AN INFINITE-HORIZON DISCRETE-TIME QUADRATIC PROGRAM AS APPLIED TO A MONOPOLY PROBLEM

BY ROBERT L. GRAVES AND LESTER G. TELSER

This article gives necessary and sufficient conditions for the existence of a finite maximum of a quadratic functional. The functional is the present value of a revenue stream in discrete time over an infinite horizon. Both scalar and vector versions of the problem are solved. It is shown that the problem always has a solution for a sufficiently high or finite discount rate. Some conditions to ensure the nonnegativity of the solution are also presented. The origin of the problem is finding the sequence of outputs that will maximize the present value of the net return of a monopolist who sells \( k \) related products with related demands described by a set of \( n \)th order linear difference equations.

1. INTRODUCTION

THE ECONOMIC PROBLEM which inspires our work is that of finding a sequence of prices and outputs over time which maximizes the present value of the profits of a monopolist if the demand function depends on past as well as current purchases. The problem is more complicated if there are \( k \) commodities to consider which have related demands and it is gratifying to report that many of the results for the one product case apply to \( k \) products after suitable modification.

Possibly the most natural way to formulate the problem is in terms of continuous time. Then the mathematical problem of maximizing the present value of the revenue stream can be described as follows:

\[
\max_{q(t)} \int_0^\infty p(t)q(t)e^{-rt} dt, \quad r > 0,
\]

where

\[
q(t) = f(t) - bp(t) + \int_0^t h(s)q(t-s)ds, \quad b > 0,
\]

and

\[
f(t), q(t) \geq 0 \text{ for all } t \geq 0,
\]

\[
= 0 \text{ otherwise}.
\]

While the functions \( f(t) \) and \( h(t) \) are given, part of the problem is to find conditions which \( h(t) \) must satisfy. Since \( p(t) \) is a linear function of \( q(t) \), the objective function (1.1) is quadratic in \( q(t) \). The choice of an optimal nonnegative \( q(t) \) as a function of \( t \) is constrained by the condition that \( q(t) \) satisfies the integral equation (1.2).\(^2\)

\(^1\) This research is part of a project analyzing competition which is being supported by a National Science Foundation grant, Gs-365, to the University of Chicago. The authors worked out several parts of this material jointly. The second author wrote the paper while a guest of the Cowles Foundation for Research in Economics at Yale University.

\(^2\) For one of the earliest formal statements of a related problem see Evans [3].
We deal with an approximation to this problem which replaces the integral
criterion (1.1) with an infinite series (1.4) and the integral equation (1.2) with a
difference equation (1.5). Thus the discrete approximation to the original problem
is as follows:

\begin{equation}
(1.4) \quad \max_{t=0}^\infty \sum_{t=0}^\infty p_t q_t \beta^t, \quad 0 < \beta < 1,
\end{equation}

subject to

\begin{equation}
(1.5) \quad q_t = f_t - bp_t - \sum_{j=1}^n a_j q_{t-j}, \quad b > 0,
\end{equation}

where \( a_j \) corresponds to an ordinate of \( h(s) \) in the neighborhood of \( s=j \). We shall
not consider goodness of approximation of the discrete to the continuous problem.
It is interesting to note, however, that the continuous analogue of the difference
equation (1.5) is the integral equation (1.2) and not a differential equation so that
our problem does not fall within the realm of classical calculus of variations.

The sequence of mathematical problems we consider largely arises from the
economic source of the problem and consists of five parts. First, we give the condi-
tions which the optimal sequence \( \{q_t\} \) must satisfy if a maximum exists. For the
existence of a maximum strict convexity of the objective function is required. This
leads to the second set of problems which is the study of certain semi-infinite
Toeplitz forms [6]. Meaningful solutions to the problem must be suitably bounded.
Thus the third set of problems arises from a study of the conditions necessary for
the existence of such solutions. Next, since the quantity variable, \( q_t \), must be
nonnegative, we are led to analyze conditions such that a linear operator carries a
nonnegative input into a nonnegative output. The study of nonnegativity is the
fourth area. Finally, in Section 3 we extend most of these results to vectors of
outputs so that \( q_t \) can be a \( k \times 1 \) vector. An appendix gives results on matrix
difference equations that are not readily found in the literature and enable this
paper to be self-contained.

The relevance of these results, which are interpreted in terms of price policies,
to the general problem of linear decision rules should be apparent.\(^3\)

\section{2. The One-Product Case}

\textit{Statement of Problem}

The monopolist is assumed to face a demand equation as follows:

\begin{equation}
(2.1) \quad q_t + a_1 q_{t-1} + \ldots + a_n q_{t-n} = f_t - bp_t, \quad b > 0,
\end{equation}

where \( q_t \) is quantity demanded in period \( t \), \( p_t \) is price in period \( t \), and \( f_t \) is an
arbitrary function of time which may represent seasonals, trends, real income, etc.
This demand equation could arise because, for instance, customers hold inventories

\(^3\) The standard references on linear decision rules are Holt, Modigliani, Muth, and Simon
so that current purchases depend on previous purchases. It is convenient to adopt a more compact notation. We use the linear lag operator $L$, and its inverse $E$, the lead operator, defined by

\begin{align*}
(2.2) \quad Lx_t &= x_{t-1}, \\
(2.3) \quad Ex_t &= x_{t+1}.
\end{align*}

This enables the demand equation to be written

\begin{equation}
(2.4) \quad (1 + a_1 L + \ldots + a_n L^n)q_t = A(L)q_t = f_t - bp_t.
\end{equation}

Let $c_t$ denote total cost in period $t$. Assume that $c_t$ is a quadratic function of current quantity, $q_t$, and a finite number of lagged quantities. Thus net revenue in period $t$ is $r_t = p_t q_t - c_t$ and the present value ($PV$) of net revenue is

\begin{equation}
(2.5) \quad PV = \sum_{0}^{\infty} \beta^t r_t = \Sigma \beta^t p_t q_t - \Sigma \beta^t c_t.
\end{equation}

From now on we adopt the convention that where the indexes of summation are omitted they shall be understood to run from zero to infinity. The discount factor, $\beta$, is the reciprocal of one plus the discount rate so that

\begin{equation}
(2.6) \quad 0 < \beta < 1.
\end{equation}

In certain formal expressions, we may allow $\beta = 1$.

Clearly, the cost and gross revenue terms in the expression for present value have the same formal structure because both are quadratic expressions in current and lagged quantities. Therefore, nothing substantive is lost by confining attention to the gross revenue term in (2.5) and completely neglecting $c_t$. The more complicated situation in which there are inventories, so that production does not equal sales, leads to a study of constrained maxima, however, and requires a somewhat different analysis.

We seek necessary and sufficient conditions on the polynomial operator $A(L)$ so that sequences $\{p_t\}$ and $\{q_t\}$ can be found such that $PV$ has a finite maximum where $PV$ is given by

\begin{equation}
(2.7) \quad PV = \Sigma \beta^t p_t q_t,
\end{equation}

and the constraints given by (2.4) are satisfied for all $t$.

Nonnegativity constraints are more difficult. The nonnegativity of the optimal sequences $\{p_t\}$ and $\{q_t\}$ clearly depend not only on the nature of the demand operator $A(L)$ and the optimal operator to be derived below, but also on the properties of the forcing function $f_t$. This greatly complicates the necessary and sufficient conditions that guarantee nonnegativity. The results given below depend only on the properties of the operators and not on the specific properties of the forcing function $f_t$ other than its nonnegativity.
Necessary and Sufficient Conditions for Existence of a Finite Maximum $PV$

Although the problem has several avenues of approach, the one we give generalizes most easily to the multi-product case. Because there are an infinite number of time periods, the conditions for a finite optimum are most securely based on first principles.

**Definition:** A $\beta$-sequence, $\{x_t\}$, is one which satisfies $\Sigma \beta^t|x_t|^2 < \infty$.

For a given $\beta$, the set of all $\beta$-sequences form a Hilbert space. We assume that $\{f_t\}$ is an element of this space. Our task is to find an element of this space for which (2.7) is a maximum and which meets prescribed initial conditions. Suppose that $\{q_t\}$, a $\beta$-sequence satisfying $q_t = 0$ for all $t < 0$, is the optimal quantity path. Consider the $\beta$-sequence $\{q_t + d_t\}$. Without losing generality, we may set $b = 1$.\(^4\) In an obvious notation,

$$
PV\{q_t + d_t\} = \Sigma \beta^t(f_t - A(L)(q_t + d_t))
$$

(2.8)

$$
= \Sigma \beta^t q_t(f_t - A(L)q_t) + \Sigma \beta^t d_t(f_t - A(L)q_t)
$$

$$
- \Sigma \beta^t q_t A(L)d_t - \Sigma \beta^t d_t A(L)d_t .
$$

The following easily verified lemma allows the regrouping of terms.

**Lemma 1:**

(2.9) \[ \Sigma \beta^t q_t A(L)d_t = \Sigma \beta^t d_t A(\beta E)q_t , \]

provided $0 \leq \beta < 1$ and $\{q_t\}$ and $\{q_t + d_t\}$ are $\beta$-sequences.

Lemma 1 applied to (2.8) yields

(2.10) \[ PV\{q_t + d_t\} = PV\{q_t\} + \Sigma \beta^t d_t[f_t - A(L)q_t - A(\beta E)q_t] - \Sigma \beta^t d_t A(L)d_t . \]

Since any $\beta$-sequence may be obtained by proper choice of $\{d_t\}$, we obtain the first theorem.

**Theorem 1:** In order that the $\beta$-sequence $\{q_t\}$ maximize $PV$ uniquely, it is necessary and sufficient that $q_t$ be the solution of

(2.11) \[ [A(L) + A(\beta E)]q_t = f_t , \]

which is a $\beta$-sequence and satisfies the condition that $q_t = 0$ for $t < 0$, and that

(2.12) \[ \Sigma \beta^t d_t A(L)d_t > 0 \]

for any $\beta$-sequence $\{d_t\}$ not identically zero.

At this point several remarks are pertinent. First, it is clear that quantity paths need not be bounded since the present value is finite even when quantity grows over

---

\(^4\) In the subsequent development we shall occasionally find it helpful to display $b$ explicitly and this will always be clear from the context.
time provided the growth rate factor is eventually less than $\beta^{-\frac{1}{2}}$.
Second, we have so far imposed no restrictions on the demand operator $A(L)$ and in particular we have said nothing about its stability. Finally, the forcing function $f$, can grow over

and it is clear from (2.11) that this imparts a growth to quantity.

It is also instructive to write (2.12) in matrix form and the case $n=2$ is sufficiently
general. Thus

$$
(2.13) \quad \Sigma \beta^i d_i A(L) d_i = [d_0 d_1 d_2 \ldots \beta a_1 \beta a_2 \beta^2 \beta^2 a_1 \beta^2 a_2 \ldots]
$$

where the matrix of the form has a countable infinity of rows and columns. By modifying the

infinite triangular matrix of the quadratic form in (2.13) in a simple

way, it becomes possible to apply the Herglotz Lemma and obtain necessary and

sufficient conditions on the demand operator $A(L)$ that will tell when condition

(2.12) is satisfied. It is helpful to state a lemma whose proof is again immediate.

**Lemma 2:** The expressions

$$
(2.14) \quad \Sigma \beta^i d_i A(L) d_i \quad \text{and} \quad \Sigma \beta^{i+1} d_i [A(\beta^i L) + A(\beta^i E)] \beta^{i+2} d_i
$$

are equal. Hence if one is positive so is the other.

The matrix associated with the second expression in (2.14) has a special character

as illustrated for $n=2$:

$$
(2.15) \quad \left[ \begin{array}{cccc}
2 & \beta a_1 & \beta a_2 & 0 & \ldots \\
\beta a_1 & 2 & \beta a_1 & \beta a_2 & \ldots \\
\beta a_2 & \beta a_1 & 2 & \beta a_1 & \ldots \\
0 & \ldots & & & \\
\vdots & & & & \\
\end{array} \right]
$$

It is a Toeplitz or band symmetric matrix to which the Herglotz Lemma applies. The

Herglotz Lemma asserts that the sum

$$
(2.16) \quad \sum_{u, v} g(u-v) h(v) \overline{h(v)} > 0
$$

for all finite integers $S$ if and only if

$$
(2.17) \quad g(u) = \int_{-\pi}^{\pi} e^{ix} dF(x), \quad -\pi \leq x \leq \pi \quad (u = 0, \pm 1, \pm 2, \ldots)
$$

\footnote{Hence revenue (in money terms) may grow at a rate which is less than the discount rate. See also p. 243.}

\footnote{For a statement and proof of the Herglotz Lemma and Bochner’s Theorem, see Loève [12, pp. 207–210]. See also Bochner [2, p. 325].}
where $F(x)$ is a non-decreasing, bounded, nonnegative function of the real variable $x$. Thus $F(x)$ may be thought of as a distribution function and $g(u)$ as its characteristic function. Some readers may prefer to make the interpretation in terms of Fourier transforms. Notice the bar over $h(v)$ which denotes the complex conjugate. Hence the form in (2.16) is to be positive definite over the complex numbers.

