need (48) and (49) for the sake of regularity (not in the sense defined in Section II but in the general sense). If (48) does not hold, then it is possible that the number of switching time-points (= q) changes when θ varies near 0 or even that the optimum solution $z(t,\theta)$ is no more regular (in the sense defined in Section II) for a small θ. If (49) is not satisfied, then it is possible that $z(t,\theta)$ is no more unique for a small θ so that the derivative $z_\theta(t)$ may be undefined. We, however, know that usually these singularities do not arise in applications. In simple maximization (2), a regularity condition ($\left| f_{xx} \right| \neq 0$) was required.

Third, the formulas (56) through (59) are regarded as a succession of $(q+1)$ initial-value problems of linear differential equations. Comments like those immediately after LEMMA 2 may be mentioned here on (56)-(60).

PROOF OF THEOREM 1:

We shall prove the theorem for the case $q = 1$ by using

LEMMA 2. A proof for the case $q = 0$ (no switching) may be obtained simply by replacing LEMMA 2 in the proof below with LEMMA 1 and eliminating all statements concerning switchings. A proof for the case $q > 1$ (multiple switchings) may be obtained first by extending LEMMA 2 to the case of multiple switchings (this is easy; all we need to do is to use LEMMA 2 successively q times) and then by applying the result thus obtained to the theorem in the same way as shown below for the case $q = 1$.

Thus, for the rest of the proof we assume that $q = 1$. For simplicity we write $s_1 = s$, $s_1(\theta) = s(\theta)$, $h^1 = h$, etc.

We now proceed to prove the theorem for the case $q = 1$. First of all, choose any $\tau (t_0 \leq \tau \leq t_1)$ at which the optimum is regular. Then, from LEMMA 2 and the assumptions of the theorem ((18), (19), (48), and others), we can assert that there exist functions $\sigma(\eta,\theta)$ and $y(t,\eta,\theta)$ satisfying

(62) $\dot{y}(t,\eta,\theta) = F(y(t,\eta,\theta),\theta)(t_0 \leq t \leq t_1, \ t \neq \sigma(\eta,\theta))$,

(63) $y(\tau,\eta,\theta) = \eta$,

(64) $h(y(\sigma(\eta,\theta),\eta,\theta),\theta) = 0$,

where y is a $(2n)$-vector (state variables), η a $(2n)$-vector representing the initial value for y (parameters), and θ a number (a parameter). The assertion above can be shown by applying LEMMA 2 $(2n+1)$-times for the $(2n)$ components of η and θ, each being successively substituted for the parameter θ of LEMMA 2. Furthermore, LEMMA 2 shows that the derivatives

(65) $\sigma_\eta \equiv \dfrac{\partial}{\partial \eta} \sigma(z(\tau),0)$,

$\sigma_\theta \equiv \dfrac{\partial}{\partial \theta} \sigma(z(\tau),0)$,

(66) $y_\eta(t) \equiv \dfrac{\partial}{\partial \eta} y(t,z(\tau),0),$

$Y_\theta(t) \equiv \dfrac{\partial}{\partial \theta} y(t,z(\tau),0)$ $(t \neq s)$

exist and satisfy

(67) $\dot{Y}_\eta(t) = F_z(z(t)) \cdot y_\eta(t)$ $(t \neq s)$

$Y_\eta(\tau) = I,$

$\sigma_\eta = \dfrac{h_z(z(s)) \cdot y_\eta(s\pm 0)}{h_z(z(s)) \cdot \dot{y}(s\pm 0)},$

$Y_\eta(s+0) = y_\eta(s-0) - [\dot{y}(s+0) - \dot{y}(s-0)]\sigma_\eta \; ;$

(68) $\dot{Y}_\theta(t) = F_z(z(t)) \cdot y_\theta(t) + F_\theta(z(t))$ $(t \neq s),$

$Y_\theta(t_0) = 0,$

$\sigma_\theta = \dfrac{h_z(z(s)) \cdot y_\theta(s\pm 0) + h_\theta(z(s))}{h_z(z(s)) \cdot \dot{y}(s\pm 0)}$

$Y_\theta(s+0) = y_\theta(s-0) - [\dot{y}(s+0) - \dot{y}(s-0)] \cdot \sigma_\theta \; ,$

where σ_η is considered as a $(2n)$-vector, σ_θ as a scalar, $y_\eta(t)$ as a $(2n,2n)$-matrix, $y_\theta(t)$ as a $(2n)$-vector, and I is an identity matrix.

Let us next turn to the conditions (20). For simplicity, we shall write as

(69) $\psi_z^i \equiv \psi_z^i(z(t_i)), \quad \psi_\theta^i \equiv \psi_\theta^i(z(t_i)),$

$y_\eta^i \equiv y_\eta(t_i), \quad y_\theta^i \equiv y_\theta(t_i),$ etc.

Then, by (20), (49), and the implicit function theorem, a continuously differentiable (2n)-vector-valued function $\eta(\theta)$ satisfying

(70) $\qquad \psi^i(y(t_i,\eta(\theta),\theta),\theta) = 0 \qquad (i=0,1)$

exists for θ sufficiently small, and satisfies

(71) $\qquad \eta_\theta \equiv \eta_\theta(0) = -A^{-1} \cdot B,$

where

(72) $\qquad \eta_\theta(0) \equiv \left. \frac{d\eta(\theta)}{d\theta} \right|_{\theta = 0}.$

The derivative (71) may be obtained by differentiating (70) with respect to θ and then setting θ equal to 0.

Now define

(73) $\qquad s(\theta) \equiv \sigma(\eta(\theta),\theta),$

$\qquad\qquad z(t,\theta) \equiv y(t,\eta(\theta),\theta).$

Then, substitution of (73) into (62), (64), and (70) yields (50)-(52). We have thus shown that a solution of (50)-(52) exists for θ sufficiently small.

