# SPECTRAL UTILITY FUNCTIONS AND THE DESIGN OF A STATIONARY SYSTEM

Roger J. BOWDEN[1)]
University of Auckland

Research Memorandum No. 94
October 1975

The conventional approach to the design of stochastic systems operates, either directly or indirectly, by minimising the variance of the model. This procedure can be regarded as a natural extension of the stability analysis of a deterministic system, according to which the degree of stability is inversely related to the absolute value of its largest characteristic root. Very often, however, the policymaker is not indifferent to the frequency composition of economic fluctuations. He may, for example, have a marked dislike for short-term fluctuations. We formalise this notion by setting up a spectral utility function as a criterion for steady-state optimisation, and show that the results from such an optimisation may conflict with those yielded by the conventional approach. Large characteristic roots may not necessarily be bad!

The scheme of the paper is as follows. Because the ideas involved may be unfamiliar we shall spend some time on a rather intuitive motivation for what follows. This is done in section I. In section II the notion of a spectral utility function is introduced and its evaluation discussed. We then return to the example of section I to give it a more precise treatment. Section III contains extensions, principally to the multivariate case.

## I. Motivation

Consider a linear system perturbed from its equilibrium path. Assuming this system to be stable, we might write the disequilibrium dynamics in terms of deviations $\underset{\sim}{y}$ from its equilibrium position as:

(1) $$\underset{\sim}{y}_t = A\underset{\sim}{y}_{t-1} .$$

The eigenvalues of the matrix A, and in particular that of greatest absolute value, constitute the standard measure of the degree of stability of the model. If the eigenvalues of A are all small in absolute value, the system will converge very quickly to its new equilibrium position; if they are not so small, although still inside the unit circle, the model will converge, but not so quickly. Now consider the stochastic analogue of (1):

(2) $$\underset{\sim}{y}_t = A\underset{\sim}{y}_{t-1} + \underset{\sim}{\varepsilon}_t ,$$

where the disturbance vector $\underset{\sim}{\varepsilon}_t$ is supposed serially uncorrelated with constant covariance matrix $E\underset{\sim}{\varepsilon}_t\underset{\sim}{\varepsilon}_t' = Q$. Define the steady-state covariance matrix $V_o = E\underset{\sim}{y}_t\underset{\sim}{y}_t'$. Then it is a standard result that:

$$V_o = AV_oA' + Q.$$

It is also well-known that if we stack the columns of $V_o$ to obtain a supervector $\underset{\sim}{v}_o$, and do the same with Q to obtain $\underset{\sim}{q}$, we can write:

(3) $$\underset{\sim}{v}_o = (I - A \otimes A)^{-1}\underset{\sim}{q} ,$$

where $\otimes$ signifies the Kronecker product. This can be expressed as:

$$(4) \qquad \underset{\sim}{v_o} = \sum_{i,j} \frac{1}{1-\lambda_i\lambda_j} (E_i \otimes E_j) \underset{\sim}{q} \quad ,$$

where $\lambda_i$ are the eigenvalues of A (assumed distinct) and $E_i$ its projection matrices (see section III). The relationship of the variance of the system to its eigenvalues exhibited in equation (4) has motivated the interpretation of the steady-state variance of the system as the natural generalisation of the deterministic stability criterion. For if the characteristic roots of A are all small in absolute value, so that the model is stable in the deterministic sense, then from (4) the elements of its covariance matrix in the stochastic context can be expected to be small. To be sure, the correspondence is not perfect. But at least we can assert that the presence of a large (i.e. close to 1) positive root will have a substantial adverse effect upon any norm of the covariance matrix $V_o$. The use of the steady-state variance as a criterion for stochastic stability is too widespread to give individual references, but we should mention the elegant formalisation of the control problem in these terms by G. Chow /3/.