To see how this Lemma gives the desired conditions on $A(L)$, we begin by defining $a_{-u} = a_u$ for $u = 1, \ldots, n$ and $a_0 = 1$ where $a_u$ is the coefficient of $L^u$ in the polynomial $A(L)$. Suppose we allow the $d_i$ 's in (2.14) to be complex. Evidently if the form in (2.14) is positive over complex numbers then it will be positive over the real numbers. Thus write

\[(2.14') \quad \sum \beta^t \bar{d}_t A(L) d_t = \frac{1}{2} \sum \beta^{1/2} \bar{d}_t \left[ A(\beta^{1/2} L) + A(\beta^{1/2} E) \right] \beta^{1/2} d_t \]

and make the correspondences as follows:

\[(2.18) \quad g(u - v) = a_{u - v} \beta^{1/2} \]

\[(u, v = 0, 1, 2, \ldots), \quad h(u) = \beta^{u/2} d_u. \]

For any finite integer $S$, the right side of (2.14') becomes

\[(2.19) \quad \sum_{u, v = 0}^{S} \beta^{u/2} \bar{d}_u \beta^{u-v/2} a_{u-v} \beta^{v/2} d_v. \]

The Herglotz Lemma, however, applies to finite sums whereas the sum in (2.14') is over all of the positive integers $t$. Therefore, we need the following:

**Lemma 3:** For all $\beta$-sequences $\{d_u\}$,

\[(2.20) \quad \sum \beta^{u/2} \bar{d}_u \beta^{u-v/2} a_{u-v} \beta^{v/2} d_v > 0 \]

if and only if the finite sum (2.19) is positive for all finite $S$.

**Proof:** Assume (2.20) is positive for all $\beta$-sequences $\{d_u\}$. Obviously, the finite sums in (2.19) must also be positive for otherwise there would be a subsequence of the infinite sum which would be negative contrary to the hypothesis. Conversely, assume that the finite sum is positive for all bounded sequences. Since the infinite sum is absolutely convergent, there is an integer $S$ sufficiently large that allows us to approximate the infinite sum by a finite sum as closely as we please. Therefore, if the finite sums are positive for all choices of $S$ then the infinite sum must also be positive for all $\beta$-sequences.

Now we can prove

**Theorem 2:** $\sum \beta^t \bar{d}_t A(L) d_t > 0$ for all $\beta$-sequences $\{d_t\}$ with $0 < \beta \leq 1$ if and only if

\[(2.21) \quad A(e^{i\alpha}) + A(e^{-i\alpha}) > 0 \text{ for all } \alpha \in [-\pi, \pi]. \]
**Proof:** According to (2.17) and the correspondence set up in (2.18),

\[ \beta^{\frac{|u|}{2}} a_u = \int_{-\infty}^{\infty} e^{iux} dF(x), \quad u \neq 0, \quad 2a_0 = \int_{-\infty}^{\infty} dF(x). \]

Since \( \sum |a_u| \beta^{\frac{|u|}{2}} \) is finite, the function \( F(x) \) is differentiable.\(^7\) Therefore,

\[ F'(x) = \frac{1}{2\pi} \left[ 1 + \sum_{u=1}^{n} \beta^{\frac{|u|}{2}} a_u e^{iux} + 1 + \sum_{u=1}^{n} \beta^{\frac{|u|}{2}} a_u e^{-iux} \right] \]

or, more compactly,

\[ F'(x) = \frac{1}{2\pi} \left[ A(\beta^x e^{ix}) + A(\beta^{-x} e^{-ix}) \right]. \]

The right side of (2.14) is positive for \( \beta \)-sequences \( \{a_u\} \) and \( 0 < \beta \leq 1 \) if and only if (2.20) holds. Appealing to Lemma 3, the Herglotz Lemma, and (2.23), inequality in (2.20) holds if and only if \( F'(x) > 0 \) for all \( 0 < \beta \leq 1 \) and all \(-\pi \leq x \leq \pi\). Since \( A(\beta^x e^{ix}) + A(\beta^{-x} e^{-ix}) = 2 \text{Re} A(\beta^x e^{ix}) \),

the left side is a harmonic function of \( \beta^x e^{ix} \) defined on the closed unit disk.\(^8\) Therefore, it attains its extrema on the unit circle where \( \beta = 1 \). This completes the proof of the theorem.

It is worth noting that since \( a_u = a_{-u} \) (2.21) is equivalent to the following:

\[ (2.21') \quad 1 + a_1 \cos x + \ldots + a_n \cos nx > 0. \]

We illustrate the conditions imposed on \( A(L) \) for \( n = 2 \). In this case we require

\[ (2.24) \quad \phi(x) = 1 + a_1 \cos x + a_2 \cos 2x > 0 \quad \text{for} \quad -\pi \leq x \leq \pi. \]

This condition is met if and only if \( \min \phi(x) > 0 \). Since

\[ \phi'(x) = -\sin x (a_1 + 4a_2 \cos x), \]

\[ \phi'(x) = 0 \quad \text{for} \quad x = 0, \pi, \] and \( \cos x = -a_1/4a_2 \) provided \( |a_1| < 4|a_2| \). Hence

\[ \phi(0) = 1 + a_1 + a_2 > 0, \]

\[ \phi(\pi) = 1 - a_1 + a_2 > 0, \] and

\[ \phi \left( \cos^{-1} \left( -\frac{a_1}{4a_2} \right) \right) = \frac{8 - a_2^2 - 8(a_2 - \frac{1}{2})^2}{8a_2} > 0. \]

The admissible region in the \((a_1, a_2)\)-space is shown in Figure 1. Any point in the

\(^7\) We use herein a well known result on characteristic functions stated in a Corollary, Loeve [12, p. 188].

\(^8\) A harmonic function is the real part of an analytic function. See Ahlfors [1] for a fine exposition of the properties of harmonic functions. The unit disk is the interior of the unit circle; the closed unit disk is the closure of the unit disk.
cone shaped region satisfies (2.24) and the associated matrix (2.15) will be positive definite.

Condition (2.21) permits further simplification in terms of the reciprocal equation

\[(2.25) \quad A(z) + A(z^{-1}) = 0 \]

that plays a fundamental role in much of the development. For \(|z| = 1\), there is an obvious correspondence between \(A(z) + A(z^{-1})\) and \(A(e^{ix}) + A(e^{-ix})\) because \(z^{-1} = e^{-ix}\) when \(z = e^{ix}\).

**Theorem 3:** Let

\[ A(z) = 1 + a_1 z + \ldots + a_n z^n, \quad a_n \neq 0, \]

have real coefficients. The reciprocal equation (2.25) has \(n\) roots inside the unit circle if and only if

\[(2.26) \quad A(z) + A(\bar{z}) > 0 \quad \text{for} \quad |z| = 1. \]

**Proof:** Assume \(A(z) + A(\bar{z}) > 0\) for \(|z| = 1\). Since \(z^{-1} = \bar{z}\) on the unit circle, the reciprocal equation (2.25) cannot have any roots on the unit circle. Therefore, it has \(n\) roots inside the unit circle (and \(n\) reciprocal roots outside the unit circle) which proves sufficiency.

Conversely, assume that (2.25) has \(n\) roots inside the unit circle. Form

\[(2.27) \quad A(z) + A(z^{-1}) = z^{-n} g(z) \]

so that \(g(z)\) is a reciprocal polynomial with real coefficients of degree \(2n\). Thus it is possible to represent \(g(z)\) as follows:

\[(2.28) \quad g(z) = a_n \prod_{j=1}^{n} (z - r_j)(z - r_j^{-1}), \]

where the \(r_j\)'s are the roots of \(g(z) = 0\) inside the unit circle, i.e., \(|r_j| < 1\). Define the polynomial

\[(2.29) \quad C(z) = \Pi(z - r_j) = z^n + b_1 z^{n-1} + \ldots + b_n. \]
If any root \( r_j \) is complex, then its conjugate is also a root inside the unit circle and included in (2.29). Therefore, \( C(z) \) has real coefficients. Moreover

\[
(2.30) \quad b_n = (-1)^n \prod r_j,
\]

\[
z^{-n} g(z) = a_n C(z) \prod (1 - z^{-1} r_j^{-1}).
\]

Therefore,

\[
(2.31) \quad A(z) + A(z^{-1}) = a_n b_n C(z) C(z^{-1}).
\]

It follows from (2.31) and the fact that the \( b \)'s are real that \( 2 = (a_n b_n)(1 + b_1^2 + \ldots + b_n^2) \). This implies that \( a_n b_n > 0 \). Therefore, for \(|z| = 1\),

\[
(2.32) \quad A(z) + A(\bar{z}) = a_n b_n |C(z)|^2 > 0,
\]

which completes the proof of the theorem.

The factor \( C(z) \) is uniquely determined by the condition that all of its roots are inside the unit disk.\(^9\) The importance of this theorem lies in the fact that it becomes possible to ascertain whether a symmetric band matrix is positive definite in a finite number of steps. One needs only to calculate the roots of the reciprocal polynomial and see whether any have modulus 1. Using this theorem we can prove an important corollary about the stability of \( A(z) \).

**Corollary:** If the reciprocal equation (2.25) has no roots on the unit circle, then the polynomial \( A(z) = 0 \) has no roots in the closed unit disk.

**Proof:** If \( A(z) = 0 \) had a root \( z_0 \) on the unit circle then \( A(z_0) = 0 \) and \( A(z_0^{-1}) = 0 \) so that (2.25) would vanish for \( z_0 \). Since (2.26) is a harmonic function, it does not vanish inside the unit circle if it is always positive on the unit circle. Therefore, \( A(z) = 0 \) has no roots in the closed unit disk.

**Conditions for Existence of a Solution**

We now consider conditions to ensure that the solution of (2.11) is a \( \beta \)-sequence; note that Theorem 1 does not assert this. We pose the problem as follows: Given a particular sequence \( \{f_1\} \) and a particular demand operator \( A(L) \), find the largest value \( \beta_0 \), \( 0 < \beta_0 \leq 1 \), so that \( 0 < \beta < \beta_0 \) implies that the solution of (2.11) is a \( \beta \)-sequence. We set \( \beta_0 = \min(\beta_\alpha, \beta_f, 1) \) where \( \beta_\alpha \) is determined by \( A(L) \) and \( \beta_f \) is determined by \( \{f_1\} \). First, \( \beta_\alpha \) is the largest number having the property that \( \beta < \beta_\alpha \) implies \( \text{Re} A(\beta^2 w) > 0 \) for \(|w| = 1\). To see that this is a sensible definition of \( \beta_\alpha \),

\(^9\) Theorem 3, but not this elementary proof, generalizes to hold for nonnegative integrable functions on the unit circle. See Nelson and Lowdenslager [9, 1] and Hoffman [10, especially Chapter 4]. For an elementary proof of a closely related theorem see also Grenander and Szegö [6].
recall that for a given value of \( \beta \), \( \text{Re}\ A(\beta^4 w) \) is a harmonic function of \( w \). If the function is positive for all \( w \) with \( |w|=1 \), then it is positive for all \( w \) with \( |w| \leq 1 \). Thus if the function is positive for \( |w|=1 \) for some value of \( \beta \), it is positive for \( |w|=1 \) for all smaller values of \( \beta \). There is some value of \( \beta \) for which it is positive on \( |w|=1 \) since \( A(0)=1 \) and \( A(w) \) is continuous. Clearly \( \beta_A > 0 \) and \( \text{Re}\ A(\beta^4_A w) = 0 \) for some \( w \) with \( |w|=1 \). If \( A(w)=1 \), then \( \beta_A \) is infinite. Next, \( \beta_T \) is the largest number with the property that \( \beta < \beta_T \) implies \( \sum_0^\infty \beta^t |f_t|^2 < \infty \). We rule out the possibility that \( \beta_T=0 \) (this corresponds to an infinite discount rate). It may or may not be true that \( \sum_0^\infty \beta^t |f_t|^2 < \infty \). It is possible that \( \beta_T \) is infinite; for example if \( f_t=0 \) for \( t \geq T \).

Before proving that the given value of \( \beta_T \) has the desired property we make a few remarks about the meaning of the result. Most important is the fact that there is always a finite maximum present value for a sufficiently small positive discount factor \( \beta \). The quantity path \( \{q_t\} \) is a \( \beta \)-sequence but need not be bounded; it is necessary, however, that \( \beta^t q_t \) approach zero. Finally if for \( \beta=\beta_T \) it is known that \( \{f_t\} \) is a \( \beta \)-sequence and \( \text{Re}\ A(\beta^4 w) > 0 \), the methods used here will only guarantee that \( \{q_t\} \) is a \( \beta \)-sequence for \( \beta < \beta_T \); nothing can be said about \( \{q_t\} \) for the case \( \beta = \beta_T \).

We phrase the problem formally as

**Theorem 4:** Let \( \{f_t\} \) and \( A(L) \) be given together with the corresponding values of \( \beta_T \) and \( \beta_A \). Let \( \beta_T = \min(\beta_A, \beta_T, 1) \). If \( \beta_0 < \beta_T \) then the problem: max \( PV = \sum_0^\infty \beta_0^t q_t P_t \), where \( P_t = f_t - A(L)q_t \), has a solution \( \{q_t\} \) which is a \( \beta \)-sequence for \( \beta = \beta_0 \).

**Proof:** To simplify the analysis we make the substitutions \( F_t = \beta_0^{-t^2} f_t, Q_t = \beta_0^{-t^2} q_t \), and \( P_t = \beta_0^{-t^2} P_t \). Then the problem becomes: max \( PV = \sum_0^\infty Q_t P_t \), where \( P_t = F_t - A(\beta_0^2 L)Q_t \). It is clear that \( \{F_t\} \) is a square summable sequence. In fact, since \( \beta_0 < \beta_T \), it is true that

\[
|F_t| \leq M \delta^t
\]

for some \( M > 0 \) and some \( \delta < 1 \). Theorems 1 and 2 apply directly to this problem. Necessary and sufficient conditions that it have a solution are that \( \{Q_t\} \) be a \( \beta \)-sequence (with \( \beta=1 \)) satisfying

\[
[A(\beta_0^2 L) + A(\beta_0^2 E)] Q_t = F_t
\]

and

\[
\text{Re}\ A(\beta_0^2 w) > 0 \quad \text{for} \quad |w|=1 .
\]

The truth of the second statement is an immediate consequence of the fact that \( \beta_0 < \beta_A \). The first statement is merely a rewriting of equation (2.11). It remains only to show that the solution of (2.34) is a \( \beta \)-sequence (with \( \beta=1 \)) or, in other words, is square summable. A (stronger) statement to this effect is given (without a complete proof) in Theorem 2 of [16]. If \( \beta_T < \beta_A \), this theorem states that \( \{q_t\} \) is a \( \beta \)-sequence.
even when $\beta = \beta_0 = \beta_r$. The constructive proof given here is useful because it is closer to the usual treatment of difference equations.

