Social Science Research Institute
University of Wisconsin

THE GENERALIZED MAXIMUM PRINCIPLE

W. A. Brock
W. D. Dochert
8316

December 1983
ABSTRACT

This paper develops qualitative analytical techniques for infinite horizon optimal control problems with time stationary current value Hamiltonians that are not concave in the state variable. In this case, given initial state \( x_0 \), there typically exists a set \( P(x_0) \) of initial costates such that solutions of the state-costate differential equations satisfy the transversality condition at infinity (TV(\( \infty \))). We use the Hamilton-Jacobi equation and convexity of the optimized Hamiltonian to show that an optimal costate maximizes the Hamiltonian over \( P(x_0) \) for each initial state \( x_0 \). This is the generalized Maximum Principle.

This technique is applied to optimal growth theory, dynamic Ramsey pricing, and general isoperimetric problems.
I. Introduction

In this paper we present a result for an infinite horizon optimal control problem in which the objective function is not concave in the state variable. Economists have found that there are many applications of dynamic growth problems where there are non-convexities in the production set. This can arise in Ramsey growth problems where per capita output is not a concave function of the capital labor ratio, but rather there is a region over which increasing returns to scale are present. In the optimal harvesting of a renewable resource, there are several models where the stock-recruitment function exhibits an increasing returns to scale portion when the species is relatively small in number. In optimal investment theory, firms that have a natural monopoly advantage are typically modeled as having decreasing costs (at least over some range of output) which are caused by increasing returns to scale in the production function.

In all of the applications of control theory to intertemporal problems in economics, one of the most powerful techniques used has been the Hamiltonian dynamics and the corresponding phase diagram analysis. However, most of the technical analysis (especially for infinite time horizon problems) has been developed for the case that the Hamiltonian is concave in both the state and control variables. In the examples cited above, none of the models would have a concave Hamiltonian for all values of the state variable. Our results in this paper extend the work for the concave case to non-concave cases as well.

The case of analysis in the concave case is well known. Benveniste and Scheinkman [1] have shown that the necessary and sufficient condition for optimality is that the Euler equation and a transversality condition be satisfied. This then reduces the problem to finding a solution to a second order differential equation which satisfies the 'boundary' value given by the transversality condition. Michel [11] and Ekeland and Scheinkman [8] have shown that the transversality conditions are also necessary when the Hamiltonian is not concave, which along with the Euler equation still reduces the problem (in part) to solving a second order differential equation. However, there may be many solutions which
satisfy both the Euler equation and the transversality conditions, and the
question is: which are optimal? Heretofore the technique has been to evaluate
the objective functional along all of these solutions and to find which one gives
the maximum value. (This is somewhat analogous to the problem of maximizing a
function and finding that there are many relative maxima. In order to determine
which is the global maximum, one evaluates the function at all the solutions and
compares these values with each other.) Unfortunately this procedure can only be
carried out when there is a concrete objective function which has been specified.
In most economic applications, objective functions are usually utility functions,
profit functions, etc., which are specified only by general properties rather than
specific functional forms.

In this paper we derive the useful property that when the Hamiltonian is
concave in the control variable but not necessarily concave in the state variable,
then when there is more than one solution to the Euler equation and the trans-
versality conditions, the optimal solution corresponds to an extreme value of the
Corresponding set of costate variables. In one dimensional problems, this in
effect reduces the set of solutions to at most two which might be optimal. In
Casons where time-stationarity is present (which we discuss in this paper) this
means that we can readily use the phase diagram, which allows us to analyze cases
when Hamiltonian is not concave even though the objective function is not given by
a specific functional form.

In terms of the existence of solutions for this type of problem, Romer [8] has
shown that the arguments of Ekeland and Temam [4] can be adapted to the type of
infinite horizon problems which we are considering in this paper. Since we are
primarily interested in characterizing optimality, we assume throughout that
solutions exist. Nevertheless, in section II we show how our methods can be used
to derive a condition which guarantees the existence of a solution for the one
sector growth problem with a production function which is everywhere convex.

In section II we apply our techniques to the Ramsey growth problem as well as
to an adjustment cost model of investment. We also develop results for these
models with an isoperimetric constraint, and show that it is not necessarily the
case that the constraint can be incorporated into the objective functional with a
Lagrange multiplier. It turns out that this is related to the fact that when the objective function is not concave in the state variable, then there are certain values of the state variable for which the optimal solution is not unique. Finally, in section III we generalize our results to multisector problems.
II. The Generalized Maximum Principle

Consider the following problem in the calculus of variations:

\[ V(x_0) = \sup \int_0^\infty e^{-rt} f(x, \dot{x}) \, dt \]

subject to: \( x(0) = x_0 > 0, \ x(t) \geq 0 \) \hspace{1cm} (1)

where \( r > 0 \), and \( f \) is assumed to be continuously differentiable and strictly concave in \( \dot{x} \). \( ^1 \) The maximization is over the set of functions, \( x(t) \geq 0 \), with piecewise continuous derivatives, \( \dot{x}(t) \), and we shall assume that the integrand is such that the supremum in (1) is finite over this set. \( ^2 \)

We are not assuming that \( f \) is concave in \( x \). Romer \[8\] has shown that one can adapt the arguments of Ekeland and Temam \[4, \text{Part 3}\] and demonstrate that optimal solutions for this type of problem exists, provided that \( f \) satisfies a growth condition, and that \( \dot{x} \) remain bounded. Since we are concerned with the characterization of optimal trajectories, we shall assume that an optimum exists.

