OPTIMAL ALLOCATION IN A PURE CAPITAL MODEL WITH WITHDRAWALS: ASYMPTOTIC PROPERTIES AND ALGORITHMS

by

Alan J. Seelenfreund

TECHNICAL REPORT NO. 95

May 31, 1967

Supported by the Army, Navy, Air Force, and NASA
under Contract Nonr-225(53) (NR-042-002)
with the Office of Naval Research*

Gerald J. Lieberman, Project Director

*This research was partially supported by the Office of Naval Research under Contract Nonr-225(50) (NR-047-004)

DEPARTMENT OF STATISTICS

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

Non-Technical Summary

The object of this report is the mathematical formulation and analysis of a class of economic processes. The processes have a particular common structure that can be exploited to facilitate computation of optimal policies and to obtain insight into the behavior of the process after it has been in operation for a period of time. Specifically, we shall analyze models for Capital Budgeting, Price Speculation, Warehouse Operation, and Economic Growth.

To illustrate the type of result obtained we first focus on a capital budgeting problem. Suppose that at a given point of time a firm has a certain amount of cash that it must allocate between present dividend payments and a number of investment opportunities which have cash-flow profiles representing future payouts to the firm. The net cash-flow from a given investment in a period subsequent to initiation of the project must be non-negative but could be represented by a random variable. In addition, the amount invested in any given project may have an upper limit placed on it, as may the amount withdrawn in any period. The object is to devise a schedule of dividends and investments to maximize the discounted sum of withdrawals. A finite computational technique is presented that finds the optimal schedule for the case of a finite number of investment periods and also for the case wherein the number of investment periods remaining can be arbitrarily large. It is also shown that the expected return from the entire process is directly proportional to the amount of resources one starts with.

The assumption placed upon the cash-profile of the investments in the model just discussed restricts an investment to be of the point-input stream-output type (i.e. net investment expenditure is not required in any period subsequent to the first). In addition, the objective criterion restricts the stockholders valuation of dividends to be linear or proportional over all magnitudes of payout and thus does not admit any decreasing marginal utility on the stockholders part. An algorithm utilizing the principles of generalized programming is derived for solving the more general problem wherein these two assumptions are relaxed. The method avoids the necessity of solving a nonlinear problem by reducing the solution technique to one involving the solution of a sequence of Linear Programming problems.

Another process considered in this report is one corresponding to a model of economic growth. This is a process wherein at each point in time a society must allocate quantities of available commodities, e.g. (steel, manpower, land) to present consumption, or to a number of industries each of which produce one of the commodities for the next time period. It is assumed that each industry has a spectrum of processes available to it, and the objective is to allocate commodities into production and consumption in such a way as to maximize the discounted sum of consumption utility derived by the society. It is shown that under certain assumptions a specific set of processes, including precisely one process from each industry, can be identified where these are the processes that must be employed when following an optimal policy if the economy has been in operation for a sufficiently long time.

CHAPTER I

INTRODUCTION TO THE CLASS OF PROCESSES TO BE STUDIED

1. Notation and Description of the Process

In order to set the stage for discussion of the multistage decision process we must define several concepts and pieces of notation.

We assume that the condition of our process may be described at any point in time by an element x of the normed linear space $\underline{\bar{X}}$ which will be referred to as the state space. Examples of interesting state spaces are the Euclidean spaces of M dimensions, (E^M), and the space of bounded infinite sequences (ℓ^∞).

At each decision point in the stream of time we are required to apply a control v to alter the future progress of the process as described by the state vector and, possibly, to receive a current benefit. We assume that the control (decision) space is a subset of E^P, i.e., the control variable v may be represented by a P-dimensional vector $1 \leq P < \infty$. We also assume that the reward (or negative of the cost) from applying control v at time t is given by $U_t(\cdot)$, a concave functional defined on E^P. Employing standard notation we let $\mathcal{L}(R,S)$ represent the set of bounded linear transformations from the normed linear vector space (NLVS) R to the NLVS S.

We thus define the transformations:

$$A_t \in \mathcal{L}(E^P,E^M) \qquad 1 \leq M < \infty$$

$$D_t \in \mathcal{L}(\underline{\bar{X}},E^M)$$

$$T_t \in \mathcal{L}(\underline{\bar{X}} \times E^P,\underline{\bar{X}})$$

and the constant vector $B_t \in E^M$. In the above the subscript t is an index of elapsed time and takes values from a set γ. If the total number of decision stages to be considered is the finite number N, then γ is the set of integers $\{1,2,3,\ldots,N\}$. If there are an infinite number of decision stages to be considered, then γ is the set of positive integers.

If we let x^t be the state of the process at time t, and v^t be the control applied at time t, we may describe the process by the difference equation:

$$x^{t+1} = T_t(x^t, v^t) \quad \text{for} \quad t \in \gamma$$

where

$$v^t \in \{v: A_t(v) = D_t(x^t) + B_t,\ v \geq 0\} = V_t(x^t) \quad \text{for} \quad t \in \gamma .$$

Each $V_t(\cdot)$ represents the feasible control region at time t. We should note that unlike many formulations of control problems the feasible region is a function both of time and state.

If the objective is to maximize the sum of the rewards received at each stage of the process, the optimization problem may be stated as:

$$\text{maximize} \quad \sum_{t \in \gamma} U_t(v_t)$$

subject to

$$x^{t+1} = T_t(x^t, v^t) \quad \text{for} \quad t \in \gamma$$

(1.1) $$v^t \in V_t(x^t) \quad \text{for} \quad t \in \gamma$$

$$x^1 = x \quad \text{the initial location in state space.}$$

It is clear that the problem formulated above includes the usual formulation of a control problem with linear state transformations, since a finite number of the values of the state variables can be included in the domain of definition of the value functional by appropriate manipulations of A_t, D_t, and B_t. Additionally the dimension of the spaces given by P and M could be made to depend on the time parameter t without changing the underlying structure. This will not be done here due to the notational complications.

The abstract formulation presented above encompasses many statements of concrete problems. Illustrations of such problems will be presented following the development of the pertinent segments of the theory. Our objective throughout this work will be to analyze versions of the general problem to arrive at the structure underlying the solution to the problem and to present computational algorithms wherever possible.

2. Formulation of a Stochastic Model

In the model formulated in Section 1 we assumed a deterministic environment existed and that, given any particular time, state, and action the new state would be precisely determined. In this section we formulate an interesting generalization of this model where the new state is specified by a probability distribution of possible states.

Retaining the notation of the previous section we introduce a sequence of independent random vectors $\langle r^t \rangle_{t \in \gamma}$. At each decision point, t, we assume that the current state of the system x^t and the feasible control region $V_t(\cdot)$, a function of the state variable, are known. The state transition rule $T_t(\cdot,\cdot)$ is now made to depend upon the outcome of the random vector r^t; thus, $x^{t+1} = T_t(x^t, v^t, r^t)$.

Using the above ideas, problem (1.1) can be restated in the form:

$$(1.2) \qquad \text{maximize} \sum_{t \in \gamma} E\{U_t(v_t)\}$$

subject to

$$x^{t+1} = T_t(x^t, v^t, r^t) \qquad t \in \gamma$$

$$v^t \in V_t(x^t) \qquad t \in \gamma$$

$$x^1 = x ,$$

where $E\{\cdot\}$ represents the expectation operator in this case taken over the joint distribution of the random vector sequence $\langle r^t \rangle_{t \in \gamma}$. The transition rule $T_t(\cdot,\cdot,\cdot)$ is required to be linear in x^t and v^t, but not in r^t.

3. Techniques for Solution of the Finite Horizon Model

If we do not wish to make any assumptions regarding the form of the various transformations it would not seem that enough structure is available to yield analytic insight into the properties of the solution. However, the main characteristics of linearity lead one to suspect that linear programming could be used to obtain numerical solutions, provided $U_t(\cdot)$ were linear for all $t \in \gamma$. To do this we need the following simple lemma:

Lemma 3.1. Let both $\underline{\overline{X}}$ and $\underline{\overline{V}}$ be a NLVS. Let $T \in \mathcal{L}(\underline{\overline{X}} \times \underline{\overline{V}}, \underline{\overline{X}})$. Then there are transformations

$$T' \in \mathcal{L}(\underline{\overline{X}}, \underline{\overline{X}})$$

and

$$T'' \in \mathcal{L}(\underline{\overline{V}}, \underline{\overline{X}})$$

such that

$$T(x,v) = T'(x) + T''(v) \quad \text{for all } (x,v) \in \underline{\overline{X}} \times \underline{\overline{V}} .$$

Furthermore, the decomposition is unique.

Proof: Let

$$T'(x) = T(x,0) \quad \text{for all} \quad x \in \overline{\underline{X}}$$

$$T''(v) = T(0,v) \quad \text{for all} \quad v \in \overline{\underline{V}} .$$

The required properties follow from the continuity and linearity of T. Q.E.D.

Thus, we see that the control problem posed in (1.1) is formally equivalent to a linear programming problem in the variables $(x^2,x^3,\ldots,x^T,v^1,v^2,\ldots,v^T)$ where $x^t \in \overline{\underline{X}}$ and $v^t \in E^P$ for all t and $\overline{\underline{X}}$ is finite dimensional.

CHAPTER II

A CLASS OF PROBLEMS WITH LINEAR UTILITY FUNCTIONS AND FINITE HORIZONS

1. An Algorithm for a Closed Form Solution

The model to be considered in this chapter is the one introduced in Chapter I with the utility function in each period a bounded linear functional defined on the control space E^P, and the time horizon finite, i.e., $\gamma = \{1,2,\ldots,N\}$. The utility function at time t is a linear function denoted by the P-dimensional vector C_t.

We shall now explore the possibilities of obtaining analytic solutions to problem (1.1) if we specialize the structure of the transformations.

Consider the LP problem:

$$\begin{aligned} \text{Max } & c(v) \\ & A(v) = b \\ & v \geq 0. \end{aligned}$$

Define the ordered pairs of sets $<\mathcal{C}, \mathcal{B}>$, where $\mathcal{C} \subset \mathcal{L}(E^P,E^1)$ and $\mathcal{B} \subset E^M$ as follows:

> For all $c \in \mathcal{C}$ optimal basic feasible sets of activities are independent of b as long as $b \in \mathcal{B}$. Call such a pair $<\mathcal{C}, \mathcal{B}>$ a stable set of A.

Clearly, pairs of the form $<\mathcal{C}, \mathcal{B}>$ exist for any matrix A (they may be singletons for which an optimal solution exists) and there may be many different combinations of sets satisfying the requirements. For particular examples sensitivity analysis could be employed to determine the sets of interest. For certain classes of matrices, A, it is possible to find ordered pairs $< \mathcal{C}, \mathcal{B} >$ such that corresponding sets of the form $\mathcal{C}$ and $\mathcal{B}$ are very large. For instance, if A has Leontief structure (see the first example discussed in Section 2 of this chapter), then the set $\mathcal{C}$ can be all of $\mathcal{L}(E^P, E^1)$ and the corresponding $\mathcal{B}$ the non-negative orthant of E^M.

In describing the statement of the control problem and in relating the abstract formulation to more concrete problems it is more convenient to employ time subscripts indexing elapsed time. In proving the following theorem it is more convenient to consider the time subscripts as the number of stages remaining. Since, with a finite horizon, this merely involves a relabeling of the transformations, none of the structure is changed. Letting $f_N(x)$ be the return from an N-stage process following an optimal policy from starting position x, we apply the principle of optimality (Bellman [4]), and obtain the recursion relations:

(2.1)
$$f_N(x) = \max_{v \in V_N(x)} \{c_N(v) + f_{N-1} \circ T_N(x,v)\}$$

$$f_0(x) = 0 ,$$

where the subscripts denote the number of periods remaining in the program. In order to state the linearity theorem we need the following definitions for a process with N stages remaining.

Let the sequence $< \mathcal{C}_n, \mathcal{B}_n >_{n=1}^N$ be a sequence of stable sets of $<A_n>_{n=1}^N$, and let

$$\mathcal{R}_n = \{x \in \underline{X}: D_n(x) + B_n \in \mathcal{B}_n\}.$$

Let

$$L_N = \sum_{n=1}^{N} E_n \circ G_{n+1} \circ G_{n+2} \circ \cdots \circ G_N$$

and

$$K_N = \sum_{n=1}^{N} [c_n + L_{n-1} \circ T_n''] \circ \hat{A}_n^{-1} B_n$$

where

$$E_n = c_n \circ \hat{A}_n^{-1} \circ D_n \qquad n = 1,\ldots,N$$

$$G_n = T_n'' \circ \hat{A}_n^{-1} \circ D_n + T_n' \qquad n = 1,\ldots,N$$

and

$\hat{A}_n$ is the matrix of the optimal basis with n periods remaining where $n = 1,\ldots,N$.

The transformations T_n'' and T_n' are the canonical decompositions of T_n given by Lemma 1.3.1. We can now state the

Linearity Theorem (Theorem 1): Suppose

1. $c_1 \in \mathcal{C}_1$
2. $[c_n + L_{n-1} \circ T_n''] \in \mathcal{C}_n \qquad n = 1,2,\ldots,N$
3. $x \in \mathcal{R}_N$
4. $T_n(x,v) \in \mathcal{R}_{n-1}$ for all $x \in \mathcal{R}_n$ and $v \in V_n(x)$,

   $n = 1,2,\ldots,N.$

Then $f_N(x)$ is an affine functional on $\underline{X}$ for all $x \in \mathcal{R}_N$ and the optimal choice of positive activities during any stage is independent of x. Equivalently,

$$f_N(x) = L_N(x) + K_N \quad \text{for all} \quad x \in \mathcal{R}_N$$

where

$$L_N \in \mathcal{L}(\underline{X}, E^1) .$$

Proof: The proof is by induction on the number of periods remaining, denoted by N.

$N = 1$. From (2.1),

$$f_1(x) = \max_{v \in V_1(x)} \{c_1(v)\} .$$

Clearly, for fixed x this is the LP problem

$$\max \{c_1(v)\}$$

subject to

$$A_1(v) = D_1 x + B_1$$

$$v \geq 0 .$$

Since $(\mathcal{C}_1, \mathcal{B}_1)$ are stable for A_1, the assumption that $c_1 \in \mathcal{C}_1$ allows us to find an optimal basis $\hat{A}_1$ independent of $x \in \mathcal{R}_1$. Thus, the optimal decision is to have

$$v = \hat{A}_1^{-1} \circ D_1(x) + \hat{A}_1^{-1} B_1$$

for the activities corresponding to the columns of $\hat{A}_1$, and all other variables set equal to zero. Thus,

$$f_1(x) = c_1 \circ \hat{A}_1^{-1} \circ D_1(x) + c_1 \circ \hat{A}_1^{-1} B_1 \quad \text{for} \quad x \in \mathcal{R}_1 ,$$

and since the class of linear transformations is closed under composition and addition the result is established.

We now assume the result holds with $N-1$ stages remaining, and complete the induction. By assumptions 3 and 4 we see that $T_N(x,v) \in \mathcal{R}_{N-1}$ at all points of $\underline{X} \times \underline{V}$ that can occur. Thus, using (2.1) and the induction hypothesis we have:

(2.2) $$f_N(x) = \max_{v \in V_N(x)} \{c_N(v) + L_{N-1} \circ T_N(x,c)\} + K_{N-1}$$

$$= \max_{v \in V_N(x)} \{[c_N + L_{N-1} \circ T''_N](v)\} + L_{N-1} \circ T'_N(x) + K_{N-1}.$$

Note that L_{N-1} can be calculated independently of the decision to be made with N periods to go.

Again since $(\mathcal{C}_N, \mathcal{B}_N)$ are stable for A_N, the assumption that $[c_N + L_{N-1}T''_N] \in \mathcal{C}_N$ enables us to find an optimal basis $\hat{A}_N$ independent of $x \in \mathcal{R}_N$. Thus, for all $x \in \mathcal{R}_N$, $v = \hat{A}_N^{-1} \circ D_N + \hat{A}_N^{-1}B_N$ for the activities corresponding to $\hat{A}_N$, all other activities being operated at zero level.

In (3.2) we now substitute the optimal solution, obtaining

$$f_N(x) = c_N \circ \hat{A}_N^{-1} \circ D_N(x) + L_{N-1} \circ [T''_N \circ \hat{A}_N^{-1} \circ D_N + T'_N](x)$$

$$+ K_{N-1} + c_N \circ \hat{A}_N^{-1}B_N + L_{N-1} \circ T''_N \circ \hat{A}_N^{-1}B_N$$

or

$$f_N(x) = [E_N + L_{N-1} \circ G_N](x) + K_{N-1} + [c_N + L_{N-1} \circ T''_N] \circ \hat{A}_N^{-1}B_N \ .$$

Clearly,

$$L_N = E_N + L_{N-1} \circ G_N$$

$$= \sum_{n=1}^{N} E_n \circ G_{n+1} \circ \cdots \circ G_N$$

and

$$K_N = \sum_{n=1}^{N} [c_n + L_{n-1} \circ T''_n] \circ \hat{A}_n^{-1}B_n \ .$$ Q.E.D.

Assumption 4 of the theorem can be weakened somewhat by requiring $T_n(x,v) \in \mathcal{R}_{n-1}$ for all $x \in \mathcal{R}_n$, but only for those $v \in V_n(x)$ that could correspond to an optimal solution.