The characteristic equation of the optimal equation (2.34) is the reciprocal equation

\begin{equation}
A(\beta_0^2 w^{-1}) + A(\beta_0^2 w) = 0.
\end{equation}

The fact that Re $A(\beta_0^2 w) > 0$ for $|w| = 1$ ensures that (2.35) has $n$ roots of modulus less than 1. If these roots are denoted by $r_1, r_2, \ldots, r_n$ then the substitution $B(z) = z^\delta C(z^{-1})$ and equation (2.31) give

\begin{equation}
A(\beta_0^2 w^{-1}) + A(\beta_0^2 w) = (a_n/b_n) B(\beta_0^2 w) B(\beta_0^2 w^{-1}) ,
\end{equation}

where $B(\beta_0^2 w) = \prod_{j=1}^{n} (1 - r_j w)$. This permits the optimal operator to be represented as

\begin{equation}
A(\beta_0^2 L) + A(\beta_0^2 E) = (a_n/b_n) B(\beta_0^2 E) B(\beta_0^2 L)
\end{equation}

which in turn allows the optimal equation to be written as

\begin{equation}
B(\beta_0^2 E) B(\beta_0^2 L) Q_t = (a_n/b_n) F_t .
\end{equation}

We next note that the constant $\delta$ which appears in equation (2.33) may be chosen so that $\delta > |r_j|$ for $j = 1, 2, \ldots, n$.

It is possible to regard (2.37) as a sequence of equations of the form

\begin{equation}
(I - r E) X_t = Y_t
\end{equation}

or

\begin{equation}
(I - r L) X_t = Y_t,
\end{equation}

with $r$ having one of the values $r_1, r_2, \ldots, r_n$. For equations of this form, we show that $|Y_t| \leq M_0 \delta^t$ implies $|X_t| \leq M_1 \delta^t$ where $M_1$ depends only on $\delta$, $r$ and some arbitrary constant (or boundary condition). The solution of (2.38) is

\begin{equation}
X_t = \sum_{0}^{\infty} r^s Y_{s+t},
\end{equation}

\begin{align*}
|X_t| & \leq M_0 \sum_{0}^{\infty} |r|^s \delta^{s+t} \\
& = M_0/(1 - \delta |r|) \delta^t. 
\end{align*}

Choosing $M_1 = M_0/(1 - \delta |r|)$ gives the desired result. The solution of (2.39) is

\begin{equation}
X_t = C r^t + \sum_{0}^{t} r^{t-s} Y_s .
\end{equation}

Thus

\begin{equation}
|X_t| \leq |C||r|^t + M_0 \sum_{0}^{t} |r|^{t-s} \delta^s
\end{equation}
\[ \leq |C| |r|^f + M_0 \delta (1 - (|r|/\delta)^{f+1})(1 - |r|/\delta) \]
\[ \leq (|C| + M_0(1 - |r|/\delta)) \delta^f. \]

The fact that \(|r| < \delta\) is used in writing the last inequality. Choosing \(M_i = |C| + M_0(1 - |r|/\delta)\) gives the desired result for equation (2.39).

Since \(B(\beta^Ez) = \Pi(I - r_jz)\) and \(B(\beta^Ez) = \Pi(I - r_jL)\) with \(|r_j| < 1\), the successive applications of \((I - r_jz)^{-1}\) or \((I - r_jL)^{-1}\) to a right hand side whose modules is bounded by \(M \delta^f\) yields another right hand side of the same form. Thus finally \(|Q_1| \leq M \delta^f\) where \(M\) depends only on \(\delta\), the various \(r_j\) and the \(n\) constants which are set by boundary conditions. This is sufficient for \(\{Q_i\}\) to be square summable which gives the desired result for \(\{q_i\}\) since \(q_i = \beta^{-i/2} Q_i\). This proves the theorem.

Although the emphasis here has been on solutions which are \(\beta\)-sequences rather than on bounded solutions, it is desirable to have some statement about the latter. The result is related to Theorem 4.

**Corollary:** Let \(\{f_i\}\) be bounded and \(\Re A(z) > 0\) for \(|z| = 1\). Then (2.11) has a bounded solution \(\{q_i\}\) for all \(\beta\) with \(0 < \beta \leq 1\).

**Proof:** The characteristic equation of (2.11) is \(A(z^{-1}) + A(\beta z) = 0\). We must show that this equation has \(n\) roots of modulus less than 1. The change of variable \(z = w/\beta^k\) gives the reciprocal equation \(A(\beta^k w^{-1}) + A(\beta^k w) = 0\). We shall show that this equation has \(n\) roots of modulus less than \(\beta^k\). Suppose there were a root \(w_0\) such that \(\beta^k \leq |w_0| \leq 1\). This implies \(\beta^k |w_0| \leq 1\) and \(\beta^k |w_0^{-1}| \leq 1\). Then there would be points \(z_0 = \beta^k w_0\) and \(z_1 = \beta^k w_0^{-1}\) inside (or on) the unit circle such that \(\Re A(z_0)\) and \(\Re A(z_1)\) differ in sign (the sum is zero). At some point on the line joining these points, which is inside the unit circle, \(\Re A(z) = 0\), which contradicts the hypothesis that the harmonic function \(\Re A(z)\) is positive for \(|z| = 1\).

We note for future reference that the representation
\[ A(\beta^Ez) + A(\beta^Ez) = (a_n/b_n)(B(\beta^Ez)B(\beta^Ez)) \]
also permits us to write
\[ (2.40) \quad A(L) + A(\beta z) = (a_n/b_n) B(\beta z) B(L) \]

This allows the optimal solution to be written formally as
\[ (2.41) \quad B(L)q_i = (b_n/a_n) B(\beta z)^{-1}f_i, \]
which shows that the optimal path responds to a "present value" of the forcing function \(f_i\).
Nonnegativity Conditions

It is helpful to begin the discussion of nonnegativity with the original demand equation (2.4). In the static case, $A(L)$ reduces to the scalar 1 and there is a linear relation between $p$ and $q$ which is illustrated in Figure 2. The ordinate $0Y_0$ corresponds to $f$ and gives the quantity that would be purchased at a zero price. The abscissa $0X_0$ gives the lowest price that is just high enough to reduce the quantity demanded to zero. A positive quantity is purchased at any positive price in the interior of the interval $(0, X_0)$.

For a dynamic relation such as (2.4), past purchases influence the current demand. Hence past purchases may be so high (and consumer inventories so large) that no matter how low the current price, there would be no demand. This happens when past purchases substitute for present purchases. Thus a high buying in the past shifts the demand schedule $Y_0 X_0$ left to $Y_1 X_1$. If there are always positive purchases at low enough (positive) prices, then past buying is complementary with present buying and signifies partial adjustment of consumers to the present price. In this case positive past purchases shift the demand right to $Y_2 X_2$. Moreover, if the time period is long enough, then past purchases have less effect on the present and saturation, while possible, is less likely. To see why, note that if $A(L)$ is of order $n$, then no matter what the past purchases, in at most $n$ periods consumers are willing to purchase positive amounts at positive prices. Hence, in at most $n$ periods, the relevant demand schedule is $X_0 Y_0$ and always intersects the positive quadrant. We shall explore primarily the consequences of assuming the demand operator $A(L)$ converts nonnegative inputs into nonnegative outputs. Formally, we state

**Definition:** $A(L)$ is an unconditionally nonnegative operator if and only if there exists $\lambda > 0$ such that for any bounded nonnegative sequence $\{y_t\}$ there is a bounded nonnegative sequence $\{x_t\}$ which satisfies $A(L)x_t = y_t$, $t \geq 0$, and $x_t = 0$ for $-n \leq t \leq -1$. 

![Figure 2](image-url)
Obviously, there are operators which transform some sequences of nonnegative inputs into sequences of nonnegative outputs but which fail to have this property for all nonnegative input sequences. Such operators are conditionally nonnegative. The adjective “conditionally” is a reminder that whether the output sequence is nonnegative depends on the properties of the specific input sequence \( \{y_t\} \). For a given operator \( A(L) \), it is possible in principle to characterize the space of input sequences that would produce nonnegative output sequences. Such a characterization leads to a study of the properties of the input sequences \( \{y_t\} \) for the given \( A(L) \) which yield nonnegative outputs. For some operators \( A(L) \) there may not be any input sequences which would yield nonnegative outputs. Our analysis poses a different problem. Given the set of all bounded nonnegative sequences, it seeks to characterize the class of operators \( A(L) \) which yield bounded nonnegative outputs. Later on we give some results pertaining to conditional nonnegativity. The requirement that bounded inputs produce bounded outputs implies that the operator must be stable.

**Lemma 4:** If \( A(L) \) is unconditionally nonnegative (u.n.n.) then there is a largest \( \lambda_0 > 0 \) such that \( A(\lambda L) \) is stable for \( \lambda < \lambda_0 \).

**Proof:** Consider the special input sequence \( \{\delta_{0t}\} \) where
\[
\delta_{0t} = \begin{cases} 
1 & t = 0 \\
0 & \text{otherwise} 
\end{cases}
\]

By hypothesis the sequence \( \{x_t\} \) which satisfies
\[
A(\lambda L)x_t = \delta_{0t}, \quad t \geq 0,
\]
is bounded and nonnegative. Define the power series
\[
A(\lambda z)^{-1} = \sum \gamma_t \lambda^t z^t.
\]
Obviously, \( A(z) \sum \gamma_t z^t = 1 \) so that the coefficients of (2.43) must satisfy (2.42). That is, \( x_t = \gamma_t \lambda^t \). Since \( \lambda^t \gamma_t \) is a linear combination of all of the roots of \( A(\lambda z^{-1}) = 0 \) and is bounded, all of the roots of \( A(\lambda z^{-1}) = 0 \) must lie inside the unit circle. Therefore, all of the roots of \( A(\lambda z) = 0 \) are outside the unit circle and \( A(\lambda L) \) is stable. If \( P \) is the largest modulus of a root of \( A(z^{-1}) = 0 \) then \( \lambda_0 = P^{-1} \).

In the course of this proof we have virtually shown

**Lemma 5:** The operator \( A(L) = 1 + a_1 L + \ldots + a_n L^n \) is u.n.n. if and only if \( \gamma_t \) is nonnegative where \( \gamma_t \) is the coefficient of \( z^t \) in (2.43).

**Proof:** Since, as we have seen,
\[
A(L)\gamma_t = \delta_{0t}
\]
if \( A(L) \) is u.n.n., then \( \gamma_t \geq 0 \). Sufficiency is obvious.
Together these two Lemmas show that unconditional nonnegativity of $A(L)$ does not depend in any essential way on its stability because it is always possible to choose a positive $\lambda$ depending only on the root of largest modulus of $A(e^{-\lambda})=0$ to ensure the stability of $A(\lambda L)$ without affecting the signs of $\gamma_i$. Therefore, in order to study the u.n.n. of $A(L)$, there is no loss of generality in assuming for the remainder of this section that $A(L)$ has been converted into a stable operator.

It would be desirable to have an algorithm for calculating whether a given operator $A(L)$ is u.n.n. in a finite number of steps. The two preceding lemmas suggest examination of the reciprocal $A(z)^{-1}$ for any negative coefficients which is obviously impractical. Since $\gamma_1$ satisfies (2.44), it would be better to have criteria of nonnegativity in terms of the $n$ roots of $A(z)\equiv 0$, the $n$ coefficients, $a_1, \ldots, a_n$, or a judicious combination. Fortunately, there are some useful necessary conditions that can shorten the labor of determining whether a given operator is u.n.n. To derive these necessary conditions we exploit the relation between unconditional nonnegativity and characteristic functions of discrete valued random variables.

**DEFINITION:** The characteristic function of the discrete valued random variable $\xi$ where $P(\xi = u) = d_u$, $d_u \geq 0$, and $\Sigma d_u = 1$, is defined by

$$f(x) = \Sigma e^{iux} d_u$$

where $x$ is a real variable.\(^{10}\)

A theorem of Bochner asserts that $f(x)$ is a characteristic function if and only if it is a nonnegative function. That is, the hermitian matrix $[f(x_u - x_v)]$ ($u, v = 1, 2, \ldots, m$), is positive semi-definite for any points $x_1, \ldots, x_m$.\(^{11}\) Thus a characteristic function has properties as follows:

(i) $f(0) = 1$,
(ii) $f(-x) = \overline{f(x)}$,
(iii) $|f(x)| \leq f(0)$.

**LEMMA 6:** $A(L)$ is unconditionally nonnegative if and only if

(2.45) $f(x) = \sigma [A(e^{ix})]^{-1}$

is the characteristic function of a discrete valued random variable $\xi$ such that $P(\xi = t) = \sigma \gamma_t \geq 0$ for some $\sigma > 0$.

**PROOF:** Assume that $A(L)$ is u.n.n. By Lemma 4, $A(z) = 0$ has no roots inside the closed unit disk, and by Lemma 5, $\gamma_t \geq 0$ for all $t \geq 0$. Therefore, $A(z)^{-1}$ converges

\(^{10}\) For the basic properties of characteristic functions of discrete valued random variables, see Gnedenko and Kolmogorov [5, pp. 55–61].

\(^{11}\) Bochner [2, pp. 325–328].
on the closed unit disk. Set $\sigma = A(1) > 0$ and $f(x)$ is the desired characteristic function. The converse is obvious, since $\sigma \gamma_i \geq 0$ and $\sigma > 0$ imply $\gamma_i \geq 0$. We note also that $\sigma < 1$ since $\gamma_0 = 1$ and $\Sigma \gamma_i > 1$.