An interior optimal solution, \( x^* \), must satisfy certain necessary conditions (see for example, Hestenes \[9, \text{Chapter II}\]) which include the Euler equation:

\[ f_x(x^*(t), \dot{x}^*(t)) = -rf_x(x^*(t), \dot{x}^*(t)) + \frac{d}{dt} f_{\dot{x}}(x^*(t), \dot{x}^*(t)) \]

and the Weierstrass-Erdmann corner conditions:

\[ f_x, f - \dot{x}f_{\dot{x}} \text{ are continuous along } x^*(t) \] \hspace{1cm} (2b)

Because of the strict concavity of \( f \) in \( \dot{x} \), \( x^* \) is also continuously differentiable. (In fact, if \( f \) is \( C^n \), then so is \( x^* \). See Hestenes \[9, \text{pg. 60}\]) Solutions must also satisfy certain transversality conditions. If we define the terminal time by

\[ T^* = \inf \{ T \mid x^*(t) = 0 \text{ for } t \geq T \} \]

and \( T^* = \infty \) if this set is empty, then when \( T^* \) is finite

and \( T^* = \infty \) if this set is empty, then when \( T^* \) is finite
\[
\lim_{t \uparrow T^*} f_x(0, \dot{x}(t)) \leq 0
\]  
(4)

and

\[
\lim_{t \uparrow T^*} \{f(0, \dot{x}(t)) - \dot{x}(t)f_x(0, \dot{x}(t))\} = 0
\]  
(5)

must hold. If \( f(0,0) = -\infty \), then \( T^* \) cannot be finite, and so if \( T^* \) is finite, we shall assume that \( f \) has been scaled so that \( f(0,0) = 0 \), and that \( f \) is continuous on its boundary.\(^3\) When \( T^\ast = \infty \), Michel [11] has shown that

\[
\lim_{t \uparrow \infty} e^{-xt} \{f(x(t), \dot{x}(t)) - \dot{x}(t)f_x(x(t), \dot{x}(t))\} = 0
\]  
(6)

is a necessary condition for the infinite horizon case.

When \( f \) is concave in both \( x \) and \( \dot{x} \), an argument along the lines of Seierstad and Sydsæter [17] and Seierstad [16] can be made to show that the neccessary and sufficient conditions for a function, \( x^* \), to be a solution to (1) is that it satisfy (2a) and (4) - (6). Thus the problem of finding an optimal solution can be reduced to finding a solution which satisfies these equations. However, when \( f \) is not concave in both \( x \) and \( \dot{x} \), there may be several functions which satisfy (2b) and (4) - (6), not all of which are optimal.

Denote the set of continuously differentiable functions, \( x(t) \), which satisfy

(i) \( x(0) = x_0 \),
(ii) \( x(t) \) satisfies (2a) and (2b)
(iii) \( x(t) \) satisfies the transversality conditions, (4) - (9)
(iv) \( \int_0^\infty e^{-xt} f(x(t), \dot{x}(t)) dt > -\infty \),

as \( \Omega(x_0) \), and for each \( x(t) \in \Omega(x_0) \) let

\[ p_0 = -f_x(x_0, \dot{x}(0)). \]  
(7)

Also define the Hamiltonian function,
\[ H(x,p) = \sup_{\dot{x}} \{ f(x,\dot{x}) + p\dot{x} \} \tag{8} \]

and notice that since \( f \) is strictly concave in \( \dot{x} \), \( H \) is strictly convex in \( p \), for those \( p \) in the range of \( -f_\dot{x} \). Furthermore,

\[ H(x_0,p_0) = f(x_0,\dot{x}(0)) + p_0\dot{x}(0) \tag{9} \]

holds for \( p_0 \) given by equation (7).

The following extension of Skiba's Lemma [18, pg 537] shows how to evaluate the functional in (1) along a trajectory \( x(t) \in \Omega(x_0) \), even if it is not an optimal trajectory:

**Lemma 1:** If \( x(t) \in \Omega(x_0) \) then \( \int_0^\infty e^{-rt}f(x(t),\dot{x}(t))dt = \frac{1}{I} H(x_0,p_0) \), where

\[ p_0 = -f_\dot{x}(x_0,\dot{x}(0)). \]

Proofs are in the appendix. Now let

\[ P(x_0) = \{ p \mid x(t) \in \Omega(x_0), p = -f_\dot{x}(x_0,\dot{x}(0)) \}. \tag{10} \]

This set is not empty since it contains \( p_0^* = -f_\dot{x}(x_0,\dot{x}^*(0)) \). Furthermore it is clear that for all \( p \in P(x_0) \),

\[ H(x_0,p) \leq rV(x_0) \]

and therefore,

\[ \max_{p \in P(x_0)} H(x_0,p) = rV(x_0). \tag{11} \]

Except for the question of the existence of solutions, (11) does not depend on the concavity of \( f \) in \( \dot{x} \). When \( f \) is concave in \( \dot{x} \) (the case we are considering here) we can sharpen this result to:

**Theorem 1** (Generalized Maximum Principle) Either \( p_0^* = \sup P(x_0) \) or \( \inf P(x_0) \).
We call this the Generalized Maximum Principle because if we let

\[ p = \inf P(x_0), \quad \bar{p} = \sup P(x_0) \]

then the theorem implies that

\[ V(x_0) = \max \{ H(x_0, p), H(x_0, \bar{p}) \}. \tag{12} \]

Define \( P^*(x_0) \) as the set of optimal costates. When \( f \) is strictly concave in both \( x \) and \( \dot{x} \), this set is always a singleton. However, when \( f \) is not concave in \( x \), there can be more than one optimal trajectory starting from an initial \( x_0 \). The following characterizes the behavior of the costates at such a point:

**Theorem 2** (One Dimensional Jump Theorem): If the cardinality of \( P^*(x) \) is greater than 1 (notation: \( \#P^*(x) > 1 \)) then \( \#P^*(x) = 2 \), and

\[ p \in P^*(x^-), \quad q \in P^*(x^+) \quad \text{implies that} \quad p < q. \]

Note that the conclusion of this theorem is that as we move along the \( x \)-axis from left to right, then at points where there is more than one optimal trajectory, there are precisely two and the value of the costate jumps up.

**Applications: Ramsey Growth**

The Ramsey [12] growth problem as studied by Skiba [18] is

\[ V(x_0) = \max \int_0^\infty e^{-rt} u(c(t)) dt \tag{13} \]

subject to: \( x(0) = x_0, \quad x(t) \geq 0, \quad \dot{x}(t) = F(x(t)) - c(t) \)

where \( u \) is strictly concave and increasing on \([0, \infty)\) with \( u'(0) = \infty \), and \( F \) is increasing on \([0, \infty)\) with

\[ F''(x) \geq 0 \quad \text{as} \quad x \notin x_I \]
for some inflection point, $x_I > 0$, and

$$\lim_{x \to \infty} F'(x) = 0. \tag{14}$$

If $x(t)$ is any function which satisfies the Euler equation (2) then

$$p(t) = u'(F(x(t))) - \dot{x}(t), \tag{15}$$

satisfies the differential equation

$$\dot{p}(t) = [\lambda - F'(x(t))]p(t) \tag{16}$$

and we can plot the phase diagram (Figure 1a) for this system. We have drawn the phase diagram for the case that $\lambda$ (the point at which $F'(\lambda) = \lambda$ and $F''(\lambda) < 0$) is an optimal steady state. See Skiba [9] for other cases. The trajectories which satisfy the transversality condition (6) are plotted on the diagram, and labeled as trajectories I and II.