2. Examples of Processes Satisfying the Conditions of Section 1

In order to illustrate the preceding theorem we shall examine several examples. It is natural to first look for classes of matrices with large stable sets. One such class sometimes referred to as Leontief matrices with substitution is the class of $m \times p$ matrices λ such that $A \in \lambda$ if and only if for each $m \times m$ nonsingular submatrix A_i of A the existence of a non-zero, non-negative pair of vectors $\langle v,b \rangle$ such that

$$A_i v = b$$

implies that $A_i^{-1} \geq 0$. Thus, we can state a result essentially observed by Dantzig [7] as follows:

If $A_t \in \lambda$, $B_t \geq 0$ and both D_t and T_t are non-negative transformations for all t, then the hypotheses of the linearity theorem are satisfied with $\mathcal{C}_t = E^P$ and $\mathcal{R}_t$ the non-negative orthant of E^M for all t.

It is easy to see that the capital budgeting problem of Dorfman [9] and Manne [14], the warehouse problem of Dreyfus [11], and the price speculation model of Arrow and Karlin [2] all have the above characteristics. In analyzing the structure of each of these problems the above authors have essentially first proven a version of the linearity theorem for the particular problem under investigation. Since the structure of optimal policies in the warehouse and price speculation models has been analyzed in the above articles we will use the linearity theorem to analyze the capital budgeting problem. We will then show how a modification to the problem which circumvents the all-or-none problem raised by Manne can be attacked by the preceding theorem even though the resultant sequence of matrices $\langle A_t \rangle$ are no longer formally of the Leontief type. The first version of the capital budgeting problem to be analyzed is similar to the

Dorfman-Manne formulation and will now be described.

Capital Budgeting

A firm is presented with a sequence of sets of investment opportunities. At each decision point the firm must choose a mix of investment opportunities from the set currently available. Additionally, the firm must decide whether or not to pay some portion of its present resources in dividends.

The state space for such a firm may be described as the couple (x,Q) where x is an element of E^{N+1}, i.e., $x = (x_0, x_1, x_2, \ldots, x_N)$, and each x_i represents the net cash input i periods from the present time due to past investments. The investment possibilities are represented by the sequence $Q = (Q_1, Q_2, \ldots, Q_N)$ where Q_n is the finite set of investment opportunities available to use with n periods remaining in the decision process.

An investment opportunity $a^j \in Q_n$ will be considered to be an element of E^N, i.e.,

$$a^j = (a_1^j, a_2^j, \ldots, a_N^j)$$

where a_i^j is the cash return of investment a^j, i periods from the investment date, per unit invested. Note:

$$Q_n = \{a^1, a^2, \ldots, a^{M_n}\} .$$

The following assumptions are made about the firm and its opportunities:

A^1: The firm is self-financing.

A^2: Dividends are discounted with discount factor $\beta \geq 0$.

A^3: Returns from investments are proportional to the amount invested.

A^4: $x \geq 0$, i.e., $x_n \geq 0$, $n = 0,1,2,\ldots$

$a \in Q_n \Rightarrow a \geq 0$, i.e., $a_i \geq 0$, $i = 1,2,\ldots,N$.

A^5: The firm's utility function is a linear function of the dividends paid out.

The problem is to choose at each period of time a policy specifying the amount of current capital to be allocated to each investment opportunity and the amount to be paid out in dividends. The objective of the firm is the maximization of the sum of the discounted value of dividend utility. For simplicity in presentation we shall assume that the value of any income received beyond the planning horizon is zero. A slight modification to the period zero costing scheme would enable us to include some discounted value of returns received beyond the planning horizon.

Let us now make the identifications required to state the problem in terms of the notation developed for Theorem 1. The decision taken with n periods remaining, v_n, can be represented as a $1 \times (M_n+1)$ vector, the first component representing the amount of capital consumed in opportunity i. Similarly, $T''_n \equiv Q_n$ for each n considered as an operator mapping E^{M_n} into E^N, $B_n \equiv 0$, $c_n = (u_n,0,\ldots,0)$, and $T'_n = L$, a left-shift operator which adjusts the state variable to compensate for an increment of elapsed time; i.e., if $x = (x_0,x_1,x_2,\ldots,x_N)$, then $Lx = (x_1,x_2,\ldots,x_{N-1},0)$. Since each A_n is simply a row of ones, it is clear that each A_n is of Leontief type, justifying an application of the linearity theorem, yielding for all $x \geq 0$:

$$f_N(x) = L_N(x) \text{ a linear functional of } x.$$

Due to the nature of A_n, all extreme points of the convex set of possible decisions at any point in time are vectors with one component a plus one and all other components zero. This, combined with the linear characteristic of $f_N(\cdot)$, yields the all-or-none theorem of Manne [14] since only one of the v_j need be positive in an optimal solution. Thus, either we pay a dividend or invest in one best investment at each decision point. Consequently, if we invest during period n, $E_n = 0$, and if a dividend is paid, $E_n = u_n D$, where D is a $1 \times N$ vector with the first component a one and all other components zero. Also, $G_n = L$ if a dividend is paid in period n, and $G_n = a^{c_n} D + L$ if investment opportunity c_n is chosen.

Once the linearity of the return function is established it is easy to see that $L_N(\cdot)$, and hence $f_N(\cdot)$ take the following form:

1° $f_N(x) = \sum_{n=0}^{N} P_{N-n} u_n x_n$ for $N \geq 0$ and $P_0 = 0$.

2° $P_1 = 1$

$$P_N = \max \{u_N, \sum_{n=1}^{N} P_{N-n} u_n a_n^{c_N}\} \quad \text{for} \quad N \geq 2 .$$

3° $a_n^{c_N}$ is chosen as the investment maximizing

$$\sum_{n=1}^{N} P_{N-n} u_n a_n^i \quad \text{for} \quad i = 1,2,\ldots,M_N$$

and $P_N = u_N$ implies that dividends are paid during period N. An induction proof of this result can be found in Appendix A.

Clearly, the calculation of the sequence $\langle P_N \rangle$ is equivalent to calculating a sequence of optimal policies, pointing out the result that the optimal policy is independent of the initial state x. However, the optimum policy will certainly be dependent on the sequence of sets of investment opportunities Q.

In order to investigate properties of the solution it is clear that we must focus our attention on the sequence $\langle P_i \rangle$. We may interpret P_i as a shadow price representing the utility of an incremental unit of discounted cash with "i" periods remaining. In particular, P_N represents the utility value of an incremental dollar at the beginning of the plan. Since the P's are not dependent on the total amount available in any period, we have constant marginal utility in any period under an optimal policy. The linearity of the optimal return function yields an immediate proof of Manne's all-or-none theorem.

As observed in the paper by Manne [14], the implications of the all-or-none theorem are rather disturbing since most well-managed firms do not operate with this type of behavior. As one method of circumventing this apparent paradox we shall add another restriction to the problem. Specifically, we require that the firm pay out at least a fixed proportion p of the capital available in any period. The effect of this added restriction is to alter the region of feasible control with n periods remaining to

$$\{(v,s)\colon v_0 + v_1 + \dots + v_{M_n} = z,\ v_0 - s = pz,\ (v,s) \geq 0\} ,$$

where s is a non-negative slack variable and z represents the amount of capital resource available. This change yields an A matrix which is not of Leontief type since there is a feasible submatrix

$$\hat{A} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

such that

$$\hat{A}^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}$$

which is not a non-negative matrix. The solution corresponding to $\hat{A}$ sets $v_0 = pz$, $v_1 = (1-p)z$, $s = 0$, and all other $v_i = 0$. We will show that Theorem 1 can be applied to this problem with $p < 1$ and thereby demonstrate that the theorem's range of application is wider than the class of matrices of Leontief type.* To do this it is only necessary to verify that the assumptions 1-4 hold. Take $\mathcal{C}_i = E^{M_i+1}$ and $\mathcal{B}_i = \{y: y \in E^2, y_1 \geq y_2\}$ for $i = 1,\ldots,N$.

It is easily verified that the sequence $<\mathcal{C}_i, \mathcal{B}_i>_1^N$ is a sequence of stable sets of $<A_i>_1^N$ and that the $<\mathcal{R}_i>_1^N$ is such that each $\mathcal{R}_i$ is the non-negative orthant of E^N. Thus, assumptions 1 and 2 hold trivially, and assumptions 3 and 4 hold as long as $x \geq 0$ and $T_i(\cdot,\cdot)$ is a non-negative operator.

Another method of circumventing the conclusion of the all-or-none theorem is to impose a non-linear rather than a linear utility function of dividend payment. This approach will be formulated and discussed in Chapter III.

3. Stochastic Problems

In this section we will show that Theorem 1 can be extended to cover the stochastic problem (1.2) under appropriate additional assumptions. Let $f_N(x)$ represent the expected return following an optimal policy from state x with N periods remaining. Again following Bellman [4]

---

* However, A. Veinott has pointed out that the problem can be transformed into one with Leontief structure by subtracting the second constraint from the first, and using the derived constraint along with the constraint $v_0 - s = pz$.

we can derive the recurrence relations:

$$f_N(x) = \max_{v \in V_N} \{c_N(v) + \mathcal{E}_{r_N} f_{N-1} \circ T_N(x,v,r_N)\} \; N \geq 1,$$

$$f_0(x) = 0,$$

where $\mathcal{E}$ denotes the expectation operator. Note that the expectation is taken only over the random vector r_N.

<u>Theorem 2</u>. Suppose that conditions 1-4 of Theorem 1 hold for any possible realization of the vector sequence $<r_n>_1^N$. Then the conclusion of Theorem 1 holds with the expectation of $T_n(\cdot,\cdot,\cdot)$ (denoted $\overline{T}_n(\cdot,\cdot)$) replacing $T_n(\cdot,\cdot)$ in the evaluation of L_N and K_N, and the matrix $\hat{A}_n$ of the optimal basis with n periods remaining is evaluated by solving the linear program

$$\max_{v \in V_n(x)} c_n(v) + L_{n-1} \circ \overline{T}_n(x,v) .$$

<u>Proof</u>: The proof is identical to the one presented for Theorem 1 due to the fact that

$$\mathcal{E}_{r_N}\{L_{N-1} \circ T_N(x,v,r_N)\} = L_{N-1} \circ \overline{T}_n(x,v)$$

by the linearity of $L_{N-1}(\cdot)$. Also note that $\overline{T}_n(\cdot,\cdot)$ is linear in x and v since each realization was assumed to be a linear operator on $\overline{\underline{X}} \times \overline{\underline{V}}$.

Using Theorem 2 we can formulate a price-speculation model similar to one studied by Arrow and Karlin [2] for the deterministic case. The process can be described by a system with two-state variables, i.e., $x^t = (x_1^t, x_2^t)$ where x_1^t represents the amount of cash on hand at time t

and x_2^t represents the stock of commodity at time t measured in dollars. At each time we may withdraw funds from the speculation in the amount v_1, sell an amount of stock v_2, or buy an amount of stock v_3. Suppose commissions are paid on buying and selling, and our objective is to maximize the expected sum of discounted withdrawals from the system. Furthermore, suppose the value of the stock retained is subject to appreciation or depreciation due to a change in price of the commodity, and that this price change is of a random nature. The price change can be represented by a random factor r^t, where we assume that each factor r^t, $t = 1,\ldots,N$, has distribution function P_t such that $P_t\{r^t < 0\} = 0$ and the factors are independent from period to period. Translating this description into our standard format, we have:

$$V_t(x^t) = \{(v_1,v_2,v_3,v_4,v_5): \begin{array}{l} v_1 - a_{12}v_2 + a_{13}v_3 + v_4 = x_1^t \\ \qquad\quad v_2 + \qquad\qquad v_5 = x_2^t,\ v \geq 0\}. \end{array}$$

$$T_t(x^t,v^t,r^t) = \begin{pmatrix} v_4^t \\ r^t v_5^t \end{pmatrix} \quad , \quad U_t(v^t) = (\beta^t,0,0,0,0) \ ,$$

where a_{12} and a_{13} represent commissions; thus, $a_{13} \geq 1$ and $a_{12} \leq 1$.

It is easy to verify that the constraint matrix is of Leontief type and the transformation T_t is non-negative. Thus, Theorem 2 can be applied and we solve the problem replacing r^t by $\bar{r}^t$, its expected value. It is an interesting consequence of this result that the variance of the price change plays no part in evaluating an optimal policy. Of course, this result arises in part due to the linearity of the utility for withdrawals, and would not be true if the investor were risk-averse, i.e., possessed a strictly concave utility for withdrawals.

CHAPTER III

A GENERAL PROCESS WITH CONCAVE UTILITY AND FINITE-STATE SPACE

1. A Decomposition Type Algorithm

In this chapter we shall relax the restriction of linearity imposed on the objective functional and present an algorithm employing decomposition techniques to solve an interesting specific case of the general control problem. The vehicle for our derivation will be a version of the capital budgeting model studied in Chapter II. At that time it was pointed out that the structure of the solution was such as to preclude the necessity of both investing and distributing dividends in any one period. To circumvent this unappealing character of the optimal policy we introduced restrictions requiring a minimum percentage of available capital to be disbursed at each time. Another approach to the problem is to assume that the firm's marginal utility of withdrawals is strictly decreasing with the amount withdrawn. In other words, each utility function $U_t(\cdot)$ is strictly concave and increasing. Moreover, we will eliminate the point input stream output character of the investment opportunity profile given by assumption A4 in Section 2.2.

In terms of the notation presented in Chapter I, the process to be considered may be described as follows:

Let L be the left-shift operator described in Chapter II, i.e.,

$$L = \begin{bmatrix} 0 & & I \\ \vdots & & \\ 0 & \cdots & 0 \end{bmatrix} ;$$

then we want to maximize

$$\sum_{t=1}^{N} U_t(w_t)$$

where

$$x^{t+1} = Lx^t + I_t v^t \qquad t = 0,\ldots,N-1$$

$$x^1 = x$$

and

$$V_t(x) = \{v^t : w_t + v^t \cdot 1 = x_1,\ v \geq 0\} ,$$

with 1 the vector consisting of all ones with an appropriate dimension.

It is clear that in this formulation assumptions A1 and A3 given for the capital budgeting problem have been retained, assumption A4 has been dropped, and assumptions A2 and A5 have been changed to

A*: The firm's utility function during period t is $U_t(\cdot)$, a function of the amount of dividends paid out.

We shall assume that $U_t(\cdot)$ is concave, strictly increasing, and continuously differentiable for each t.

In a subsequent chapter we will investigate asymptotic properties of the solution for a similar class of problems. However, since this formulation of the capital budgeting model seems realistic it is desirable to have a relatively efficient technique available to solve numerically a specific finite horizon problem. For this purpose we will formulate the problem in a way that emphasizes the structure suitable for application of generalized programming techniques.

Following Baumol and Quandt [3], for any given time horizon T we may put the capital rationing problem in the form:

$$\max \sum_{t=1}^{T} U_t(w_t)$$

subject to

$$\sum_j a_{jt} v_j + w_t \leq M_t \qquad t = 1,2,\ldots,T$$

$$v_j \geq 0$$

$$w_t \geq 0$$

where

v_j is the number of units of project j constructed

M_t is the exogenous cash input at time t

$-a_{jt}$ is the net cash flow obtained from a unit of project j at time t

w_t is the amount of cash withdrawn at time t

$U_t(\cdot)$ is the utility of withdrawal at time t.

The assumption that $U_t(\cdot)$ is strictly increasing as a function of withdrawals implies that the t^{th} equation will hold with equality.

It seems that this formulation includes many of the dynamic allocation problems such as those considered in Bellman [4, Chap. 1]. Our problem then is given by:

A:
$$\max \sum_{t=1}^{T} U_t(w_t)$$

$$Av + I\overline{w} = M$$

$$v \geq 0$$

$$\overline{w} \geq 0$$

$$A = (a_{jt}),\ v = (v_j),\ \overline{w} = (w_t),\ M = (M_t)\ .$$

A problem equivalent to A is

B:
$$\max \ w_0$$
$$Av + \begin{bmatrix} w_1 \\ \vdots \\ w_T \end{bmatrix} = M \qquad v \geq 0$$
$$\begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_T \end{bmatrix} = w \in W$$

$$W = \{(w_0,w_1,\ldots,w_T): \ w_i \geq 0, \ i = 1,\ldots,T, \ \text{and} \ w_0 \leq \sum_{t=1}^{T} U_t(w_t)\}.$$

The concavity of U_t implies that W is convex. To see this, choose $w^1, w^2 \in W$ and let $w^\lambda = \lambda w^1 + (1-\lambda)w^2$; then

$$w_0^\lambda = \lambda w_0^1 + (1-\lambda)w_0^2 \leq \sum_{t=1}^{T} [\lambda U_t(w_t^1) + (1-\lambda)U_t(w_t^2)]$$

$$\leq \sum_{t=1}^{T} U_t(\lambda w_t^1) + (1-\lambda)w_t^2) = \sum_{t=1}^{T} U_t(w_t^\lambda) \ ;$$

thus $w^\lambda \in W$; hence, W is convex. The equivalence of problems A and B is clear:

If (x,w) is optimal for B, then $w_0 = \sum_{1}^{T} U_t(w_t)$; for if not, w_0 could be increased without destroying the feasibility of (x,w), thereby contradicting the assumption of optimality. Let $\bar{w} = (w_1,\ldots,w_T)$; then if (x,w) is optimal for B, $(x,\bar{w})$ is feasible for A with the same value of the objective function. Hence, we may conclude that

$$\max \sum_{t=1}^{T} U_t(w_t) \geq \max \ w_0 \ .$$

If $(v,\bar{w})$ is optimal for A, then (v,w) with $w_0 = \sum_1^T U_t(w_t)$ is feasible for B, again yielding the same value of the objective function. Thus, $\max w_0 \geq \max \sum_{t=1}^T U_t(w_t)$, proving the equivalence of problems A and B.