An unconditionally nonnegative operator $A(L)$ possesses several important properties and we collect these in the following:

**Theorem 5:** If $A(L)$ is unconditionally nonnegative then (i) the smallest root of $A(z) = 0$ must be positive, (ii) $a_i < 0$, (iii) $|A(e^{ix})| \geq A(1)$, (iv) $0 < A(1) < 1$, and (v) strict inequality obtains in (iii) unless $x = 0$.

**Proof:** (i) The asymptotic behavior of $\gamma_i$, where $\gamma_i$ is the coefficient of $x^i$ in the power series $A(z)^{-1}$, is governed by the reciprocal of the roots of the smallest modulus of $A(z) = 0$. If the smallest root (or roots) of $A(z) = 0$ were negative, then $\gamma_i$ would change sign contradicting the hypothesis that $A(L)$ is u.n.n. Next, suppose that the roots of the smallest modulus of $A(z) = 0$ were complex. Thus there would be a pair of roots, $r_1 = \rho e^{ix}$ and $\bar{r}_1 = \rho e^{-ix}$, $-\pi \leq x \leq \pi$. For large enough $t$, $\gamma_i \sim \rho^i [e^{i(x+\theta)} + e^{-i(x+\theta)}]$,

or in real terms,

$\gamma_i \sim 2\rho^t \cos(tx + \theta)$.

In order that $\cos (tx + \theta) \geq 0$ for all $t$, we must have

$-\pi/2 + 2m\pi \leq tx + \theta \leq \pi/2 + 2m\pi$,

where $m$ depends on $t$ and is integral. Therefore, $x$ must satisfy the following:

$2\pi m/t - (1/t)(\pi/2 + \theta) \leq x \leq (1/t)(\pi/2 - \theta) + 2\pi m/t$ for all $t = 1, 2, \ldots$.

Therefore, $2\pi m/t \leq x \leq 2\pi m/t$. This is satisfied only by $x = 0$. Thus the roots must be real and hence positive.

(ii) Since $\gamma_1 = -a_1 > 0$, this proves (ii). Cf. (2.44).

(iii) By Lemma 6, the hypothesis implies that $A(e^{ix})^{-1}$ is proportional to a characteristic function so that $|A(e^{ix})^{-1}| \leq A(1)^{-1}$. Therefore, $|A(e^{ix})| \geq A(1)$ for all $-\pi \leq x \leq \pi$.

(iv) Obviously, $A(1) > 0$. In Lemma 6 we observed that $\Sigma \gamma_i > 1$. But $A(1)^{-1} = \Sigma \gamma_i$. Hence $A(1) < 1$.

(v) The random variable for which $A(e^{ix})^{-1}$ is the characteristic function, takes on integral values with probabilities proportional to $\gamma_i$. Therefore, the span of the distribution is 1 and strict inequality obtains in (iii) for all $x \neq 0$ in the interval $-\pi \leq x \leq \pi$.

This completes the proof.

---

12 A span is the difference between two successive values of a discrete valued random variable. In the theorem the span is one. See Gnedenko and Kolmogorov [5] for an extensive discussion.
One of the implications of Theorem 5 is that at most a finite number of terms in 
the reciprocal of $A(z)$ require examination to ascertain unconditional nonnegativity
because eventually $\gamma_t$ is dominated by the largest root of $A(z^{-1})=0$ which must be
positive. Unfortunately, no conditions are known to us which characterize com pletely
either the roots of $A(z)=0$ or the coefficients for an unconditionally nonnegative operator. It is, however, easy to verify two important sufficient conditions.
First, if $a_1, a_2, \ldots, a_n$ are all negative, then $A(L)$ is unconditionally nonnegative.
This follows easily from (2.44). Second, if all of the roots of $A(z)=0$ are positive,
then $A(L)$ must be u.n.n. This follows from the more general proposition that
the product of characteristic functions is a characteristic function. If all of the roots
of $A(z)=0$ are positive then $A(z)=\Pi(1-r_j z)$, $0<r_j<1$, and each factor
$(1-r_j z)$ is proportional to the characteristic function of a simple random variable
distributed according to a binomial law so that the product must also be a characteris tic function.

The discussion of unconditional nonnegativity refers so far to the demand operator $A(L)$, which contains only positive powers of $L$. Hence the reciprocal $A(z^{-1})$ converges in the closed unit disk. The optimal operator, $[A(L)+A(\beta E)]$
refers to the difference equation, (2.11), which the optimal output path satisfies.
The corresponding function $[A(z^{-1})+A(\beta z)]$, where $z^{-1}$ replaces $L$ and $z$ replaces $E$, contains negative as well as positive powers of $z$ so that its reciprocal converges in an annulus where it can be represented by a Laurent series. The definition of unconditional nonnegativity can be extended to apply to the optimal operator.
Since it is assumed that the optimal operator is such that the maximum present
value is finite, the optimal output path satisfies

\[ B(L)q_t = (\beta z_0 z \alpha z_0) B(\beta E)^{-1} f_t. \]

The definition of u.n.n. applies to $B(L)$ and the operator $B(L)$ is u.n.n. if and only
if every coefficient of $B(\beta E)^{-1}$ is nonnegative. Moreover, Theorem 5 applies to
$B(L)$ if it is u.n.n.

It is convenient at this point to consider conditionally nonnegative operators.
Although the quantity path must be nonnegative because the firm is not allowed to
repurchase goods from its customers, prices may be negative in some periods.
Negative prices mean sales below cost (recalling our earlier convention about costs).
Hence it would be more natural to maximize the present value of receipts subject to
explicit nonnegativity constraints on $q_t$ but not on $p_t$. By so doing (2.11) needs
little modification since for $q_t=0$, the permissible variation of $d_t$ must be positive,
which implies that the optimal output path must satisfy $f_t - [A(L)+A(\beta E)]q_t \leq 0$ in
order that neighboring sequences have a lower $P \nu$. Moreover, there is a sequence
of nonnegative numbers $\{r_t\}$, such that

\[ q_t r_t = 0, \]

\[ [A(L)+A(\beta E)]q_t = f_t + r_t. \]
The numbers \( r_i \) represent the partial derivatives of the present value of revenue evaluated at the optimal nonnegative quantity sequence. This shows that if the optimal operator is conditionally nonnegative and the given nonnegative sequence \( \{f_i\} \) does not yield a nonnegative \( \{q_i\} \), then adding \( r_i \) to \( f_i \) does give a nonnegative \( q_i \). Moreover, for any period \( t \) in which it is necessary to increase \( f_i \), adding \( r_n \), the corresponding optimal quantity must be zero. In any event, since \( A(L) \) is of degree \( n \) so that \( q_i \) depends on, at most, \( n \) previous quantities, in any sequence of \( n+1 \) periods, there can be, at most, \( n \) members of the optimal quantity sequence which are actually zero. In other words, there must always be at least one strictly positive quantity in any sequence of \( n+1 \) periods. Unfortunately, we do not have an algorithm calculating the nonnegative sequence \( \{q_i\} \) which satisfies (2.46) for any nonnegative \( \{f_i\} \), although there are some results for special sequences which allow use of standard finite quadratic programming methods. For example, if \( f_i \) satisfies 

\[
 f_i = [A(L) + A(\beta E)]h_i, \text{ where } h_i \geq 0 \text{ for } t \geq T, \text{ clearly a nonnegative solution is readily found. This includes the cases in which } f_i \text{ is eventually constant, or grows geometrically. However, if } B(L) \text{ is u.n.n., then for any nonnegative } f_i \text{, there exists a nonnegative sequence } \{q_i\} \text{ such that (2.38) is satisfied. In other words, unconditional nonnegativity is a sufficient condition for the existence of nonnegative solutions. Indeed, it is sufficient to ensure that equality and not merely inequality holds. Therefore, u.n.n. of } B(L) \text{ implies that a linear decision rule maximizes PV. Unfortunately, even the analysis of unconditional nonnegativity is quite difficult and we can only report partial results on the relations between u.n.n. of the demand operator } A(L) \text{ and the optimal operator } B(L). 

First, there is the following useful theorem:

**THEOREM 6:** If \( A(L) \) is unconditionally nonnegative and \( A(e^{i\pi}) + A(e^{-i\pi}) > 0 \) for all \( -\pi \leq x \leq \pi \) then

\[
|B(z)| < \sqrt{\frac{\sum b_i^2}{\sum a_i}} |A(z)| \text{ for all } |z| \leq 1,
\]

and

\[
0 < \sum_{i=0}^{n} a_i < \sum_{i=0}^{n} b_i^2, \quad a_0 = b_0 = 1,
\]

where \( B(z) \) is defined in (2.36).

**PROOF:** The hypothesis ensures that \( A(z) = 0 \) has no roots in the closed unit disk by the corollary to Theorem 3, and that by Theorem 3

\[
A(z) + A(z^{-1}) = (a_n/b_n) B(z) B(z^{-1})
\]

is a valid representation such that \( B(z) = 0 \) has no roots in the closed unit disk. By Lemma 6, \( A(e^{i\pi})^{-1} \) is proportional to a characteristic function. Hence,

\[
A(e^{i\pi})^{-1} = \sum \gamma_t e^{i\pi}, \quad \gamma_t \geq 0 \text{ for all } t = 0, 1, \ldots
\]
Therefore,

\[(2.49) \quad A(e^{i\theta})^{-1} + A(e^{-i\theta})^{-1} = [A(e^{i\theta}) + A(e^{-i\theta})] / |A(e^{i\theta})|^2\]

is also a characteristic function subject to multiplication by a certain positive normalizing constant. Applying the factorization and property (iii) of characteristic functions,

\[(2.50) \quad A(e^{i\theta})^{-1} + A(e^{-i\theta})^{-1} = (a_n / b_n) |B(e^{i\theta})|^2 / |A(e^{i\theta})|^2 \leq 2/A(1)\]

Since \((a_n / b_n)(1 + b_1^2 + \ldots + b_n^2) = 2\) and \(A(1) = \Sigma a_n\), we obtain

\[(2.51) \quad |B(e^{i\theta})| < \sqrt{\frac{\Sigma b_n^2}{\Sigma a_n}} |A(e^{i\theta})| .\]

Since both \(A(z)\) and \(B(z)\) are zero-free in the closed unit disk, both \(\log |A(z)|\) and \(\log |B(z)|\) are harmonic functions in the closed unit disk. Hence, we may represent \(\log |A(z)|\) as follows:

\[(2.52) \quad \log |A(z)| = (1/2\pi) \int \text{Re} \left( \frac{e^{ix} + z}{e^{ix} - z} \right) \log |A(e^{ix})| dx\]

for all \(|z| < 1\), and similarly for \(\log B(z)\). Since

\[\text{Re} \left( \frac{e^{ix} + z}{e^{ix} - z} \right)\]

is positive, the inequality in (2.51) extends to the entire unit disk and we obtain (2.47).\(^{13}\) It follows that \(\log |B(z)| < (1/2)^m + \log |A(z)|\) where \(m = (\Sigma b_n^2) / (\Sigma a_n) > 0\). Since \(B(0) = A(0) = 1\), \(\log m > 0\), \(m > 1\) and (2.48) holds.

This theorem establishes a connection between the demand and optimal polynomials on the hypothesis that the demand operator is unconditionally nonnegative and permits a finite maximum present value. If in addition \(B(L)\) is u.n.n., then Theorems 5 and 6 together are helpful in deciding for a given problem whether the nonnegativity conditions are automatically satisfied.

The following rather special theorem illustrates the difficulties of generalizing about u.n.n.

**Theorem 7**: Given that

\[A(L) + A(E) = (a_2 / b_2) B(L) B(E)\]

where both \(A(L)\) and \(B(L)\) are quadratics with real coefficients, unconditional nonnegativity of \(B(L)\) implies u.n.n. of \(A(L)\).

**Proof**: Denote the roots of \(B(z) = 0\) by \(s_j^{-1}\) and of \(A(z) = 0\) by \(r_j^{-1}\) where \(|s_j| < 1\)

\(^{13}\) We use Poisson's formula. See Ahlfors [I, pp. 179–181 and 184].
and $|r_j| < 1, j = 1, 2$. Since $B(L)$ is u.n.n., its largest root is positive and therefore both roots are real:

$$A(z) + A(z^{-1}) = (a_2/b_2)B(z)B(z^{-1}).$$

We can equate coefficients of like powers of $z$ thereby obtaining

$$\frac{1}{a_2}a_1 = \frac{1}{b_2} \left[b_1 + b_1b_2\right] = -\frac{1}{\left[(s_1 + s_2)/s_1s_2 + (s_1 + s_2)\right]}.$$  

Therefore,

$$\frac{a_1}{a_2} = -\frac{1}{\left[(s_1 + s_2)/s_1s_2 + (s_1 + s_2)\right]}.$$  

In addition,

$$2a_2 = \frac{1}{b_2} \left[1 + b_1^2 + b_2^2\right].$$

Since $B(L)$ is u.n.n., both roots of $B(z^{-1}) = 0$ must be real and their sum is positive. Therefore, there are either two positive roots or a positive and a smaller (in absolute value) negative root. First assume there are two positive roots. Then the expression for $a_1/a_2$ in terms of these roots shows that $a_1/a_2 < 0$. Further, $b_2 > 0$ and the expression for $a_2$ in terms of $b_1$ and $b_2$ shows that $a_2 > 0$. It follows that $a_1 < 0$ and $A(z^{-1}) = 0$ has two positive roots. Hence $A(L)$ is u.n.n. Next assume that $b_2 < 0$ so that one of the roots of $B(z) = 0$ is negative and necessarily $b_1 < 0$. In this case, $1/s_1 + 1/s_2 + s_1s_2 < 0$, since it is dominated by the smaller and negative root of $B(z^{-1}) = 0$. Hence $a_1/a_2 > 0$, $a_2 < 0$, and $a_1 < 0$. Therefore, $A(z^{-1}) = 0$ has two real roots whose sum is positive. Thus in this case as well, $A(L)$ is u.n.n.