(FIGURE 1 ABOUT HERE)

For some initial values of capital there is only one possible trajectory which satisfies the transversality condition, and so $\lambda(x_0)$ consists of only one point. However, when there is more than one trajectory (as is the case at $x_0$ in Figure 1a) the Generalized Maximum Principle shows that one of $\lambda_0$ or $\lambda_0$ is the optimal initial value of $\lambda(0)$. Thus only the upper portion of trajectory I and the lower portion of trajectory II could ever be optimal. The range of capital stocks for which there is more than one trajectory which satisfies the necessary conditions is from $x_*$ to $x^*$.

At $x^*$ trajectory I is optimal since at that point trajectory II cuts the $\dot{x} = 0$ locus and so the Hamiltonian, $H(x^*, p)$ is at a minimum there. Similarly, at $x_*$, trajectory II is optimal. By a continuity argument as in Dechert and Nishimura [3] (see the appendix) it can be shown that there cannot be overlapping segments of these trajectories for which both are optimal. Thus there is a point, $x_* > 0$, ...
which we shall call a Skiba Point, so that

\[
\begin{align*}
\text{if } x_0 &\leq x_S, \quad \text{trajectory II is optimal,} \\
\text{if } x_0 &\geq x_S, \quad \text{trajectory I is optimal.}
\end{align*}
\]

(17)

Also, in spite of the non concavity in \( x \), \( V(x_0) \) is differentiable everywhere except at \( x_S \), and since \( V'(x_0) = p_0^* \) for \( x_0 \neq x_S \), we can use the phase diagram to sketch the value function. Define \( V_I \) and \( V_{II} \) by

\[
\begin{align*}
V_I(x_0) &= r^{-1}H(x_0, p_0) \quad \text{for } x_0, p_0 \text{ on trajectory I} \\
V_{II}(x_0) &= r^{-1}H(x_0, p_0) \quad \text{for } x_0, p_0 \text{ on trajectory II}
\end{align*}
\]

The value, \( V(x_0) \), as defined by the optimization problem, is then the upper envelope of these two functions,

\[
V(x_0) = \max \{ V_I(x_0), V_{II}(x_0) \}
\]

which has been sketched in Figure 1b for the case that \( u(0) = 0 \).

When \( F \) is convex for all \( x \), there is no optimal steady state \( \bar{x} \). Nevertheless, we can apply the techniques developed above to this case as well. As in Figure 1a, let \( \bar{x}^* \) be the point at which trajectory II cuts the \( \dot{x} = 0 \) locus and take \( x_0 > x^* \). Define an increasing sequence of truncated economies as follows: let \( \{x_n\} \) be any increasing sequence which diverges to \( \infty \), and let

\[
F_n(x) = \begin{cases} 
F(x) & \text{if } x \leq x_n \\
F(x_n) & \text{if } x \geq x_n
\end{cases}
\]

For each truncated economy, define the value function as:

\[
V_n(x_0) = \max \int_0^\infty e^{-rt}u(F_n(x) - \dot{x})\,dt.
\]

In spite of the fact that \( F_n \) is not differentiable at \( x_n \), we can still apply the techniques of this section. The Weierstrass–Erdmann corner conditions imply that the optimal trajectory is continuous at \( x_n \), and so we can 'paste' together the
phase diagram at \( x_n \). By the analysis for the Skiba type problem, trajectory \( I_n \) (see Figure 2) converges to \( x_n \). Let \( p_n \) be the optimal initial costate (at \( x_0 \)) on this trajectory. It is easy to see that \( \{p_n\} \) is an increasing sequence, and that

\[ rV_n(x_0) = H(x_0, p_n) \]

is also increasing in \( n \).

(FIGURE 2 ABOUT HERE)

**Theorem 3:** If \( \lim p_n = \bar{p} < \infty \), then an optimal solution exists and \( rV(x_0) = H(x_0, \bar{p}) \).

Note that when \( \bar{p} < \infty \), trajectory \( I \), the trajectory that satisfies the Euler equation and passes through the point \( (x_0, \bar{p}) \), is the upper envelope of those trajectories which cut the \( \dot{x} = 0 \) locus.

**Adjustment Cost Models**

Another application of these ideas is to a continuous time version of the model in Brock and Dechert [2],

\[
\max \int_0^\infty e^{-rt}[R(x) - C(I)] dt
\]

subject to: \( x(t) \geq 0, x(0) = x_0, \dot{x}(t) = I(t) - \delta x(t) \)

where \( R(x) \) is the (maximum) revenue a firm can raise with \( x \) units of capital stock and \( C(I) \) is the total cost (including purchase price and installation costs of investment goods, and \( \delta \) is the rate of depreciation. If \( x(t) \) is any function which satisfies the Euler equation (2) and \( x(0) = x_0 \), then

\[ p(t) = C'(\dot{x}(t) + \delta x(t)) \]

which satisfies the differential equation

\[ \dot{p}(t) = (r + \delta)p(t) - R'(x(t)). \]
As long as $C$ is convex in $I$, the Hamiltonian will be convex in $\dot{x}$. A typical phase diagram is drawn in Figure 3 for the case that $C(0) = 0$ and $C'(0) = 1$. (In this model, we can normalize prices so that $I$ is the purchase cost of $I$ units of investment goods, and $C(I) - I \geq 0$ as the cost of installing them at the firm's plant.) By reasoning along lines similar to those for the Ramsey growth problem, there is a Skiba point, $x_S$, for this problem as well so that (17) holds.

(Figure 3 About Here)

Similar to the analysis of the Ramsey growth case, we have also diagrammed the two value functions, $V_I$ and $V_{II}$, and their upper envelope, $V$. Other cases of the monopoly investment problem can be found in Dechert [4, 5].