Next we shall show that this solution is unique. Let $s(\theta)$ and $z(t,\theta)$ be an arbitrary solution of (50)-(52). Define $\eta(\theta)$ by

(74) $\qquad \eta(\theta) \equiv z(\tau,\theta).$

Then, (73) must hold, since the pair $s(\theta)$, $z(t,\theta)$ satisfies the system (50), (74), (51), and the pair $\sigma(\eta(\theta),\theta)$, $y(t,\eta(\theta),\theta)$ satisfies the system (62)-(64) with η replaced by $\eta(\theta)$, and these two systems are identical. (Note that the solutions of these system are unique by LEMMA 2 and the assumption (48) of the theorem.) If we substitute the second equation of (73) into (52), then we

know that the function $\eta(\theta)$ is determined uniquely by the implicit function theorem and the assumption (49) of the theorem. We then conclude that the solution $s(\theta)$ and $z(t,\theta)$ is uniquely determined by (50)-(52). This completes proving the first part of the theorem.

Proving the second part is straightforward. Since the right sides of (73) are differentiable with respect to θ $(t \neq s(\theta))$ (see (65), (66), and (71)), so are the left sides. Differentiation of (73) then yields (putting $\theta = 0$)

$$(75) \qquad s_\theta = \sigma_\eta \eta_\theta + \sigma_\theta ,$$

$$z_\theta(t) = y_\eta(t)\eta_\theta + y_\theta(t) \qquad (t \neq s),$$

so that

$$(76) \qquad \dot{z}_\theta(t) = \dot{y}_\eta(t)\eta_\theta + \dot{y}_\theta(t) \qquad (t \neq s).$$

If we multiply (67) by η_θ, add this to (68), and use (75) and (76), then we obtain (56) through (59). Equation (60) can be proved in an analogous way. (QED)

We have thus obtained the formulas for computing the derivatives s_θ and $z_\theta(t)$. It is straightforward (though tedious) to write down the equation (56) in terms of the original notation $f(x,u,\theta)$, $f^0(x,u,\theta)$, and $u(p,x,\theta)$ (see (17), (15), and (8)). In applications, however, it is easier to differentiate (50) directly with respect to θ and to put $\theta = 0$ than to refer to the formula of (56) stated in terms of the original notation.

In the foregoing discussions the original control problem was of the type of variable end-points; the end and the transversality conditions were expressed by (20). In the following, we shall consider two special cases:

(a) The initial point of the state variable is fixed at a point determined by the parameter, and the end point is left free: $\psi^0(z,\theta) \equiv \psi^0(p,x,\theta) = x - \emptyset(\theta)$, and $\psi^1(z,\theta) \equiv \psi^1(p,x,\theta) = p$. The conditions (20) become

(20a) $x(t_0,\theta) - \emptyset(\theta) = 0,$

$p(t_1,\theta) = 0.$

where $\emptyset$ is an n-vector-valued function of θ, which is smooth near $\theta = 0$. Let us put

(54a) $z_\theta(t) \equiv (p_\theta^*(t) \; x_\theta^*(t))$ $(t \neq s),$

(63a) $\eta \equiv (\pi,\xi), \; \tau = t_0,$

(66a) $y_\eta(t) \equiv \begin{bmatrix} p_{\pi(t)} & p_{\xi(t)} \\ x_{\pi(t)} & x_{\xi(t)} \end{bmatrix}'$

$y_\theta(t) \equiv (p_\theta(t) x_\theta(t))$ $(t \neq s).$

We then obtain

$$A = \begin{bmatrix} p_\pi(t_1) & p_\xi(t_1) \\ 0 & I \end{bmatrix}, \quad B = \begin{bmatrix} p_\theta(t_1) \\ -\emptyset_\theta \end{bmatrix},$$

so that the regularity condition (49) and the initial condition (57) for this case can be written as

(49a) $\left| p_\pi(t_1) \right| \neq 0,$

(57a) $p_\theta^*(t_0) = -p_\pi(t_1)^{-1} \cdot (p_\xi(t_1)\emptyset_\theta + p_\theta(t_1)),$

$x_\theta^*(t_0) = \emptyset_\theta ,$

where $\emptyset_\theta$ is the derivative of $\emptyset(\theta)$ at $\theta = 0$.

(b) The initial and the terminal values of the state variables are completely determined by the parameter: $\psi^i(z,\theta) \equiv \psi^i(p,x,\theta) = x - \emptyset^i(\theta)$ $(i = 0,1)$, say. For this case, we obtain

$$A = \begin{bmatrix} 0 & I \\ x_\eta(t_1) & x_\xi(t_1) \end{bmatrix}, \quad B = \begin{bmatrix} -\emptyset_\theta^0 \\ x_\theta(t_1) - \emptyset_\theta^1 \end{bmatrix},$$

and hence

(49b) $\quad \left| x_\eta(t_1) \right| \neq 0,$

(57b) $\quad p_\theta^*(t_0) = -x_\eta(t_1)^{-1} \cdot (x_\xi(t_1)\emptyset_\theta^0 + x_\theta(t_1) - \emptyset_\theta^1),$

$$x_\theta^*(t_0) = \emptyset_\theta^0 .$$

V. Sensitivity of the Objective Function

In this section, we shall examine the effect of the parameter on the maximized value of the objective function. To do this, we compute the derivative of (3) or (3a) with respect to θ at $\theta = 0$:

(77) $\quad v_\theta \equiv v_\theta(t_1) \equiv \dfrac{d}{d\theta} v(t_1,\theta) \bigg|_{\theta = 0}$.