The converse to Theorem 7 is false. To see why, consider this simple counterexample. Let $A(L) = (1 - rL)^2$ where $0 < r < 1$. Obviously $(1 - rz)^2 + (1 - rz)^{-1} = 0$ has no real roots, so that $B(L)$ cannot be u.n.n. because $B(z) = 0$ has a pair of complex conjugate roots.

It is also not possible to extend Theorem 7 to polynomials of higher degree. For example, consider cubics. As before we denote the roots of $A(z^{-1}) = 0$ by $r_j$ and of $B(z^{-1}) = 0$ by $s_j$ where the $r$'s and $s$'s are all inside the unit disk. Equate coefficients of equal powers of $z$ using

$$A(z) + A(z^{-1}) = (a_2/b_2)B(z)B(z^{-1}),$$

and obtain the following:

$$\frac{a_2}{a_3} = \frac{1}{b_3} \left[b_2 + b_1b_3\right]$$

$$= -\frac{1}{\left[(s_1 + s_2 + 1)/s_1s_2 + s_1 + s_2 + s_3\right]}.$$  

Letting $\Sigma s_1s_2$ denote the sum over all possible distinct pairs,

$$a_1/a_3 = \frac{1}{b_3} \left[b_1 + b_1b_2 + b_2b_3\right]$$

$$= \Sigma 1/s_1s_2 + \Sigma s_1s_2 + \Sigma s_1s_2 + s_1/s_1s_2,$$

$$2a_3 = \frac{1}{b_3} \left[1 + b_1^2 + b_2^2 + b_3^2\right].$$
Assume that $C(L)$ is u.n.n. and has three real roots, the largest of which is positive. Therefore, $b_1$, $b_2$, and $b_3 < 0$. Since $a_3$ has the same sign as $b_3$, $a_3 < 0$. If the two negative roots are small then

$$\Sigma 1/(s_1s_2) + \Sigma s_1s_2 + \Sigma \Sigma 1/s < 0.$$ 

Therefore, $a_1/a_2 < 0$ and $a_1 > 0$ so that $A(L)$ cannot be u.n.n.

Even if all of the roots of $B(z) = 0$ are positive, it is possible to construct an example in which only one of the roots of $A(z) = 0$ is positive, the other two roots are complex conjugates, and $A(L)$ is not u.n.n. Hence there do not seem to be any neat general results linking unconditional nonnegativity of the optimal and demand operators.

Nonnegativity of the Price Sequence

As noted above, nonnegativity of the output path is required because the firm is not allowed to repurchase the commodity from its customers. That is, if the solution gave rise to some negative $q_t$'s, this would imply that repurchases were required and without analysis of the customers' inventory levels, this procedure would not make any sense. Hence only nonnegative output paths can be optimal. The optimal price sequence, however, can include some negative prices. Negative prices would mean that the firm finds it desirable to make some sales at prices below cost in order to maximize its long run profits. To see how this can come about we proceed to obtain the optimal price path. The demand equation is given by $A(L)q_t = f_t - bp_t$, and the optimal output path is the solution of $[A(L) + A(\beta E)]q_t = f_t$. It readily follows that the optimal price path is the $\beta$-sequence satisfying

$$[A(L) + A(\beta E)]p_t = (1/b)A(\beta E)f_t.$$ 

The factored form is

$$(2.53') \quad B(L)p_t = (a_1/b_0)(1/b)B(\beta E)^{-1}A(\beta E)f_t.$$ 

Even if $B(L)$ is u.n.n. so that every coefficient of $B(\beta E)^{-1}$ is nonnegative, and even if $A(L)$ is u.n.n., at least one coefficient of $A(\beta E)$, namely $a_t$, is negative so that the right side of (2.53') may be negative for some $t$. Hence, the optimal price path may require some prices to be negative. This is more likely to be true the closer the discount factor is to one and the larger in absolute value is $a_t$. The latter condition could arise when the largest and positive root of $A(z^-1) > 0$ is close to one so that the customers of the firm have a strong propensity to repeat their purchases of the good.

Relative Stability of the Optimal and Demand Operator

The next theorem compares the stabilizing behavior of the demand and optimal operators.
THEOREM 8: Let $s_1$ denote the largest root in modulus inside the unit circle of the reciprocal equation $A(z) + A(z^{-1}) = 0$ and assume that the reciprocal equation has no roots on the unit circle. If $s_1$ is real there is a real root of $A(z^{-1}) = 0$ denoted $r_1$ such that $|s_1| < |r_1| < 1$.

PROOF: By hypothesis, $A(s_1) + A(s_1^{-1}) = 0$. Since the reciprocal equation has no roots on the unit circle, $A(z)$ has no roots in the closed unit disk by the Corollary to Theorem 3. Therefore, $A(z) > 0$ for all real $z$ such that $|z| < 1$. Therefore, $A(s_1) > 0$ and $A(s_1^{-1}) < 0$. Since $A(1) > 0$, there must be a real root $r_1^{-1}$ of $A(z) = 0$ such that $1 < |r_1^{-1}| < |s_1^{-1}|$, which proves the theorem.

This theorem has several important implications. First, if the optimal operator is u.n.n., then $s_1$ is positive so that the demand operator must also have a positive root $r_1$ such that $0 < s_1 < r_1$. This implies that the transient component of the optimal output path damps out more quickly than the transient component of the demand equation. Thus if the optimal operator is u.n.n. it is asymptotically stabilizing relative to the demand operator. It should be noted that it is not necessarily true that the largest root of $A(z^{-1}) = 0$ is real since the theorem merely asserts that there is at least one real root of this equation which exceeds the largest real root of the reciprocal equation inside the unit circle.

It can be shown that if the largest root of the reciprocal equation inside the unit circle is complex, then its modulus can exceed that of the largest root of $A(z^{-1}) = 0$. Thus unconditional nonnegativity of the optimal operator is a necessary but not a sufficient condition for the asymptotic stability of the optimal operator relative to the demand operator.

3. THE MULTI-PRODUCT CASE

We now assume that the monopolist wishes to maximize the present value of the revenue arising from the sale of $k$ products for which the demand is represented by a system of $k$ linear difference equations of order $n$. The general plan of this section is the same as that of section 2. Thus write the demand equations as follows:

$$p_t = f_t - [A_0 q_t + A_1 q_{t-1} + ... + A_n q_{t-n}]$$

(3.1)

where $p_t$ is a $k \times 1$ vector of prices, $q_t$ is a $k \times 1$ vector of quantities, $f_t$ is a $k \times 1$ vector of arbitrary functions of time, $A_k$ is a $k \times k$ matrix of real constants, $u=0, 1, ..., n$. The components of $f_t$ must be nonnegative and the sequence of components of both $f_t$ and $q_t$ are required to be $\beta$-sequences. The set of equations represented by (3.1) corresponds to (2.1) except that we here set $b=1$. In terms of the lag operator $L$, (3.1) becomes

$$p_t = f_t - A(L)q_t$$

(3.2)
where $A(L)$ is a polynomial matrix and each of the $k^2$ elements of $A(L)$ is a polynomial in $L$ of, at most, degree $n$. That is,

$$A(L) = \left[f_{hs}(L)\right] \quad (h, j = 1, \ldots, k),$$

and $f_{hs}(L)$ is a polynomial. Another way of representing $A(L)$ shows more clearly its polynomial character.

$$A(L) = A_0 + A_1 L + \ldots + A_n L^n.$$  (3.4)

Since each coefficient of $A(L)$, $A_n$, is a $k \times k$ matrix, the polynomial matrix $A(L)$ is said to be of order $k$. Hence $A(L)$ is a polynomial matrix of order $k$ and degree $n$. The degree is determined by the highest power of $L$ which has a non-null matrix coefficient. In addition we define the polynomial matrix in the complex scalar $z$ as follows:

$$A(z) = A_0 + A_1 z + \ldots + A_n z^n.$$  (3.5)

**Definition:** The matrix polynomial $A(z)$ given by (3.5) is of full rank if and only if $\det A(z)$ does not vanish identically.

It follows that if $A(z)$ is of full rank then the polynomial equation

$$\det A(z) = 0$$  (3.6)

has a finite number of roots and the polynomial matrix is singular only for those values of $z$ which are roots of (3.6).

From now on, all of the matrix polynomials considered are assumed to be of order $k$ and of full rank. Additional properties of matrix difference equations are given in the Appendix.

The present value of revenue is defined by

$$PV = \sum \beta^t q_t p_t.$$  (3.7)

This differs from the corresponding expression for the single product case only in that the revenue of period $t$ is the inner product $q_t p_t$. It would be unnecessarily tedious to repeat all of the steps which led to Theorem 1, except to point out that in the multi-product case Lemma 1 would require restatement as follows:

$$\sum \beta^t q_t A(L) d_t = \sum \beta^t d_t A'(\beta E) q_t,$$  (3.8)

and it should be noted that the operator $A(L)$ becomes transposed into $A'(\beta E)$. For the multi-product case the counterpart to Theorem 1 is

**Theorem 9:** In order that the $\beta$-sequence of vectors $\{q_t\}$ maximize $PV$ uniquely, it is necessary and sufficient that $q_t$ be a $\beta$-sequence and the solution of

$$[A(L) + A'(\beta E)]q_t = f_t,$$  (3.9)
and that

\[ (3.10) \quad \sum \beta^t d_t^* A(L)d_t > 0 \]

for any \( \beta \)-sequence \( \{d_t\} \) not identically zero. \( A(L) \) is the matrix operator defined in \( (3.4) \).

The conditions on \( A_u, u = 0, 1, \ldots, n \) which guarantee that \( (3.10) \) is satisfied is a matrix generalization of the Herglotz Lemma and Bochner's Theorem. The results are stated here for polynomials although they can be extended to cover matrix power series. As in Section 2 we give our results for a complex version of \( (3.10) \). If \( p_t \) and \( q_t \) were permitted to be complex, the maximand \( (3.7) \) would be \( \sum \beta^t q_t^* p_t \), and \( (3.10) \) would become

\[ (3.10^*) \quad \sum \beta^t d_t^* A(L)d_t > 0 \]

where the * denotes the complex conjugate so that \( d_t^* = \bar{d}_t \). Clearly, if \( (3.10^*) \) is positive for complex \( d_t \), then \( (3.10) \) is positive for real \( d_t \).

It is interesting to note that the choice of \( d_t = 0 \) for \( t \geq 1 \) in \((3.10^*)\) implies that \( A_0 \) must be positive definite and symmetric as is easily verified. However, condition \( (3.10) \) implies only that \( A_0 \) must be positive definite, not that it must be symmetric as well. Thus the extension to complex sequences narrows the class of admissible matrix operators at least with respect to \( A_0 \).

Define

\[ (3.11) \quad A_{-u} = A_u^*. \]

The elements of the matrices \( A_u \) are always assumed to be real. If the \( A \)'s were allowed to be complex, than \( A_{-u} \) would be the complex conjugate of \( A_u \). It is readily verified that

\[ (3.12) \quad \sum \beta^u d_u^* A(L)d_u = (\frac{1}{\beta}) (\sum \beta^u A(L) \beta^{-u} d_u) \]

(compare with Lemma 2), and the latter can be written out as follows:

\[ (3.13) \quad \sum \beta^{u-v} d_u^* A_{u-v} d_v + \sum \beta^{u} d_u^* A_0 d_v \]

(Cf. 2.14°). We used here the fact that \( A_0 \) is symmetric. The band matrix of the form given by \((3.12)\) is illustrated for \( n = 2 \):

\[ (3.14) \begin{bmatrix} 2A_0 & \beta^1 A_1 & \beta A_2 & 0 & \ldots \\ \beta^1 A_{-1} & 2A_0 & \beta^1 A_1 & \beta A_2 & \ldots \\ \beta A_{-2} & \beta^1 A_{-1} & 2A_0 & \beta^1 A_1 & \ldots \end{bmatrix} \]

The theorem which follows refers to a finite sum of the form \((3.13)\).

**Theorem 10:** The form

\[ (3.15) \quad \left(\frac{1}{\beta} \right) \left[ \sum_{u,v=0}^{s} d_u^* A_{u-v} d_v + \sum_{u=0}^{s} d_u^* A_0 d_u \right] \]
is positive for all choices of finite integers \( S > 0 \) and finite sequences \( \{d_u\} \) not identically zero, if and only if for \( |z| = 1 \), the form

\[(3.16) \quad \eta^* [A'(z) + A(z)] \eta > 0 \]

for any vector \( \eta \neq 0 \).

**Proof:** Define

\[(3.17) \quad W(x) = A(e^{ix}) + A^*(e^{ix}) \]

where \( z = e^{ix} \). Taking the Fourier transform of \( W(x) \) gives

\[(3.18) \quad A_u = (\frac{1}{2\pi}) \int W(x) e^{-iux} dx \quad (u = 1, ..., n) , \]

and

\[(3.19) \quad A_0 = (\frac{1}{2\pi}) \int W(x) dx . \]

The integration of a matrix is understood to be done by individual elements, over the interval \([-\pi, \pi]\). Hence (3.15) can be written as follows:

\[(3.20) \quad (\frac{1}{2\pi}) \left[ \sum_{u=0}^{S} d_u^* \left( \int W(x) e^{-(u-x)e^{ix}} dx \right) a_u \right] = (\frac{1}{2\pi}) \left[ \sum_{u=0}^{S} d_u^* e^{-iux} \right] W(x) \left[ \sum_{u=0}^{S} d_u e^{iux} \right] dx . \]

Hence, if (3.16) is positive, then (3.15) must be positive, which proves sufficiency.

