BUREAU OF THE CENSUS

STATISTICAL RESEARCH DIVISION REPORT SERIES

SRD Research Report Number: CENSUS/SRD/RR-82/08

ON BACKSHIFT-OPERATOR POLYNOMIAL TRANSFORMATIONS TO STATIONARITY FOR SINGLE- AND MULTI-COMPONENT TIME SERIES

by

David F. Findley
U.S. Bureau of the Census

Recommended by: Myron J. Katzoff

Report completed: July 15, 1982

Report issued: March 7, 1983

ON BACKSHIFT-OPERATOR POLYNOMIAL TRANSFORMATIONS TO STATIONARITY FOR SINGLE- AND MULTICOMPONENT TIME SERIES

By David F. Findley

U.S. Bureau of the Census

It is demonstrated that any two backshift operator polynomials which transform a given non-stationary time series into stationary series with continuous spectral distributions must have a common divisor which has this property. It follows that the lowest-degree polynomial with this property is unique to within a constant multiple. Using this result some derivations are given, under varying assumptions, of a transformation formula used in non-stationary signal extraction. Counter-examples are presented to show that continuity assumptions on the spectral distribution functions involved are necessary to obtain these results.

1. Introduction. It is a common practice in the analysis of a nonstationary time series to look for a polynomial in the backshift operator B, such as $p(B) = (1-B)(1-B^{12})$, normalized by $p(0)=1$, which transforms the given series into a stationary series (Box and Jenkins, 1976). It appears to have been taken for granted that when such a polynomial can be found, then after superfluous factors (e.g., from overdifferencing) have been eliminated, this polynomial provides the only normalized polynomial transformation to stationarity for the series. As we show, after verifying some calculational formulas, this result can be established quite generally, but not universally. We apply it to validate a transformation formula used in non-stationary signal extraction by Pierce (1979) and Bell (1982).

Except in a discussion in section 4 below, all (covariance) stationary time series which appear in this paper will be assumed to have mean zero.

2. Preliminaries. The proofs of the uniqueness results given in Section 3 require some simple manipulations with backshift-operator polynomial filters and their inverses as they apply to stationary time series. Here we establish the validity of these manipulations.

Let the Cramér representation of the stationary time series x_t be

$$x_t = \int_{-\pi}^{\pi} e^{-it\lambda} dz(\lambda) \quad ,$$

and let B denote the backshift operator, $Bx_t = x_{t-1}$. Recall [Hannan (1970, page 115)] that a function $f(e^{i\lambda})$ defines a linear time-invariant filter $f(B)$ which can be applied to x_t if and only if

$$\int_{-\pi}^{\pi} |f(e^{i\lambda})|^2 dF(\lambda) < \infty \tag{2.1}$$

holds, where $F(\lambda)$ is the spectral distribution of x_t. If (2.1) is satisfied, then the series which results from applying $f(B)$ to x_t is given by

$$f(B)x_t = \int_{-\pi}^{\pi} e^{-i\lambda t} f(e^{i\lambda}) dz(\lambda) \quad .$$

In this case, we say that $f(B)x_t$ is defined. Often, $f(B)$ will be obtained from a rational function

$$f(z) = \tilde{\phi}(z)/\phi(z) \tag{2.2}$$

whose denominator polynomial $\phi(z)$ may have roots on the unit circle $\{|z| = 1\}$. Then the condition (2.1) is equivalent to

$$\int_{-\pi}^{\pi} |\phi(e^{i\lambda})|^{-2} dF(\lambda) < \infty \quad , \tag{2.3}$$

because $|\phi(e^{i\lambda})|$ is bounded and is bounded away from zero near the roots of $\tilde{\phi}(e^{i\lambda})$. Henceforth, it will be convenient to refer to zeros of polynomials such as $\phi(z)$ as roots of $\phi(B)$.

3. The uniqueness of polynomial transformations to stationarity.
Suppose that the nonstationary time series w_t $(t=0,\pm 1,\ldots)$ is such that a backshift operator polynomial $p(B)$ exists having the property that $p(B)w_t$ is a *stationary time series with continuous spectral distribution*. Let C $(=C\{w_t\})$ denote the class of all polynomials $p(B)$ defining such operators. We shall now show that all polynomials in C have a common divisor, which is then necessarily of minimal degree 4 in C. Consequently, the lowest degree polynomial transforming w_t into a stationary time series

with continuous spectral distribution is *unique* up to a constant scalar multiple.