If the derivatives $x_\theta(t)$ and $p_\theta(t)$ have already been obtained by means of the comparative dynamics as explained in the previous section (for simplicity, we write $x_\theta(t)$, $p_\theta(t)$ for $x_\theta^*(t)$, $p_\theta^*(t)$, respectively), we may calculate (77) directly by differentiating (3) (Note that $p_\theta(t)$, $x_\theta(t)$, and $u_\theta(t)$ exist on (t_0,t_1) except at a finite number of points.):

(78) $\quad v_\theta = \displaystyle\int_{t_0}^{t_1} [f_x^0(t)x_\theta(t) + f_u^0(t)u_\theta(t) + f_\theta^0(t)]\,dt,$

where $f_x^0(t) \equiv f_x^0(x(t),\ u(t))$,

$$u_\theta(t) \equiv u_x(p(t),x(t))x_\theta(t) + u_p(p(t),x(t))p_\theta(t) + \frac{\partial u(p(t),x(t))}{\partial \theta}$$

(see (15)), etc. In the following, we shall abbreviate arguments of a function as done above.

It is possible to obtain, from Pontryagin's condition, further properties of the derivative v_θ. In order to do this, define

$$(79) \qquad w(t) = H_u(t)u_\theta(t) + H_\theta(t) \qquad (t \neq s_j)$$

$$W(t) = \int_{t_0}^{t} w(\tau)d\tau.$$

We shall prove

THEOREM 2: If the condition assumed for THEOREM 1 is satisfied, then

$$v_\theta(t_1) = p(t_0)x_\theta(t_0) - p(t_1)x_\theta(t_1) + W(t\ ).$$

Furthermore, the term $p(t_i)x_\theta(t_i)$ vanishes, if the function ϕ^i in the end condition (6) of the original control problem is independent of θ: $\phi^i(x,\theta) = \phi^i(x)$, say.

PROOF: Differentiation of (7) with respect to θ yields

$$(81) \qquad \dot{v}_\theta(t) = f_x^0(t)x_\theta(t) + f_u^0(t)u_\theta(t) + f_\theta^0(t) \qquad (t \neq s_j).$$

In view of (17), we write down the second half of (56):

$$(82) \qquad \dot{x}_\theta(t) = f_x(t)x_\theta(t) + f_u(t)u_\theta(t) + f_\theta(t). \qquad (t \neq s_j)$$

The auxiliary equation (11) is (see (8))

$$(11a) \qquad \dot{p}(t) = - [f_x^0(t) + p(t)f_x(t)] \qquad (t \neq s_j).$$

Forming $[(81) + p(t) \cdot (82) + (11a) \cdot x_\theta(t)]$ and using (8), (79), we get

(83) $\quad \dfrac{d}{dt} (v_\theta(t) + p(t)x_\theta(t)) = w(t) \qquad (t \neq s_j).$

We then obtain (80) by integrating (83) over (t_0, t_1), since $v_\theta(t_0) = 0$ (see (7)).

Next, assuming that $\emptyset^i$ is independent of θ, we write down (6):

(6a) $\quad \emptyset^i(x(t_i, \theta)) = 0.$

Differentiate (6a) with respect to θ and set $\theta = 0$, obtaining

(84) $\quad \emptyset_x^i(x(t_i))x_\theta(t_i) = 0.$

This implies that the vector $x_\theta(t_i)$ is orthogonal to the row vectors of $\emptyset_x^i(x(t_i))$. In view of (12), we conclude that $p(t_i)x_\theta(t_i) = 0$.

$$\text{(QED)}$$

Equation (80) states that v_θ is the sum of three terms, the first and the second being the derivative $x_\theta(t_i)$ multiplied by the auxiliary variable $p(t_i)$ and the third $W(t_1)$. Since $W(t_1)$ is obtained from the Hamiltonian (see (79)), which is also the product of the functions f^0, $f(= \dot{x}(t))$ with $p(t)$, we can recognize that a role of the auxiliary variables $p(t)$ is to relate the state variables to the objective function.

Furthermore, let x^k be the k-th component of the vector x. If we assume that

(85) $\quad x^k(t_0) = \theta$

for the end condition (6) for i=0, and that (85) is the only place that the parameter θ appears in the control problem, then (80)

yields

$$(86) \qquad v_\theta = p^k(t_0),$$

since the last two terms of the right side of (80) vanish. Thus, we have shown:

COROLLARY: if the condition assumed for THEOREM 1 is satisfied, then

$$(87) \qquad \frac{dv}{dx^k(t_0)} = p^k(t_0) \qquad (k=1,\ldots,n),$$

where the derivative takes into account all the adjustments needed to keep the solution optimal. Similarly,

$$(88) \qquad \frac{dv}{dx^k(t_1)} = - p^k(t_1) \qquad (k=1,\ldots,n).$$

The property (87), (88) is well known; $p(t)$ is interpreted to be the marginal contribution of the state variables to the objective function. It is not true, however, that this property holds without qualification.[12/] The corollary above shows a sufficient condition for (87), (88).

VI. <u>An Example (Comparative Dynamics in a Model of Optimal Capital Accumulation)</u>

In the present section, we shall attempt to show how the comparative dynamics developed in the previous sections can be used for economic analyses. A standard model of optimal capital accumulation developed by Cass [3] will be used as an example.[13/] For short, we shall follow his notation and avoid repreating the definition of variables.

To begin with, write Pontryagin's condition for the model

of Cass:[14]/

(18c) $\dot{q} = (\delta - f'(k))q$,

  $\dot{k} = f(k) - c(q)$,

(20c) $k(0) = k^0$, $k(T) = k^T$,

where $u'(c(q)) \equiv q$. For simplicity, it is assumed that $\lambda = 0$ and that $z > 0$ throughout the planning period $(0,T)$ (i.e., no switching).