We now consider the problem:

$$* \qquad \max \sum_{i=1}^K w_0^i \lambda_i$$

subject to

$$Av + \sum_{i=1}^K \bar{w}^i \lambda_i = M$$

$$\sum_{i=1}^K \lambda_i = 1$$

$$w^i = (w_0^i, \bar{w}^i) \in W \text{ for all } i, \lambda = (\lambda_i) \geq 0,\ v \geq 0.$$

Since W is convex, any point of the form $\sum_1^K w^i \lambda_i$ is an element of W. This implies that any solution to * for a given number of vectors $w^i \in W$ is primal-feasible for B.

A feasible solution to * is

$$v = 0$$

$$\lambda_i = \frac{1}{T} \qquad i = 1,2,\dots,T$$

$$\bar{w}_j^i = \begin{cases} TM_i & \text{if } j = i \\ 0 & \text{o.w.} \end{cases}$$

$$w_0^i = U_i(TM_i)\ .$$

Thus, we can solve an initial master program * by the simplex method.

Following the usual generalized programming formulation, see Dantzig

[6], we form a subprogram which tests solutions to * for dual feasibility in B. Since each solution to * is primal-feasible for B, dual feasibility of the solution implies that it is optimal for B and hence for the original problem A.

Let $\bar{\pi}^k = (\bar{\pi}^k_1, \bar{\pi}^k_2, \ldots, \bar{\pi}^k_{T+1})$ be the optimal dual variables at the K^{th} stage of the master problem, where $\bar{\pi}^k_{T+1}$ is the shadow price corresponding to the $\Sigma\lambda_i$ constraint. Let

$$Z = \min_{w \in W} \left[\sum_{i=1}^{T} w_i \bar{\pi}^k_i - w_0 \right] .$$

If $Z \geq - \bar{\pi}_{T+1}$, then the current optimal solution to * is dual-feasible for B and thus optimal for B.

We thus have at each stage the subproblem:

Minimize

$$\sum_{i=1}^{T} w_i \bar{\pi}^k_i - w_0$$

subject to

$$\sum_{1}^{T} U_t(w_t) - w_0 \geq 0 \quad \text{and} \quad w_i \geq 0, \qquad i = 1, \ldots, T .$$

The Lagrangian for this problem is:

$$F(w,\mu) = \sum_{t=1}^{T} w_t \bar{\pi}^k_t - w_0 - \mu\left(\sum_{1}^{T} U_t(w_t) - w_0\right)$$

$$w \geq 0, \quad \mu \geq 0 .$$

Examining the single constraint equation for the subproblem, we observe that $\sum_{1}^{T} U_t(w_t) - w_0$ is a concave function of $w = (w_0, w_1, \ldots, w_T)$. Also, by choosing w_0 small enough we can find a point $\hat{w}$ for which the restriction holds with strict inequality. Thus, the Kuhn-Tucker constraint qualification holds, and since the objective function is linear the Kuhn-Tucker-Lagrange (KTL) conditions are both necessary and sufficient for optimality.

The KTL conditions for the K^{th} subproblem are:

$\mu = 1$ since w_0 is unrestricted in sign.

$\bar{\pi}_t^k \geq \frac{\partial}{\partial w_t} U_t(w_t)$, $\quad t = 1,\ldots,T$ with equality holding if $w_t > 0$.

$\sum_1^T U_t(w_t) = w_0$ since $\mu = 1$.

From our assumption postulating the strictly decreasing marginal utility of withdrawals we have

$$z_1 > z_2 \Rightarrow \frac{\partial}{\partial z_1} U_t(z_1) < \frac{\partial}{\partial z_2} U_t(z_2), \qquad t = 1,2,\ldots,T \ .$$

Thus, if

$$\bar{\pi}_t^k \geq \frac{\partial}{\partial w_t} U_t(0), \quad \text{then} \quad w_t = 0,$$

and if

$$\bar{\pi}_t^k < \frac{\partial}{\partial w_t} U_t(0), \quad \text{then} \quad w_t > 0.$$

Hence, for any t,

$$w_t > 0 \Leftrightarrow \bar{\pi}_t^k < \frac{\partial}{\partial w_t} U_t(0) \ ,$$

and otherwise $w_t = 0$.

We now use the above results to solve the subproblem. We know that $\bar{\pi}_t^k \geq 0 \ \forall t$ and $\forall k$ since the $\bar{\pi}^k$ represent a sequence of money shadow prices. Let L be the subset of the integers $1,\ldots,T$ for which

$$0 \leq \bar{\pi}_t^k < \frac{\partial}{\partial w_t} U_t(0), \qquad t \in L \ .$$

Then the KT conditions hold $\Leftrightarrow$

$$\frac{\partial}{\partial w_t} U_t(w_t) = \bar{\pi}_t^k \ , \quad t \in L$$

and

$$w_t = 0 \quad , \quad t \notin L \ .$$

Since the KTL conditions are necessary and sufficient there will be a unique w vector solving the K^{th} subproblem with

$$w_0 = \sum_{t=1}^{T} U_t(w_t)$$

$$w_t = v_t(\bar{\pi}_t^k), \quad t \in L$$

$$w_t = 0 \qquad , \quad t \notin L \ ,$$

where v_t is the inverse function of $\frac{\partial}{\partial w_t} v_t(w_t)$ which exists since we have assumed a strictly decreasing marginal utility of withdrawals. Thus we have found a method to solve the subproblem at each stage by inspection. Inasmuch as the generalized programming procedure has been proven to be convergent, we have developed an algorithm to solve problem B and thus problem A.

2. Application to Capital Budgeting

The Baumol-Quandt formulation introduced in Section 1 is also useful for investigating the structure of optimal policies. We will prove a theorem for the capital budgeting version of the general control problem with the quite general concave utility function postulated in Section 1 that is analogous to the all-or-none theorem of Manne's obtained for a linear utility function.

Theorem 2.1. Suppose an optimal solution exists for the capital budgeting problem stated in Section 1. Then there is an optimal solution to that problem employing at most T of the available investment projects where T is the number of decision periods. (The conditions imposed on the functions $U_t(\cdot)$ are not required for this theorem.)

Proof: Consider problem B of the previous section, and recall that any solution optimal for B is also optimal for A. Since we assume the existence of an optimal solution to A, one also exists for B by the equivalence established for the problems. Let

$$\hat{w} = \begin{bmatrix} \hat{w}_0 \\ w_1 \\ \vdots \\ \hat{w}_T \end{bmatrix} \in W$$

be the w-component of an optimal solution to problem B. Let

$$\hat{\bar{w}} \quad \begin{bmatrix} \hat{w}_1 \\ \hat{w}_2 \\ \vdots \\ \hat{w}_T \end{bmatrix}$$

and consider the system

$$1) \qquad Av = M - \hat{\bar{w}} \quad .$$

Since $\hat{\bar{w}}$ is derived from an optimal solution to problem B, there exists some non-negative vector $\hat{v}$ such that

$$A\hat{v} = M - \hat{\bar{w}}, \quad \hat{v} \geq 0 \ .$$

But this means that we can find some basic feasible solution v^* to system 1 by applying the phase I procedure of the simplex method.

Inasmuch as system 1 has T equations, any basic feasible solution will have at most T non-zero components. Thus, $< v^*, \hat{w} >$ is an optimal solution to problem A, using at most T of the available investment projects. Q.E.D.

Examining this result with respect to the one obtained by Manne for the case of a linear utility function, we observe that by introducing the non-linear utility function we can no longer say that both investment and withdrawal need not occur simultaneously. However, we have established the result that there is an optimal solution such that if more than one project is initiated in any one period, there must be a corresponding period or periods during which no project is undertaken. This result is similar to the one that would be obtained by the alternative method described in Chapter II of requiring a fixed proportion of the available capital to be disbursed. It is also interesting to observe that Manne's all-or-none result depended upon the point input stream output character of the investment opportunity profile, while Theorem 2.1 does not require any such assumption and thus holds for the more general Baumol-Quandt model.

CHAPTER IV

A CLASS OF PROBLEMS WITH LINEAR UTILITY FUNCTIONALS AND INFINITE HORIZONS

1. Extension of the Linearity Theorem

The substance of this chapter will be an investigation of the control problem studied in Section 1 of Chapter II, with the time horizon assumed to be infinite. The linearity theorem obtained for the finite horizon case will be extended to cover the infinite horizon case. The notation for the state variables, control variables, and transformations remains as described in Chapter I. The return or utility function will be linear as in Chapter II. It is interesting to note that the problem we pose is a programming problem with an infinite number of control variables since an infinite number of decisions must be made, and an infinite-dimensional state variable since we may assume that $x \in \ell^{\infty}$. The infinite character of the problem notwithstanding, we are still able to arrive at a computational algorithm. Of course, the assumptions imposed on the relationship of the constraint set to the state transformations are rather severe. In the next chapter we will investigate the results obtainable when these assumptions and the objective linearity assumptions are relaxed.

In discussing the infinite-horizon model we will assume stationarity of all transformations and look at the sum of discounted returns. Thus, the functional equation for the process becomes

$$f(x) = \max_{v \in V(x)} \{c(v) + \beta f \circ T(x,v)\} \tag{4.1}$$

where

$$V(x) = \{v : Av = Dx + B,\ v \geq 0\}.$$

We shall make the basic assumption that $V(x)$ is compact for each $x \in \mathcal{R}$ and that $(\mathcal{C}, \mathcal{B})$ is stable for A. Recall that $\mathcal{R} = \{x:(Dx + B) \in \mathcal{B}\}$. Define the sequence $< f_N(x) >_{N=1}^{\infty}$ for $x \in \mathcal{R}$ by:

$$f_N(x) = \max_{v \in V(x)} \{c(v) + \beta f_{N-1} \circ T(x,v)\} \ . \tag{4.2}$$

Under the assumptions of Theorem 2.1, which in this case reduce to

H1) 1. $c \in \mathcal{C}$

2. $[c + L_{n-1} \circ T''] \in \mathcal{C}$ $\qquad n = 1,2,3 \cdots$

3. $x \in \mathcal{R}$

4. $T(x,v) \in \mathcal{R}$ for all $x \in \mathcal{R}$ and $v \in V(x)$,

we obtain as before

$$f_N(x) = L_N(x) + K_N \qquad \forall \ x \in \mathcal{R}$$

where

$$L_N(x) = \sum_{n=1}^{N} \beta^{N-n} E_n \circ G_{n+1} \circ \ \cdots \circ G_N(x)$$

$$E_n = c \circ \hat{A}_n^{-1} \circ D$$

$$G_n = T'' \circ \hat{A}_n^{-1} \circ D + T'$$

and

$$K_N = \sum_{n=1}^{N} \beta^{N-n}[c + \beta L_{n-1} \circ T''] \circ \hat{A}_n^{-1} B \ .$$

Let us now assume that we are dealing with a "productive" system, i.e., that with $N+1$ periods to go we can always do as well as we could with N periods to go. Clearly sufficient conditions for this assumption to hold are either $c \geq 0$ or the "do nothing" solution (slack) is feasible.

Under this assumption $< f_N(x) >$ will be an increasing function of N for each $x \in \mathcal{R}$. In case $B = 0$ this assures us that $< L_N(x) >$

is increasing in N for each $x \in \mathcal{R}$ and even if $B \neq 0$ it is likely that this result still holds.

In any case we make the assumption:

H2) $< L_N(x) >$ is an increasing sequence for each $x \in \mathcal{R}$, and $< K_N >$ is an increasing sequence of real numbers. We now wish to show that $< L_N(x) >$ is a convergent sequence for each $x \in \mathcal{R}$.

Since the total number of bases that can be chosen from the matrix A is finite, the number of different operators G_n in the sequence $< G_n >_{n=1}^{\infty}$ is also finite. Thus, we may write:

$$\|G\| = \max_{1 \leq n < \infty} \|G_n\| \quad \text{and} \quad \|E\| = \max_{1 \leq n < \infty} \|E_n\|$$

where the norm $(\|\cdot\|)$ of a linear operator is defined in the usual way. Thus,

$$\begin{aligned} |L_N(x)| &\leq \sum_{n=1}^{N} \beta^{N-n} |E_n \circ G_{n+1} \circ \cdots \circ G_N(x)| \\ &\leq \sum_{n=1}^{N} \beta^{N-n} \|E_n\| \, \|G_{n+1}\| \cdots \|G_N\| \, \|x\| \\ &\leq \|E\| \, \|x\| \sum_{n=1}^{N} \beta^{N-n} \, \|G\|^{N-n} \\ &\leq \|E\| \, \|x\| \sum_{n=o}^{N-1} (\beta \|G\|)^n \qquad \text{for} \quad x \in \mathcal{R} . \end{aligned}$$

If we now make the assumption:

H3) $\beta \|G\| < 1$

we have

$$|L_N(x)| \leq \|E\| \, \|x\| / (1 - \beta \|G\|) \qquad \text{for} \quad x \in \mathcal{R} ,$$

and since the right-hand side is independent of N the sequence $< L_N(x) >$ is bounded above for each $x \in \mathcal{R}$, and thus due to H2 approaches a limit, say $L(x)$.

Lemma 4.1. If a sequence of linear functionals $< L_N(x) >$ approaches a limit at each point of a closed set $\mathcal{R}$, then the sequence converges at each point of the smallest closed linear subspace, S, containing $\mathcal{R}$.

Proof:

$$S = \{y;y = \sum_{i=1}^{P} a_i x_i \quad \text{with} \quad x_i \in \mathcal{R} \qquad \text{for} \quad i = 1,\ldots,P\}.$$

Thus, for $y \in S$,

$$L_n(y) = L_n(\sum_1^P a_i x_i) \quad \text{with} \quad x_i \in \mathcal{R} \qquad \text{for each} \quad i$$

$$= \sum_{i=1}^{P} a_i L_n(x_i) \to \sum_{i=1}^{P} a_i L(x_i)$$

and thus

$$\lim_{n \to \infty} L_n(y) = \sum_{i=1}^{P} a_i L(x_i) \quad \text{where} \quad y \in S \quad \text{and} \quad y = \sum_1^P a_i x_i$$

$$\text{with} \quad x_i \in \mathcal{R}, \ i = 1,\ldots,P.$$

Q.E.D.

Since S is a closed linear subspace of the Banach Space X, S is also a Banach space, and

$$L_n \in \mathcal{L}(S,E^1) \qquad \text{for} \quad n = 1,2,\ldots .$$

Furthermore, we have shown by the Lemma that $\lim_{n \to \infty} L_n(x)$ exists for all $x \in S$. Thus, defining

$$L(x) = \lim_{n \to \infty} L_n(x)$$

we have satisfied the conditions of the Banach-Steinhaus Theorem [8] and may conclude that $L \in \mathcal{L}(S,E^1)$, and also that for some C, $\|L_n\| < C$ for all n.

We now turn our attention to the sequence of constant terms $< K_N >$, again assuming that it is an increasing sequence (H2). To show that $< K_N >$ is bounded, we employ the same argument as before to write

$$\|\hat{A}^{-1}\| = \max_{1 \leq n < \infty} \|\hat{A}_n^{-1}\|$$

and

$$|K_N| \leq \sum_{n=1}^{N} \beta^{N-1}[\,|C \circ \hat{A}_n^{-1}B| + \beta|L_{n-1} \circ T'' \circ \hat{A}_n^{-1}B|$$

$$\leq [\,\|C\| + \beta K\|T''\|\,]\ \|\hat{A}^{-1}\|\ \|B\|\ (\frac{1}{1-\beta})$$

and thus the sequence $< K_N >$ is convergent and we may write

$$K = \lim_{n \to \infty} K_n .$$

It is now possible to state the main theorem.

Theorem 4.1: Existence and Characterization of Solutions.

Under assumptions H1, H2, and H3 the sequence of return functions $< f_N(x)$ converges monotonically to a function $f(x)$ for all $x \in S$, where

$$f(x) = L(x) + K$$

and

$$L \in \mathcal{L}(S,E^1) .$$

Furthermore, f satisfies the required functional equation, i.e.,

$$f(x) = \max_{v \in V(x)} \{c(v) + \beta f \circ T(x,v)\} \quad \text{for all} \quad x \in \mathcal{R} .$$

Proof: It remains necessary only to prove the last statement. The proof uses a monotonicity argument adapted from Bellman [1]. Let

$$P(f,v,x) = c(v) + \beta f \circ T(x,v)$$

and let M denote the maximum operator over $V(x)$. Then

$$f_N(x) = MP(f_{N-1},v,x)$$

from (4.2), and by monotonicity, for a fixed $x \in \mathcal{R}$, we have

$$\begin{aligned} f(x) &\geq MP(f_{N-1},v,x) \\ &\geq P(f_{N-1},v,x) \quad \text{for all} \quad v \in V(x) . \end{aligned}$$

Since this inequality holds for every N, letting $N \to \infty$ yields

$$f(x) \geq P(f,v,x) \quad \text{for all} \quad v \in V(x),$$

and thus

$$f(x) \geq \sup_{v \in V(x)} P(f,v,x) .$$

Similarly,

$$f_N(x) \leq \sup_{v \in V(x)} P(f,v,x) \quad \text{for every} \quad N$$

and thus

$$f(x) \leq \sup_{v \in V(x)} P(f,v,x) ,$$

yielding

$$f(x) = \sup_{v \in V(x)} P(f,v,x) .$$

Since f is continuous from our previous result, the compactness of $V(x)$ allows us to conclude that

$$f(x) = MP(f,v,x) \quad \text{for all} \quad x \in \mathcal{R} .$$

Q.E.D.