That all members of C have a common divisor follows immediately from

(3.1). *Let $p(B)$ and $\tilde{p}(B)$ be any two polynomials in C and let $\phi(B)$ be their greatest common divisor. Then $\phi(B)$ also belongs to C.*

PROOF. We have $p(B) = \psi(B)\phi(B)$ and $\tilde{p}(B) = \tilde{\psi}(B)\phi(B)$, where the polynomials $\psi(z)$ and $\tilde{\psi}(z)$ have no roots in common. We define the stationary series

$$x_t = \tilde{p}(B)w_t \tag{3.2a}$$

and

$$y_t = p(B)w_t \quad . \tag{3.2b}$$

Then

$$\tilde{\psi}(B)y_t = \psi(B)x_t \quad . \tag{3.3}$$

By hypothesis, the spectral distributions of x_t and y_t are continuous at any λ for which $e^{i\lambda}$ is a zero of $\psi(B)$ or $\tilde{\psi}(B)$, so we can recover x_t and y_t from (3.3) by applying $\psi^{-1}(B)$, respectively, $\tilde{\psi}^{-1}(B)$ thereby obtaining $x_t = \psi^{-1}(B)\tilde{\psi}(B)y_t$ and $y_t = \tilde{\psi}^{-1}(B)\psi(B)x_t$. We now observe from (2.4) that $\tilde{\phi}^{-1}(B)y_t$ and $\phi^{-1}(B)x_t$ are defined, so we have

(3.4a) $$x_t = \phi^{-1}(B)\tilde{\phi}(B)y_t = \tilde{\phi}(B)\phi^{-1}(B)y_t$$

and

(3.4b) $$y_t = \tilde{\phi}^{-1}(B)\phi(B)x_t = \phi(B)\tilde{\phi}^{-1}(B)x_t \quad .$$

Using (3.2a) and (3.4a) it follows that $\tilde{\phi}(B)\{\phi(B)w_t - \phi^{-1}(B)y_t\} = 0$. Hence $\phi(B)w_t = \phi^{-1}(B)y_t + a_t$, where a_t is some series satisfying $\tilde{\phi}(B)a_t = 0$. Similarly, using (3.2b) and (3.4b), we obtain $\phi(B)w_t = \tilde{\phi}^{-1}(B)x_t + b_t$ with $\phi(B)b_t = 0$. Our proof will be completed by showing that

(3.5) $$a_t = b_t = 0 \ .$$

The first equality in (3.5) holds because (3.4a) implies that $\phi^{-1}(B)y_t$ and $\tilde{\phi}^{-1}(B)x_t$ coincide. Consequently b_t solves two homogeneous difference equations whose characteristic polynomials have no roots in common. This is possible only if the series b_t is zero [cf. Henrici (1974, page 586)]. Hence, (3.5) holds, and the proof is complete.

We shall call that minimal-degree polynomial p(B) in C for which p(0) = 1 <u>the polynomial transformation to stationarity for</u> w_t.

To appreciate the role played by the continuity requirement on the spectral distributions of the transformed series in (3.1), consider the following

EXAMPLE. Let A be a random variable with mean zero and finite variance. Then the series $z_t = (1/2)\{1 + (-1)^t\}A$ is nonstationary, but the transformed series $(1-B)z_t = (-1)^t A$ and $(1+B)z_t = A$ are both stationary. Thus z_t admits two polynomial transformations to stationarity having no non-constant common divisor.

4. Series admitting improper transformations to stationarity and mean functions which satify a homogeneous difference equation. Suppose that $p_w(B)$, the polynomial transformation to stationarity for w_t as defined in section 3, has a divisor $\phi(B)$ such that $\phi^{-1}(B)$ can be applied to the stationary series $\tilde{w}_t = p_w(B)w_t$, yielding a stationary series $\phi^{-1}(B)\tilde{w}_t$. (From (2.1), we observe that this will happen if there are roots of $p_w(B)$ off the unit circle; or, in case $\tilde{w}_t$ is an ARMA process, also if the arguments of some unit circle roots of $p_w(B)$ are zeroes of the spectral density of $\tilde{w}_t$.)
Then, since $\phi(B)\{[p_w(B)/\phi(B)]w_t - \phi^{-1}(B)\tilde{w}_t\} = 0$, we have

$$[p_w(B)/\phi(B)]w_t = \phi^{-1}(B)\tilde{w}_t + a_t \quad , \tag{4.1}$$

where the series a_t, which need not be stationary, satisfies $\phi(B)a_t = 0$. For a given realization $w_t(\omega)$, the right hand side of (4.1) will therefore be indistinguishable from what is obtained by adding to the stationary realization $[\phi^{-1}(B)\tilde{w}_t](\omega)$ a mean function consisting of a sum of terms of the form