We shall consider the discount rate δ as the parameter. We then write: $q \equiv q(t,\delta)$, $k \equiv k(t,\delta)$, $c \equiv c(q(t,\delta)) \equiv c(t,\delta)$, etc. In view of (57b) appearing at the end of Section IV, we know that we need $k_{q^0}(T)$ and $k_\delta(T)$, where $q(0) \equiv q^0$. To do this, let us differentiate (18c) and (20c) with respect to q^0 to obtain

$$(67c) \quad \begin{bmatrix} \dot{q}_{q^0} \\ \\ \dot{k}_{q^0} \end{bmatrix} = \begin{bmatrix} \delta - f' & -f''q \\ \\ -c' & f' \end{bmatrix} \cdot \begin{bmatrix} q_{q^0} \\ \\ k_{q^0} \end{bmatrix} ,$$

$$q_{q^0}(0) = 1, \; k_{q^0}(0) = 0 .$$

The sign of each element in the matrix appearing in (67c) is determined from Cass's assumption as follows:[15]/

$$\begin{bmatrix} \dot{q}_{q^0} \\ \\ \dot{k}_{q^0} \end{bmatrix} = \begin{bmatrix} ? & + \\ \\ + & + \end{bmatrix} \cdot \begin{bmatrix} q_{q^0} \\ \\ k_{q^0} \end{bmatrix} .$$

To investigate the path of $q_{q^0}(t)$ and $k_{q^0}(t)$, we construct a phase diagram (see Figure 3). Each arrow in the phase diagram indicates a possible direction of the path. Since the path starts at a point
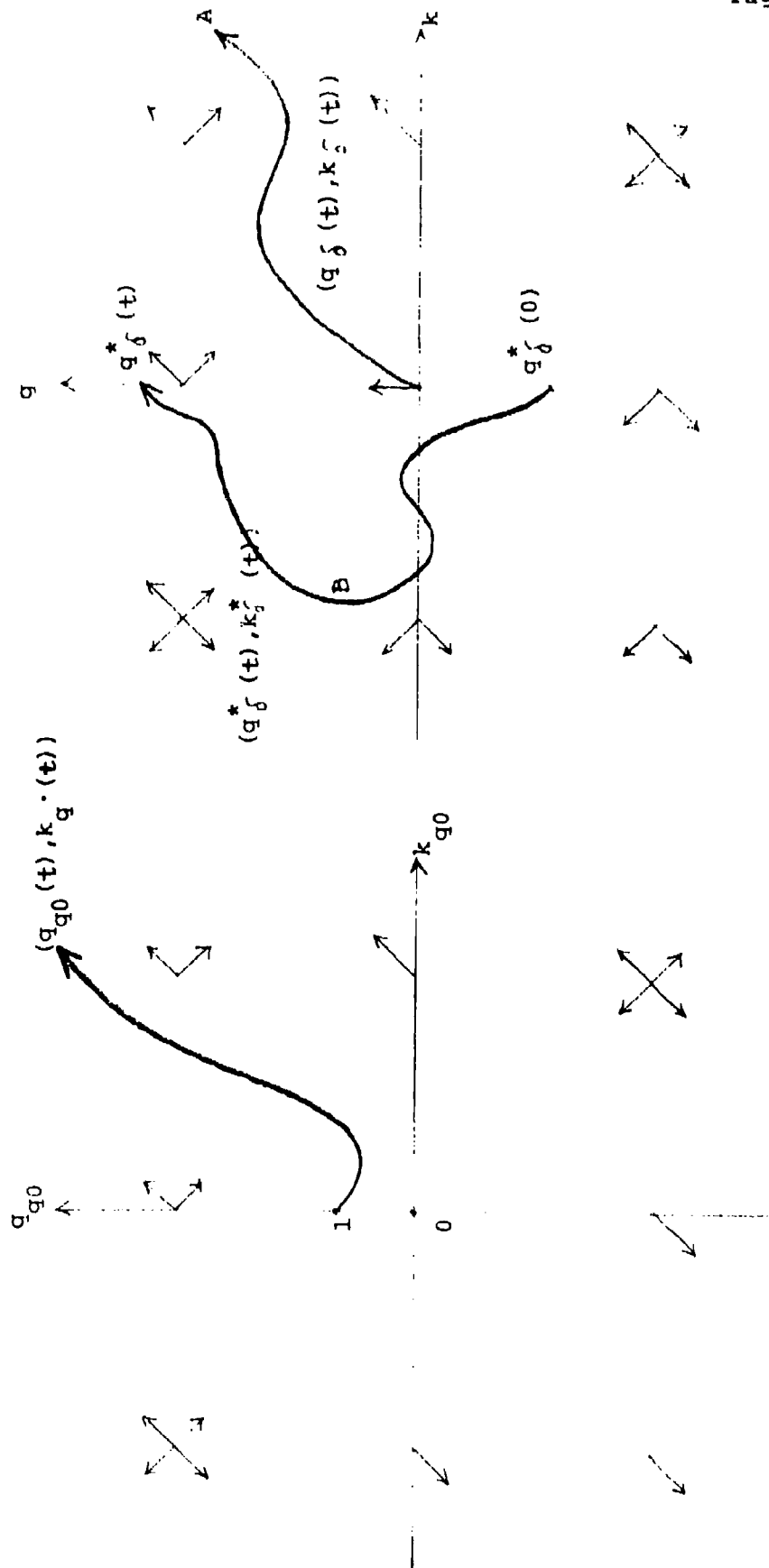
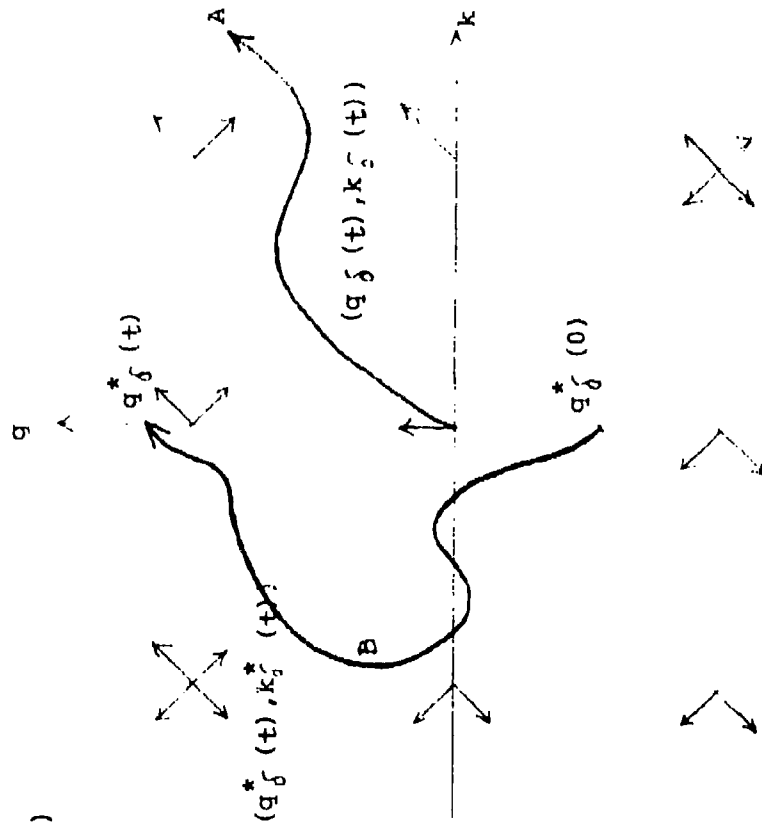Figure 3: Phase-diagram of $(q_{q0}(t), k_{q0}(t))$.