In order to make use of the previous theorem we should like to know that the solution to 4.1 is unique. Let us consider the control problem with $B = 0$, i.e., $\hat{V}(x) = \{v : Av = Dx,\ v \geq 0\}$. If the other transformations remain unchanged it can easily be seen that all of the previous results hold except that $K_N = 0$, $N = 1,2,\ldots,$. Thus, it is clear that the limit function $f(\cdot)$ is an element of $\mathcal{L}(S,E^1)$ and satisfies the functional equation

$$f(x) = \max_{v \in \hat{V}(x)} P(f,v) \quad \text{for fixed} \quad x \in \mathcal{R}\,.$$

Theorem 4.2: Uniqueness Theorem.

Let $F \in \mathcal{L}(S,E^1)$ such that

$$F(x) = \max_{v \in \hat{V}(x)} P(F,v) \quad \text{for all} \quad x \in \mathcal{R}\,.$$

Then $F(x) = f(x)$ for all $x \in S$.

Proof: Choose any $x \in \mathcal{R}$. Since $\hat{V}(x)$ is compact, there exist points y and w in $\hat{V}(x)$ such that $f(x) = P(f,y)$ and $F(x) = P(F,w)$. Then by the usual argument we obtain

$$\begin{aligned} |f(x) - F(x)| &\leq \max \{|P(f,y) - P(F,y)|,\ |P(f,w) - P(F,w)|\} \\ &\leq \beta \max \{|(f-F) \circ T(x,y)|,\ |(f-F) \circ T(x,w)|\}. \end{aligned}$$

Since both f and F are linear, we have

$$y = \hat{A}_y^{-1} D \quad \text{and} \quad w = \hat{A}_w^{-1} D\,.$$

Thus, letting

$$G_y = T' + T'' \circ \hat{A}_y^{-1} \circ D \quad \text{and} \quad G_w = T' + T'' \circ \hat{A}_w^{-1} \circ D$$

yields

$$|f(x) - F(x)| \leq \beta \max \{ |(f-F) \circ G_y(x)|, |(f-F) \circ G_w(x)| \} .$$

From H3 we know that $\beta\|G\| = k < 1$, where the norm is considered as defined on S. Let us now examine the subspace $\hat{S} = \{z : z = \lambda x, \lambda$ in the scalar field$\}$. Clearly, $\hat{S} \subset S$ and thus we may restrict our operators to $\hat{S}$ and consider norms on $\hat{S}$ denoted by $\|\cdot\|_*$. Then

$$\beta\|G_j\|_* \leq \beta\|G_j\| \leq \beta\|G\| = k < 1 \quad \text{for} \quad j = y,w ,$$

and

$$\beta|(f-F) \circ G_j(x)| \leq k\|f-F\|_*\|x\| \quad \text{for} \quad j = y,w .$$

Therefore,

$$|f(x) - F(x)| \leq k \|f-F\|_*\|x\| .$$

However, since $\hat{S}$ is a one-dimensional subspace, it is also true that

$$|f(x) - F(x)| = \|f-F\|_*\|x\| ,$$

yielding the conclusion

$$|f(x) - F(x)| = 0 \quad \text{since} \quad k < 1.$$

Inasmuch as x was an arbitrary point in $\mathcal{R}$ we have shown that

$$f(x) = F(x) \quad \text{for all} \quad x \in \mathcal{R} .$$

Since $y \in S \Rightarrow y = \sum_1^P a_i x_i$ for $x_i \in \mathcal{R}$, it follows that

$$f(x) = F(x) \quad \text{for all} \quad x \in S.$$

Q.E.D.

Corollary 4.1. The function $f(x) = L(x) + K$ obtained as the limit of the $N-$period return functions is the unique affine functional on S satisfying equation 4.1 whenever $x \in \mathcal{R}$.

Proof: The existence theorem shows that $f(\cdot)$ does satisfy equation 4.1, and the uniqueness theorem proves that $L(\cdot)$ is unique. It remains to show that K is unique.

Suppose the contrary; i.e., for some $x \in \mathcal{R}$,

$$L(x) + K_1 = MP(L,v) + \beta K_1$$

and

$$L(x) + K_2 = MP(L,v) + \beta K_2 .$$

Subtracting:

$$|K_1 - K_2| = \beta |K_1 - K_2| \quad \text{and thus} \quad K_1 = K_2 \ .$$

It is now possible to use the uniqueness and existence theorems to state an algorithm for the determination of an optimum infinite-horizon policy and the optimal return function.

Choose $x \in \mathcal{R}$ $(x \neq 0)$. Our first concern will be to find the unique return function $f(x)$ satisfying 4.1.

2. A Computational Algorithm

Algorithm for the Determination of f. From 4.1 we have

$$L(x) + K = MP(L,v) + \beta K \ .$$

Suppose we knew that $\hat{A}_*$ was a dual feasible basis for the programming problem $MP(L,v)$ (note that we do not as yet know L); then since $x \in \mathcal{R}$ implies the optimal basis is independent of x, an optimal policy could be expressed as

$$v_* = \hat{A}_*^{-1} \circ D(x) + \hat{A}_*^{-1} B \ .$$

This result yields the functional equation

$$L(x) + K = [c \circ \hat{A}_*^{-1} \circ \beta L \circ T' + \beta L \circ T'' \circ \hat{A}_*^{-1} \circ D](x)$$
$$+ c \circ \hat{A}_*^{-1}B + \beta L \circ T'' \circ \hat{A}_*^{-1}B ,$$

and thus the functional L must satisfy

$$L[I - \beta G_*](x) = E_*(x) \quad \text{for} \quad x \in S$$

where, corresponding to the previous notation,

$$G_* = T' + T'' \circ \hat{A}_*^{-1} \circ D \quad \text{and} \quad E_* = c \circ \hat{A}_*^{-1} \circ D.$$

Now assumption H3 shows that $\beta\|G_*\| < 1$ and therefore $1/\beta > \|G_*\|$, and hence is greater than the spectral radius of the operator G_*. Therefore, $I - \beta G_*$ is invertible, and

$$L = E_* \circ [I - \beta G_*]^{-1} .$$

Similarly,

$$K = \frac{1}{(1-\beta)} [C \circ \hat{A}_*^{-1}B + \beta L \circ T'' \circ \hat{A}_*^{-1}B] ,$$

and it is clear that a solution to the infinite horizon problem would be achieved if we knew an optimal basis $\hat{A}_*$. In order to calculate an optimal basis we observe that the $M \times P$ matrix A has only a finite number K of possible bases, $[\hat{A}_1, \hat{A}_2, \ldots, \hat{A}_K]$, where $K \leq \binom{P}{M}$, and proceed as follows:

1) Examine $\hat{A}_1$ for primal feasibility; i.e.,

$$\hat{A}_1^{-1} \circ D(x) + \hat{A}_1^{-1}B \geq 0 \quad \text{for some} \quad x \in \mathcal{R} .$$

If $\hat{A}_1$ is primal-feasible, proceed to step 2; otherwise repeat step 1 with $\hat{A}_2$. Since by assumption the optimal basis is independent of the right-hand side for all $x \in \mathcal{R}$, we eventually will reach step 2 with a primal-feasible basis $\hat{A}_j$.

2) Compute L_j corresponding to the basis $\hat{A}_j$ where

$$L_j = E_j \circ [I - \beta G_j]^{-1} .$$

Note that the inversion operation may be difficult due to the possibly infinite character of the transformations involved.

3) Compute the dual variables Π_j corresponding to $\hat{A}_j$ where

$$\Pi_j = \gamma_j \hat{A}_j^{-1} + \beta L_j \, T'' \, \hat{A}_j^{-1},$$

and γ_j includes those components of C corresponding to $\hat{A}_j$, and similarly for T'' as indicated previously.

4) Check the solution for dual feasibility; i.e., if a_i is any column of A not included in $\hat{A}$, then dual feasibility has been achieved if and only if

$$\Pi_j \cdot a_i \geq (C + \beta L_j \, T'')_i$$

where the component $(c + \beta L_j \, T'')_i$ corresponds to the activity represented by a_i. If this test fails, repeat step one with $\hat{A}_{j+1}$. Since an optimal solution to the problem exists by the existence theorem, this test must ultimately be satisfied yielding an optimal basis $\hat{A}_*$ in a finite number of iterations. However, the finiteness of the computational process does depend on the feasibility of the inversion operation described in step 2.

The uniqueness theorem guarantees that the return function calculated in this manner will be the pointwise limit on S of the sequence of N-period return functions. It should be noted that neither $\hat{A}_*$ nor the associated optimal policy v_* need be unique, but the return function f is unique on S. Q.E.D.

3. The Structure of the Solution

It is interesting to analyze the results obtained so far from the viewpoint of the following three questions posed by Bellman [5]:

1. When does it make sense to consider an infinite-horizon model?

2. When it does, are the optimal trajectories in state space and the optimal policies limits of the corresponding quantities for the finite-horizon problem as the horizon becomes large?

3. What is the effect of using steady-state optimal policies for a finite-horizon problem?

The first question has clearly been answered, the pertinent sufficient condition being H3, i.e., $\beta\|G\| < 1$. This condition gives the process the required contraction characteristic that allows an infinite-stage generalization.

Attention will now be turned to points 2 and 3 to gain some insight into the effect of using the easily calculated infinite-horizon policies to approximate optimal policies for the finite-horizon problem.

If we fix $x \in \mathcal{R}$ and define

$$F_n(v) = c(v) + \beta f_{n-1} \circ T(x,v)$$

and

$$F(v) = c(v) + \beta f \circ T(x,v) \quad \text{for all} \quad v \in V(x),$$

then, under the hypothesis of the existence theorem, both F_n and F can be expressed as continuous linear functionals of v plus a term independent of v; and furthermore,

$$F_n(v) \to F(v) \quad \text{for all} \quad v \in V(x) \; .$$

The desired results on policy convergence will be deduced from the following theorem.

A policy v^n will be called ϵ-optimal for F_n if $F_n(v^n) > F_n(v) - \epsilon$ for all $v \in V(x)$. A policy v^n will be called optimal for F_n if $F_n(v^n) \geq F_n(v)$ for all $v \in V(x)$.

Theorem 4.3: Policy Convergence

If $V(x)$ is compact for each x, $F_n(v) \to F(v)$ for $v \in V(x)$, and each F_n and F can be expressed as a continuous linear function of v plus a term not dependent on v, then:

1. Given $\epsilon > 0$, $\exists N \ni \forall n \geq N$, v^* optimal for F implies that v^* is ϵ-optimal for F_n.
2. Let v^* be a basic policy. Then $\exists N > 0 \ni$ if for some $n \geq N$, v^* is optimal for F_n, then v^* is optimal for F.

Proof: The proof depends on the observation that at each stage, and also in the infinite horizon, the problem

$$\max_{v \in V(x)} F_i(v)$$

is a linear programming problem. Thus, if v^* is optimal for F, there exists a basic policy $\hat{v}$ such that $F(v^*) = F(\hat{v})$. There are only a finite number of basic policies $\{v_1, v_2, \ldots, v_K\}$.

1) By hypothesis, $F(v^*) \geq F(v) \; \forall v \in V(x)$. Since there are a finite number of basic policies, given $\epsilon > 0$, $\exists N_0 \ni \forall n \geq N_0$,

$$|F_n(v_i) - F(v_i)| < \epsilon/2 \quad \text{where } i = 1, \ldots, K, K+1,$$

indexes the basic policies for $i = 1, \ldots, K$ and $v_{k+1} = v^*$.

Let v_k be a basic policy optimal for F_n. Then

$$F_n(v_k) - F_n(v^*) < F(v_k) + \epsilon - F(v^*) ;$$

but $F(v_k) - F(v^*) \leq 0$, and therefore

$$F_n(v_k) - F_n(v^*) < \epsilon .$$

Since

$$F_n(v_k) \geq F_n(v), \quad \forall v \in V(x),$$

$$F_n(v) - F_n(v^*) < \epsilon , \forall v \in V(x) ,$$

and v^* is ϵ-optimal for F_n.

2) Suppose v^* is not optimal for F, and that v^* is basic. Then $\exists$ a basic policy v_k such that

$$F(v_k) > F(v^*) , \text{ or for some } \delta > 0 ,$$

$$F(v_k) = F(v^*) + \delta .$$

Now choose N such that $\forall n \geq N$,

$$|F(v_i) - F_n(v_i)| < \delta/2 \text{ for every basic polic.}$$

Thus, for any $n \geq N$,

$$\begin{aligned} F_n(v_k) + \delta/2 &> F(v^*) + \delta \\ &> F_n(v^*) - \delta/2 + \delta , \end{aligned}$$

and therefore

$$F_n(v_k) > F_n(v^*) ,$$

implying that there is no $n \geq N$ for which v^* is optimal for F_n. The desired result follows by contraposition.

Q.E.D.

Corollary 4.2. If v^* is the unique basic policy which is optimal for F, then $\exists\ N \ni \forall n \geq N$, v^* is the unique basic policy optimal for F_n.

Proof: Follows directly from 2) and the fact that for any n there exists a basic optimal policy.

We can now relate these results to the original problem by the following statements:

1. If the policy v^* corresponds to the optimal basis obtained in step 4 of the algorithm, then given $\epsilon > 0$, for N large enough, v^* is ϵ-optimal for the policy decision with N periods to go. In other words,

$$c(v^*) + \beta f_{N-1} \circ T(x,v^*) + \epsilon > \max_{v \in V(x)} \{c(v) + \beta f_{N-1} \circ T(x,v)\} .$$

2. If there is only one basis, $\hat{A}$, that satisfies the tests in steps 1 and 4 of the computation algorithm, then there exists a horizon N such that for all greater horizons the policy $\hat{v} = \hat{A}^{-1} \circ Dx + \hat{A}^{-1}B$ is the unique basic optimal first-period decision; i.e.,

$$f_N(x) = c(\hat{v}) + \beta f_{N-1} \circ T(x,\hat{v}) ,$$

and if v is a different basic policy,

$$f_N(x) > c(v) + \beta f_{N-1} \circ T(x,v) \quad \text{for all} \quad x \in \mathcal{R} .$$

Since the form of an optimal infinite horizon policy has been determined for the infinite horizon case, it seems reasonable to investigate the asymptotic properties of the state vector when an optimal infinite horizon policy is followed for a number of periods. Suppose that the optimal basis $\hat{A}_*$ is unique, and that starting from the state x_0 the sequence $< x_n >_{n=0}^{N}$ represents the successive states visited by a trajectory employing at each state the optimal basis $\hat{A}_*$. Using notation consistent with that presented previously, we have

$$x_{n+1} = (T'' \circ \hat{A}_*^{-1} \circ D + T')x_n$$

or

$$x_N = (T'' \circ \hat{A}_*^{-1} \circ D + T')^N x_0 \ .$$

Thus, the asymptotic question can be answered by determining when the sequence of linear operators, $< G^N >_{N=1}^{\infty}$, converges.

CHAPTER V

ASYMPTOTIC PROPERTIES OF A MODEL OF ECONOMIC GROWTH

1. Optimality Conditions

In this chapter we investigate the control problem formulated in Chapter I, with the restriction that the state space X be a finite-dimensional Euclidean space. We first derive conditions which an optimal policy must satisfy, analogous to the Pontryagin conditions [17] for continuous time. These conditions are then applied to specific formulations of the general problem to obtain results about the structure and asymptotic properties of optimal policies and trajectories.

Assume the original state space X is a subset of E^m and, as is usual in this type of investigation, add a component x_o^t to the state vector x^t representing the total utility accumulated through time t. Then the problem to be solved may be stated as follows:*

Find a trajectory $(x_o, x) = \langle (x_o^t, x^t) \rangle_{t=1}^{T+1}$ and a control sequence $v = \langle v^t \rangle_{t=1}^{T}$ which maximize

$$x_o^{T+1}$$

subject to

$$1. \quad x^{t+1} \leq T_t' \, x^t + T_t'' \, v^t \quad , \qquad t = 1,\ldots,T,$$

$$x_o^{t+1} \leq x_o^t + U_t(v^t) \quad , \qquad t = 1,\ldots,T.$$

$$2. \quad A_t v^t \leq x^t \; , \; v^t \geq 0 \quad , \qquad t = 1,\ldots,T.$$

---

* This derivation of a discrete maximum principle was suggested in private communication by Professor A. Veinott.

3. $x_o^1 = 0$

x^1 is a given point in state space.