Isoperimetric Constraint Problems

In the adjustment cost model of the previous section, if the firm is a monopoly then the revenue function is $R(x) = \max \{qP(q) \mid q \leq f(x)\}$, where $P(q)$ is the demand price for $q$ units of output delivered to the market. If we define the area under the demand curve by

$$A(Q) = \int_0^Q P(q) dq$$

then the first best problem of maximizing the sum of consumer plus producer surpluses is

$$V(x_0) = \max \int_0^\infty e^{-\lambda t}[A(Q) - C(I)] dt$$

subject to: $x(0) = x_0$, $x(t) \geq 0$, $\dot{x} = I - \delta x$, and $Q \leq f(x(t))$. (18)

However, just as in the static case, along an optimal solution to (18) the firm will not necessarily cover its fixed costs because of the increasing returns to scale. In this case a second best problem consists of solving (18) subject to
the isoperimetric constraint

\[ \int_0^\infty e^{-rt}[P(Q)Q - C(I)]dt \geq \pi. \]  \hspace{1cm} (19a)

This type of constraint can arise from a cash flow constraint,

\[ C(I(t)) + ra(t) \leq \dot{x}(t) + R(x(t)) \]  \hspace{1cm} (19b)

where \( a(t) \) are the firm's obligations (i.e., bonds), which are assumed to satisfy an asymptotic indebtedness criterion of \( \lim_{t \to \infty} e^{-rt}a(t) = 0 \). By multiplying both sides of (19b) by \( e^{-rt} \) and integrating we get (19a) with \( \pi = a(0) \). In the spirit of the modern finance literature, if it is assumed that there are perfect capital markets, then the two constraints are economically equivalent.

In order to analyze the properties of solutions to this problem, consider the following auxiliary problem,

\[ L(\alpha, x_0) = \max \int_0^\infty e^{-rt}[\alpha P(Q)Q + (1 - \alpha)A(Q) - C(I)]dt \]  \hspace{1cm} (20)

subject to: \( x(0) = x_0, \ x(t) \geq 0, \ \dot{x} = I - \delta x, \) and \( Q \leq f(x(t)) \).

for \( 0 \leq \alpha \leq 1 \). Note that when \( \alpha = 0 \), problem (20) corresponds to a first best pricing problem, while \( \alpha = 1 \) corresponds to a monopoly pricing problem. Since \( A(Q) \) is always larger than \( QP(Q) \), \( L \) is a decreasing function of \( \alpha \), and since the objective functional in (20) is linear in \( \alpha \), \( L \) is convex in \( \alpha \). In order for the isoperimetric constraint problem (18) - (19) to be non-trivial, we shall assume that \( \pi < L(1, x_0) \), and that the inequality in (19a) fails to be satisfied along a first best solution, i.e., when \( \alpha = 0 \).

Under these circumstances, the question which naturally arises is, Does there exist an \( \alpha_0 \) so that a solution to the isoperimetric problem (18) - (19) solves the auxiliary problem (20) for this value of \( \alpha \), and vice-versa? This is asking whether or not the isoperimetric constraint (19a) can be incorporated into the objective function (18) with a Lagrange multiplier. The answer depends on whether
or not there is a unique solution to the isoperimetric problem.

Consider the construction in Figure 4a: draw a support line through \( \pi \) on

\[ (\text{FIGURE 4 ABOUT HERE}) \]

the \( \alpha = 1 \) line tangent to \( L(\alpha, x_0) \), and label the point of tangency \( \alpha_0 \). If \( \tilde{x}(t), \tilde{q}(t) \) is a solution to (20) for \( \alpha = \alpha_0 \), and \( x(0) = x_0 \), then

\[
- \int_0^\infty e^{-rt}[A(\tilde{\alpha}) - \tilde{q}P(\tilde{\alpha})]dt
\]

is a subgradient of \( L(\alpha, x_0) \) at \( \alpha_0 \). When there is a unique solution, then this subgradient is unique, and the present value of profits along this solution is

\[
\int_0^\infty e^{-rt}[\tilde{q}P(\tilde{\alpha}) - C(\tilde{t})]dt = \int_0^\infty e^{-rt}[\alpha_0 \tilde{q}P(\tilde{\alpha}) + (1 - \alpha_0)A(\tilde{\alpha}) - C(\tilde{t})]dt
\]

\[
= (1 - \alpha_0) \int_0^\infty e^{-rt}[A(\tilde{\alpha}) - \tilde{q}P(\tilde{\alpha})]dt
\]

\[
= L(\alpha_0, x_0) + (1 - \alpha_0) \int_0^\infty e^{-rt}[A(\tilde{\alpha}) - \tilde{q}P(\tilde{\alpha})]dt
\]

\[
= \pi.
\]

Thus the solutions to the isoperimetric constraint problem and the auxiliary problem coincide.

Now consider the case that at \( \alpha_0 \) the Skiba point for problem (20) is \( x_0 \). In that case Theorems 1 and 2 imply that there are two solutions to (20), one of which, \( \bar{x}(t) \), converges to 0, while the other, \( \underline{x}(t) \), converges to a positive steady state. Let the consumer's surplus (the negative of (21) evaluated along these two solutions) be \( \bar{S} \) and \( \underline{S} \), respectively. The set of subgradients of \( L(\alpha, x_0) \) at \( \alpha_0 \) satisfies:

**Lemma 2:** If \( \lambda \) is a subgradient of \( L(\alpha, x_0) \) at \( \alpha_0 \), then \( -\bar{S} \leq \lambda \leq -\underline{S} \).

Note that the left- and right-hand derivatives of \( L(\alpha, x_0) \) at \( \alpha_0 \) are \( -\bar{S} \) and \( -\underline{S} \) respectively, and the value of profits along the two solutions are
\[ \pi = \int_0^\infty e^{-rt} [Q(\dot{q}) - C(\ddot{q})] dt = L(a_0, x_0) - (1 - a_0) S \]
\[ \bar{\pi} = \int_0^\infty e^{-rt} [\bar{Q}(\dot{\bar{q}}) - C(\ddot{\bar{q}})] dt = L(a_0, x_0) - (1 - a_0) \bar{S} \]

and therefore \( \pi > \bar{\pi} \). Now if \( \pi > \bar{\pi} > \pi \) there is no value of \( a \) for which the solution to (20) has (19) holding with equality. See Figure 4b.