Yet there is a residual nagging discomfort in the above generalisation. Consider the special case where the system is one-dimensional. We can write:

$$(5) \qquad y_t = a y_{t-1} + \varepsilon_t \quad , \qquad E\,\varepsilon_t^2 = \sigma^2$$

$$y_{t+1} = a^2 y_{t-1} + a\varepsilon_t + \varepsilon_{t+1} \quad ,$$

and so on. Now fix $y_{t-1}$. Consider the differences $y_t - y_{t-1}$, $y_{t+1} - y_{t-1}$ and so on. Denoting expectations conditional on $y_{t-1}$ by a semicolon, we have:

$$(6) \qquad E(y_t - y_{t-1})^2 ; y_{t-1} = E(y_t - Ey_t)^2 + (Ey_t - y_{t-1})^2 ; y_{t-1} ,$$

$$= \sigma^2 + (1-a)^2 y_{t-1}^2 \; .$$

(7) $$E(y_{t+1}-y_{t-1})^2 = E(y_{t+1}-Ey_{t+1})^2 + (Ey_{t+1}-y_{t-1})^2 ; y_{t-1},$$
$$= \sigma^2 + a^2\sigma^2 + (1-a^2)y_{t-1}^2 ,$$

and so on. In general,

(8) $$E(y_{t+\tau}-y_{t-1})^2 ; y_{t-1} = \sigma^2 \frac{1-a^{2(\tau+1)}}{1-a^2} + (1-a^{\tau+1})^2 y_{t-1}^2 .$$

As $\tau \to \infty$ , the second term on the RHS $\to y_{t-1}^2$, provided of course that $|a|<1$ . Now compare two values of a, one small and positive, the other close to unity. The smaller value will give the smaller long-run mean-square difference as defined by (8). But if we look at (6) and (7) for proximal values, we can see that the reverse is true. The higher the value of a, the closer can one expect adjacent values of y to lie. As we go further into the future, the influence of the conditioning on $y_{t-1}$, as manifested in the second terms on the RHS of (6) - (8), wears off.

Let us now turn from the sophistry of arithmetic to the business of Governments. Being, as they are, all things to all people, politicians may dislike change, and particularly any change that manifestly disadvantages one group of voters to the enrichment of another. Let us take as a concrete example the housing market in New Zealand. As in many other countries, house prices showed a truly staggering increase over the short space of not much more than a year from 1972-73. There is not too much doubt that in doing so, they were adjusting to a new equilibrium position caused inter alia by a remarkable increase in the volume of immigration.[2] Nor is this best treated as a new deterministic equilibrium, for immigration - as did other variables involved - subsequently dropped sharply. After the accompanying outcry, it is clear that the government of the day would have looked with a jaundiced eye upon any mathematician who claimed that in fact this market was evidently a highly stable one. They would have preferred a smoother market, in the sense of smaller proximal changes, even at the expense of larger long-run changes. According to this philosophy, change is more acceptable if it is done by gradual increments, and one can accept a gradual larger change more readily than a more sudden one of smaller magnitude.

There are several possible reasons why this might be the case. For one thing the active life of no (human) politician or decision-maker is infinite, and like everybody else, he has his own system of time discounting. The preference for smoother short-run changes may also reflect the existence of a recognition or action lag in policy-making. One could with justice argue that if the nature of the system is accurately known, one should be able to design an automatic feedback which would obviate such a preference. Usually however this is not the case. Finally, there is the argument mentioned above, namely that the utility effects of sudden change are rarely symmetric between gainers and losers.

Obviously, however, there is a limit to the extent to which policy-makers would prefer to swap short-run fluctuations for longer-run movements. Our argument must now be given a more precise expression than is possible with the simple and rather intuitive example incorporated in equations (5) - (7), although we have not yet finished with this as an example. The formalisation of our policy-maker's preferences is the subject of the next section.

## II. The Spectral Utility Approach

Suppose that $(y_t)$ constitutes a scalar stationary[3] stochastic process with spectral distribution function F(w), which for the sake of expositional ease we shall assume continous, with density function f(w). As a way of describing the policy-maker's preferences as between short, medium or long-term fluctuations, we can define a real, bounded, semipositive function U(w) which represents his relative aversion to different frequencies. Thus we assume U(w) has period $2\pi$, i.e. $U(w+2\pi) = U(w)$, that $U(-w) = U(w)$, and that

(9) $$\int_{-\pi}^{\pi} U(\omega) = 1 .$$

In conventional discussions of steady-state stochastic systems, the criterion function can usually be cast in terms of the variance of the system, which as a matter of frequency decomposition is given by

(10) $$v_c = \int_{-\pi}^{\pi} f(\omega)\,d\omega = 2\int_{0}^{\pi} f(\omega)\,d\omega .$$