We shall assume that each utility function $U_t(\cdot)$ has the properties required to insure the necessary and sufficient characteristics of the Kuhn-Tucker (KT) theorem for this problem. Applying KT, the following conditions are both necessary and sufficient for $< (x_o^t, x^t) >_{t=1}^{T+1}$ to be an optimal trajectory, and $< v^t >_{t=1}^{T}$ to be an optimal control sequence:

a) There exist m+1-dimensional multipliers $< (\psi_o^t, \psi^t) >_{t=1}^{T}$ and m-dimensional multipliers $< p^t >_{t=1}^{T}$ with $p^t \geq 0$ and $\psi_o^t = 1$ for all $t = 1,\ldots,T$.

b) $\psi^t T_t' - \psi^{t-1} + p^t = 0, \qquad t = 2,\ldots,T,$

$\psi^T = 0.$

c) Primal conditions 1, 2, and 3 hold.

d) $\nabla U_t(v^t) + \psi^t T_t'' - p^t A_t \leq 0 , \qquad t = 1,\ldots,T,$

with equality in the i^{th} equation of group t if $v_i^t > 0$.

e) $p^t(A_t v^t - x^t) = 0, \;\; p^t \geq 0, \;\; t = 1,\ldots,T.$

These conditions will be applied to a particular formulation of an economic growth model which will now be described. The model to be considered will be stationary over time in the sense that all time subscripts on the transformations will be eliminated and the time subscript on the utility function will be replaced by a sum-of-discounted-utility criterion. Transitions will be assumed to be independent of the current state of the system, i.e., T' is the null operator, and utility will be derived solely by withdrawals of goods from the system. Denoting these withdrawals by a sequence of m-dimensional vectors $< w^t >_{t=1}^{T}$, we state the problem as follows:

Primal Conditions

1. $x^{t+1} \leq T''v^t$

   $x_o^{t+1} \leq \sum_{i=1}^{t} \beta^{i-1}U(w^t), \quad t = 1,\ldots,T$

2. $Av^t + w^t \leq x^t$, $v^t \geq 0$, $w^t \geq 0$, $\quad t = 1,\ldots,T$

3. $x_o^1 = 0$, $x^1 \geq 0$ is fixed.

Adjoint Conditions

a) $\beta^{t-1} \nabla U(w^t) \leq p^t, \quad t = 1,\ldots,T$

   with equality holding in the i^{th} equation if $w_i^t > 0$.

b) $p^{t+1} T'' \leq p^t A, \quad t = 1,\ldots,T-1$

   with equality holding in the i^{th} equation of group t if $v_i^t > 0$.

c) $p^t \geq 0, \quad t = 1,\ldots,T$.

Relationships a, b, and c are easily obtained from the KT conditions, previously stated for the more general problem, upon noting that $T_t' = 0$ implies $\psi^{t-1} = p^t$. We will make the usual assumption that A and T are non-negative, and that the utility function is concave, non-decreasing, and continuously differentiable as a function of the consumption vectors w^t. Under these assumptions it is clear that we may replace the inequalities in condition 1 and the first group of condition 2 by equalities since an increase in the corresponding consumption activity can be used without penalty to represent disposal of a given commodity. The model will be called a production-consumption model since we have explicitly separated a sequence of utility-deriving consumptions $< w_t >_{t=1}^{T}$ and a sequence of production decisions $< v_t >_{t=1}^{T}$ that provide goods for future periods.

2. Existence and Convergence of Optimal Policies for Processes of Long Duration

At this point we shall examine conditions that are sufficient to insure the existence of optimal policies if the time horizon is considered to be infinite. The main purpose of this investigation is to show the naturalness of certain conditions that we must place on the transformation of the finite horizon problem in order to derive the desired asymptotic properties. The method used in proving the existence of optimal policies is adopted from Karlin [13]. Employing the notation developed previously, we assume that all transformations are stationary over time and that we wish to maximize the discounted sum of utilities received at each decision stage of the process. The problem is stated as:

Problem 5.1.

$$\max \sum_{t=1}^{\infty} \beta^{t-1} U(v^t)$$

$$v^t \in V(x^t)$$

$$x^{t+1} \leq T(x^t, v^t), \qquad t = 1,2,3,\ldots$$

$$x^1 \text{ fixed and } x^1 \in X .$$

We make the basic assumption that $V(x)$ is compact for all $x \in X$. For the moment we shall assume that the utility function $U(\cdot)$ is concave and homogeneous of degree α, with $\alpha \leq 1$. In other words, $U(kv) = k^{\alpha}U(v)$. It is clear that if $U(\cdot)$ is a linear functional on V, then $\alpha = 1$. The spaces X and V are subsets of finite-dimensional Euclidean spaces of m and p dimensions, respectively. Consider a sequence of feasible decisions $v = < v^t >_{t=1}^{\infty}$, where each $v^t \in V(x^t)$ for a sequence of states $x = < x^t >_{t=1}^{\infty}$ satisfying the above constraints.

Let $R(x^t)$ be the set of states reachable from the state x^t in one step. Symbolically,

$$R(x^t) = \{x : x \leq T(x^t, v^t),\ v^t \in V(x^t)\}.$$

Since $V(x^t)$ is compact, and $T(\cdot,\cdot)$ is a bounded operator on $X \times V$, the set of states $R(x^t)$ will be compact, and by the Tychonoff theorem the product space $V(x^t) \times R(x^{t-1})$ will also be compact. Since x^1 is fixed, we may represent a policy by the sequence $s = < v^t, x^t >_{t=1}^{\infty}$, where $s \in S = < V(x^1) \times x^1 > \times < V(x^2) \times R(x^1) > \times < V(x^3) \times R(x^2) > \ldots$. Again employing the Tychonoff theorem we see that the product space S is compact. For any given feasible policy $s = < v^t, x^t >_{t=1}^{\infty}$ the total yield for a given initial state x^1 may be represented as

$$\phi(x^1, s) = \sum_{t=1}^{\infty} \beta^{t-1} U(v^t) ,$$

and our problem is to maximize $\phi(x^1,s)$ for all $s \in S$. Since S is compact, the maximum will exist if $\phi(x^1,s)$ is continuous in s for each $x^1 \in X$. This will be the case provided

1) $$\phi_K(x^1, s) = \sum_{t=1}^{K} \beta^{t-1} U(v^t)$$

converges uniformly in S for each $x^1 \in X$.

In order to establish conditions yielding uniform convergence in 1) we make use of work done to investigate the structure of von Neumann type growth models. To do this we shall assume that the transition operator $T(\cdot,\cdot)$ is non-negative. The development uses some of the results obtained by Winter [23]. Following Karlin [12] and Winter [23] we define the coefficient of expansion of a pair (x,y) as

$$\lambda(x,y) = \max_{y \in R(x)} [c \mid y \geq cx]$$

Using the above scheme it is now possible to state

Theorem 2.1.

(A) There exist M-dimensional semipositive vectors $\hat{x}$ and $\hat{p}$, and a positive scalar ρ_o such that

1) $\rho_o \hat{x} \in R(\hat{x})$

2) $\rho_o \geq \lambda(x,y)$ for all pairs (x,y) such that $y \in R(x)$.

3) $p_o \cdot y \leq \rho_o p_o \cdot x$ for all pairs (x,y) such that $y \in R(x)$.

(B) If $p_o \gg 0$, then there exists a constant k such that for any feasible sequence of states $< x^t >_{t=1}^{\infty}$,

$$\|x^t\| \leq \rho^t k \|x^1\| \quad \text{for all} \quad t = 1,2,\ldots$$

and for any $\rho > \rho_o$, where k does not depend on ρ.

(C) If $\beta\rho_o^{\alpha} < 1$, and $p_o \gg 0$, then $\phi_K(x^1,s)$ converges uniformly in S for each $x^1 \in X$, and there exists an optimal policy for problem 5.1. Furthermore, if we let

$$f(x) = \max_{s \in S} \phi(x,s) ,$$

then $f(\cdot)$ is a continuous concave function on X and satisfies the "Principle of Optimality,"

$$f(x) = \max_{v \in V(x)} \{U(v) + \beta f[T(x,v)]\}.$$

Proof: Parts (A) and (B) follow from Theorems 1 and 2 of Winter [23] upon verification that his conditions (A1)-(A4) are satisfied by the model presented as Problem 5.1. Defining the cone T in 2m-dimensional Euclidean space as

$$T = \{(x,y): x \in X,\ y \in R(x)\},$$

we see that T is a closed convex cone since X is a closed convex cone, and $v \in V(x) \Rightarrow \alpha v \in V(\alpha x)$ for all $x \in X$, and $\alpha \geq 0$. Thus, assumption (A1) of Winters[23] holds and assumptions (A2), (A3), and (A4) are immediate. To prove part (C) we define the functions

$$M(x) = \max_{v \in V(x)} U(v) \quad \text{and} \quad m(x) = \min_{v \in V(s)} U(v)$$

on the set $W = \{x: \|x\| \leq k\|x^1\|,\ x \in X\}$.

Since $V(x)$ is compact for each $x \in W$ and $U(\cdot)$ is a concave function on V, we have $|M(x)| < \infty$ and $|m(x)| < \infty$ for each $x \in W$, and thus $M(x)$ is concave on W and $m(x)$ is continuous on W. Let

$$\overline{M} = \max_{x \in W} M(x) \quad \text{and} \quad \underline{M} = \min_{x \in W} m(x)$$

which exists finitely due to the compactness of W and the continuity of $m(\cdot)$ and the concavity of $M(\cdot)$ on W. Using the fact from (B) that $\|x^t\| \leq \rho^t k\|x^1\|$ for any $\rho > \rho_o$, and the result that $V(\cdot)$ is scalar homogeneous on X for non-negative scalars, we have

$$v^t \in V(x^t) \Leftrightarrow \frac{v^t}{\rho^t} \in V\left(\frac{x^t}{\rho^t}\right)$$

and

$$\left\|\frac{x^t}{\rho^t}\right\| \leq k\|x^1\| \ .$$

Thus, since $U(\cdot)$ is scalar homogeneous of degree α ,

$$|U(v^t)| = (\rho^\alpha)^t \left|U\left(\frac{v^t}{\rho^t}\right)\right|$$

$$\leq (\rho^\alpha)^t M^*, \qquad \text{for} \quad t = 1,2,\ldots$$

where

$$M^* = \max\ [\,|\overline{M}|, |\underline{M}|\,]\ .$$

Using this result we obtain from equation 1,

$$\left| \sum_{t=1}^{K} \beta^{t-1}\, U(v^t) \right| \leq \frac{1}{\beta} \sum_{t=1}^{K} (\beta\rho^{\alpha})^t\ \rho^{\alpha} M^*$$

for x^1 fixed in X. The uniformity of convergence can be established by noting that $\beta\rho_o^{\alpha} < 1$ implies we can find a $\rho > \rho_o$ such that $\beta\rho^{\alpha} < 1$, causing the series on the right to converge, and thus yielding uniform convergence in 1. Hence, $\phi(x,s)$ is continuous for each x on the compact set S, the optimal return function

$$f(x) = \max_{s \in S} \phi(x,s)$$

exists finitely and the existence of an optimal policy is established. The fact that $f(\cdot)$ satisfies the Principle of Optimality was demonstrated in Karlin [13]. Q.E.D.

Although the proof of the theorem above depended on the scalar homogeneity of the utility function $U(\cdot)$, it is clear that the same proof would suffice for a concave utility function dominated by some scalar homogeneous function of the required type. In particular, a simpler version of the proof could be constructed for bounded utilities.

An additional consequence of the uniform convergence demonstrated above is that $f_n(x) \to f(x)$ for each $x \in X$ where $f_n(\cdot)$ is the optimal return function with n decision stages remaining. Using this result, it can be shown that $f(\cdot)$ is concave since each $f_n(\cdot)$ is concave. Defining $v_n(x)$ as an optimal decision with n stages remaining, and present resource position x and $v(x)$ as the corresponding quantity for the infinite horizon problem, we should like to prove that $v_n(x) \to v(x)$. Unfortunately, a uniqueness question arises here and the result cannot be obtained in general. However, if $U(\cdot)$ is <u>strictly</u> concave, the optimal decision at each stage is unique. Thus, for a strictly concave utility function we can indeed show that $v_n(x) \to v(x)$ for each $x \in X$, where $v(x)$ is some optimal decision for the infinite horizon problem. Moreover, it can be shown that each optimal decision function $v_n(\cdot)$ is continuous on the state space. Since many interesting utility functions will not be strictly concave in all of the decision variables, we will need a convergence-type result dealing with the case of $U(\cdot)$ concave. To accomplish this we need the following results describing characteristics of optimal policies:

<u>Lemma 2.1</u>. Fix an initial point $x \in X$. Let $< v_n(x) >_{n=1}^{\infty}$ represent any sequence of optimal first-step decisions with the number of periods remaining indexed by n. Then "eventually" every element $v_n(x)$ is "near" an optimal infinite horizon element $v(x)$. More specifically, if $\overline{v}(x)$ is the set of optimal first-step decisions for an infinite horizon program starting at point $x \in X$, then for every $\delta > 0$ there is an integer N such that for all $n \geq N$, $d(v_n(x), \overline{v}(x)) < \delta$, where the distance function $d(\cdot,\cdot)$ is the usual point to set distance.

<u>Proof:</u> We first show that if $< v_n >_{n=1}^{\infty}$ is any sequence from a compact set V, then ultimately each point v_n must be near a cluster point of the sequence. Let F be the set of cluster points of the sequence $< v_n >_{n=1}^{\infty}$. Then we have to show that, given $\delta > 0$,

$\exists N > 0 \ni \forall n \geq N,\ d(v_n,F) < \delta$. We know that F is not empty by the Bolzano-Weierstrass theorem and we proceed by assuming the contrary result; i.e., suppose $\exists \delta > 0 \ni \forall N > 0, \exists n \geq N \ni d(v_n,F) > \delta$. Cover each point f of F by an open sphere O_f of radius $\delta/2$, and let $O = \bigcup_{f \in F} O_f$. Then $V \sim O$ is closed and compact since V is compact, also $F \cap V \sim O = \emptyset$. But by the construction and the contrary hypothesis, there are an infinite number of elements of the sequence $< v_n >_{n=1}^{\infty}$ in the set $V \sim O$ and, hence, by the Bolzano-Weierstrass theorem the sequence must have a cluster point in $V \sim O$, providing a contradiction and establishing the result.

We show now that each cluster point of any sequence of optimal decisions $< v_n(x) >_{n=1}^{\infty}$ is an optimal solution to the infinite horizon problem; i.e., if $F(x)$ is the set of cluster points of all optimal sequences $< v_n(x) >_{n=1}^{\infty}$, then $F(x) \subset \bar{v}(x)$. Again suppose the contrary and let $\hat{v} \in F(x)$ be such that $\hat{v} \notin \bar{v}(x)$. Thus, we can find a point $v^* \in \bar{v}(x)$ such that for some $\epsilon > 0$,

$$U(\hat{v}) + \beta f[T(x,\hat{v})] + \epsilon \leq U(v^*) + \beta f[T(x,v^*)] \ .$$

Since $f_n(x) \to f(x)$ for each $x \in X$, we can find an integer N such that for all $n \geq N$,

$$U(\hat{v}) + \beta f_n[T(x,\hat{v})] + \frac{\epsilon}{2} \leq U(v^*) + \beta f_n[T(x,v^*)] \ .$$

As $\hat{v}$ is a cluster point of some optimal sequence $< v_n(x) >_{n=1}^{\infty}$, we can find a subsequence of optimal policies $< v_{n_i}(x) >_{i=1}^{\infty}$ converging to $\hat{v}$, and since $\hat{v} \neq v^*$, $\exists K > 0$ such that $\forall i \geq K$, $v_{n_i}(x) \neq v^*$. Thus, using the continuity of $U(\cdot)$ and $f_n(\cdot)$ we can find an n_i such that

$$U(v_{n_i}(x) + \beta f_{n_i}[T(x,v_{n_i}(x))] + \frac{\epsilon}{4} \leq U(v^*) + \beta f_{n_i}[T(x,v^*)] ,$$

which means that $v_{n_i}(x)$ is not an optimal decision with n_i periods remaining. This establishes the contradiction and yields the desired result.

Combining the two results, we see that for a given optimal sequence $< v_n(x) >_{n=1}^{\infty}$, given $\delta > 0$, $\exists N > 0 \ni \forall n \geq N$, $d(v_n(x),F(x)) < \delta$, and since $F(x) \subset \overline{v}(x)$, $d(v_n(x), \overline{v}(x)) < \delta$. Q.E.D.

It should be noted that the integer N may be a function of the particular sequence $< v_n(x) >_{n=1}^{\infty}$ that was chosen. We shall now state a corollary to Lemma 2.1 which will extend the results of the lemma to cover uniformly all sequences of optimal first-step decisions from a given point. In other words, the integer N will not be a function of the choice of a particular sequence but will apply to all possible sequences.

Corollary 2.1. Let $x \in X$ be the initial point of a program. Then for every $\delta > 0$ there is an integer N such that for all $n \geq N$, $d(v_n(x), \overline{v}(x)) < \delta$ for any optimal first-step decision $v_n(x)$.

Proof: Suppose the contrary. Then $\exists \delta_1 > 0 \ni \forall$ integers N $\exists n_1 \geq N$ and an optimal first-step decision with n_1 periods remaining $v_{n_1}(x)$, such that $d(v_{n_1}(x), \overline{v}(x)) \geq \delta_1$. But this means that we can find a sequence $< v_{n_j}(x) >_{j=1}^{\infty}$ of optimal first-step decisions such that

$$d(v_{n_j}(x), \overline{v}(x)) \geq \delta_1 \text{ for all } j = 1,2,3,\ldots .$$

This sequence is clearly a subsequence of some sequence $< v_n(x) >_{n=1}^{\infty}$ of optimal first-step decisions, and it does not get near $\overline{v}(x)$, thus contradicting Lemma 2.1. Q.E.D.