The above technique for handling isoperimetric constraint problems of the form

\[ V(x_0) = \max \int F(t, x, \dot{x}) dt \]
subject to:
\[ x(0) = x_0 \]
and:
\[ \int G(t, x, \dot{x}) dt \geq \pi \]

is quite general, provided that \( F \) and \( G \) are strictly concave in \( \dot{x} \). The corresponding auxiliary problem is

\[ L(a, x_0) = \max \int [(1 - a)F(t, x, \dot{x}) + aG(t, x, \dot{x})] dt \]

which is convex in \( a \). The Hamiltonian for (23) is

\[ H(t, a, x, p) = \max \frac{[(1 - a)F(t, x, \dot{x}) + aG(t, x, \dot{x}) + p\dot{x}]}{\dot{x}} \]

which is convex in \( p \). When the range of integration (i.e., the horizon) is infinite, then Theorems 1 and 2 apply to this problem, and so there are at most two solutions to (23). Suppose for some \( a_0 \) there are two solutions. Then \( L(a_0, x_0) \) will have a kink at \( a_0 \), and there will be a range of values, \( \pi > \bar{\pi} > \bar{\pi} \), for which the solutions to (23) will not correspond to a solution of the isoperimetric constraint problem (22). Even when the range of integration is finite, the above construction shows that if there is more than one solution to (23) at some \( a_0 \), then \( L \) will have a kink at \( a_0 \) and so the solution to the two problems may not coincide.
III. The Generalized Maximum Principle: The Multidimensional Case

In this section, we develop a generalization of the results for the one dimensional case of section II to multidimensional problems where the Hamiltonian is strictly convex in the costate variables, but not necessarily everywhere concave in the state variables.

In keeping with the notation of section II, we will use $\Omega(x)$, $P(x)$ and $P^*(x)$ as the multidimensional analogues of the one dimensional case. Notice that the conclusion of Theorem 1 could be restated as: the elements of $P^*(x)$ cannot be represented as proper convex combinations of the elements of $P(x)$. This is the concept that generalizes to multidimensional problems:

**Definition:** An extreme point of a set $S$ is a point which cannot be represented as a proper convex combination of other points of $S$.

**Lemma 3:** Let $p_0 \in P(x_0)$, and assume that there is a neighborhood $N(p_0)$ and a function $g(t)$ which is integrable with respect to $e^{-rt}$ such that $|f(x(t),\dot{x}(t))| \leq g(t)$ for all $x(t) \in \Omega(x_0)$ with $-f_x(x_0,\dot{x}(0)) \in N(p_0)$. Then $P(x_0)$ is compact, and the elements of $P^*(x_0)$ are extreme points of $P(x_0)$.

**Theorem 4:** Let $P(x)$ be compact, and let $K$ be its convex hull. Then

$$\max_{p \in K} H(x,p) = \max_{p \in P(x)} H(x,p)$$

and $P^*(x) \subseteq \partial K$, the boundary of $K$.

This theorem is not quite as sharp as Theorem 2 in that we cannot specify the cardinality of $P^*(x_0)$. Consider for example,

$$\max \int_0^\infty e^{-rt}[u(F(x^1) - \dot{x}^1) + u(F(x^2) - \dot{x}^2)]$$

subject to: $(x^1(0),x^2(0)) = (x_0^1,x_0^2)$, $x^1 \geq 0$, $\dot{x}^1 \leq F(x^1)$

In this example the objective function is separable in the two coordinates, so we
can use the analysis for the one dimensional case to each problem separately. In Figure 5 we have drawn a phase diagram in \((x^1,x^2)\) space. Along the lines \(x^1 = x^2\), there are two optimal costate values, except at the point \((x^*_S,x^*_S)\) where there are four.

(FIGURE 5 ABOUT HERE)

In this example the set of \((x^1,x^2)\) points at which there is more than one optimal path divides the plane into regions (four in this case) and along the common boundaries between these regions, the value function is not differentiable. This example does not depend on the separability of the objective function. If we add \(\gamma G(x^1,x^2,\dot{x}^1,\dot{x}^2)\) as an additional term to (24), where \(G\) is concave in \(\dot{x}^1\) and \(\dot{x}^2\), then for small \(\gamma\) the structure of the diagram in Figure 5 is preserved. It is also clear from the diagram that the four steady states, \((0,0), (0,\bar{x}), (\bar{x},0), (\bar{x},\bar{x})\), are locally asymptotically stable.\(^4\)

In terms of generalizing the jump theorem, notice that if \(p,q \in P^*(x)\), then \(H(x,p) = H(x,q)\) and by the convexity of \(H\) in the second variable,

\[
0 = H(x,q) - H(x,p) \geq H_2(x,p)(q - p),
\]

with strict inequality if \(p \neq q\). This implies that if \(x(0) = x\) and \(-f(x,\dot{x}(0)) = p\), then for any other optimal costate \(q\) at \(x\),

\[
\dot{x}(0)(p - q) > 0.
\]

This leads to the following multidimensional version of Theorem 2: Denote by \(A_i\) the regions of \(\mathbb{R}^n\) in which the optimal solution is unique. Then along the boundaries of these regions, \(\partial A_i\), there is more than one optimal value of the costate. Denote by \(p_{A_i}\) that optimal costate value such that if \(\dot{x}(0)\) corresponds to this costate value then \(x(t) \in A_i\). Since we are assuming that \(f\) is continuously differentiable and strictly concave in \(\dot{x}\), the Principle of Optimality and the Weierstrass-Erdmann corner conditions imply that if \(x(t)\) is an optimal trajectory,
then the unique optimal path starting from \( x(s) \) for any \( s > 0 \) is \( x(t) \) for \( t \geq s \).

So far, little is known about the structure of the regions \( \Lambda_i \). For \( x \in \Lambda_i \), define \( V_{\Lambda_i}(x) \) in the obvious way. At a boundary point, \( x \in \partial \Lambda_i \), there are at least two optimal solutions, one of which enters \( \Lambda_i \). By solving the Euler equation backwards in \( t \), with \(-f(x, \dot{x}(0)) = p_{\Lambda_i}\), we can extend the domain of definition of \( V_{\Lambda_i} \) to those \( x \) which are reached by one of these backwards solutions. In the one dimensional Ramsey growth problem depicted in Figure 1, the regions are \( \Lambda_1 = (0, x_S) \) and \( \Lambda_2 = (x_S, \infty) \). The backwards solution starting at \( x_S \) with costate on trajectory II extends the domain of definition of \( V_{\II} \) to \((0, x^*)\). Similarly, \( V_{\I} \) is extended to \([x^*, \infty)\). Denote these extended regions by \( \Lambda_i \).