Our proposal is that this criterion should be replaced by

(11) $$E.U. = -\int_{-\pi}^{\pi} U(\omega)\,f(\omega)\,d\omega = -2\int_{0}^{\pi} U(\omega)\,f(\omega)\,d\omega ,$$

which is to be maximised with respect to the parameters of interest. We call this the expected spectral utility associated with a given process. Our objective in maximising it is simply a rather more convenient way of saying that we are minimising the weighted sum of the power at each frequency, where the weights involved constitute the aversion function U(w). Note that the expression (11) is a true expectation of the utility function - U(w), bearing in mind the fundamental correspondence between frequency w and the underlying probability space of the stochastic process (see e.g. J. Doob /5/). On the other hand it need not be true that the utility function involved has the particular characteristics, such as concavity, of those employed in the theory of choice.

Before we can proceed to discuss further the implications of the expected utility approach, we shall have to investigate the determination of the integral in (11). The evaluation of this integral is in general a difficult matter, especially where the question of optimisation with regard to parameters of interest is involved. In one leading[4] case, however, this is not true, and the resulting interpretation is both natural and interesting.

Introduce the lag-operator z such that $zy_t = y_{t-1}$, or $z^{-1}y_t = y_{t+1}$ (for standard accounts of its calculus in different contexts, see P. Whittle /7/ and E.I. Jury /6/). When reference to frequencies is involved we shall employ the standard substitution $z = e^{-iw}$, in terms of the value of z on the unit circle $|z| = 1$.

Suppose now that U(w) can be expressed in terms of the quotient of polynomials ("rational polynomial" form) as:

$$U(w) = \left| \frac{h(e^{-iw})}{g(e^{-iw})} \right|^2 . \qquad (12)$$

Now write the moving-average representation of the process as:

$$y_t = b(z)\,\varepsilon_t,$$

where $\varepsilon_t$ is a white noise process, and $b(z) = b_o+b_1z+b_2z^2+\ldots$ .

Then

$$(13a) \qquad E.U. = -\frac{\sigma^2}{2\pi}\int_{-\pi}^{\pi}\left|\frac{h(e^{-i\omega})}{g(e^{-i\omega})}\,b(e^{-i\omega})\right|^2 d\omega \;,$$

which can alternatively be expressed in terms of contour integration around the unit circle as

$$(13b) \qquad E.U. = -\frac{\sigma^2}{2\pi i}\oint \frac{1}{z}\,\frac{h(z)\,b(z)}{g(z)}\,.\,\frac{h(z^{-1})\,b(z^{-1})}{g(z^{-1})}\,dz \;.$$

The interpretation of equation (13a) is clear. Maximising expected utility amounts to minimising the variance of a new process defined by

$$(14) \qquad x_t = \frac{h(z)}{g(z)}\,y_t \;.$$

This idea is analogous to the process of filtering. The crucial difference is that instead of choosing the filter for a given process, we are choosing the process for a given filter. The integration of (13) can then be carried out by noting that

$$(15) \qquad Var.\ x_t = Abs.\left\{\sigma^2\,\frac{h(z)\,b(z)}{g(z)}\,.\,\frac{h(z^{-1})\,b(z^{-1})}{g(z^{-1})}\right\} \;,$$

where Abs. denotes the process of finding the constant term in the expansion in powers of z of the quantity inside the brackets. The same result can be obtained independently of this interpretation by applying the Cauchy residue theorem directly to the contour integration (13b), noting that under the boundedness assumption on U, the only pole inside the unit circle is z = 0. The fact that the formula on the RHS of (15) is independent of the interpretation in terms of a derived process $x_t$ will be useful in section III, where difficulties arise with such a concept.

Let us now proceed with an example which will clarify the above procedures while continuing with the discussion of section I. Suppose that our policy-maker's spectral utility function is such that the higher frequencies are always less preferred. We might capture this monotonicity by the following function:

$$(16) \qquad U(\omega) = (1-\delta^2)\left|\frac{1}{1-\delta e^{-i\omega}}\right|^2 = \frac{1-\delta^2}{1+\delta^2-2\delta\cos\omega} ,$$

with $-1 < \delta < 0$. This has a minimum at $w = 0$, and increases to a maximum at $w = \pi$ .

Suppose that the underlying process is the first-order Markov process specified by equation (5). We imagine that the policymaker has control over the parameter a, perhaps after the fashion described by Chow /3/. Our problem is then to choose a to maximise expected spectral utility. As we have seen in section I, a = 0 yields the system with minimum variance. Is zero a still optimum?