Conversely, assume that (3.15) and hence (3.20) are positive for all finite integers \( S > 0 \) and sequences \( \{d_u\} \). For arbitrary \( x \) and nonzero \( \eta \) we wish to prove that

\[(3.21) \quad \eta^* W(x) \eta > 0 . \]

It is sufficient to show that

\[(3.22) \quad \left( \frac{1}{2\pi} \right) \int \left[ \frac{\sin N(s-x)}{s-x} \right]^2 \eta^* W(x) \eta dx = \int F(s-x) \eta^* W(x) \eta dx > 0 \]

since, for large \( N \), (3.22) approximates (3.21). This is a standard use of a Fejer kernel. Let the Fourier series of \( F(s) \) be

\[(3.23) \quad F(s) = \sum_{k=-\infty}^{\infty} c_k e^{iks} , \]

where the \( c_k \) are scalars. Since \( F \) is real and even, \( c_k = c_{-k} \) and \( c_k \) is real. The series can be approximated uniformly by

\[(3.24) \quad \sum_{k=-S}^{S} c_k e^{iks} \]

and the latter in turn can be written as

\[(3.25) \quad \left( \sum_{a=0}^{S} g_a^* e^{-iax} \right) \left( \sum_{a=0}^{S} g_a e^{iax} \right) . \]

\[\text{14 It simplifies notation to define } A^*(e^{ix}) \text{ as the conjugate transpose of } A(e^{ix}). \text{ The conjugate transpose of } A(z) \text{ is obtained by transposing the polynomial elements of } A(z) \text{ and substituting } \bar{z} \text{ for } z.\]
To obtain the representation given by (3.25) in which the \( g_u \) are complex scalars, it is necessary to factor (3.24) as a reciprocal equation on the unit circle as shown by (3.25). In more detail write (3.24) as follows:

\[
(3.26) \quad \sum_{\gamma} c_\gamma z^\gamma.
\]

If \( r \) is real and \( z \rightarrow r \) is a factor of (3.26), then \( z^{-1} - r \) is also a factor. If \( r \) is complex and \( z \rightarrow r \) is a factor, then so are \( z \rightarrow \bar{r}, z^{-1} - r, \) and \( z^{-1} - \bar{r} \). This verifies the representation (3.25). Now choose \( d_u = g_u \eta e^{iz} \). This gives the desired result that (3.22) is arbitrarily close to a positive quantity.

The theorem gives necessary and sufficient conditions for the finite sum (3.15) to be positive. The maximization, however, is taken over an infinite sequence. Hence it is necessary to obtain conditions that ensure the positivity of

\[
(3.27) \quad \beta^{\nu/2} d_u \beta^{\nu/2} A_u - \beta^{\nu/2} d_v + \Sigma d_u^* A d_u
\]

For \( \beta = 1 \), (3.27) is the same as (3.13). The argument used to prove Lemma 3 applies to the multi-product case and we state

**Corollary:** The form (3.27) is positive for all nonzero \( \beta \)-sequences \( \{d_u\} \) if and only if (3.16) is positive for all nonzero vectors \( \eta \).

It is possible to extend Theorem 2 to the multi-product case.

**Theorem 11:** The form (3.13) is positive for all \( \beta \)-sequences \( \{d_u\} \) and \( 0 < \beta \leq 1 \) if and only if \( W(x) \) defined in (3.17) is positive definite for \( -\pi \leq x \leq \pi \).

**Proof:** First, notice that for any vector \( \eta \), the scalar inner product

\[
(3.28) \quad \eta^* [A(z) + A^*(z)] \eta = 2 \Re \eta^* A(z) \eta.
\]

Therefore, the inner product is a harmonic function of \( x \).

To prove sufficiency we assume that \( W(x) \) is positive definite for all \( x \). Therefore, (3.28) is positive on the unit circle and being a harmonic function, attains its extrema on the boundary of the closed unit disk. Hence (3.18) is positive for all \( |z| \leq 1 \). Since any point inside the unit circle can be represented by \( \beta^j e^{ix} \), this completes the proof of necessity. Conversely, if (3.13) is positive for all \( 0 < \beta \leq 1 \), then Theorem 10 implies that (3.16) must hold.

Part of Theorem 3 also generalizes to vectors.

**Theorem 12:** \( W(x) = A(e^{ix}) + A^*(e^{ix}) \) is positive definite for \( -\pi \leq x \leq \pi \) if and only if \( A(z) + A^*(z^{-1}) \) is nonsingular for all \( |z| = 1 \) and

\[
(3.29) \quad \eta^* [A(0) + A'(0)] \eta > 0
\]

for all vectors \( \eta \neq 0 \).

**Proof:** Evidently, \( A(z) + A'(z^{-1}) \) and \( W(x) \) are the same for all \( |z| = 1 \). If \( W(x) \)
is positive definite, then it must be nonsingular so that the reciprocal equation
\[ \det [A(z) + A'(z^{-1})] = 0 \]
cannot have any roots on the unit circle. Since \( \eta^* [A(z) + A^*(z)] \eta \) is a harmonic function in the closed unit disk and positive on the boundary, it is also positive in the interior. Thus (3.29) is true.

Conversely, assume that (3.30) has no roots on the unit circle and that (3.29) holds. We must show that for all \( x \) and \( \eta \neq 0 \), \( \eta^* W(x) \eta > 0 \). Suppose the contrary; that is for some \( z_0 \) with \( |z_0| = 1 \)
\[ \eta^* [A(z_0) + A'(z_0^{-1})] \eta \leq 0. \]
Now for \( |z| = 1 \), (3.31) has the same values as (3.28). Since (3.28) is harmonic it must take a value for \( |z| = 1 \) at least as large as its value for \( z = 0 \). Thus it must be positive for some \( z_1 \) with \( |z_1| = 1 \). Thus (3.31) is positive for \( z = z_1 \) and negative for \( z = z_0 \) and must vanish for some \( z \) with \( |z| = 1 \). This gives a contradiction since (3.30) has no roots on the unit circle.

**Corollary:** If \( W(x) = A(e^{ix}) + A^*(e^{ix}) \) is positive definite for all \( -\pi \leq x \leq \pi \) then the matrix operator \( A(L) \) is stable.

**Proof:** The matrix operator \( A(L) \) is stable if and only if the matrix polynomial \( A(z) \) is nonsingular for all \( |z| \leq 1 \). Since the positive definiteness of \( W(x) \) implies \( \Re \eta^* A(z) \eta > 0 \) for all \( |z| \leq 1 \), it follows that the only solution of \( A(z) \eta = 0 \) for all \( |z| \leq 1 \) is the trivial one \( \eta = 0 \).

Theorem 12 gives relatively simple criteria for deciding the positive definiteness of \( W(x) \). These criteria are necessary and sufficient for maximum present value of revenue if \( p_t \) and \( q_t \) are allowed to be complex. Hence they are sufficient but not necessary for real sequences of prices and quantities. Of course, only the real sequences are economically meaningful. Therefore, it is desirable to give a necessary condition for real sequences. Formally since \( W(x) \) is a hermitian matrix, \( \eta^* W(x) \eta > 0 \) for all \( \eta \neq 0 \) implies \( \eta^* W(x) \eta > 0 \) for all \( \eta \neq 0 \). The converse, however, is false.

Obviously, if \( \eta^* W(x) \eta > 0 \) for all real \( \eta \neq 0 \) and \( W(x) = W'(x) \) so that \( W(x) \) is a real symmetric matrix then the criteria of Theorem 12 are both necessary and sufficient for maximum present value and real sequences of prices and quantities. In other words if \( W \) is symmetric then every element of \( W \) is real for all \( x \). Hence positivity of the quadratic form for real \( \eta \) is necessary and sufficient for complex \( \eta \). 

The assumption of symmetry has strong economic consequences which we do not wish to accept. Hence we do not pursue the matter. In general, finding weaker necessary conditions for real sequences involves a complicated comparison between the symmetric and the skew symmetric parts of \( W(x) \).
In Theorem 3 it was shown that the reciprocal polynomial could be factored into the product of two polynomials such that one of the factors has no roots inside the unit circle and the other factor's roots are the reciprocals of the roots of the first factor. It is considerably more difficult to factor the matrix \( A(z) + A(z^{-1}) \). Moreover, the problem of factoring \( W(x) \) is of considerable importance in function theory and is the subject of considerable current research by mathematicians. The existence of factorizations for positive semi-definite \( W(x) \) has been shown in two papers, the first by Wiener and Masani and the second by Helson and Lowdenslager.\(^{15}\) However, if \( W(x) \) is not of full rank, so that \( \det W(x) = 0 \) identically, then the existence of factors has been shown only for the special case of \( k = 2 \) by Masani and Wiener [13]. For general \( W(x) \) not of full rank, it is not known whether factors exist. Even if \( W(x) \) is positive semi-definite so that factors exist, algorithms for calculating the factors are not known. As far as we know, it is possible to calculate the factors of \( W(x) \) only when \( A(e^{ix}) \) is a polynomial matrix and \( W(x) \) is positive definite and not merely positive semi-definite. In our factorization we follow a suggestion of Whittle.\(^{16}\)

**Lemma 7:** Let \( W(x) = A(e^{ix}) + A^*(e^{ix}) \) where \( A(e^{ix}) = A_0 + A_1 e^{ix} + \ldots + A_n e^{inx} \) and assume that \( W(x) \) is positive definite for all \( -\pi \leq x \leq \pi \). If \( W(x) \) can be factored such that \( W(x) = B(e^{ix})B^*(e^{ix}) \) where \( B(z) \) is a matrix polynomial then (i) \( B(z) \) is nonsingular for all \( |z| \leq 1 \) (ii) \( A(z) \) and \( B(z) \) must be of the same degree and the matrix coefficients of \( B(z) \) must be real.\(^{17}\)

**Proof:** Again we exploit the fact that \( e^{ix} [A(z) + A^*(z)] e^{-ix} \) is a harmonic function which must be positive on the closed unit disk because the positive definiteness of \( W(x) \) implies it is positive on the boundary. Therefore, if a factorization by matrix polynomials exists then \( e^{ix} B(z) B^*(z) e^{-ix} \) is also a harmonic function for all \( |z| \leq 1 \) and \( B(z) \) must be nonsingular for all \( |z| \leq 1 \). In particular, \( \det B(0) = \det B \neq 0 \).

To prove the second assertion, suppose that the degree of \( B(z) \) is \( m \) and that \( m > n \). Since

\[
A(e^{ix}) + A^*(e^{ix}) = B(e^{ix}) B^*(e^{ix}),
\]

\(^{15}\) See Wiener and Masani [17] and Helson and Lowdenslager [9]. There are also pertinent articles in Russian which we have not examined. A non-technical summary of the current status of the problem by Lowdenslager is given in the Appendix to Yaglom [18]. For a technical exposition of the most recent developments see Helson [8].

\(^{16}\) Whittle suggests the use of the Yule-Walker relations to calculate the factors (Whittle, 15, p. 101). However, he gives no precise statement of the conditions which would allow the use of these relations and does not prove the validity of the algorithm. With a charitable interpretation, his statement at the bottom of p. 100 is obscure. The Yule-Walker relations are our equations (42) and are well known in regression theory.

\(^{17}\) Although our Lemma 8 suffices for our purposes, a stronger result is given in Theorem 20, Helson [8].
we can multiply out the right side and equate the matrix coefficients of like powers of \( e^{ix} \). In particular, \( B_0B_0^* = 0 \) since \( A(e^{ix}) \) is of degree \( n \). Hence \( B_0^* = 0 \), since \( B_0 \) is nonsingular. Similarly,

\[
B_0^* B_0^* = B_0^* = \ldots = B_n^* = 0
\]

which proves that \( B(z) \) must be of degree \( n \) as asserted. That the \( B \)'s are real follows from the maintained hypothesis that the matrix coefficients of \( A(z) \) are real.

**Lemma 8:** If \( W(x) \) is positive definite then

\[
(A(z) + A'(z^{-1}))^{-1} = \sum_{\omega} \frac{G_{-\omega} z^{-\omega}}{G_0 + \sum_{\omega} G_\omega z^\omega}
\]

is the uniquely defined Laurent expansion for the annulus \( |r| < |z| < |r^{-1}| \), \( |r| < 1 \) where \( |r| \) is the largest modulus of the roots of the reciprocal equation

\[
\det [A(z) + A'(z^{-1})] = 0
\]

that are inside the unit circle, and the \( G_\omega \)'s are certain linear combinations of \( r_j^\omega \) which denote the distinct roots of (3.33) inside the unit circle. In addition \( G_u \) is real and \( G_{-u} = G_{-\omega} \).

**Proof:** The positive definiteness of \( W(x) \) and Theorem 12 imply that the reciprocal equation (3.33) has no roots on the unit circle. Hence the inverse of \( A(z) + A'(z^{-1}) \) has a representation in an annulus including the unit circle as follows:

\[
(A(z) + A'(z^{-1}))^{-1} = \text{adj} [A(z) + A'(z^{-1})] / \det [A(z) + A'(z^{-1})]
\]

\[
= \frac{\text{adj} [A(z) + A'(z^{-1})]}{q_0 \prod (1 - r_j)(1 - r_j z^{-1})}
\]

and \( |r_j| < 1 \). The denominator of (3.34) has a partial fraction expansion in terms of \( (1 - r_j z) \) and \( (1 - r_j z^{-1}) \) (and their powers in case of repeated roots) while the numerator is a matrix polynomial in powers of \( z \) and \( z^{-1} \). The representation of (3.32) is an immediate consequence of the expansion of (3.34) by means of partial fractions. The matrix coefficients are uniquely determined for the annulus

\[
|r_j| < |z| < |r_j^{-1}|
\]

where \( |r_j| < |r_j| < 1 \). It is clear from the construction of the Laurent series that \( G_u \) is a certain linear combination of all \( r_j^\omega \). Since \( A_u \) is real, it implies \( G_u \) is real. To prove that

\[
G_u = G_{-u}
\]

take transposes in (3.32) and substitute \( z^{-1} \) for \( z \). Since the result is an identity in \( z \) and \( z^{-1} \) by the uniqueness of the Laurent series representation, the matrix coefficients of like powers of \( z \) must be identical and this yields (3.36).