Corollary 2.2. If the first k components of every element $v \in \bar{v}(x)$ are positive for a given x, then there is an integer N such that for $n \geq N$ the first k components of each optimal decision vector $v_n(x)$ in every sequence of optimal first-period decisions with n periods to go are positive.

Proof: The set $\bar{v}(x)$ is closed since if it were not, there would be a sequence $< v^i >_{i=1}^{\infty}$ such that each $v^i \in \bar{v}(x)$ and $v^i \to v^*$, where $v^* \notin \bar{v}(x)$. But $v^i \in \bar{v}(x)$ for each i implies that

$$f(x) = U(v^i) + \beta f[T(x,v^i)] \quad \text{for every } i$$

and thus, by the continuity of $U(\cdot)$ and $f(\cdot)$,

$$f(x) = U(v^*) + \beta f[T(x,v^*)] ,$$

which implies that $v^* \in \bar{v}(x)$.

Thus, we can find an open set O such that $\bar{v}(x) \subset O$, and each element $v \in O$ has its first k components positive. From Corollary 2.1 we know that eventually every element of $< v_n(x) >_{n=1}^{\infty}$ is in O for every sequence of optimal finite horizon decision vectors, and thus the desired result holds. Q.E.D.

Let $\hat{x}_t(x_1)$ be the set of elements in state space that may be attained following an optimal infinite horizon policy for t periods starting in state x_1. Let P be the set of components that are strictly positive for all elements

$$v \in \bigcup_{x \in \hat{x}_t(x_1)} \bar{v}(x) .$$

Note that P may be empty. The next theorem to be proven will be useful when taken in conjunction with the results in the next section. It should be emphasized that Corollary 2.2 refers to optimal first-period decisions from a single initial point x, while Theorem 2.2 refers to decisions taken after a certain number of time periods have elapsed. The non-uniqueness of optimal policies implies that the state obtained at any time following an optimal policy may not be unique, which is precisely the reason for considering all states in the set $\hat{x}_t(x_1)$.

<u>Theorem 2.2</u>. Suppose the decision process has proceeded T_1 steps, where T_1 is arbitrary. Then there is an integer T_2 such that if more than T_2 stages remain in the process, every optimal decision vector has positive components corresponding to those components which are in the set P.

<u>Proof:</u> The set of states reachable in T_1 steps from a given initial state is bounded, and thus the closure of $\hat{x}_{T_1}(x_1)$ denoted by $\overline{\hat{x}_{T_1}(x_1)}$ is compact. From Corollary 2 we know that for each $x \in \overline{\hat{x}_{T_1}(x_1)}$ we can find an integer $T_2(x)$ such that the result holds. The theorem will be established if we can find an integer T_2 such that

$$T_2 \geq \sup_{x \in \hat{x}_{T_1}(x_1)} T_2(x) .$$

Using Corollary 2.1 we can find an integer N such that for all horizons not less than N every optimal policy is near an optimal infinite horizon policy for the first T_1 steps. N depends on T_1, x_1, and the nearness required. Since the state transition function is continuous in the decision variable, this implies that, given $\delta > 0$, we can find an integer N depending only on T_1 and x_1 such that, following an optimal policy, the state vector at time T_1 must be within a distance, δ, of the set $\hat{x}_{T_1}(x_1)$ if the time horizon is not less than N.

Define $\bar{v}_n(x)$ as the set of optimal first-step decisions for a program with n stages remaining, starting at point $x \in X$. Let x be an arbitrary point in the set $\hat{x}_{T_1}(x_1)$. Then, for any number of remaining stages n, given $\varepsilon > 0$, $\exists\, \delta_1 > 0$, such that $\|x-y\| < \delta_1$ implies the existence of a decision vector $v \in \bar{v}_n(x)$ such that $d(v, \bar{v}_n(y)) < \varepsilon$.

If this statement were not true, then we could find an n and an $\varepsilon > 0$ such that $\forall \delta > 0$, $\exists\, y$ with $\|x-y\| < \delta$ such that for all $v_n(y) \in \bar{v}_n(y)$,

$$d(v_n(y), \bar{v}_n(x)) \geq \varepsilon .$$

Thus, since $\bar{v}_n(x)$ is the set of optimal decisions starting at x with n stages remaining $\exists\, \gamma > 0$ such that

$$U(v_n(y)) + \beta f_n(T(x, v_n(y)) \leq U(v) + \beta f_n(T(x,v)) - \gamma$$

$$\text{for all } v \in \bar{v}_n(x) ,$$

where γ does not depend on the particular element $v_n(y)$ that was chosen as long as $d(v_n(y), \bar{v}(x)) \geq \varepsilon$. As $f_n \circ T(\cdot,\cdot)$ is continuous on the compact set $\hat{x}_{T_1}(x_1)\, x[V(x) \cup V(y)]$, it is uniformly continuous there. Thus, we can find a $\delta_2 > 0$ such that for $\|x-y\| \leq \delta_2$,

$$\|f_n(T(x, v_n(y)) - f_n(T(y, v_n(y))\| < \frac{\gamma}{4}$$

and

$$\|f_n T(x,v)) - f_n(T(y,v))\| < \frac{\gamma}{4} \text{ for all } v \in \bar{v}_n(x).$$

But this means that

$$U(v_n(y)) + \beta f_n(T(y, v_n(y)) \leq U(v) + \beta f_n(T(y,v)) - \frac{\gamma}{2} \text{ for all } v \in \bar{v}_n(x).$$

Since the feasible set $V(x)$ is a continuous set function of the state vector x, given $\gamma_1 > 0$, we can find a $\delta_3 > 0$ such that for all y with $\|x-y\| < \delta_3$, $v \in V(x) \Rightarrow d(v,V(y)) < \gamma_1$. Thus, choosing δ_3 small enough and using the continuity of $U(\cdot)$ and $f_n \circ T(\cdot,\cdot)$, we can find a $\hat{v} \in V(y)$ which yields

$$U(v_n(y)) + \beta f_n(T(y,v_n(z)) \leq U(\hat{v}) + \beta f_n(T(y,\hat{v})) - \frac{\gamma}{4} .$$

Consequently, $v_n(y) \notin \bar{v}_n(y)$, contradicting the hypothesis. Thus, for a given set of positive numbers $\{\mathcal{E}(x)\}_{x \in \hat{x}_{T_1}(x_1)}$, we can cover the compact set $\overline{\hat{x}_{T_1}(x_1)}$ by a collection of open spheres centered at each point $x \in \hat{x}_{T_1}(x_1)$ and with radius $\delta_1(x)$ such that:

1) There is an integer N such that for all time horizons not less than N the state vector at time T_1, denoted by x_{T_1}, must satisfy $x_{T_1} \in \bigcup_{x \in \hat{x}_{T_1}(x_1)} N_{\delta_1(x)}(x)$ if an optimal policy was followed for the first T_1 stages.

2) If $y \in N_{\delta_1(x)}(x)$ for some $x \in \hat{x}_{T_1}(x_1)$, then for any number n of stages remaining and for each $v_n(y) \in \bar{v}_n(y)$ we have $d(v_n(y), \bar{v}_n(x)) < \mathcal{E}(x)$.

Applying the Heine-Borel theorem, we can choose a finite number, k, of spheres from the collection that cover $\hat{x}_{T_1}(x_1)$, have property 2) and property 1) with the union taken over these spheres alone. Suppose the covering spheres have centers $z_1, z_2, \ldots, z_k$; then by Corollary 2 there are integers $N_1, N_2, \ldots, N_k$ such that for $n \geq N_j$, $v_n(z_j)$ has positive components corresponding to those components in P with $j = 1,2,\ldots,k$.

Since the elements $\mathcal{E}(x)$ of the set $\{\mathcal{E}(x)\}_{x \in \hat{x}_{T_1}(x_1)}$ can be chosen small enough so that each $v_n(y) \in \bar{v}_n(y)$ has positive components corresponding to P if $\bar{v}_n(x)$ has them, the result holds by choosing $T_2 = \max\{N_1, N_2, \ldots, N_k\}$. Q.E.D.

3. Asymptotic Analysis of the Production-Consumption Model

In this section we shall apply the conditions derived for a fairly general formulation of the control problem to a specific problem dealing with optimal paths of economic growth. Problems of this type have been considered in the literature, beginning with von Neumann's model of an expanding economy and progressing to the turnpike theorems of Dorfman, Samuelson, and Solow [10], Radner [18], McKenzie [15], and Morishima [16]. For a more detailed historical description the reader is referred to Chapter VI of the book by Morishima [16] and to the paper by McKenzie [15].

The present work fits into the historical scheme of things in the following way: After von Neumann [26] formulated his model and established the existence of a maximal balanced growth path, the other authors mentioned above showed that for certain formulations of the general growth model efficient paths of accumulation were near (in an appropriate sense) the maximal balanced growth path (von Neumann ray) for most of the programming period. The criterion for the "goodness" or desirability of a feasible-state trajectory in these studies is the concept of efficient production. A one-period production plan $\{x^t, x^{t+1}\}$ is said to be efficient if $x^{t+1} \in R(x^t)$, $x^t > 0$, and there is no $y \in R(x^t)$ such that $y \neq x^{t+1}$ and $x^{t+1} \leq y$. (For a detailed description of general balanced growth models and characterizations of efficient production sets, see Chapter 9 of the book by Karlin [12]). In the above models, consumption of goods was not explicitly considered and consumers' preference functions (utility) functions were not introduced. In this work we formulate a model technologically similar to the one studied by Morishima [16] except that we explicitly introduce consumption activities and our criterion for the goodness of a program is the maximization of the discounted sum of consumer-derived utility, rather than the concept of efficient production. Recently Gale [25] considered a similar type of model pioneered by Ramsey [19], which assumes a labor force growing at an exogenously given rate. Instead of introducing a discount factor they make use of an overtaking sequence criterion for optimality of a program. The concepts pertinent to this investigation will be stated more explicitly in the subsequent material.

Following Morishima [16, Ch. 6], we concern ourselves with an economy consisting of m industries whose outputs are used as factors of production during the next period. We assume that each industry can be related to a single product (no joint production). We also assume that there is a demand by consumers for the available goods at each stage and that we wish to allocate goods between present consumption and activities corresponding to industries which produce goods for the next period. Let us define the possible activity set A_i available to industry i for production of commodity i as a set of non-negative m-dimensional column vectors, $a^i = \{a^i_{1j}, a^i_{2j}, \ldots, a^i_{mj}\}$, where a^i_{kj} represents the amount of commodity k required to produce one unit of commodity i if industry i chooses its j^{th} activity. Let $\mathcal{a}$

represent the set of all possible matrices A of the form $A = [a^1, a^2, \ldots, a^m]$, with a^i a convex combination of the columns composing A_i for $i = 1,2,\ldots,m$. Several restrictive assumptions will be required of the elements of $\mathcal{A}$. These will be given as needed to proceed with the development, the first being

A1: The activity sets A_i are compact.

Define the real-valued function $\lambda(\cdot)$ on $\mathcal{A}$, such that $A \in \mathcal{A}$ $\Rightarrow \lambda(A)$ is the maximum positive eigenvalue of A. Since each A is non-negative we know that there will be at least one non-negative real eigenvalue of maximum modulus.

<u>Lemma 3.1</u>. Under Assumption A1, the set $\mathcal{A}$ is compact and convex. The function $\lambda(\cdot)$ is continuous on $\mathcal{A}$ and assumes its minimum on $\mathcal{A}$.

<u>Proof</u>: Since each column of any $A \in \mathcal{A}$ is a convex combination of a given set of columns, it is clear that $\mathcal{A}$ is a convex set. Let $\mathcal{H}(A)$ represent the convex hull of a set A. $\mathcal{H}(A)$ is compact if A is compact, and $\mathcal{A} = \mathcal{H}(A_1) \times \mathcal{H}(A_2) \times \ldots \times \mathcal{H}(A_m)$. Thus, $\mathcal{A}$ can be represented as a product of compact sets and is consequently compact by the Tychonoff theorem (Royden [20]).

Since each A is non-negative, the maximum positive eigenvalue $\lambda(A)$ is not less than the modulus of any other eigenvalue, and in fact is equal to the spectral radius of A. Thus, for $A \in \mathcal{A}$ we have

$$\lambda(A) = \lim_{n \to \infty} \|A^n\|^{1/n}$$

where

$$\|A\| = \sup_{\|x\|=1} \|Ax\| ,$$

the norm on the right-hand side being the usual Euclidean norm. Since the norm is a continuous function on $\mathcal{a}$, $\lambda(\cdot)$ is also continuous on $\mathcal{a}$ and therefore attains its minimum since $\mathcal{a}$ is compact. Q.E.D.

Since the element of $\mathcal{a}$ at which $\lambda(\cdot)$ assumes its minimum plays a central role in the following development, it is useful to investigate the properties of $\lambda(\cdot)$ in more detail. For instance, one might hope that in addition to $\lambda(\cdot)$ being continuous it is also concave, thereby assuring that the minimum is attained at an extreme point* of $\mathcal{a}$. Unfortunately, simple examples with 2×2 diagonal matrices illustrate that $\lambda(\cdot)$ is not in general concave, nor does it generally assume its minimum on $\mathcal{P}(\mathcal{a})$. Although $\lambda(\cdot)$ turns out to be convex for 2×2 diagonal matrices; matrices of the form $\begin{pmatrix} a_{11} & a_{12} \\ 0 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & 0 \\ b_{21} & b_{22} \end{pmatrix}$, with a_{12} and b_{21} large with respect to a_{11} and b_{22} serve to show that convexity does not hold in general either. The following two lemmas describe some results of a more positive nature.

Lemma 3.2. If $\lambda(\cdot)$ has a unique minimum on $\mathcal{a}$, say A_o, and if A_o is indecomposable, then there is a vector $p_o \gg 0$ such that $\lambda(A_o)p_o = p_oA_o$ and $\lambda(A_o)p_o \ll p_oA$ for all $A \in \mathcal{a}$ such that no column of A is a column of A_o. Furthermore, $A_o \in \mathcal{P}(\mathcal{a})$.

Proof: The first part of the proof is similar to one given by Morishima [16, p. 160]. Choose for p_o the left eigenvector of A_o corresponding to the root $\lambda(A_o)$. Since A_o is indecomposable, $p_o \gg 0$ (Debreu and Herstein [24]). Suppose that the first j

---

* The set of points $\mathcal{P}(Z)$ for any set z will be called extreme points of Z. They are defined in the usual way; i.e., $\hat{z} \in \mathcal{P}(Z) \Longleftrightarrow \hat{Z} = \gamma z_1 + (1-\gamma)z_2$ for $0 < \gamma < 1$ implies that either $z_1 \notin Z$ or $z_2 \notin Z$.

components of $\lambda(A_o)p_o$ are greater than or equal to the corresponding components of p_oA' for some $A' \epsilon \mathcal{A}$, none of whose columns are columns of A_o. Then, replacing the last m-j columns of A' by the corresponding columns of A_o, we have a non-negative matrix $A'' \epsilon \mathcal{A}$ such that

$$\lambda(A_o)p_o \geq p_oA'' \ .$$

But this implies (Karlin [12]) that $\lambda(A_o)$ is an upper bound to the spectral radius of A'', and thus $\lambda(A_o) \geq \lambda(A'')$, contradicting the uniqueness of A_o.

Now suppose that $A_o \notin \mathcal{P}(\mathcal{A})$. Then there is a set of matrices $\{A(i)\}_{i=1}^{r}$ such that $A(i) \epsilon \mathcal{P}(\mathcal{A})$, and non-negative weights $\{\gamma_i\}_{i=1}^{r}$ such that $\sum_1^r \gamma_i = 1$, yielding

$$A_o = \sum_1^r \gamma_i \ A(i) \quad ,$$

where at least one of the $A(i)$ corresponding to a $\gamma_i > 0$ is not identical to A_o. Then

$$\lambda(A_o)p_o = p_oA_o = \sum_1^r \gamma_i p_o \ A(i) > \lambda(A_o)p_o \ ,$$

a contradiction. Q.E.D.

Lemma 3.3. Let $\lambda(\cdot)$ have a unique minimum on $\mathcal{P}(\mathcal{A})$ at A_o; i.e., $A_o \epsilon \mathcal{P}(\mathcal{A})$ and $\lambda(A_o) < \lambda(A)$ for all $A \epsilon \mathcal{P}(\mathcal{A})$ and $A \neq A_o$. Assume that A_o is indecomposable. Then there is a strictly positive vector, p_o, such that $\lambda(A_o)p_o = p_oA_o$ and $\lambda(A_o)p_o \ll p_oA$ for all $A \epsilon \mathcal{P}(\mathcal{A})$, such that no column of A is a column of A_o. Furthermore, $\lambda(\cdot)$ attains its minimum uniquely over the entire set $\mathcal{A}$ and this minimum is attained at A_o; i.e., $\lambda(A_o) < \lambda(A)$ for all $A \epsilon \mathcal{A}$ such that $A \neq A_o$.