According to (14) and (16), the policy-maker now acts as if he were minimising the variance of the process defined by

$$x_t = (1-\delta^2)^{\frac{1}{2}}(y_t + \delta y_{t-1} + \delta^2 y_{t-2} + \dots) .$$

Since the weights $\delta$ are negative, the weight pattern is one of alternating sign. We obtain the required variance from equation (15) as follows:

$$\begin{aligned}
\mathrm{Var}\ x_t &= \sigma^2(1-\delta^2)\ \mathrm{Abs.} \left\{ \frac{1}{(1-\delta z)(1-\delta z^{-1})(1-az)(1-az^{-1})} \right\}, \\
&= \sigma^2(1-\delta^2)\ \mathrm{Abs.} \{ (1+\delta z+\delta^2 z^2+\dots)(1+\delta z^{-1}+\delta^2 z^{-2}+\dots)(1+az+a^2z^2+\dots)\,. \\
&\qquad .(1+az^{-1}+az^{-2}+\dots) \}, \\
&= \frac{\sigma^2(1-\delta^2)}{(1-\delta^2)(1-a^2)}\ \mathrm{Abs.} \{ (1+\delta z+\delta z^{-1}+\delta^2 z^2+\delta^2 z^{-2}+\dots)(1+az+az^{-1}+ \\
&\qquad +a^2z^2+a^2z^{-2}+\dots) \}, \\
&= \frac{\sigma^2}{1-a^2} \{ 1 + 2\delta a + 2\delta^2 a^2 + \dots \}, \\
&= \frac{\sigma^2}{1-a^2}\ \frac{1+\delta a}{1-\delta a}\,.
\end{aligned}$$

To choose the value of a which maximises E.U., we differentiate with respect to a, to obtain

$$\frac{d}{da}\,E.U = \frac{d}{da}\left(-\mathrm{Var.}\ x_t\right) = \frac{2\sigma^2}{(1-\delta a)^2(1-a^2)^2}\,(a^3\delta^2+a^2\delta-a-\delta)\,.$$

Write $f(a) = a^3\delta^2+a^2\delta-a-\delta$. Observe that $f(1) < 0$, $f(0) > 0$, $f(-1) > 0$ and that $f(a)$ has stationary points at $a = \frac{1}{3\delta}, \frac{-1}{\delta}$.

That at $a = \frac{-1}{\delta}$ need not concern us. That at $a = \frac{1}{3\delta}$ represents a maximum which must always lie in the negative half line. Thus bearing in mind the value of f at -1, 0 and 1, it will follow that only one stationary point for f exists in the interval (-1, 1) and that this root must in fact lie between 0 and 1. Denoting its value by $a^*$, it is easy to show that

$$\frac{d^2}{da^2}\,(E.U.) < 0 \text{ at } a = a^*,$$

so that our stationary point represents maximum expected utility. Finally one can show that if we decrease $\delta$, i.e. increase our policy-maker's aversion to short-run fluctuations, we obtain a value for $a^*$ closer to unity.

The notion of a spectral utility function has therefore enabled a useful formalisation of the notions introduced in section I. In the stochastic context, a decision-maker may well prefer his system to have higher characteristic roots than the deterministic notion of stability would suggest, simply because he is relatively averse to short-run fluctuations.

## III. Extensions

Consider now a vector process $\underset{\sim}{y}_t$, with spectral density matrix denoted by $\Phi(w)$. The question arises as to the generalisation of the function U(w) of section II. Rather than attempt to define an arbitrary generalisation, the best course is to define the function in terms of the implied operations on the $\underset{\sim}{y}_t$ process, on the supposition that whatever our welfare effects, they can at least be approximated by a "rational polynomial" form. Thus one might define, as an extension of equation (14)

(17) $$\underset{\sim}{x}_t = \text{Diag.}\left( \frac{h_i(z)}{g_i(z)} \right) \underset{\sim}{y}_t ,$$

where Diag. indicates a matrix with diagonal elements as given, for i = 1....n, and zeros off the diagonal.