Figure 4: Phase-diagram of $(q_{\delta}(t), k_{\delta}(t))$ and $(q_{\delta}^{*}(t), k_{\delta}^{*}(t))$.

on the upper half of the vertical axis, it is clear from the diagram that

(89) $\quad q_{q^0 0}(t) > 0, \ k_{q^0 0}(t) > 0$ for all t.

Following a similar method, we obtain

$$(68c) \quad \begin{bmatrix} \dot{q}_\delta \\ \dot{k}_\delta \end{bmatrix} = \begin{bmatrix} \delta - f' & -f''q \\ -c' & f' \end{bmatrix} \cdot \begin{bmatrix} q_\delta \\ k_\delta \end{bmatrix} + \begin{bmatrix} q \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} ? & + \\ + & + \end{bmatrix} \cdot \begin{bmatrix} q_\delta \\ k_\delta \end{bmatrix} + \begin{bmatrix} + \\ 0 \end{bmatrix} ;$$

$$q_\delta(0) = 0, \ k_\delta(0) = 0.$$

From this and a phase diagram (see path A in Figure 4), we conclude that

(90) $\quad q_\delta(t) > 0, \ k_\delta(t) > 0$ for all t.

Then, (57b), (89), and (90) yield

(57c) $\quad q^*_\delta(0) = -k_\delta(T)/k_{q^0 0}(T) < 0,$

$k^*_\delta(0) = 0.$

In addition, since $k^*(T, \delta) = k^T$,

(91) $\quad k^*_\delta(T) = 0.$

Let us finally consider variational equations for $q^*_\delta(t)$ and $k^*_\delta(t)$ (which is a special case of (56)). In fact, these equations are obtained by substituting q^*_δ and k^*_δ into (68c). We may use Figure 4 to describe the path of $q^*_\delta(t)$, $k^*_\delta(t)$. In view of the end conditions (57c), (91), we know that the path must start at a

point on the lower half of the vertical axis, and must terminate
at a point on the upper half of it. Hence, it stays within the
second and the third quadrants. It is seen that the path may
cross the horizontal axis more than once, but never the vertical
axis. Path B in Figure 4 is a typical one.

From the above consideration, we can conclude that

$$k^*_\delta(t) < 0, \text{ for all } t,$$

(92)
$$c^*_\delta(t) = c'(q(t)) \cdot q^*_\delta(t) > 0 \text{ for small } t,$$
$$< 0 \text{ for large } t,$$

summarizing the effect of a change in δ on the optimum path.

APPENDIX

Proof of LEMMA 2

1. Let $x(t)$ and $s \in T$ be given as stated in the first part of the LEMMA. We may extend the solution $x(t)$ of (31) slightly beyond the switching time s; [16] i.e., there exists a number $s^1 \in T$ and a function $y(t)$ defined on $[0, s^1]$ such that

(A1) $s < s^1 < 1$,

(A2) $y(t) = x(t)$,   on $[0, s]$

(A3) $\dot{y}(t) = F(y(t))$,   on $[0, s^1]$,

         $y(\tau) = \xi$.

Likewise, for the solution $x(t)$ of (33), there exists a number $s^2 \in T$ and a function $z(t)$ defined on the interval $[s^2, 1]$ such that

(A4) $0 < s^2 < s$,

(A5) $z(t) = x(t)$,   on $[s, 1]$,

(A6) $\dot{z}(t) = G(z(t))$,   on $[s^2, 1]$,

         $z(s) = x(s)$.

2. Consider the system of differential equations

(A7) $\dot{y} = F(y, \theta)$,   on $[0, s^1]$,

and the initial condition

(A8) $y(\tau(\theta)) = \xi(\theta)$.

Then, LEMMA 1 and the construction of the solution $y(t)$ by (A3) immediately yields that, for a sufficiently small θ, the solution $y(t, \theta)$ of (A7), (A8) exists uniquely on the interval $[0, s^1]$ and satisfies (i), (ii), (iii) of LEMMA 1 with x, T replaced respectively by y, $[0, s^1]$.

3. Let us next consider the equation

(A9) $h(y(\sigma,\theta),\theta) = 0,$

where σ is an unknown variable.