Proof: Observe that $A \in \mathcal{P}(\mathcal{A})$ if and only if the i^{th} column of A is an element of $\mathcal{P}(A_i)$ for each i. For suppose that the i^{th} column of A is not in $\mathcal{P}(A_i)$ for some i. Then we can find positive weights to attach to more than one element of A_i to form the i^{th} column of A. These same weights attached to the matrices formed from A by replacing the i^{th} column by the appropriate elements of A_i, and retaining all the other columns, give a representation of A as a proper convex combination of points of $\mathcal{A}$ and thus $A \notin \mathcal{P}(\mathcal{A})$. If each column of A is an element of $\mathcal{P}(A_i)$ for the appropriate i, it is clear that $A \in \mathcal{P}(\mathcal{A})$.

As demonstrated in Corollary 3.2 we can find $p_o \gg 0$ such that $\lambda(A_o)p_o = p_oA_o$. To show that $\lambda(A_o)p_o \ll p_oA$ for all $A \in \mathcal{P}(\mathcal{A})$ such that $A \neq A_o$, we use the proof given in Corollary 3.2, noting that if $A' \in \mathcal{P}(\mathcal{A})$, then $A'' \in \mathcal{P}(\mathcal{A})$ by the observation just stated, and thus the same contradiction serves to prove this result.

Now choose any $A^* \in \mathcal{A}$ such that $A^* \neq A_o$. Choose matrices $A(i) \in \mathcal{P}(\mathcal{A})$ and non-negative weights γ_i such that

$$A^* = \Sigma\, \gamma_i\, A(i)\ .$$

This can be done since the compact set $\mathcal{A}$ is the convex hull of its extreme points (Royden [20]). Then

$$p_oA^* = \Sigma\, \gamma_i p_o\, A(i) > \lambda(A_o)p_o$$

since $A_o \neq A^*$ implies that at least one $A(i)$ corresponding to a $\gamma_i > 0$ has a column different from the corresponding columns of A_o. Thus, $\lambda(A^*) \geq \lambda(A_o)$.

Let x^* be the right eigenvector of A^* corresponding to the root $\lambda(A^*)$. Then $x^* > 0$ since $A^* > 0$ (Debreu [24]), and furthermore there is some j for which strict inequality holds in the j^{th} equation above, and for which $x_j^* > 0$. Note that strict inequality holds for all equations corresponding to columns of A^* that are not identical to corresponding columns of A_o. Thus, if no column of A^* corresponds identically to a column of A_o, the assertion is true since $x^* > 0$. Now suppose for some k that the last $m-k$ columns of A^* correspond to the last $m-k$ columns of A_o, and that the first k do not where $1 \leq k < m$. We will show that $x_i^* > 0$ for some i between one and k.

Suppose the contrary; i.e., $x_i^* = 0$ for $1 \leq i \leq k$. Then

$$\lambda(A^*)x^* = A^*x^* = A_o x^* ,$$

which means that $\lambda(A^*)$ is a root of A_o corresponding to the eigenvector x^*. However, $\lambda(A^*) \neq \lambda(A_o)$, since any eigenvector corresponding to $\lambda(A_o)$ must be strictly positive (Debreu-Herstein [20]). But $\lambda(A_o)$ is the spectral radius of A_o, and thus $\lambda(A^*) < \lambda(A_o)$, a contradiction to the result already proven that $\lambda(A^*) \geq \lambda(A_o)$.

With the result that $x_j^* > 0$ for some j corresponding to an equation holding with strict inequality, it is easy to see that

$$\lambda(A^*)(p_o,x^*) = p_o A^*x^* > \lambda(A_o)(p_o,x^*) ,$$

and thus $\lambda(A^*) > \lambda(A_o)$. This establishes the result that $\lambda(\cdot)$ attains its minimum uniquely on $\mathcal{a}$ at the point A_o. Q.E.D.

There is a simple consequence of Lemma 3.3 for the case of convex, compact and polyhedral industry activity sets A_i. In this case, each of the sets $\mathcal{P}(A_i)$ contain only a finite number of points, and thus $\mathcal{P}(\mathcal{a})$

also contains only finitely many points. Therefore, one need only verify that $\lambda(\cdot)$ has a unique minimum over a finite set of points, and that this minimum occurs at an indecomposable point, to conclude that $\lambda(\cdot)$ is uniquely minimized at that point over all of $\mathcal{A}$. We now make the assumption:

A2: A_o is the unique element of $\mathcal{P}(\mathcal{A})$ minimizing $\lambda(\cdot)$ on $\mathcal{P}(\mathcal{A})$, and A_o is indecomposable. We also assume that A_o is primitive (Debreu and Herstein [20]), although we do not need this until Lemma 3.8. Let

$$\lambda_o = \lambda(A_o) = \min_{A \in \mathcal{P}(\mathcal{A})} \lambda(A) .$$

The relationship of λ_o, p_o, and A_o to the von Neumann balanced growth path and prices will be demonstrated in the next lemma. To do this, we first describe the following notation. Let the m-vector x represent the state of the system (x_i is the amount of commodity i available) at the start of a given time period, and y the state of the system at the start of the next time period. Let the m-vector w represent withdrawals (consumption) of commodities, and the m-vector v represent production levels of the industries. Recall that the production technology may be any element $A \in \mathcal{A}$. Then we have

$$* \qquad \begin{aligned} Av + w &\leq x \\ v &\geq y \end{aligned}$$

with $x, y, v,$ and w non-negative, and $A \in \mathcal{A}$. Let $\mathcal{J}$ be the set of all $\{x,y\}$ which satisfy $*$ for some $v \geq 0$, $w \geq 0$, and $A \in \mathcal{A}$. Let $\lambda(x,y) = \max\{\lambda \mid y \geq \lambda x\}$.

<u>Lemma 3.4</u>. Under assumptions A1 and A2 for the model described

above there is a balanced growth rate, ρ_o, a balanced growth path $\{x_o, y_o\}$, and a von Neumann price vector, p_o, such that

1) $\rho_o = 1/\lambda_o$

2) $y_o = \rho_o x_o$, $\rho_o = \lambda(x_o, y_o)$

3) $x_o \gg 0$, $p_o \gg 0$

4) $A_o y_o = x_o$ and no other $A \in \mathcal{A}$ satisfies this equation

5) $\rho_o \geq \lambda(x,y)$ for all $\{x,y\} \in \mathcal{T}$ with $x > 0$

6) $(p_o, y) \leq \rho_o (p_o, x)$ for all $\{x,y\} \in \mathcal{T}$.

Before proving the lemma, several comments are in order. The existence of a balanced growth rate ρ_o, a non-negative balanced growth path $\{x_o, y_o\}$, and a non-negative price vector p_o satisfying 2), 5), and 6), is a consequence of Theorems 9.10.1 and 9.10.2 in Karlin [12], as it is easy to see that his assumptions $T_1 - T_4$ are valid for the model being considered. We shall now show that under assumptions A1 and A2 Lemma 3.3 gives a simple proof of Karlin's result, and indeed allows the strengthening of the result given by 1), 3), and 4). It should be observed, however, that this model is more restrictive than Karlin's.

<u>Proof</u>: Make the identifications

$\rho_o \triangleq 1/\lambda_o$

$x_o \triangleq$ the right eigenvector of A_o corresponding to λ_o

$p_o \triangleq$ the left eigenvector of A_o corresponding to λ_o

$y_o \triangleq \rho_o x_o$.

Since A_o is indecomposable, it is clear that 1), 3), and the first part of 2) hold. Since

$$\lambda(x_o,y_o) = \max \{\lambda | y_o \geq \lambda x_o\},$$

it is clear that $\lambda(x_o,y_o) \geq \rho_o$. Suppose $\lambda(x_o,y_o) = \hat{\rho} > \rho_o$. Then $y_o \geq \hat{\rho}x_o$, and thus $\rho_o x_o \geq \hat{\rho}x_o$ since $y_o = \rho_o x_o$. This contradiction implies that 2) holds. Clearly,

$$A_o y_o = \rho_o A_o x_o = \rho_o \lambda_o x_o = x_o \ .$$

Suppose $Ay_o = x_o$ for some $A \in \mathcal{A}$, $A \neq A_o$. Then $x_o = Ay_o = \rho_o Ax_o$ or $Ax_o = \lambda_o x_o$. But this means that $\lambda_o = \lambda(A)$ since $x_o \gg 0$, which contradicts the conclusion of Lemma 3.3 and proves 4).

Now choose any $\{x,y\} \in \mathcal{T}$ with $x > 0$. From * we have $Ay \leq x$ for some $A \in \mathcal{A}$. Suppose $\lambda(x,y) > \hat{\rho} > \rho_o$. Then $\hat{\rho}\lambda_o > 1$ and $y \geq \hat{\rho}x$. Thus, $Ay \geq \hat{\rho}Ax$, and $\hat{\rho}Ax \leq x$, $\hat{\rho}p_o Ax \leq (p_o,x)$ and, consequently, $\hat{\rho}\lambda_o(p_o,x) \leq (p_o,x)$ since $p_o A \geq \lambda_o p_o$ by Lemma 3.3. But $p_o \gg 0$ and $x > 0$ imply $(p_o,x) > 0$ and thus $\hat{\rho}\lambda_o \leq 1$, a contradiction. Thus, 5) holds, and ρ_o is the maximum rate of balanced growth.

Observing that $\{x,y\} \in \mathcal{T}$ implies $Ay \leq x$ and thus $p_o Ay \leq (p_o,x)$ Lemma 3.3 applied as above yields $\lambda_o(p_o,y) \leq (p_o,x)$ or $(p_o,y) \leq \rho_o(p_o,x)$ for any $\{x,y\} \in \mathcal{T}$, establishing 6). Q.E.D.

Lemma 3.4 implies that there is a unique element of $\mathcal{A}$ called A_o which we shall call the von Neumann activity set. Corresponding to A_o is the maximal rate of balanced growth $\rho_o = 1/\lambda(A_o)$, the strictly positive von Neumann price vector p_o, and the strictly positive balanced growth ray x_o.

In order to apply the adjoint conditions developed in Section 1 to the capital model under consideration we shall have to assume that each A_i has only a finite number of extreme points. To show that this assumption is not overly restrictive we state a lemma:

Lemma 3.5. Let K be a compact convex set. Then, given $\varepsilon > 0$, there is a compact convex set K_ε with only finitely many extreme points such that $K_\varepsilon \subset K$ and for any $k \in K$, $d(k,K_\varepsilon) < \varepsilon$.

Proof: Cover K by a set of spheres of radius ε centered at each point k of K. By the Heine-Borel theorem there are a finite number of these spheres that serve as an open covering of K. Suppose there are t of them, with centers at $\{k_i\}_{i=1}^{t}$. Since K is compact and convex it is spanned by its extreme points (Karlin [12], Lemma 13.2.4). Thus, $k_i \in K$ implies $\exists \{x_{i_f}\}_{f=1}^{t_f}$ and non-negative weights γ_i such that $\sum_{f=1}^{t_f} \gamma_{i_f} = 1$, $x_{i_f} \in \mathcal{P}(K)$ for each f, and $k_i = \sum_{f=1} \gamma_{i_f} x_{i_f}$. Now let K_ε be the closed convex hull of $\{ \bigcup_{i=1}^{t} \bigcup_{f=1}^{t_f} \{x_{i_f}\} \}$. Clearly, K_ε is compact, convex, has only a finite number of extreme points, and $K_\varepsilon \subset K$. By construction for each $k \in K$, $\exists$ $k_i \in K_\varepsilon$ such that $d(k,k_i) < \varepsilon$ since each of the original k_i are convex combinations of extreme points of K_ε and thus included in K_ε . Q.E.D.

We now make another assumption concerning the technology.

A3: Each industry's activity possibility set A_i, $i = 1,\ldots,m$, is approximated by convex combinations of a finite number of activities drawn from A_i, including the activities that make up the matrix A_o. Let $\bar{A}_i$ be the matrix whose columns consist of the finite set of activities drawn from A_i.

In examining assumption A3 we note that if any A_i is a polyhedral convex set, no approximation would be required since such a set may be expressed as a convex combination of a finite number of points. If A_i

is a general compact, convex set, then Lemma 3.5 may be invoked to demonstrate that A_i can be closely approximated by convex combinations of a finite number of points. Let $\bar{\mathcal{A}}$ represent the set of matrices from $\mathcal{A}$ obtainable by approximating the A_i by the $\bar{A}_i$ for $i = 1,2, \ldots,m$.

A4: The utility function $U(\cdot)$ is a real-valued function defined on E^m, which is concave, non-decreasing, continuously differentiable, and dominated by a scalar homogeneous function of degree α, where $\beta/\lambda_o^\alpha < 1$, $0 \leq \alpha \leq 1$, and β is the factor discounting future returns.

Before discussing optimality conditions for long-run programs we must establish that an infinite horizon solution does indeed exist for the model presented under assumptions A1-A4. To do this we make use of Theorem 2.1, proven in Section 2. It has been shown that the von Neumann growth rate, ρ_o, in the hypothesis of Theorem 2.1 is equal to $1/\lambda_o$ in the model under discussion. Consequently, there is, by A2 and A4, a utility function which dominates $U(\cdot)$, which is scalar-homogeneous of degree α, and $\beta \rho_o^\alpha < 1$. Part (C) of Theorem 2.1 assures us that an optimal infinite horizon solution exists for the dominating utility. Thus, the proof of the theorem for the utility function $U(\cdot)$ holds because, for any feasible policy $< v_t >_{t=1}^{\infty}$, $\sum_1^\infty \beta^t U(v_t)$ is dominated by a uniformly convergent sum. Therefore, an optimal infinite horizon program will exist. However, referring to the discussion in Section 2, neither the optimal infinite horizon solution nor the optimal solution for any finite horizon program need be unique. It is clear that in this model, $U(\cdot)$ cannot be strictly concave in all the decision variables since it is a function of consumptions alone and not of allocation to production. It is for this reason that the policy convergence results proven in the last part of Section 2 are interesting. In order to relate the model we are discussing to the one for which the Kuhn-Tucker conditions were derived in Section 1 of this chapter, we let

$$\bar{A} = [\bar{A}_1, \bar{A}_2, \ldots, \bar{A}_m]$$

and

$$\bar{T} = [\bar{T}_1, \bar{T}_2, \ldots, \bar{T}_m] ,$$

where $\bar{T}_i$ is a matrix with m rows, the same number of columns as $\bar{A}_i$, and with all its elements zero except for those in row i, which are all ones. This form for $\bar{T}$ is a consequence of the no-joint production assumption which enables industry i to be identified with the production of product i.

To proceed with the analysis and develop long-run properties of optimal policies, it appears to be necessary to create conditions that will imply that each product is produced during each period. The need for this type of result will become obvious when we attempt to derive the asymptotic path of the prices. The key to demonstrating this property is to show that the first period-prices, $p^1(T)$, are bounded as the time horizon T becomes large. As we shall see, this will ensure that eventually, for large enough T, each good is consumed in each period.

<u>Lemma 3.6</u>. Let $p^1(N)$ represent an optimal first-period price vector for a program with N periods. Then, for fixed $x_1 \epsilon X$ with $x_1 \gg 0$, there exists a constant K^1 such that any sequence of optimal first-period price vectors $< \hat{p}^1(N) >_{N=1}^{\infty}$ satisfying the adjoint conditions satisfies:

$$\|\hat{p}^1(N)\| \leq K^1 \quad \text{for all} \quad N = 1,2,\ldots \quad .$$

Proof: We may rewrite the primal conditions, given in Section 3.1, in the following form for a program with N periods:

$$\max \sum_1^N \beta^{t-1} U(w_t)$$

$$(1) \qquad \begin{aligned} &\bar{A}v_1 + w_1 \leq x_1 \\ &\bar{A}v_t + w_t \leq \bar{T}v_{t-1} , \qquad t = 2,\ldots,N \\ &v_t \geq 0,\ w_t \geq 0, \qquad t = 1,2,\ldots,N \ . \end{aligned}$$

Let $v(N) = (v_1,v_2,\ldots,v_N)$, $w(N) = (w_1,w_2,\ldots,w_N)$, and form the Lagrangian:

$$\phi_N(v,w,p) = \sum_1^N \beta^{t-1}U(w_t) + (p^1(x_1-\bar{A}v_1-w_1)) + \sum_2^N (p^t(\bar{T}v_{t-1}- \bar{A}v_t-w_t)) \ .$$

Then if $\hat{v}(N)$ and $\hat{w}(N)$is an optimal N-period program, it is both necessary and sufficient that there exist an optimal price sequence $< \hat{p}^t >_{t=1}^N$, which satisfies the adjoint conditions of Section 3.1. However, this implies that $\hat{v}(N)$, $\hat{w}(N)$ and $\hat{p}(N)$ constitute a saddle point of the Lagrangian in the sense that

$$(2) \quad \phi(v,w,\hat{p}(N)) \leq \phi(\hat{v}(N),\hat{w}(N),\hat{p}(N)) \leq \phi(\hat{v}(N),\hat{w}(N),p))$$

for all $p \geq 0$ and feasible programs (v,w). It is clear that there is an optimal program for which the constraints in (1) hold with equality since $U(\cdot)$ is non-decreasing. Thus,

$$x_1 - \bar{A}\hat{v}_1(N) - \hat{w}_1(N) = 0$$

and

$$\bar{T}\hat{v}_{t-1}(N) - \bar{A}\hat{v}_t(N) - \hat{w}_t(N) = 0, \quad t = 2,3,\ldots,N.$$

Thus, from the first inequality of (2), we have

$$\sum_1^N \beta^{t-1} U(w_t) + \hat{p}^1(N)(x_1 - \bar{A}v_1 - w_1) + \sum_2^N \hat{p}^t(N)(\bar{T}v_{t-1} - \bar{A}v_{t-1} - w_t) \leq \sum_1^N \beta^{t-1} U(\hat{w}_t(N))$$

for all feasible v and w. In particular, we may take $v \equiv 0$ and $w \equiv 0$ for any starting point x_1 yielding

$$\hat{p}^1(N)\, x_1 \leq \sum_1^N \beta^{t-1} U(\hat{w}_t(N)) \ .$$

Observing that this result must hold for any horizon N, and noting that the existence of an optimal infinite horizon program implies that $\sum_1^N \beta^{t-1} U(\hat{w}_t(N))$ approaches the finite limit $f(x_1)$, the value of the infinite horizon program, we have

$$\hat{p}^1(N)\, x_1 \leq f(x_1) \quad \text{for all} \quad N = 1,2,\ldots$$

and $x_1 \gg 0$ implies $\exists\, M$ independent of the horizon such that

$$\|\hat{p}^1(N)\| \leq M \quad \text{for} \quad N = 1,2,\ldots \ .$$

Q.E.D.