$$(4.2) \qquad \{C_0 + C_1 t + \ldots + C_{m-1} t^{m-1}\} r^{-t}$$

times $\cos\lambda t$ or $\sin\lambda t$, one such term for each root $re^{i\lambda}$ of $\phi(B)$ of multiplicity m. The factor $\phi(B)$ of $p_w(B)$ can thus be regarded as having the function of eliminating a possibly non-stationary mean, at the expense, when $\phi(B)$ has unit circle roots, of inducing zeros in the spectrum of $p_w(B)w_t$.

There will be occasions in the next section when such divisors must be excluded. For ease of reference, we shall call a non-constant divisor $\phi(B)$ of $p_w(B)$ *improper for* w_t if $\phi^{-1}(B)[p_w(B)w_t]$ is defined. We say that $p_w(B)$ is *proper for* w_t if it has no improper divisors. The existence or nonexistence of such divisors is, of course, a property of the series w_t.

5. Nonstationary aggregates. Suppose the nonstationary series w_t is the sum of two component series s_t and n_t,

$$w_t = s_t + n_t$$

which have minimal-degree polynomial transformations to stationarity $p_s(B)$ and $p_n(B)$, respectively, such that the transformed series are jointly stationary. Then the product $p_s(B)p_n(B)$ transforms w_t to stationarity, but it will obviously not coincide with the minimal-degree transformation, $p_w(B)$, if $p_s(B)$ and $p_n(B)$ have a (non-constant) common divisor. Suppose that $\phi_c(B)$ is the greatest common divisor of $p_s(B)$ and $p_n(B)$ and that $p_s(B) = \phi_s(B)\phi_c(B)$, and $p_n(B) = \phi_n(B)\phi_c(B)$. By the definition of $\phi_c(B)$, there are no roots common to $\phi_s(B)$ and $\phi_n(B)$. It seems reasonable to expect, and it is sometimes asserted, e.g. (Pierce, 1979), that

$$(5.1) \qquad p_w(B) = \phi_s(B)\phi_n(B)\phi_c(B) \quad .$$

We shall continue to be concerned only with the case in which $p_w(B)w_t$, $p_s(B)s_t$ and $p_n(B)n_t$ have continuous spectral distributions. This will enable us to prove that (5.1) holds if the time series s_t and n_t are uncorrelated, and the roots of $p_s(B)$ and $p_n(B)$ are on the unit circle. Bell (1982) has shown, however, that this assumption of no correlation between components, which is common in stationary signal extraction, conflicts with other attractive assumptions in the nonstationary case, in a way that the weaker assumption, that $p_s(B)s_t$ and $p_n(B)n_t$ are uncorrelated, does not. We first present some results requiring only this weaker assumption.

(5.2) <u>Assume that the stationary series</u> $p_s(B)s_t$ <u>and</u> $p_n(B)n_t$ <u>are uncorrelated, and consider the following two conditions</u>:

(i) <u>The polynomials</u> $p_s(B)$ <u>and</u> $p_n(B)$ <u>are proper for</u> s_t <u>and</u> n_t, <u>respectively</u>.

(ii) <u>The polynomial</u> $p_w(B)$ <u>is proper for</u> w_t <u>and is described by</u> (5.1).

If (i) *holds, then so does* (ii). *Conversely, if* $p_s(B)$ *and* $p_n(B)$ *have no unit circle roots in common, then* (ii) *implies* (i).

PROOF. Let w_t' be the stationary series defined by

$$w_t' = \phi_s(B)\phi_n(B)\phi_c(B)w_t \quad ,$$

and set $\tilde{s}_t = p_s(B)s_t$ and $\tilde{n}_t = p_n(B)n_t$. Then,

$$w_t' = \phi_n(B)\tilde{s}_t + \phi_s(B)\tilde{n}_t \quad . \tag{5.3}$$

If $\tilde{F}_s(\lambda)$ and $\tilde{F}_n(\lambda)$ denote the spectral distributions of $\tilde{s}_t$ and $\tilde{n}_t$, the spectral distribution $F_w'(\lambda)$ of w_t' is given by

$$F_w'(\lambda) = \int_{-\pi}^{\lambda} |\phi_n(e^{i\theta})|^2 d\tilde{F}_s(\theta)$$