Since we have [see (32), (A2)]

(A10) $h(y(s,0),0) \equiv h(y(s)) = h(x(s)) = 0,$

and [see (35)]

(A11) $\left. \dfrac{dh(y(\sigma,0),0)}{d\sigma} \right|_{\sigma=s} = h_x(x(s)) \cdot F(x(s)) \neq 0$

it follows from the implicit function theorem that, for a sufficiently small θ, there exists a function $s(\theta)$ such that

(A12) $s(0) = s,$

(A13) $h(y(s(\theta),\theta),\theta) \equiv 0,$

(A14) $s(\theta)$ is continuously differentiable.

Differentiation of the identity (A13) with respect to θ at $\theta = 0$ yields [see (35)]

(A15) $s_\theta = - \dfrac{h_x(y(s)) \cdot y_\theta(s) + h_\theta(y(s))}{h_x(y(s)) \cdot \dot{y}(s)}$

Furthermore, we have, in view of (A1), (A4),

(A16) $s^2 < s(\theta) < s^1,$

for a sufficiently small θ.

4. We shall next show that, for a sufficiently small θ,

(A17) $h(y(t,\theta),\theta) > 0,$ if $t_0 \leq t < s(\theta).$

For, suppose, on the contrary, that, for each θ^i converging to 0, there exists a t^i such that

(A18) $0 \leq t^i < s(\theta^i)$,

(A19) $h(y(t^i, \theta^i), \theta^i) \leq 0$.

Since $h(y(0,0),0) > 0$ [see (31)], we have $h(y(0,\theta^i),\theta^i) > 0$ for a sufficiently small θ^i. Since $h(y(t,\theta^i),\theta^i)$ is continuous in t, we may redefine t^i so that inequality (A18) and

(A20) $h(y(t^i, \theta^i), \theta^i) = 0$

hold for all i. The range of t^i, as being contained in T, is bounded, so that we may choose a converging subsequence. Let us once more redefine t^i to be such a converging subsequence and de-note its limit by t^*. Passing to the limit of inequality (A18) [see (A12), (A14)], we have exactly two alternatives: $t^* < s$, or $t^* = s$. But if $t^* < s$, (A20) implies that $h(y(t^*,0),0) = h(x(t^*)) = 0$, a contradiction to (31). Hence, let $t^* = s$. In view of (A13), we have $h(y(s(\theta^i),\theta^i),\theta^i) = 0$. Then, (A20) and the mean value theorem yield that

(A21) $h(y(s(\theta^i),\theta^i),\theta^i) - h(y(t^i, \theta^i),\theta^i)$

$$= \left. \frac{dh(y(\sigma,\theta^i),\theta^i)}{d\sigma} \right|_{\sigma=\sigma^i} \cdot [s(\theta^i) - t^i] = 0,$$

where $t^i \leq \sigma^i \leq s(\theta^i)$. Since $t^i \neq s(\theta^i)$ [see (A18)], (A21) implies that

(A22) $\left. \dfrac{dh(y(\sigma,\theta^i),\theta^i)}{d\sigma} \right|_{\sigma=\sigma^i} = h_x(y(\sigma^i,\theta^i),\theta^i) \cdot F(y(\sigma^i,\theta^i),\theta^i) = 0.$

We then obtain, taking the limit of (A22),

(A23) $h_x(y(s,0),0) \cdot F(y(s,0),0) = h_x(x(s)) \cdot F(x(s)) = 0$,

which contradicts to the assumption (35). Therefore, inequality

(A17) holds.

5. Let us next consider the system of differential equations

(A24) $\quad \dot{z} = G(z,\theta)$, $\quad$ on $[s^2,1]$,

with the initial condition

(A25) $\quad z(s(\theta)) = y(s(\theta),\theta)$.

For a sufficiently small θ, $y(s(\theta),\theta)$ is continuously differentiable with respect to θ. Furthermore, $z(t)$ is the solution of (A24), (A25) on the interval $[s^2,1]$ when $\theta = 0$ [see (A6)]. Therefore, all the assumptions of LEMMA 1 are again satisfied when x, F, $\tau(\theta)$, $\xi(\theta)$, T of LEMMA 1 are replaced respectively by z, G, $s(\theta)$, $y(s(\theta),\theta)$, $[s^2,1]$. Then, we may state that, for a sufficiently small θ, the solution $z(t,\theta)$ of (A24), (A25) exists uniquely on the interval $[s^2,1]$ and satisfies (i), (ii), (iii) of LEMMA 1. In particular, (22a) becomes

(A26) $\quad z(s(\theta),\theta) = y(s(\theta),\theta)$.

Also, we obtain, in view of (A25), (26),

(A27) $\quad z_\theta(s) = [y_\theta(s) + \dot{y}(s)s_\theta] - \dot{z}(s)s_\theta$.

Furthermore, it can be shown, as was done in 4., that, for a sufficiently small θ,

(A28) $\quad h(z(t,\theta),\theta) < 0$, $\quad$ if $s(\theta) < t \leq t_1$.

6. Let us now choose a sufficiently small $\hat{\theta}$ for which all the foregoing constructions are valid. Let us define [see (A26)]

(A29) $\quad x(t,\theta) \equiv y(t,\theta)$, $\quad (t_0 \leq t < s(\theta))$,

$\qquad\qquad x(s(\theta),\theta) \equiv y(s(\theta),\theta) = z(s(\theta),\theta)$,

$\qquad\qquad x(t,\theta) \equiv z(t,\theta)$, $\quad (s(\theta) < t \leq t_1)$,

for each θ such that $|\theta| < \hat{\theta}$. We then conclude that the number $\hat{\theta}$ and the functions $s(\theta)$, $x(t,\theta)$ satisfy (i)-(iv) of LEMMA 2. Note that

(A30) $\dot{x}(s-0) = \dot{y}(s)$,

$x_\theta(s-0) = y_\theta(s)$,

$\dot{x}(s+0) = \dot{z}(s)$,

$x_\theta(s+0) = z_\theta(s)$,

and that (44) and (45) can be derived from (A27), (A15), (A30), and (43).