Introduce the moving-average representation

$$\underset{\sim}{y}_t = B(z)\underset{\sim}{\varepsilon}_t,$$

where $\text{Cov}(\underset{\sim}{\varepsilon}_t) = Q$. Write the diagonal matrix in (17) as $\Lambda(z)$. Then we can write our criterion as:

(18) $$E.U. = -\frac{1}{2\pi i} \oint \frac{1}{z} \text{ trace } C\Lambda(z)B(z)QB'(z^{-1})\Lambda(z^{-1})dz ,$$

after the fashion of equation (13b) for the univariate case. Here C is a matrix of constants referring to the weights to be assigned to the different variables. Equation (18) is our generalisation of the conventional criterion $Ey_t'Cy_t$ = trace $CV_o$.

The generalisation of our discussion in sections I and II would require $B(z) = (I-Az)^{-1}$. Suppose once more that the eigenvalues of A are distinct, and write $U = (\underset{\sim}{u}_1, \underset{\sim}{u}_2 \dots \underset{\sim}{u}_n)$, where $\underset{\sim}{u}_i$ is the eigenvector corresponding to the root $\lambda_i$. Write $V = U^{-1}$ and let $\underset{\sim}{v}_j'$ denote the j th row of V. Then the projection matrices $E_i$ referred to in section I are defined as $E_i = \underset{\sim}{u}_i \underset{\sim}{v}_i'$, and we can write[5]

$$B(z) = \sum_i \frac{1}{1-\lambda_i z} E_i .$$

Thus

$$B(z)QB'(z^{-1}) = \sum_{i,j} \frac{1}{(1-\lambda_i z)(1-\lambda_j z)} E_i Q E_j' .$$

Suppose that C is diagonal, so that we are interested only in the individual spectra. Absorb the constants $c_k$ in the functions $\frac{h_k(z)}{g_k(z)}$ each of which will now normalise to $c_k^2$ in place of equation (9). We thus obtain

$$\text{trace } \Lambda(z) B(z) Q B'(z) \Lambda(z) = \sum_{i,j} \frac{1}{(1-\lambda_i z)(1-\lambda_j z)} \left\{ \sum_k \frac{h_k(z)}{g_k(z)} \frac{h_k(z^{-1})}{g_k(z^{-1})} (E_i Q E_j')_{kk} \right\} ,$$

where $(E_i Q E_j')_{kk}$ denotes the k th diagonal element of the matrix in brackets. Recalling that $E_i = \underset{\sim}{u}_i \underset{\sim}{v}_i'$, and distributing the contour integral $\oint$ inside the brackets, we obtain:

$$(19) \quad E.U. = -\sum_{i,j,k} (\underset{\sim}{v}_i' Q \underset{\sim}{v}_j) u_{ik} u_{jk} \text{ Abs.} \left\{ \frac{h_k(z)}{g_k(z)(1-\lambda_i z)} \cdot \frac{h_k(z^{-1})}{g_k(z^{-1})(1-\lambda_j z^{-1})} \right\} .$$

Determination of the absolute term inside the brackets will usually be facilitated by a partial factions expansion of the terms in the denominator. Thus suppose

$$\frac{h_k(z)}{g_k(z)} = \frac{c_k(1-\delta_k^2)^{\frac{1}{2}}}{(1-\delta_k z)} .$$

We can write:

$$\frac{1}{(1-\delta_k z)(1-\lambda_i z)} = \frac{\delta_k}{\delta_k-\lambda_i}\,\frac{1}{1-\delta_k z} + \frac{\lambda_i}{\lambda_i-\delta_k}\,\frac{1}{1-\lambda_i z},$$

$$= \frac{\delta_k}{\delta_k-\lambda_i}\left(1+\delta_k z+\delta_k^2 z^2+\cdots\right) + \frac{\lambda_i}{\lambda_i-\delta_k}\left(1+\lambda_i z+\lambda_i^2 z^2+\cdots\right).$$

The term in $z^{-1}$ and $\lambda_j$ can be similarly treated. Multiplying the two together we easily obtain the required absolute part as:

$$\text{Abs.}\ \frac{c_k(1-\delta_k^2)^{\frac{1}{2}}}{(1-\delta_k z)(1-\lambda_i z)} \cdot \frac{c_k(1-\delta_k^2)^{\frac{1}{2}}}{(1-\delta_k z^{-1})(1-\lambda_j z^{-1})} = \frac{c_k^2(1-\delta_k^2)}{(\lambda_i-\delta_k)(\lambda_j-\delta_k)}\left(\frac{\lambda_i\lambda_j}{1-\lambda_i\lambda_j} - \frac{\delta_k\lambda_i}{1-\delta_k\lambda_i} - \frac{\delta_k\lambda_j}{1-\delta_k\lambda_j} + \frac{\delta_k^2}{1-\delta_k^2}\right).$$