**Theorem 5:** (Generalized Jump Theorem): Let \( x \in (\partial \Lambda_i) \cap (\partial \Lambda_j) \), and let

\[
\xi : [-1, 1] \to \mathbb{R}^n \text{ be any differentiable function such that } \xi(0) = x,
\xi(t) \in A_i \cap B_j \text{ for } t < 0, \text{ and } \xi(t) \in A_j \cap B_j \text{ for } t > 0.
\]

Then under the conditions of Lemma 3,

\[
(p_{\Lambda_j} - p_{\Lambda_i}) \xi'(0) \geq 0.
\]

If at \( x \) it is possible to choose \( \xi'(0) \) to be equal to the \( k \)th coordinate unit vector, then the conclusion of this theorem implies that the \( k \)th coordinate of the optimal costate jumps up at \( x \) along \( \xi \).

Another result on the nature of the regions is that these extended regions cannot completely overlap the original regions, \( \Lambda_i \) in the following sense. If \( \partial \Lambda_i = (\partial \Lambda_i) \cap (\partial \Lambda_j) \) is a compact manifold for some \( i \neq j \), then it is not the case that \( \Lambda_i \subseteq B_j \). A sketch of this assertion is: for \( x \in \Lambda_i \cap \partial \Lambda_i \), define \( F(x) = p_{\Lambda_i}(x) - p_{\Lambda_j}(x) \), where \( p_{\Lambda_i} \in P(x) \) for \( x \in \Lambda_i \) and \( p_{\Lambda_j} \in P(x) \) is such that if \( x(0) = x \) and \(-f(x, \dot{x}(0)) = p_{\Lambda_j}(x) \), then \( x(t) \in \Lambda_j \) for sufficiently large \( t \). Similarly, \( p_{\Lambda_i} \) for \( x \in \partial \Lambda_i \). The differential equation, \( \dot{x} = F(x) \) for \( x \in \Lambda_i \cup \partial \Lambda_i \) has an inward pointing vector field along the boundary of \( \Lambda_i \), and hence there is a rest point inside \( \Lambda_i \). That is, for some \( x \in \Lambda_i \), \( 0 = F(x) = p_{\Lambda_i}(x) - p_{\Lambda_j}(x) \). But then \( V_{\Lambda_i}(x) = V_{\Lambda_j}(x) \) for some \( x \in \Lambda_i \), which is a contradiction.

The technique for handling an isoperimetric constraint in the previous section
can also be generalized to the multidimensional problem. For example, when the objective and constraint functions are separable,

\[ F(t, x, \dot{x}) = \lambda(t) + e^{-rt} f_i(x^i, \dot{x}^i) \]
\[ G(t, x, \dot{x}) = \mu(t) + e^{-rt} g_i(x^i, \dot{x}^i) \]

and the \( f_i, g_i \) are concave in the \( \dot{x}^i \), then the \( L(\alpha, x_0) \) function has as many as \( n \) kinks, each corresponding to a critical level, \( X_S^i \). Brock and Dechert [2] analyze a multiservice monopoly in a discrete time framework with this type of model. For more general functions \( F \) and \( G \), the remarks at the end of section II would apply to the multidimensional case as well.
IV. Appendix

**Lemma 1:** For \( T < T^* \) and \( x^*(t) > 0 \),

\[
\begin{align*}
 f(x_0, \dot{x}(0)) - \dot{x}(0)f_x(x_0, \dot{x}(0)) & - e^{-rt}[f(x(T), \dot{x}(T)) - \dot{x}(T)f_x(x(T), \dot{x}(T))] \\
 & = \int_0^T - \frac{d}{dt}[e^{-rt}[f(x(t), \dot{x}(t)) - \dot{x}(t)f_x(x(t), \dot{x}(t))]] dt \\
 & = \int_0^T [e^{-rt}r[f(x, \dot{x}) - \dot{x}f_x(x, \dot{x})] - e^{-rt}\frac{d}{dt}[f(x, \dot{x}) - \dot{x}f_x(x, \dot{x})]] dt \\
 & = \int_0^T e^{-rt}[rf - r\dot{x}f_x - [\dot{x}f_x + xf_x - xf_x - \frac{d}{dt}f_x]] dt \\
 & = r \int_0^T e^{-rt}f(x, \dot{x}) - \int_0^T e^{-rt}\dot{x}[f_x(x, \dot{x}) + rf_x(x, \dot{x}) - \frac{d}{dt}f_x(x, \dot{x})] dt \\
 & = r \int_0^T e^{-rt}f(x, \dot{x}) dt
\end{align*}
\]

where the last equality follows since the Euler equation is satisfied for \( x(t) \in \Omega(x_0) \). By applying the transversality condition (5) for \( T^* < \infty \) and (6) for \( T^* = \infty \), the result follows. \(/\!\!/\!\!/\)

**Theorem 1:** It is clear that \( p_0^* \in \mathcal{P}(x_0) \). If

\[
p = \inf \mathcal{P}(x_0) < p_0^* < \sup \mathcal{P}(x_0) = \overline{p}
\]

then there are sequences, \( \{p_n\} \) and \( \{\overline{p}_n\} \), of elements of \( \mathcal{P}(x_0) \) which converge to \( p \) and \( \overline{p} \), respectively. Thus there is an index, \( N \), so that

\[
p_N < p_0^* < \overline{p}_N
\]

and functions, \( x_N(t) \) and \( \overline{x}_N(t) \), both elements of \( \Omega(x_0) \) such that

\[
p_N = - f_x(x_0, \dot{x}_N(0)) \quad \quad \overline{p}_N = - f_x(x_0, \dot{\overline{x}}_N(0)).
\]

By the strict convexity of \( H \) in \( p \), it follows that

\[
H(x_0, p_0^*) < \max \{H(x_0, p_N), H(x_0, \overline{p}_N)\}. \quad \text{(A.1)}
\]
But this is a contradiction since \( V(x_0) = r^{-1}H(x_0, p_0^*) \), and (A.1) implies that there is a function \( x(t) \) such that

\[
V(x_0) < \int_0^\infty e^{-rt} f(x(t), \dot{x}(t)) dt
\]

contrary to the assumption that \( x^*(t) \) is an optimal solution. 

\[\text{THEOREM 2: If } \#P^*(x) \geq 3, \text{ then there are } p_1 < p_2 < p_3 \in P^*(x) \text{ with } \]

\[H(x, p_1) = H(x, p_2) = H(x, p_3), \text{ which cannot be since } H \text{ is strictly convex in } p. \text{ Hence } \#P^*(x) \leq 2. \]