By the method with which the \( G_u \) can be constructed, it follows that the matrix coefficients of the Laurent series obey a certain recurrence relation (see (2) in the
Appendix). Thus for $t \geq 0$,
\begin{equation}
(A(L) + A'(E))G_t = \Delta_{ot}
\end{equation}
where $\Delta_{ot}$ is a matrix generalization of the Kronecker delta so that
\[
\Delta_{ot} = \begin{cases} I & t=0, \\ 0 & \text{otherwise}. \end{cases}
\]
In brief, $G_t$ is the bounded solution of
\begin{equation}
(A(L) + A'(E))G_t = 0, \quad t \geq n,
\end{equation}
subject to the initial conditions
\begin{equation}
(A(L) + A'(E))G_t = \Delta_{ot}, \quad 0 \leq t < n,
\end{equation}
where
\[
L'G_{t-u} = G_{t-u}' = G'_{t-u}, \quad u \geq 0.
\]
The problem of factoring the optimal operator reduces to that of calculating the difference equation of smallest degree which will generate the matrix coefficients of the Laurent expansion. Recalling that $E = L^{-1}$, we see that (3.38) is a difference equation of degree equal to $2n$, and $G_t$ is its bounded solution so that it is a certain linear combination of the roots of the characteristic polynomial which are inside the unit circle. Since there are $n$ initial conditions as shown in (3.39), we seek the $n$th degree difference equation which meets these given initial conditions. Provided $W(x)$ is positive definite, it is possible to give a constructive proof of the factorability of $A(z) + A'(z^{-1})$.

**Theorem 13:** Assume that $A(z) + A'(z^{-1})$ is nonsingular for $|z| = 1$. First, there is a polynomial matrix $B(z)$ with real matrix coefficients
\begin{equation}
B(z) = B_0 + B_1 z + \ldots + B_n z^n, \quad \det B_0 \neq 0,
\end{equation}
such that
\begin{equation}
(A(z) + A'(z^{-1}))^{-1} B(z)
\end{equation}
contains no positive powers of $z$ if and only if $C_1, \ldots, C_n$ satisfy the following equations
\begin{equation}
\begin{align*}
G_1 + (G_0 + G_0') C_1 + \ldots + G_{n+1} C_n &= 0, \\
G_2 + G_1 C_1 + (G_0 + G_0') C_2 + \ldots + G_{n+2} C_n &= 0, \\
& \quad \vdots \\
G_n + G_{n-1} C_1 + \ldots + (G_0 + G_0') C_n &= 0, \\
G_{t+n} + G_{t+n-1} C_1 + \ldots + G_t C_n &= 0, \quad t > 0,
\end{align*}
\end{equation}
where
\begin{equation}
C_j = B_j B_0^{-1}.
\end{equation}
Second, $B_0$ is determined up to multiplication by an arbitrary orthogonal matrix; $C_1, \ldots, C_n$ is the unique solution of (3.42), and
(3.45) \[ A(z) + A'(z^{-1}) = B(z)B'(z^{-1}) \]

where \( B(z) \) is nonsingular for all \( |z| \leq 1 \).

**Proof:** By Lemma 8, the hypothesis that \([A(z) + A'(z^{-1})]\) is nonsingular for \(|z| = 1\), assures the absolute convergence of the uniquely determined Laurent series in the annulus defined in (3.35). Clearly, (3.42) through (3.44) hold if and only if there is an \( n \)th degree polynomial matrix such that

(3.46) \[ [A(z) + A'(z^{-1})]^{-1} B(z) = D(z^{-1}) \]

where \( D(z^{-1}) \) is a power series in nonpositive powers of \( z \). Therefore, it is only required to show there is a nontrivial set of \( C_1, \ldots, C_n \) that will satisfy (3.42) and (3.43).

By the method of constructing the matrix coefficients of the Laurent series (3.32), it is clear that the power series

(3.47) \[ G(z) = G_0 + G_1 z + \cdots + G_{\omega_1} z^{m} \]

is a recurrent series which is absolutely convergent in the disk \( |z| < |r_1^{-1}| \). Therefore, by Lemma 1 (in the Appendix), \( G(z) \) is a rational matrix function of the scalar \( z \) and there are polynomial matrices

(3.48) \[ C(z) = I + C_1 z + \cdots + C_n z^n \]

(3.49) \[ P(z) = P_0 + P_1 z + \cdots + P_{n-1} z^{n-1} \]

such that

(3.50) \[ G(z) C(z) = P(z) \]

It only remains to choose the matrix coefficients of \( P(z) \) to satisfy (3.42) and (3.43). The following choice of \( P_n \) accomplishes the purpose:

\[
\begin{align*}
P_0 &= -[G_{-1} C_1 + G_{-2} C_2 + \cdots + G_{-n} C_n], \\
P_1 &= -[G_{-1} C_2 + \cdots + G_{-n} C_n], \\
P_{n-1} &= -[G_{-1} C_n].
\end{align*}
\]

As remarked above, the degrees of the polynomials \( P(z) \) and \( C(z) \) are fixed by the requirement that \( n \) initial conditions must be satisfied.

The hypothesis that \([A(z) + A'(z^{-1})]\) is nonsingular on the unit circle permits the application of Theorem 12 and Theorem 10. The equation, \([G_{u-v}] \) \((u, v = 0, 1, \ldots, n-1)\), is positive definite so that there is a unique set of nontrivial real \( C_1, \ldots, C_n \) which satisfies (3.42).

We have shown that \( D(z^{-1}) \) defined in (3.46) contains no positive powers of \( z \). Take transposes in (3.46) and substitute \( z^{-1} \) for \( z \). This yields

(3.51) \[ B'(z^{-1})[A'(z^{-1}) + A(z)] = D'(z) \]

Multiply (3.46) on the left by \( B'(z^{-1}) \) and (3.51) on the right by \( B(z) \) and obtain

(3.52) \[ B'(z^{-1}) D(z^{-1}) = D'(z) B(z) \]
which is an identity in z. Since the right side has no negative powers of z and the left side no positive powers of z,

\[(3.53) \quad D'(z)B(z) = K\]

where \(K\) is a nonsingular matrix of scalars. Therefore,

\[(3.54) \quad D(z^{-1}) = B'(z^{-1})^{-1}K,\]

and substituting into (3.46) this yields,

\[A(z) + A'(z^{-1}) = B(z)K^{-1}B'(z^{-1}).\]

Clearly, \(K\) must be positive definite but is otherwise arbitrary. It follows that \(K\) can be factored by classical algebraic tools into the product of a matrix and its transpose. Therefore, (3.55) is equivalent to the desired representation (3.45).

Since \(B(z)\) is a polynomial factor, Lemma 7 applies and shows that \(B(z)\) must be nonsingular on the closed unit disk. This is evident directly from the fact that (3.43) is a stable difference equation. Write

\[\left[A(z) + A'(z^{-1})\right]^{-1}B(z) = B'(z^{-1})^{-1}\]

and equate the coefficients of the constant term. This gives

\[(G_0 + G'_0)B_0 + G_{-1}B_1 + G_{-2}B_2 + \ldots + G_{-n}B_n = B'_0\]

which reduces to

\[(3.56) \quad G_0 + G'_0 + G_{-1}C_1 + \ldots + G_{-n}C_n = (B_0B'_0)^{-1}.\]

Since the \(G\)'s are the uniquely determined coefficients of the Laurent expansion and the \(C\)'s are the unique solutions of (3.42), the matrix product \(B_0B'_0\) is uniquely determined. Let \(M\) be any orthogonal matrix. Then, \(B_0B'_0 = B_0MM'B_0\) since \(MM' = I\). Therefore, \(B_0\) can be any nonsingular matrix and is fixed only up to multiplication by an arbitrary orthogonal matrix.

We now reintroduce \(\beta\) and establish the factorization pertinent to the problem. The first corollary corresponds to Theorem 4.

**Corollary 1:** Let \(\{f_i\}\) and \(A(L)\) be given. Let \(\beta_f\) be the largest value for which \(\beta < \beta_f\) implies that each component of \(f_i\) is a \(\beta\)-sequence. Let \(\beta_A\) be the largest value for which \(\beta < \beta_A\) implies that \(Re \eta^\ast A(\beta z)\eta > 0\) for all \(|z| = 1\) and all nonzero vectors \(\eta\). Set \(\beta_q = \min(\beta_f, \beta_A, 1)\). Then for \(\beta_0 < \beta_q\) the sequence \(\{q_t\}\) which satisfies (3.9) is a \(\beta\)-sequence with \(\beta = \beta_0\) and maximizes \(PV\).

**Proof:** The main idea of the proof is like that of Theorem 4. The formal substitutions,

\[F_t = \beta^tf_t, \quad Q_t = \beta^tq_t,\]

allow (3.9) to be written as

\[(3.57) \quad [A(\beta^tL) + A'(\beta^tE)]Q_t = F_t.\]
Since (3.10) is satisfied (because $\beta_0 < \beta_4$) it suffices to show that the solution to (3.57) is a $\beta$-sequence for $\beta = \beta_0$. As in Theorem 4, there is a $\delta < 1$ such that each component of $F_t$ is bounded by $M\delta^t$. To show that similar inequality holds for the components of $Q_t$, we should like to use a sequential solution as before. However, there is the (apparent) problem that matrix operators do not commute. This problem is overcome by using the operator version of equation (3.34). The equation must be modified to account for the way we have introduced $\beta$. It becomes

$$\begin{align*}
A(\beta^t w^{-1}) + A'(\beta^t w)^{-1} &= \text{adj} [A(\beta^t w^{-1})] + A'(\beta^t w)/a_0 \Pi (1-r_j w)(1-r_j w^{-1}),
\end{align*}$$

where

$$\det [A(\beta^t w^{-1}) + A'(\beta^t w)] = a_0 \Pi (1-r_j w)(1-r_j w^{-1}).$$

In the operator form (3.57) may be written

$$\Pi (I-r_j E)(I-r_j L)Q_t = \text{adj} [A(\beta^t L) + A'(\beta^t E)] (F_t/a_0).$$

Now the operator which acts on $F_t$ is complicated, but it consists of a finite number of matrix multiplications and shifts. Thus each component of the resultant sequence is bounded by $M_1 \delta^t$ for some sufficiently large $M_1$ (not depending on $t$). But then the argument used in Theorem 4 applies provided, as there, $\delta$ is chosen larger than $|r_j|$. This argument showed that the sequence resulting from each application of $(I-r_j E)^{-1}$ or $(I-r_j L)^{-1}$ gave a sequence bounded by $M\delta^t$ for some $M$. This concludes the proof.

**Corollary 2:** Let $\beta_4$ be defined as in Corollary 1 and $\beta$ be given with $0 < \beta < \beta_4$. The optimal matrix operator can be factored so that

$$A(L) + A'(\beta E) = B'(\beta E) B(L)$$

where $B(\beta^t L)$ is stable and of degree $n$ if and only if

$$W(\beta^t x) = A(\beta^t e^{ix}) + A^*(\beta^t e^{ix})$$

is positive definite.

**Proof:** Suppose $W(\beta^t x)$ is positive definite. Consider the matrix polynomial

$$A'(\beta z) + A(z^{-1})$$

which corresponds to the optimal operator $A'(\beta E) + A(L)$ after the substitution of $z$ for $E$ and $z^{-1}$ for $L$. Let $z = w^\beta$ so that (3.61) can be written as a reciprocal polynomial matrix\(^{18}\)

$$A'(\beta^t w) + A(\beta^t w^{-1}).$$

Since $W(\beta^t x)$ is positive definite it follows that (3.62) is positive definite for $|w| = 1$. In particular, for given $\eta \neq 0$,

$$\eta^* [A(\beta^t w) + A^*(\beta^t w)] \eta$$

\(^{18}\) In our problem we wish to factor $A'(\beta) + A(z^{-1})$ whereas Theorem 12 refers to $A(\beta) + A'(z^{-1})$. Clearly, the first is factorable if and only if the second is.
is positive for \(|w| = 1\), so (3.62) is nonsingular for \(|w| = 1\). Equation (3.45) of Theorem 13 yields the representation

\[(3.63) \quad A(\beta^t w) + A'(\beta^t w^{-1}) = B'(\beta^t w) B(\beta^t w^{-1}) \]

where \(B'(\beta^t w)\) is of degree \(n\) and is nonsingular for \(|w| \leq 1\). The substitution \(z = w/\beta^t\) yields

\[(3.64) \quad A'(\beta z) + A(z^{-1}) = B'(\beta z) B(z^{-1}) \cdot \]

By the substitution of \(E\) for \(z\) and \(L\) for \(z^{-1}\) (3.64) gives the desired representation (3.60). The fact that \(B(\beta^t L)\) is stable follows because \(B'(\beta^t w)\) is nonsingular for \(|w| \leq 1\).

Conversely if the factorization (3.60) holds, then the substitution \(z = w/\beta^t\) permits the representation given by (3.63). Since \(B(\beta^t L)\) is stable, the right side of (3.63) is nonsingular for \(|w| = 1\). Since it is also self-adjoint, it is positive definite for \(|w| = 1\). This is the desired statement about \(W(\beta^t x)\) and concludes the proof.