Define a set of goods $\mathcal{C}$ as primary consumption goods such that $i \in \mathcal{C}$ iff there exist $k_i > 0$ such that

$$\frac{\partial U}{\partial w_i}(w) \geq k_i$$

whenever $i \in \mathcal{C}$ and w is a vector of consumptions whose i^{th} component is zero. In order to establish the property that each good must be produced in each period t for t large enough, we make two further assumptions:

A5: The set $\mathcal{C}$ of primary consumption goods has the property that, for each good $j \notin \mathcal{C}$, there is a good $i \in \mathcal{C}$ such that j is directly indispensable for the production of i.

A6: $\beta/\lambda_o > 1$, where β is the discount factor, and $\lambda_o = \lambda(A_o)$.

As an alternative to A6, which is an assumption placed jointly on the technology and preference functions, we may substitute an assumption that pertains more directly to the preference function and does not involve the rate of discount or maximal growth.

A6': For each $i \in \mathcal{C}$, $k_i = +\infty$.

Some remarks are probably in order about the assumptions. First, the requirement that $U(\cdot)$ be dominated by a certain type of scalar homogeneous function can be dispensed with if $U(\cdot)$ is bounded, since in this case an optimal infinite horizon solution exists by the remarks made following Theorem 2.1. If this is the case, then no restrictions are placed on β and λ_o by A4. However, if we do need a restriction of the type $\beta/\lambda_o^{\alpha} < 1$ in A4, it does not necessarily contradict the condition of A6 that $\beta/\lambda_o > 1$ as long as $\alpha < 1$. Since β is a discount factor, i.e., $\beta < 1$, the $\beta/\lambda_o > 1$ requirement of A6 assures us that $\lambda_o < 1$. But since the von Neumann growth factor for this system is $\rho_o = 1/\lambda_o$, we see that $\rho_o > 1$, implying that the system is capable of expanding proportionately. This kind of assumption was made in most of the previous work on growth models cited earlier. In our case, A6 requires the stronger assumption $\beta\rho_o > 1$. This may be interpreted as implying that the system is capable of discounted proportional expansion.

One consequence of the definition of a primary consumption good is to say that a good for which there exists a perfect substitute with respect to the utility function $U(\cdot)$ cannot be a primary consumption good. However, this concept does not rule out the existence of a substitution effect between these and other goods, nor does it restrict complementarity relationships. In fact, goods in $\mathcal{C}$ may have perfect complements in both $\mathcal{C}$ and $\sim\mathcal{C}$. Furthermore, since $U(\cdot)$ is concave, its partials are non-increasing, and thus we may choose the $\{k_i\}_{i=1}^{m}$ so that

$$\lim_{\substack{w \to \infty \\ w_i = 0}} \frac{\partial U(w)}{\partial w_i} \to k_i > 0 \quad \text{for all} \quad i \in \mathcal{C}$$

If all goods were in $\mathcal{C}$, then of course assumption A5 would not be necessary and the development would be simpler. However, it seems advantageous to build a framework that does not require all goods to be desirable for consumption, and in fact the utility function may be independent of the level of consumption of certain commodities. Both assumptions, A6 and A6', serve to ensure that eventually all consumption goods are consumed in each period of an optimal program.

Lemma 3.7. Under assumptions A1-A5 and A6 or A6', there is a time horizon T^* and a number t_1^* such that for all horizons $T \geq T_1^*$ every primary consumption good is consumed in every period $t \geq t_1^*$ in every optimal policy.

Proof: We first prove the result under assumptions A1-A6.* The adjoint condition (b) of Section 3.1 implies that any sequence of optimal prices $\langle p^t \rangle_{t=1}^{T}$ satisfies

---

* The method used is similar to one used by Morishima [16].

$$\bar{\pi}\, p^{t+1} \leq p^{t}\bar{A} \quad \text{for} \quad t = 1,2,\ldots,T-1,$$

thus

$$p^{t+1} \leq p^{t} A_{o}.$$

Let the i^{th} component of the von Neumann price vector p_o be denoted by p_{oi} and let $\hat{P}_o$ be the $m \times m$ diagonal matrix whose diagonal elements are $p_{o1},\ldots,p_{om}$. Define

$$Z(t) = p^{t} \lambda_{o}^{-t}\, \hat{P}_{o}^{-1}$$

and note that if $<p^{t}>$ satisfies the adjoint conditions, we have

$$Z(t+1) \leq Z(t)\, \hat{P}_{o}\, A_{o} \hat{P}_{o}^{-1}\, \lambda_{o}^{-1} .$$

Let $L = [1,1,\ldots,1]$, an m-dimensional vector, and observe that

$$L\, \hat{P}_{o} = p_{o} ,$$

and consequently,

$$L = L\, \hat{P}_{o}\, A_{o} \hat{P}_{o}^{-1}\, \lambda_{o}^{-1} .$$

Now define

$$C(t) = \max[Z_{1}(t),\ldots,Z_{m}(t)] .$$

Since $Z(t) \leq C(t)L$, we have

$$Z(t+1) \leq C(t)\, L\, \hat{P}_{o}\, A_{o}\, \hat{P}_{o}^{-1}\, \lambda_{o}^{-1} = C(t)L ,$$

and thus $<C(t)>$ is non-increasing and we may write

$$Z(t) \leq C(1)L \quad \text{for all} \quad t = 1,2,\ldots,T ,$$

where T is the time horizon.

We now view $C(1)$ as a function of the time horizon T and apply Lemma 3.6 to establish the existence of a number C^{*} independent of the time horizon T, such that $C(1) \leq C^{*}$ for any T. This yields

$$Z(t) \leq C^{*}\, L ,$$

or

$$p^t \leq C^* P_o \lambda_o^t \quad \text{for all} \quad t = 1,2,\ldots,T,$$

where the right-hand side is independent of the time horizon T.

Since A6 implies that $\lambda_o < 1$, we have established that each component of $p^t(T)$ goes to zero at geometrical rate λ_o, uniformly with respect to the time horizon T. Now look at any good $i \in \mathcal{C}$. From adjoint condition (a) of Section 3.1 and the result just obtained, we have shown that in any optimal policy,

$$(3) \qquad C^* p_{oi} \lambda_o^t \geq p_i^t(T) \geq \beta^{t-1} \frac{\partial U(w^t)}{\partial w_i} \quad \text{for} \quad t = 1,2,\ldots,T.$$

Now choose $k = \min_{i \in \mathcal{C}} \{k_i\}$. Then

$$\frac{\partial U(w)}{\partial w_i} \geq k > 0 \quad \text{for all} \quad i \in \mathcal{C}$$

and all consumption vectors w with $w_i = 0$. However, $\frac{\lambda_o}{\beta} < 1$ implies that $K^*(\frac{\lambda_o}{\beta})^{t-1} \to 0$ as $t \to \infty$, and thus $\exists\, t_1^*$ such that for all $t \geq t_1^*$,

$$(4) \qquad K^*(\frac{\lambda_o}{\beta})^{t-1} < k.$$

Thus, if $w_i^t = 0$ for some $i \in \mathcal{C}$, the result in (4) would contradict equation (3) if $T \geq T_1^*$, and hence this condition cannot exist in an optimal policy. Consequently, $w_i^t > 0$ for all $i \in \mathcal{C}$ and $t \geq t_1^*$ as long as $T \geq T_1^*$. Q.E.D.

If assumption A6' is substituted for A6, then adjoint condition (a) is immediately contradicted if $w_i = 0$ for some $i \in \mathcal{C}$ since $\frac{\partial U(w)}{\partial w_i} = +\infty$ for all $i \in \mathcal{C}$ with $w_i = 0$. Thus, $w_i^t > 0$ for all $i \in \mathcal{C}$ and $t = 1,2,\ldots,T$.

In the remainder of this section we shall assume that assumptions

A1-A5 and either A6 or A6' hold, and thus all previously developed results may be used freely.

Corollary 3.1. For all time horizons T greater than T_1^* every good is produced during each period t for $t_1^* \leq t < T$ under every optimal policy.

Proof: Lemma 3.7 tells us that each good in $\mathcal{C}$ is consumed during every period $t \in [t_1^*, T]$. Thus, each good in $\mathcal{C}$ must be produced during each of these periods. But assumption A5 tells us that for each good $j \notin \mathcal{C}$, there is a good $i \in \mathcal{C}$ for whose production j is indispensable; consequently, each good must be produced in each of the above periods.

Q.E.D.

Corollary 3.2. There are integers T_1^* and t_1^* such that, for all time horizons $T \geq T_1^*$ and all periods $t_1^* \leq t < T$, we can find a square matrix $A_t \in \bar{\mathcal{A}}$, a subset of the optimal activity set at time t, and an optimal price sequence $< p^t >_{t=1}^{T}$ such that

$$(5) \qquad p^{t+1} = p^t A_t \leq p^t A \quad \begin{cases} \text{for } & t_1^* \leq t \leq T, \text{ and for all} \\ & A \in \bar{\mathcal{A}}. \end{cases}$$

Proof: Using adjoint condition (b) of Section 1 we have $p^{t+1} \leq p^t A$ for all t and $A \in \bar{\mathcal{A}}$. Since Corollary 3.1 tells us that each good is produced when $t \geq t_1^*$, the same adjoint conditions gives the equality

$$p^{t+1} = p^t A_t,$$

where A_t contains at least one column from each industry's activity set since each good must be produced.

Lemma 3.8 (Morishima [16]): Given any cone N containing the price ray p_o in its interior, and any initial price p^1, there exists a neighborhood of p^1 and an integer $t(p^1)$ such that any path of prices starting in the neighborhood of p^1 and satisfying (5) remains within N for all $t \geq t(p^1)$.

Proof: The proof is given by Morishima in Lemmas 3 and 4, pages 164-69. The development makes use of the fact, proven in Lemma 3.3, that $\lambda_o p_o \ll p_o A$ for all $A \in \mathcal{Q}$ such that no column of A is a column of A_o, assumption A2, and Corollary 3.2.

Theorem 3.1. Under Assumptions A1-A5 and A6 or A6', if the time horizon T is sufficiently large, there exists an integer T^* such that for $t \geq T^*$ the optimal activity set at time t denoted by A_t is the set A_o.

Proof: We have shown that $\lambda_o p_o \ll p_o A$ if A does not contain any of the activities comprising A_o. Since $\lambda_o p_o = p_o A_o$, we have

$$p_o A_o \ll p_o A$$

for all activity sets A that do not include any of the activities in A_o. Thus, we can find a cone K containing p_o in its interior for which $pA_o \ll pA$ for all $p \in K$ and A, as above. For any initial price p^1, we can find a time $t(p^1)$ such that any price sequence $\langle p^t \rangle$ satisfying the adjoint conditions with initial element in a neighborhood of p^1 remains within K for all $t \geq t(p^1)$. As the set of optimal first-period prices is bounded with a bound independent of the time horizon, we can find a finite set of covering neighborhoods by the Heine-Borel theorem. If these covering neighborhoods have centers

$p^{11},\ldots,p^{1k}$, we choose $T^* = \max[t(p^{11}),\ldots,t(p^{1k})]$. Thus, if $T > T^*$, every optimal price sequence remains in K for all $t \geq T^*$. But $pA_o \ll pA$ for $p \in K$ implies that none of the activities comprising A can be operated at positive levels since the adjoint conditions require equality in the appropriate price equation whenever a corresponding primal variable is operated positively. Thus, no activity other than those included in A_o will be operated positively at time t where $T^* \leq t \leq T$.

Q.E.D.

REFERENCES

1. Arrow, K. J., L. Hurwicz, and H. Uzawa, Studies in Linear and Non-Linear Programming. Stanford, Calif.: Stanford University Press, 1958.

2. Arrow, K. J., and S. Karlin, "Price Speculation Under Certainty," ibid., Chapter 13.

3. Baumol, W. J., and R. E. Quandt, "Investment and Discount Rates Under Capital Rationing: A Programming Approach," The Economic Journal (June, 1965), 317-30.

4. Bellman, R. E., Dynamic Programming. Princeton, N. J.: Princeton University Press, 1957.

5. Bellman, R. E., and J. Bucy, "Asymptotic Control Theory," SIAM Journal on Control, Series A, Vol. 2, No. 1 (1964).

6. Dantzig, G. B., Linear Programming and Extensions. Princeton, N. J.: Princeton University Press, 1957.

7. ____________., "On the Status of Multistage Linear Programming Problems," Management Science, Vol. 2, No. 1 (Oct., 1959).

8. de Leeuw, K., "Lecture Notes," Mathematics 261, Stanford University (unpublished, 1965).

9. Dorfman, R., "Regional Allocation of Investment: Comment," Quarterly Journal of Economics (Feb., 1963).

10. Dorfman, R., P. A. Samuelson, and R. M. Solow, Linear Programming and Economic Analysis. New York: McGraw-Hill Book Co., 1958.

11. Dreyfus, S., "An Analytic Solution of the Warehouse Problem," Management Science, Vol. 4, No. 1 (Oct., 1957).

12. Karlin, S., Mathematical Methods and Theory in Games, Programming, and Economics, Vol. 1. Reading, Mass.: Addison-Wesley, 1959.

13. __________., "The Structure of Dynamic Programming Models," Naval Research Logistics Quarterly, No. 2 (1955), 285-95.

14. Manne, A. S., "Optimal Dividend and Investment Policies for a Self-Financing Business Enterprise," Working Paper No. 93R, Graduate School of Business, Stanford University (1966).

15. McKenzie, L. W., "Turnpike Theorems for a Generalized Leontief Model," Econometrica (January-April, 1962) 165-80.

16. Morishima, M., Equilibrium Stability and Growth. Oxford at the Clarendon Press, 1964.

17. Pontryagin, L. S., V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mischenko, The Mathematical Theory of Optimal Processes. New York: Interscience Publishers, 1962.

18. Radner, R., "Paths of Economic Growth that are Optimal with Regard Only to Final States: A Turnpike Theorem," Review of Economic Studies (February, 1961), 98-104.

19. Ramsey, F. P., "A Mathematical Theory of Savings," Economic Journal, Vol. 38 (1928), 543-59.

20. Royden, H. L., Real Analysis. New York: Macmillan, 1963.

21. Veinott, A. F. (ed.), Mathematical Studies in Management Science. New York: Macmillan, 1965.

22. ____________, "Lecture Notes, Non-Linear Programming," Operations Research Department, Stanford University (unpublished, 1964).

23. Winter, S. G., "Some Properties of the Closed Linear Model of Production," International Economic Review, Vol. 6, No. 2 (1965).

24. Debreu, G., and I. N. Herstein, "Non-negative Square Matrices," Econometrica, Vol. 21 (1953), 597-607.

25. Gale, D., "On Optimal Development in a Multi-Sector Economy," Operations Research Center, University of California at Berkeley, 1966.

26. Von Neumann, J., "A Model of General Economic Equilibrium," Review of Economic Studies, Vol. 13 (1945-46), 1-9.

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) Stanford University Department of Statistics Stanford, California	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP

3. REPORT TITLE
Optimal Allocation in a Pure Capital Model with Withdrawals: Asymptotic Properties and Algorithms

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
Technical Report, May 1967

5. AUTHOR(S) (Last name, first name, initial)
Seelenfreund, Alan J.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
May 31, 1967	83	26

8a. CONTRACT OR GRANT NO. Nonr-225(53)	9a. ORIGINATOR'S REPORT NUMBER(S) Technical Report No. 95
b. PROJECT NO. NR-042-002	
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) Technical Report No. 149 Nonr-225(50) (NR-047-004)
d.	

10. AVAILABILITY/LIMITATION NOTICES
Distribution of this document is unlimited.

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

13. ABSTRACT

The purpose of this report is to analyze a class of control problems that represent mathematical models of economic processes. Properties of optimal solutions will be investigated from both asymptotic and short-run viewpoints, and specific algorithmic results are presented whenever possible. In addition, a model of economic growth is formulated as a control process and it is shown that the activities associated with the von Neumann equilibrium must be chosen in each period if the process has proceeded for a sufficient length of time.

UNCLASSIFIED

Security Classification

KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Economic growth Asymptotic properties Capital budgeting Price speculation						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8a. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, 8c, & 8d. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through ________________."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through ________________."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through ________________."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS)*, *(S)*, *(C)*, or *(U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, roles, and weights is optional.