Finally, the two derivatives

(A31) $\left.\dfrac{dy(s(\theta),\theta)}{d\theta}\right|_{\theta=0} = \dot{y}(s)\,s_\theta + y_\theta(s)$

$\left.\dfrac{dz(s(\theta),\theta)}{d\theta}\right|_{\theta=0} = \dot{z}(s)\,s_\theta + z_\theta(s)$,

are defined and coincide [see (A27)]. Equations (A29), (A30), then lead one to (46), (47) immediately.   (QED)

*Acknowledgment

An earlier version of Section III and the Appendix of this paper
was included in the author's doctoral dissertation submitted to
Stanford University, 1968. I owe much to Professor Kenneth J.
Arrow, especially for the results obtained in Section IV.

FOOTNOTES

1/ See, e.g., Koopmans [6], pp. 119-122, a graphical approach to
comparative dynamics in a model of optimal economic growth.
Cass [3], p. 844, also dealt with a comparative dynamics in
optimal growth. Jorgenson [5], pp. 147-151, stated a method
for comparative dynamics in investment theory. Oniki [7] and
[8] made extensive use of the method of this paper for problems
of optimal growth and human investment, respectively.
Outside the economics literature, Barriere [2] set forth in-
vestigating the present subject. His viewpoint, however, is
narrower than ours: first, he analyzed the effects on the ob-
jective function only, while ours covers those on the optimum
control and the state variables as well; second, he did not
state any condition sufficient for the objective function to
be differentiable with respect to a parameter, while ours does

2/ In fact, the maximum principle contains more than what is sup-
plied by the first-order condition in calculus of variations.
In this paper, however, we do not use the implications of
the maximum principle beyond those expressed in terms of first-
order derivatives.

3/ LEMMA 1 of Section III.

4/ Cass [3] used this theorem for comparative dynamics (his foot-
note 5). In his case, however, the optimal control is continu-
ous in state variables; the difficulty stated in the text does
not arise there.

5/   Formally speaking, the assumption that the differential
equations are autonomous is not restrictive, since a non-
autonomous system can always be converted into an autonomous
system by introducing an additional state variable: $x_{n+1} \equiv t$,
say.

6/   For simplicity we assume that the constraints on controls
are fixed so that g in (6) is independent of x. The main
results of the paper will continue to hold if $g(u,\theta)$ is re-
placed by $g(x,u,\theta)$ (with appropriate modifications, some of
which will be mentioned later in footnotes).

7/   See Pontryagin [9], pp. 66-69, 189-191. If the time-horizon
is infinite so that $t_1 = + \infty$, then we assume that $\lim_{t_1 \to +\infty} x(t_1)$

exists. (Similarly, for the case $t_0 = -\infty$.) Furthermore, an
optimum solution of the infinite-horizon problem might not
satisfy Pontryagin's condition (see Arrow and Kurz [1], p. 46).
In the following, we exclude such cases from our consideration.

8/   If the function g depends not only on (u,θ) but also on x,
then (9) is modified accordingly, and (10) is replaced by
$\dot{p}(t) = -[H_x(p(t),x(t),u(t) + \lambda(t)g_x(x(t),u(t))]$, where $\lambda(t)$
is the Lagrangian multipliers associated with g in (9). With
these modifications, the discussions through Section IV will
continue to hold.

9/   The situation explained in the text may arise even in the
simple maximization problem discussed in Section I. Figures
a and b illustrate. The function $f(x,\theta)$ has two local maxima
near $\theta = 0$: $x^1(\theta)$ and $x^2(\theta)$ (Figure a). The overall maximum
$x(\theta)$ is given by $x^1(\theta)$ if $\theta \leq 0$ and by $x^2(\theta)$ if $\theta \geq 0$ (Figure
b). One needs to combine the derivatives of $x^1(\theta)$ and $x^2(\theta)$,

$x(\theta)$

$x^1(\theta)$

$x^1(0)$

$x(\theta)$

$x^2(0)$

$x^2(\theta)$

x

θ''

θ'

0

Figure b.