These two corollaries mean that there is always a \(\beta\)-sequence \(\{q_t\}\) that satisfies arbitrary initial conditions and gives a finite maximum \(PV\) for all positive values of \(\beta\) less than \(\min \{\beta_\ell, \beta_\delta, 1\}\). The optimal output path must satisfy

\[(3.65) \quad B(L) q_t = B'(\beta E)^{-1} f_t \cdot \]

In general \(B(L)\) need not be stable; nevertheless, it is possible to calculate the optimal output path recursively by means of (3.65). The characteristic polynomial of \(B(L)\) is \(\det B(z^{-1}) = 0\). Consider

\[(3.66) \quad z^n \det B(z^{-1}) = 0 \cdot \]

Since \(B_0\) is nonsingular, Theorem 2 (of the Appendix) applies and the degree of (3.66) is \(kn\). However, some of the roots of (3.66) can be zero. There are zero roots of (3.66) if and only if \(B_n\) is singular. Moreover, \(B_n\) is singular if and only if \(A_n\) is singular. Therefore, singularity of \(A_n\) implies that

\[(3.67) \quad \det B(z) = 0 \cdot \]

is of lower degree than (3.66) and has fewer roots. This possibility is mentioned in the analysis of matrix difference equations given in the Appendix. In general, \(A_n\) can be expected to be singular, which would mean that the components of \(q_t\) obey difference equations of diverse degrees. For example, it is possible for some components of \(q_t\) to satisfy trivial difference equation of zero degree, that is, to respond to price without any lag whatever.

Finally, there is the problem of nonnegativity of \(q_t\). Although it is more complicated for the multi-product case, given the hypothesis of unconditional nonnegativity of the matrix operator, many of the results for scalar difference equations apply to vectors. In particular, the definition of unconditional nonnegativity becomes:

**DEFINITION:** The matrix operator \(A(L)\) is unconditionally nonnegative provided there is a \(\lambda > 0\) such that for any bounded nonnegative sequence \(\{\eta_t\}\), \(\eta_t = 0\) for \(t < 0\) and \(\xi_t = 0\) for \(-n \leq t < -1\), \(A(L) \xi_t = \eta_t\) has a bounded nonnegative solution \(\xi_t\).
Lemma 5 generalizes to matrix operators. Thus $A(L)$ is u.n.n. if and only if the recurrent matrix power series

$$(3.68) \quad A(\lambda z)^{-1} = \sum D_i \lambda^i z^i$$

has nonnegative coefficients. Therefore, every element of every $D_i$ is nonnegative. Since

$$A(\lambda z)^{-1} = \text{adj} A(\lambda z)/\text{det} A(\lambda z),$$

$A(L)$ is u.n.n. if and only if every element of $A(\lambda z)^{-1}$ is a positive constant times a characteristic function. Since the denominator of every element of $A(\lambda z)^{-1}$ is the same, namely $\text{det} A(\lambda z)$, Theorem 5 applies to the rational functions of $A(\lambda z)^{-1}$ which are the elements of $A(\lambda z)^{-1}$. Thus every one of the $k^2$ elements of $A(\lambda z)^{-1}$ has the properties of a characteristic function. In particular, the smallest root of $\text{det} A(z) = 0$ must be positive and every element of $A_1$ must be negative. There are two sufficient conditions for u.n.n. matrix operators. If $A_1, \ldots, A_k$ have nonpositive elements or if all of the roots of $\text{det} A(z) = 0$ are positive then $A(L)$ is u.n.n.

A variant of Theorem 6 carries over for vectors. Thus if $A(L)$ is u.n.n. and $A(z) + A'(z^{-1})$ is nonsingular on the circle, then every element of

$$A(e^{i\gamma})^{-1} + A^*(e^{i\gamma})^{-1} = A(e^{i\gamma})^{-1} \left[ A^*(e^{i\gamma}) + A(e^{i\gamma}) \right] A^*(e^{i\gamma})^{-1} = A(e^{i\gamma})^{-1} B^*(e^{i\gamma}) B(e^{i\gamma}) A^*(e^{i\gamma})^{-1}$$

is a characteristic function and is subject to certain bounds in modulus similar to those described in Theorem 6.

It is remarkable that even the stabilization property of Theorem 8 carries over to vectors. This fact is worth a formal statement.

**THEOREM 14:** Assume that $A(z) + A'(z^{-1})$ is nonsingular on the unit circle so that

$$A(z) + A'(z^{-1}) = B(z) B(z^{-1}).$$

Let $s_1$ denote the largest root of $\text{det} B(z^{-1}) = 0$. If $s_1$ is real then there is a real root, $r_1$, of $\text{det} A(z^{-1}) = 0$ such that $|s_1| < |r_1| < 1$.

**Proof:** Since $s_1$ is real, there is a real vector $\eta \neq 0$ such that

$$[A(s_1) + A'(s_1^{-1})] \eta = 0.$$ 

Therefore,

$$\eta' [A(s_1) + A'(s_1^{-1})] \eta = 0.$$ 

By the positive definiteness of $A(z) + A^*(z)$ for all $|z| < 1,$

$$\eta' [A(s_1) + A'(s_1^{-1})] \eta > 0$$

so that $\eta' A(s_1) \eta > 0$. Therefore, $\eta' A'(s_1^{-1}) \eta < 0$ and $\eta' A(s_1^{-1}) \eta < 0$. Thus there is a point $r_1^{-1}$ on the line joining $s_1$ and $s_1^{-1}$ with $\eta' A(r_1^{-1}) \eta = 0$. Now $|r_1^{-1}| > 1$ since $\eta' A(r) \eta > 0$ for $|r| < 1$. Thus $r_1^{-1}$ is a singularity of $A(z)$ and $1 < |r_1^{-1}| < |s_1^{-1}|$. This proves the theorem. Therefore, if the optimal operator is u.n.n. so that its largest
root inside the unit circle must be positive, then it is asymptotically stabilizing relative to the demand operator.

*University of Chicago*

**APPENDIX**

**MATRIX DIFFERENCE EQUATIONS**

A matrix power series is an expression of the form

\[(A.1) \quad M(z) = \sum M_k z^k\]

where each of the \(k^2\) elements converges in some disk. The disk of convergence of \(M(z)\) is determined by the smallest disk of convergence of each of the \(k^2\) elements of \(M(z)\).

**DEFINITION:** A recurrent matrix power series of the form (A.1) is a power series whose matrix coefficients satisfy a recurrence relation of the form

\[(A.2) \quad M_{i+n} + R_1 M_{i+n-1} + \ldots + R_n M_i = 0, \quad t > 0.\]

Although systems of difference equations have much in common with single difference equations, they present certain complications which deserve our attention. The solution of a system of difference equations is handled most expeditiously by means of rational matrix functions. Let

\[(A.3) \quad P(z) = P_0 + P_1 z + \ldots + P_{n-1} z^{n-1}.\]

**DEFINITION:** \(Y(z)\) is said to be a proper rational matrix function of the complex scalar \(z\) if it can be represented as follows:

\[(A.4) \quad Y(z) = P(z) A(z)^{-1}\]

provided both \(A_0\) and \(P_0\) are nonsingular and \(A(z)\) and \(P(z)\) are matrix polynomials of degree \(n\) and at most \(n - 1\). A rational matrix function is the sum of a polynomial matrix and a proper rational matrix function.

We now prove that every rational matrix can be represented as a recurrent power series.

**LEMMA 1:** A necessary and sufficient condition that a matrix power series in \(z\) should be a recurrent matrix power series is that it should be the expansion of a proper rational matrix function of \(z\).

**PROOF:** For proof of sufficiency,

\[A(z)^{-1} = \text{adj } A(z)/\det A(z) = \text{adj } A(z)/a_0 I_1 (1 - r_1 z),\]

where \(\text{adj } A(z)\) means the adjoint of \(A(z)\) and \(r_1^{-1}\) denotes the \(j\)th root of \(\det A(z) = 0\). The adjoint of \(A(z)\) is an ordinary matrix polynomial and \(\det A(z)\) is a scalar polynomial. Therefore, \(A(z)^{-1}\) can be represented in a matrix power series by means of a partial fraction expansion of the denominator \(\det A(z)\). The power series is absolutely convergent in the disk \(|z| < |r_1^{-1}| \leq |r_j^{-1}|\).

\[^{19}\text{See Frazer, Duncan and Collar [4] for a discussion of some properties of polynomial matrices.}\]

\[^{20}\text{The definition of a recurrent matrix power series and Lemma 7 generalize Hardy's results for a scalar power series (Hardy [7 pp. 392–393]).}\]
Hence \( P(z)A(z)^{-1} \) is also a power series which converges in the same disk as \( A(z)^{-1} \) and
\[
Y(z) = \sum Y_t z^t = P(z)A(z)^{-1}
\]
is a formal identity in \( z \). Form
\[
P(z) = Y(z)A(z),
\]
equate the matrix coefficients of equal powers of \( z \), and obtain
\[
Y_0 A_0 = P_0,
\]
\[
Y_1 A_0 + Y_0 A_1 = P_1,
\]
\[
\vdots
\]
\[
Y_{n-1} A_0 + \ldots + Y_0 A_{n-1} = P_{n-1},
\]
\[
Y_{1+n} A_0 + \ldots + Y_t A_n = 0, \quad t > 0.
\]
This proves that \( Y(z) \) is a recurrent power series.

Conversely, if the coefficients of the power series \( Y(z) = \sum Y_t z^t \) satisfy a recurrence relation of the form (A.8), then form the product \( Y(z)A(z) \) and observe that all of the terms of degree \( t > n \) vanish by virtue of (A.8) so that \( Y(z) \) has a representation (A.6) and is, therefore, a proper rational matrix function.

The system (A.8) is a difference equation in the matrices \( Y_t \). Hence it includes as special cases difference equations in vectors. We may now prove

**Theorem 1:** If \( A(z) \) is nonsingular for all \( |z| \leq 1 \) then
\[
A(L)Y_{1n} = 0, \quad t > 0,
\]
has a non-trivial bounded solution where \( Y_t \) is a linear combination of the powers of the distinct roots of \( \det A(z^{-1}) = 0 \).

**Proof:** By hypothesis, the matrix power series \( Y(z) = \sum Y_t z^t \) obeys a recurrence relation so that there is a representation of \( Y(z) \)
\[
Y(z) = P(z)A(z)^{-1} = P(z)\text{adj}A(z)/a_0 \Pi(1 - r_j z).
\]
A partial fraction expansion of the right side shows that \( Y_t \) is a certain linear combination of powers of \( r_j \) whose form is more or less complicated depending on whether the roots are repeated or simple. The solution can be bounded only if \( |r_j| < 1 \) for all \( j \), so that \( Y(z) \) converges in the disk \( |z| < |r_j^{-1}| \leq |r_j|^{-1} \) which includes the unit circle.

By analogy with scalar difference equations, a homogeneous matrix difference equation which has a bounded non-trivial solution, meeting suitable initial conditions, is said to be stable. Hence the hypothesis of the preceding theorem could have required \( A(L) \) to be a stable matrix operator in place of \( A(z) \) being nonsingular on the closed unit disk. The general difference equation
\[
(A.9) \quad A(L) Y_t = M_t,
\]
where \( M_t \) is a sequence of matrices which has a solution given by
\[
(Y_t = A(L)^{-1} M_t).
\]
The simplest computational procedure is to use the recursions displayed in (A.7) and (A.8).

For the development in part 3 we need certain additional facts about matrix polynomials. First, the polynomial \( \det A(z) = 0 \) has a root \( z = 0 \) if and only if \( A_0 \) is singular. This is clear since \( \det A(0) = \det A_0 = 0 \) and the second equality holds if and only if \( A_0 \) is singular.

The second fact we state and prove as
THEOREM 2: A necessary and sufficient condition for a $k$-order $n$th degree polynomial matrix to have $kn$ singularities is that the matrix coefficient $A_n$ be nonsingular.\footnote{1}

PROOF: To prove sufficiency, assume $A_n$ is nonsingular so that

$$A_n^{-1}A(x) = M_0 + M_1x + \ldots + M_{n-1}x^{n-1} + Ix^n.$$ 

Therefore,

$$A_n^{-1}A(x) = [g_n(x)]$$

where every element on the diagonal is of degree $n$ and every element off the diagonal is of degree at most $n - 1$. Hence $\det A_n^{-1}A(x) = 0$ is a polynomial equation of degree $kn$ and $A(x)$ has $kn$ singularities.

To prove necessity assume that $\det A(x) = 0$ has $kn$ roots and suppose that $A_n$ is singular of rank $r < k$. There exist nonsingular scalar matrices $P$ and $Q$ such that

$$PA_nQ = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix},$$

where $I_r$ is the $r \times r$ identity matrix. Therefore, $PA(z)Q = [g_n(x)]$ has $r$ polynomials on the diagonal of degree $n$ and $k - r$ on the diagonal of degree at most $n - 1$. All of the off-diagonal polynomials are of degree at most $n - 1$. The equation

$$\det PA(z)Q = \det A(x) = 0$$

is a polynomial of degree at most

$$rn + (k - r)(n - 1) = kn - (k - r) < kn$$

which contradicts the hypothesis that $\det A(x) = 0$ has $kn$ roots.

In the course of proving this theorem it was shown that if the rank of $A_n$ is $r < k$, then the maximum degree of the polynomial $\det A(x)$ equals $kn - (k - r)$. If in addition $A_0$ is nonsingular, then the minimum degree is $kr$. These results can be improved by taking into consideration the ranks of the other matrix coefficients but the complications are considerable and the matter is not worth pursuing. The degree of the $\det A(x)$ gives the number of arbitrary constants which enter the solution of the system of difference equations. It should be noted, however, that (i) $\det A(x) = 0$, and (ii) $\det A(x^{-1}) = 0$ do not necessarily have the same degree. If $A_0$ is nonsingular then (ii) will be of degree $kn$ as shown by Theorem 2. However, (i) will not be of degree $kn$ unless $A_n$ is nonsingular. If $A_n$ is singular then some of the roots of (ii) will equal zero. Therefore, (i) and (ii) will be of the same degree if and only if $A_n$ is nonsingular or, equivalently, all of the roots of (ii) are nonzero.

REFERENCES


\footnote{1} A partial statement and proof of what amounts to Theorem 2 is given in Frazer, Duncan, and Collar [4, pp. 162-163].