To show the second part of the theorem, suppose that \( p_1, p_2 \in P^*(x) \) with \( p_1 < p_2 \). Then,

\[
0 = H(x, p_1) - H(x, p_2) > H_2(x, p_2)(p_1 - p_2) = \dot{x}_2(0)(p_1 - p_2)
\]

implies that \( \dot{x}_2(0) > 0 \), where \( -f_\dot{x}(x, \dot{x}_2(0)) = p_2 \). Similarly, \( \dot{x}_1(0) < 0 \). Hence, by the strict concavity of \( f \) in \( \dot{x} \), \( -f_\dot{x} \dot{x} \) is increasing in \( \dot{x} \) and so

\[
p_1(t) = -f_\dot{x}(x_1(t), \dot{x}_1(t)) < -f_\dot{x}(x_2(s), \dot{x}_2(s)) = p_2(s)
\]

for sufficiently small \( t, s > 0 \). 

\[\text{Monotonicity of optimal paths (Dechert and Nishimura[3]): Let } x(t) \text{ be optimal starting from } x_0. \text{ Without loss in generality, we may assume that } x_0 \leq x \text{ (see Figure 1). Suppose that for some } t_0 > 0 \text{ it is the case that } x(t_0) = x_0 \text{ and } x < x_0 \text{ for } 0 < t < t_0. \text{ Define } x^* \text{ by}
\]

\[
x^*(nt_0 + t) = x(t) \quad \text{for } n=0,1,2,... \text{ and } 0 \leq t \leq t_0
\]

which, by the Principle of Optimality is also optimal. By Jensen's inequality,

\[
\int_0^\infty re^{-rt} u(f(x^*) - \dot{x}^*) dt \leq u\left(\int_0^\infty re^{-rt} [f(x^*) - \dot{x}^*] dt\right) \quad (A.1)
\]

and since \( x^*(t) \leq x_0 \leq x \),
\[ \int_0^\infty e^{-rt}[f(x^*) - \dot{x}^*] = x_0 + \int_0^\infty e^{-rt}[f(x^*) - rx^*]dt \leq x_0 + \int_0^\infty e^{-rt}[f(x_0) - rx_0]dt = x^{-1}f(x_0). \quad (A.2) \]

Hence,
\[ \int_0^\infty re^{-rt}u(f(x^*) - \dot{x}^*)dt \leq f(x_0) \quad (A.3) \]

where strict inequality holds in (A.1) and (A.2) by the non constancy of \( x^*(t) \).

Therefore, by combing (A.1) and (A.3),
\[ \int_0^\infty e^{-rt}u(f(x^*) - \dot{x}^*)dt \leq x^{-1}u(f(x_0)) = \int_0^\infty e^{-rt}u(f(x_0))dt \]

and so \( \tilde{x}(t) = x_0 \) for all \( t \) is strictly better than \( x^*(t) \), which by construction is an optimal path. This same contradiction can be derived for a path which starts at \( x_0 \) and increases \( (\dot{x}(0) > 0) \) before returning to \( x_0 \), since we can find an \( x_0' = x(t) \) for some \( t > 0 \) with \( \dot{x}(t) < 0 \). The conclusion of this argument can be succinctly stated as: \( \dot{x}(t)\dot{x}(0) \geq 0 \) for all \( t \) along any optimal path. \\

**THEOREM 3**: Let \( x_n(t) \) satisfy the Euler equation and the Weierstrass–Erdmann corner conditions, and let \( p_n(t) \) be the corresponding costate variable for the truncated economies with \( x_n(0) = x_0 \) and \( p_n(0) = p_0 \) (see Figure 2). Similarly, let \( x(t), p(t) \) be the corresponding state and costate for the original economy with the same initial conditions. Also, let \( T_n \) be the times such that \( x(T_n) = x_n \). Note that
\[ x_n(t) = x(t), \ p_n(t) = p(t) \quad 0 \leq t \leq T_n \]

and
\[ \lim_{n \to \infty} x_n(t) = x(t), \ \lim_{n \to \infty} p_n(t) = p(t) \quad 0 \leq t. \]

It is obvious that \( p_{n+1}(t) \leq p_n(t) \), and since \( p_n(t) = u'(F_n(x_n(t) - \dot{x}_n(t))) \),
\[ u(c_n(t)) \leq u(c(t)) \]
where \( c_n(t) = F_n(x_n(t)) - \dot{x}_n(t) \) and \( c(t) = F(x(t)) - \dot{x}(t) \). Denote the value of the truncated and original economies by \( V_n(x_0) \) and \( V(x_0) \). Then by construction,

\[
\int_0^\infty e^{-rt}u(c_n(t))dt \leq V_n(x_0) \leq V(x_0)
\]

and \( rV_n(x_0) = H(x_0, p_n) \). Hence,

\[
\int_0^\infty e^{-rt}u(c(t))dt \leq r^{-1}H(x_0, \overline{p}).
\]

To show that equality holds, suppose that \( \dot{x}(t) \) satisfies the Euler equation and the Weierstrass–Erdmann corner conditions with initial conditions \( x(0) = x_0 \) and \( \dot{p}_0 = u'(f(x_0) - \dot{x}(0)) < \overline{p} \). For some \( n, \dot{p}_0 < p_n \) and hence \( x(t) < x_n \) for all \( t \). Therefore,

\[
\int_0^\infty e^{-rt}u(c(t))dt \leq V_n(x_0) \leq \int_0^\infty e^{-rt}u(c(t))dt
\]

where \( c(t) = F(x(t)) - \dot{x}(t) \). On the other hand, if \( \dot{p}_0 > \overline{p} \) then by construction \( x(t), c(t), x(t), c(t) \) are all monotonic functions of \( t \) with \( c(0) < c(0) \), which implies that \( \dot{x}(0) > \dot{x}(0) \), and since \( F'(x) \) is increasing in \( x \), this in turn implies that \( c(t) < c(t) \) for all \( t \). Hence \( u(c(t)) \leq u(c(t)) \) for all \( t \), and the conclusion of the theorem follows.