$f(x, \theta)$

$f(x, \theta')\ (\theta' < 0)$

$f(x, 0)$

$f(x, \theta'')\ (\theta'' > 0)$

$> x$

$x^1(\theta')\ x^1(0)\ x^1(\theta'')$
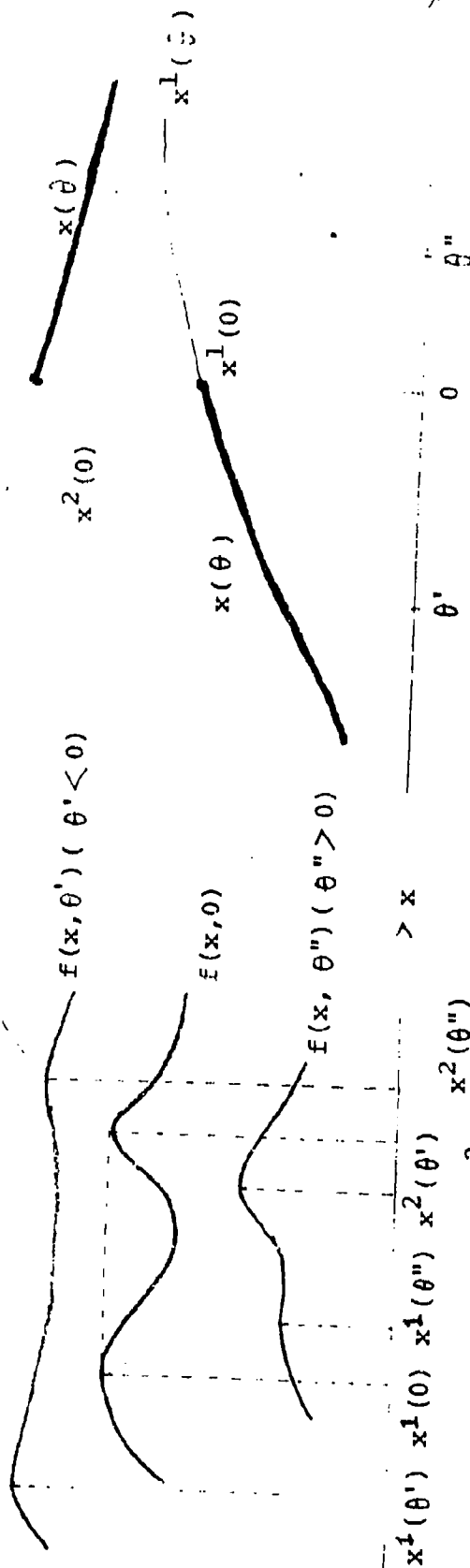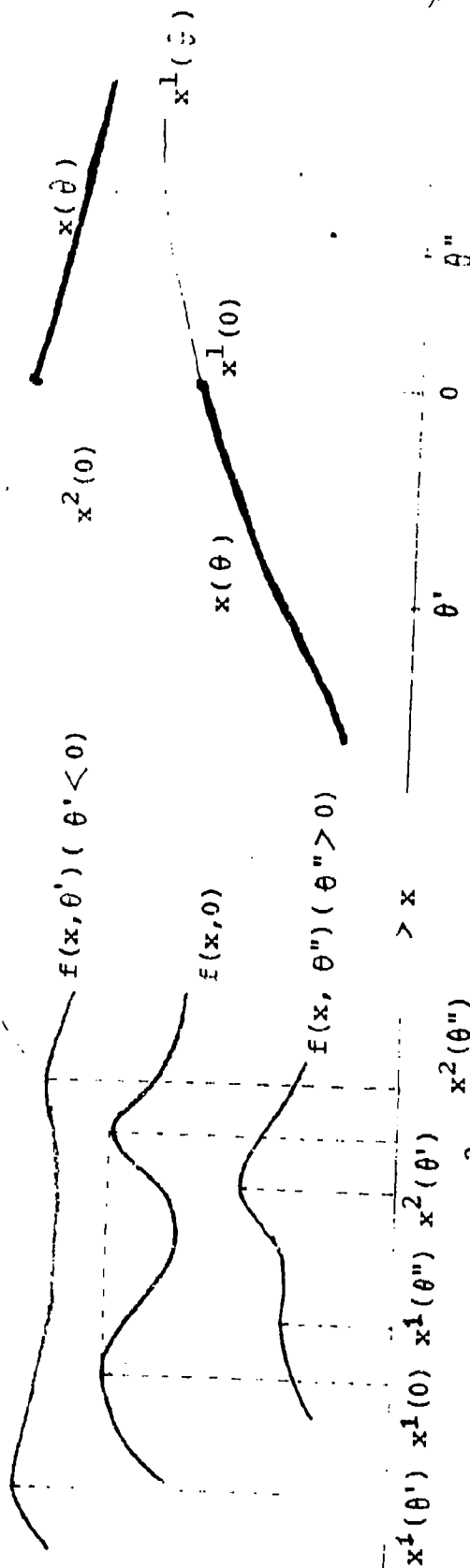
$x^2(\theta')$

$x^2(0)$

$x^2(\theta'')$

Figure a.

each being computable as in (2), in order to describe the
shift of the optimum near θ = 0.

10/   For proof, see Pontryagin [10], pp. 170-177, 194, 198, or
Hartman [4], pp. 93-94, 95-100.

11/   For proof, see the Appendix.

12/   For a counterexample, see Pontryagin [9], p. 73.

13/   The method presented in this section was once introduced
by Oniki [7].

14/   See Cass [3], pp. 837-838.

15/   See, ibid., pp. 834-835.

16/   See, e.g., Pontryagin [10], pp. 163-166.

REFERENCES

1. K. J. Arrow and M. Kurz, Public Investment, the Rate
    of Return, and Optimal Fiscal Policy, Johns Hopkins
    Press, Baltimore, Md., 1970.

2. R. P. Barriere, "On the Cost of Constraints in Dynamic
    Optimization," in Mathematical Theory of Control, ed.
    by A. V. Balakrishnan and L. W. Neustadt, Academic Press,
    New York, 1967, 246-250.

3. D. Cass, "Optimum Growth in an Aggregative Model of
    Capital Accumulation: A Turnpike Theorem," Econometrica,
    34 (1966), 833-850.

4. P. Hartman, Ordinary Differential Equations, John
    Wiley, New York, 1964.

5. D. W. Jorgenson, "The Theory of Investment Behavior," in
    Determinants of Investment Behavior, ed. by R. Ferber,
    Columbia Univ. Press, New York, 1967, 129-155.

6. T. C. Koopmans, "Intertemporal Distribution and 'Optimal'
    Aggregate Economic Growth," in Ten Economic Studies in
    the Tradition of Irving Fisher, by W. Fellner and
    others, John Wiley, New York, 1967.

7. H. Oniki, "Comparative Dynamics in the Theory of Optimal
    Growth," Keizaigaku (Tohoku Economic Journal), 30 (1969),
    48-57.

8. H. Oniki, "On Deriving the Individual Demand Function for
    Educational Investment," Harvard Institute of Economic
    Research, Discussion Paper, No. 154, 1970.

9. L. S. Pontryagin and others, The Mathematical Theory of Optimal Processes, John Wiley, New York, 1962.

10. L. S. Pontryagin, Ordinary Differential Equations, Addison Wesley, Reading, Ma., 1962.