$$+ \int_{-\pi}^{\lambda} |\phi_s(e^{i\theta})|^2 d\tilde{F}_n(\theta) \quad ,$$

which makes it clear that a filter can be applied to w_t' if and only if it can be applied to $\phi_n(B)\tilde{s}_t$ and $\phi_s(B)\tilde{n}_t$. If $\tilde{\phi}_s(B)$, $\tilde{\phi}_n(B)$

and $\tilde{\phi}_c(B)$ are (possibly constant) divisors of $\phi_s(B)$, $\phi_n(B)$ and $\phi_c(B)$, respectively, one can show, using (2.4), that $\tilde{\phi}_s^{-1}(B)\tilde{\phi}_n^{-1}(B)\tilde{\phi}_c^{-1}(B)w_t'$ is defined if and only if $\tilde{\phi}_s^{-1}(B)\tilde{\phi}_c^{-1}(B)\tilde{s}_t$ and $\tilde{\phi}_n^{-1}(B)\tilde{\phi}_c^{-1}(B)\tilde{n}_t$ are defined.
By (3.1), $p_w(B)$ is a divisor of $\phi_s(B)\phi_n(B)\phi_c(B)$, so that $p_w(B)w_t$ can be written in the form $\tilde{\phi}_s^{-1}(B)\tilde{\phi}_n^{-1}(B)\tilde{\phi}_c^{-1}(B)w_t'$. The assertions of (5.2) follow immediately from this observation.

It is of interest to have a derivation of (5.1) which does not exclude $p_w(B)$ from having roots inside the unit circle. To this end, we now assume that the series $p_w(B)w_t$, $p_s(B)s_t$ and $p_n(B)n_t$ are not purely deterministic. Then each will, by the Wold decomposition, be the sum of two uncorrelated components, a (possibly zero) purely deterministic component and a purely non-deterministic component $(nd)_t$ having an innovations representation of the form

$$\{nd\}_t = e_t + \sum_{j=1}^{\infty} c_j e_{t-j} \quad ,$$

where e_t is a stationary series of uncorrelated, zero mean, random variables (the innovations), and where the analytic function defined by

$$f(z) = 1 + \sum_{j=1}^{\infty} c_j z^j \qquad (|z|<1)$$

has no zeroes inside the unit circle [Hannan (1970, page 147)]. Having in mind the sort of calculation done in the proof of (5.2), we note that if $\tilde{\phi}(z)$ is a polynomial whose roots are in $\{|z|<1\}$ and are distinct from the roots of the polynomial $\phi(z)$, then $\tilde{\phi}^{-1}(z)\phi(z)f(z)$ will have poles in $\{|z|<1\}$, which means that its Laurent expansion,

$$\tilde{\phi}^{-1}(z)\phi(z)f(z) = \sum_{j=-\infty}^{\infty} d_j z^j \quad ,$$

will have non-zero coefficients d_k for some values $k < 0$. Hence

$$\tilde{\phi}^{-1}(B)\phi(B)\{nd\}_t = \sum_{j=-\infty}^{\infty} d_j e_{t-j}$$

will be correlated with e_{t+k} for some $k > 0$. Using this observation, it is easy to adapt arguments from the proof of (5.2) to verify the following result, which illuminates the role played by (5.1) when improper divisors having roots inside the unit circle are allowed.

(5.4) *Let* e_t^s *and* e_t^n *denote the innovations series for* $p_s(B)s_t$ *and* $p_n(B)n_t$, *respectively, which are assumed to be non-zero. Suppose that* $p_s(B)$ *and* $p_n(B)$ *have no improper divisors for* s_t *and* n_t, *respectively, with roots on or outside the unit circle. Then* (5.1) *holds if and only if for each* t *and for each* $k > 0$, *the innovations* e_{t+k}^s *and* e_{t+k}^n *are uncorrelated with* $p_w(B)w_t$.

Bell (1982) shows that the innovations condition described in (5.4) is a useful one.

Now we give a general result.

(5.5) <u>If the series</u> s_t <u>and</u> n_t <u>are uncorrelated, and are such that the roots of</u> $p_s(B)$ <u>and</u> $p_n(B)$ <u>lie on the unit circle, then</u> (5.1) <u>holds</u>.