---

**Lemma 2:** Let \( x^*, Q^* \) be a solution to (20) for initial values of \( x_0, a_0 \). Then,

\[
L(a, x_0) \geq \int_0^\infty e^{-rt}[\alpha P(Q^*)Q^*] + (1 - \alpha)A(Q^*) - C(I^*)]dt
\]

\[
L(a_0, x_0) = \int_0^\infty e^{-rt}[\alpha_0 P(Q^*)Q^*] + (1 - \alpha_0)A(Q^*) - C(I^*)]dt
\]

Therefore, by subtracting these two equations, we get

\[
L(a, x_0) - L(a_0, x_0) \geq (\alpha - \alpha_0) \int_0^\infty e^{-rt}[P(Q^*)Q^* - A(Q^*)]dt
\]

and hence \( [-\overline{S}, -\underline{S}] \subseteq \partial_a L(a_0, x_0) \).

---

**Lemma 3:** First, let us show that \( P(x) \) is closed. Let \( p_n \in P(x) \) and suppose that \( p_n \rightarrow p \). If \( x_n(t) \) satisfies (2a) with \( x_n(0) = x, -f_\dot{x}(x, \dot{x}_n(0)) = p_n \), and \( x(t) \) satisfies (2a) with \( x(0) = x, -f_\dot{x}(x, \dot{x}(0)) = p \), then \( x_n(t) \rightarrow x(t) \) and \( \dot{x}_n(t) \rightarrow \dot{x}(t) \) for all \( t \). By the continuity of \( H \) and the dominated convergence theorem,
\[ H(x,p) = \lim_{n \to \infty} H(x, p_n) = \lim_{n \to \infty} r \int_0^\infty e^{-rt} f(x_n(t), \dot{x}_n(t)) \, dt \]

\[ = r \int_0^\infty e^{-rt} f(x(t), \dot{x}(t)) \, dt \]

and hence \( p \in P(x) \). To show that \( P(x) \) is bounded, suppose that there are \( p_n \in P(x) \) with \( |p_n| \geq n \). For any \( p^* \in P^*(x) \),

\[ f(x, u) + p_n u \leq H(x, p_n) \leq H(x, p^*) \]  \hspace{1cm} (A.4)

holds for all \( u \). In particular for \( u_n^i = \text{sign}(p_n^i) \)

\[ f(x, u_n) + p_n u_n \geq |p_n| + \min_k f(x, e_k) \]

where the \( e_k^i \) range over all combinations of \( \pm 1 \). Since \( |p_n| \) are unbounded, this contradicts (A.4). Hence \( P(x) \) is compact. By the strict convexity of \( H \) in \( p \), it is clear that an optimal costate, \( p^* \), cannot be a convex combination of the elements of \( P(x) \). ///

**THEOREM 4:** Since \( P(x) \) is compact, so is \( K \), and since \( H \) is convex in \( p \), the maximizing value of \( p \in K \) must lie in the boundary of \( K \), \( \partial K \). By the Krein–Milman Theorem (Rudin [14, pg. 70]) \( K \) is the convex hull of the set of its extreme points. By the argument in the last sentence of Lemma 3, the maximizing \( p \in K \) must lie in this set of extreme points which lies in \( \partial K \). ///

**THEOREM 5:** For this proof we closely follow the development in Hestenes [9, Ch. 3] of Meyer field theory. Since we are assuming that \( f_{x^i_T} \neq 0 \), the solution to (1) is non-singular and has no corners. For \( x \in A^i_1 \), this implies that the finite horizon fixed endpoint problem of connecting \((0, x)\) with \((T, x^*(T))\) has no point which is conjugate to 0 between 0 and \( T \). [9, Theorem 3.1, pg 124] This holds for all \( T \), and implies that \( x^* \) is an extremal of a Meyer field. [9, Theorem 8.1, pg 137] By the construction of a Meyer field [9, section 7, pg 136] there is a slope function \( U(x) \) which is \( C^1 \), such that the optimal costate at \( x \) is \( p^*(x) = -f_{x^i_T}(x, U(x)) \). By our Lemma 1 \( rV_{A^i_1}(x) = H(x, p(x)) \), and hence \( V_{A^i_1} \) is \( C^1 \) at \( x \). By hypothesis, \( x(t) \) is
in $B_i \setminus B_j$ for $-1 \leq t \leq 1$, and $V_{A_j}(\xi(t)) \leq V_{A_i}(\xi(t))$ for $t \leq 0$. Since $V_{A_j}(\xi(0)) = V_{A_i}(\xi(0))$, we have that

$$\frac{V_{A_j}(\xi(t)) - V_{A_j}(\xi(0))}{t} \leq \frac{V_{A_i}(\xi(t)) - V_{A_i}(D(0))}{t} \quad t < 0$$

and a similar inequality holds for $t > 0$. By taking limits as $t \to 0$, and noting that $\xi(0) = x$, $p^*(\xi(t^-)) = p_{A_1}$, and $p^*(\xi(t^+)) = p_{A_j}$, the conclusion follows.
Footnotes

1 To be more specific we are assuming that $f$ is an extended real valued function and that there is an open connected subset, $\Delta = (0, \infty) \times \mathbb{R}$, such that for each $x$, $\{i \mid (x, i) \in \Delta\}$ is convex, and $f$ is $C^2$ and strictly concave in $i$ on $\Delta$. Moreover, in the interior of the complement of this set $f$ takes on the value of $-\infty$. We are assuming that the problem is 'normal' so that in Hamiltonian form the multiplier on the objective function is non zero. In this way we can apply Michel's Theorem [11, pg 981] to deduce that the Hamiltonian goes to zero as $t \to \infty$.

2 Other sets of functions may be more appropriate for specific types of problems. For example, it may be required that $0 \leq x(t) \leq X$, in which case our arguments go through without any alterations at all.

3 By this we mean that if $(x_n, i_n) \in \Delta$ (footnote 1) converges to $(x, i) \in \partial \Delta$, then $f(x_n, i_n)$ converges to $f(x, i)$.

4 For more general problems, one can apply the techniques of Scheinkman [15] when the Hamiltonian is separable, or the techniques of Brock and Scheinkman [3] and Magill and Scheinkman [10] to obtain conditions for the asymptotic stability of steady states.

5 In discrete time Brock and Dechert (1982) showed that $[-S, -S] = \partial_a I(c_0, x_0)$.

6 What is needed here is that for some $\varepsilon > 0$, $|u| \leq \varepsilon$ is an allowable control. Then set $u_n^i = \varepsilon \cdot \text{sign}(p_n^i)$. 
References


FIGURE 2
FIGURE 4