If any of the eigenvalues $\lambda_i$ are complex, the resulting expression will naturally also be complex. But the fact that complex roots - and their associated eigenvectors - occur in conjugate pairs will mean that the grand total defined by (19) will be real. The above treatment can be interpreted in terms of a set of elementary Markov processes defined by the roots $\lambda_i$ (as in Chow /4/), taken in conjunction with equation (14) and (15) above. However where the $\lambda_i$ are complex, such an interpretation is not satisfactory from the viewpoint of formal proof. We have therefore cast the above demonstration directly in terms of the appropriate contour integration.

Evidently the above expressions are complicated functions of the underlying parameters $a_{ij}$ and $q_{ij}$. Thus the maximisations involved will usually have to be carried out numerically. This may not be true, however, for simple systems. Thus a univariate second-order autoregressive process could either be cast as a first-order two-dimensional system (by defining the appropriate companion matrix) or else could be treated directly by employing the appropriate partial fraction expansion in terms of the characteristic roots of the system. In either case it may be possible to differentiate E.U. with respect to the basic parameters of interest.

The analysis of section II could also be extended to allow the utility function to be expressed in the form $U(w,v_o)$, where $v_o$ denotes the variance of the process. (The normalisation constraint (9) must nevertheless be satisfied). Thus the policy-maker may be indifferent to the frequency decomposition of fluctuations if the overall power is small. But if $v_o$ is large, this indifference may disappear. Incorporating $v_o$ into the utility function is a way of handling such interaction effects. In this case, the maximisation of expected utility will take place subject to equation (10), and Lagrangean methods are appropriate.

The present paper has been devoted to exposition of the philosophy and techniques of expected-utility maximisation in the frequency domain. It would now be interesting to choose a particular stabilisation context and to compare the results from the conventional variance minimisation with those obtained by assuming something about the preference structure of policy-makers over the frequency decomposition of fluctuations.

## References

/1/ BOWDEN, R.J.: "Nonlinearity and Nonstationary in Dynamic Econometric Models" Review of Economic Studies, Vol. 41, 1974, 173-179.

/2/ BOWDEN, R.J.: "Disequilibrium and Speculation in the Housing Market", Economic Record, forthcoming.

/3/ CHOW, G.C.: "Optimal Stochastic Control of Linear Economic Systems" Journal of Money, Credit and Banking, Vol. 2, 1970, 291-302.

/4/ CHOW, G.C.: Analysis and Control of Dynamic Economic Systems, J. Wiley, New York, 1975.

/5/ DOOB, J.: Stochastic Processes. J. Wiley, New York, 1953.

/6/ JURY, E.I.: Theory and Application of the z-Transform Method, J. Wiley, New York, 1964.

/7/ WHITTLE, P.: Prediction and Regulation by Linear Least Square Methods, EUP/Van Nostrand, 1963.

## Footnotes

[1] This paper was conceived and written while I was on leave at the Institute for Advanced Studies, Vienna. I am most grateful to the Institute for their generous financial and material support.

[2] Some of the dynamic aspects of this episode are discussed in Bowden /2/.

[3] There do not appear to be any difficulties in generalising this to mildly non-stationary processes. The connection with the characteristic roots of the system matrix can be approached through the method of moving linearisation suggested by Bowden /1/.

[4] The question arises as to the conditions under which an arbitrary $U(w)$ can be at least approximated in the form (12). This question may be closely related to the problem of achieving a spectrum factorisation. For if U is continuous, it can itself be regarded as a spectral density function corresponding to some process. Let $\beta(z)$ be the corresponding moving-average generating function. The question can then be reduced to the closeness of approximation of $\frac{h(z)}{g(z)}$ to $\beta(z)$ on the unit circle, once $\beta(z)$ has been found by factorising U. We shall not pursue further here the subtle and interesting problems that arise.

[5] Such expressions for the spectral matrix in terms of those for elementary autoregressive processes are extensively discussed in Chow /4/.