PROOF. From (3.1), $p_w(B)$ is a divisor of $\phi_s(B)\phi_n(B)\phi_c(B)$. Suppose that $\tilde{\phi}_s(B), \tilde{\phi}_n(B)$ and $\tilde{\phi}_c(B)$ are divisors of $\phi_s(B), \phi_n(B)$ and $\phi_c(B)$, respectively, such that

$$p_w(B) = [\phi_s(B)/\tilde{\phi}_s(B)][\phi_n(B)/\tilde{\phi}_n(B)][\phi_c(B)/\tilde{\phi}_c(B)] \quad .$$

Obviously, $p_w(B)w_t = p_w(B)s_t + p_w(B)n_t$, and it follows from the proof of (5.2) that the stationary series $\{\tilde{\phi}_s^{-1}(B)\tilde{\phi}_c^{-1}(B)\}p_s(B)s_t$ and $\{\tilde{\phi}_n^{-1}(B)\tilde{\phi}_c^{-1}(B)\}p_n(B)n_t$ are defined. Therefore,

$$p_w(B)s_t = \tilde{\phi}_n^{-1}(B)\phi_n(B)[\{\tilde{\phi}_s^{-1}(B)\tilde{\phi}_c^{-1}(B)\}p_s(B)s_t+a_t]$$

with $\tilde{\phi}_s(B)\tilde{\phi}_c(B)a_t=0$, and

$$p_w(B)n_t = \tilde{\phi}_s^{-1}(B)\phi_s(B)[\{\tilde{\phi}_n^{-1}(B)\tilde{\phi}_c^{-1}(B)\}p_n(B)n_t+b_t]$$

with $\tilde{\phi}_n(B)\tilde{\phi}_c(B)b_t=0$.

Since the series s_t and n_t are uncorrelated, the same is true of the trigonometrical polynomial series $[\phi_n(B)/\tilde{\phi}_n(B)]a_t$ and $[\phi_s(B)/\tilde{\phi}_s(B)]b_t$. These two series must sum to zero, however, because the spectral distribution of $p_w(B)w_t$ is continuous. Thus they themselves must be zero. Consequently, $P_w(B)s_t$ and $p_w(B)n_t$ are stationary, as are

$$\{\tilde{\phi}_s^{-1}(B)\tilde{\phi}_c{}^{-1}(B)\}p_s(B)s_t = \phi_n^{-1}(B)\tilde{\phi}_n(B)p_w(B)s_t$$

and

$$\{\tilde{\phi}_n^{-1}(B)\tilde{\phi}_c^{-1}(B)\}p_n(B)n_t = \phi_s^{-1}(B)\tilde{\phi}_s(B)p_w(B)n_t \ .$$

The fact that $p_s(B)$ and $p_n(B)$ have minimal degree now implies that the $\tilde{\phi}$-polynomials are constant. Thus $p_w(B)$ is given by (5.1), as asserted.

REMARK 1. If polynomial transformations to stationarity are considered for which the transformed series are not required to have continuous spectral distributions, then the analogue of (5.5) can fail when the component transformations have a common root: Let λ satisfying $0 < \lambda < \pi$ be given, along with uncorrelated, zero-mean, non-zero random variables A_1, A_2, B_1, and B_2, whose variances satisfy $\text{var } A_j \neq \text{var } B_j$ $(j=1,2)$ and $\text{var}(A_1+A_2) = \text{var}(B_1+B_2)$. Then the component series x_t and y_t defined by $x_t = A_1\cos\lambda t + B_1\sin\lambda t$ and $y_t = A_2\cos\lambda t + B_2\sin\lambda t$

are nonstationary, uncorrelated with each other, and have the polynomial transformation to stationarity given by $p(B) = 1-2\cos\lambda B + B^2$. However, $w_t = x_t + y_t$ is stationary.

REMARK 2. The uncorrelatedness assumption in (5.5) can be replaced by an independence assumption to obtain (5.1) when z_t has infinite variance. (Bell (1982) addresses the widespread misconception that this must always or even usually be the case when $p_w(B)$ has a unit circle root.)

We note, finally, that the results of this section extend readily to the situation in which w_t is the sum of more than two components.

Acknowledgement. The investigation of the formula (5.1) was suggested by William Bell, to whom the author is indebted for valuable comments on a preliminary version of this paper.

REFERENCES

Bell, W. R. (1982). Signal extraction for nonstationary time series. To appear in *Ann. Stat.*

Hannan, E. J. (1970). *Multiple Time Series*, Wiley, New York.

Henrici, P. (1974). *Applied and Computational Complex Analysis*. Wiley, New York.

Pierce, D. A. (1979). Signal extraction error in nonstationary time series. Ann. Stat. 7, 1303-1320.

Statistical Research Division
U.S. Bureau of the Census
Room 3524, FB-3
Washington, DC 20233