STATISTICAL RESEARCH REPORT No.3
University of Oslo May 1981
Institute of Mathematics

# NOTES ON AUTOREGRESSIVELY DEFINED STATIONARY PROCESSES (ARMA-PROCESSES) BASIC MATHEMATICAL PROPERTIES

by

Erling Sverdrup

## Preface

These notes were prepared in the course of giving a rather extensive course in Time Series Analysis, following T.W. Anderson's book. As often happens when lectures are based on a certain text, one wants to deviate from it on certain points. The deviations are contained in the present notes.

If there should be anything novel in them, it would be the study of the stationarity and the "anteimpulse" nature of the autoregressive processes, which is treated more extensively. (I take "autoregressive process" in a general sense with few restrictions on the "noice", the ARMA process is a special case.) I have made a point of giving an elementary presentation, since I have felt that the subject is by nature elementary. Only elementary results about convergence in quadratic mean are used. For comparison I have also included a derivation based on the famous spectral theorem of Stone and Cramér, see Cramér (1967). That is similar to derivations given by Grenander & Rosenblatt (1956) and Gihmann & Skorohod (1974).

I have found it useful to expand upon and deviate in details from Anderson's derivation of the distribution of the spectrum. A self-contained proof is presented in Chapter VI.

## Table of contents

# I INTRODUCTION

## A. Structural cycles

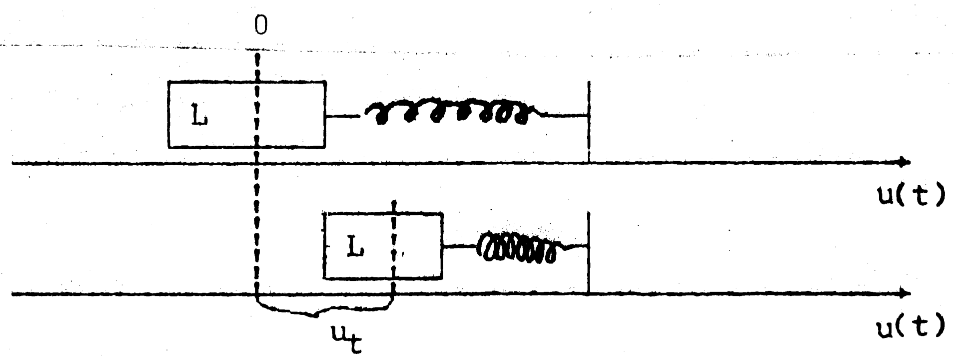
In the study of time series the attention will in many cases be directed toward oscillary motions, i.e. more or less periodical regularities. This is true both when studying empirical observations u(t) of a variable u over time t, or when making models to explain the structures of the time series u(t). The study of periodicities and frequences will appear to be useful even if the phenomenon we are studying do not have conspicuous periodicities.

Let us take a brief look at these phenomenon. Sometimes the oscillations are forced on a time series u(t) from the outside. It may be realistic to assume that

$$u(t) = \rho_0 + \sum_{j=1}^{m} \rho_j \sin\left(\frac{2\pi}{p_j} t + \phi_j\right) + V_t$$

where the $V_t$ are independent. One of the sine components may e.g. represent seasonal variation (with $p_j = 12$ months). However, more often the oscillations are structural, arising out of the system itself. Thus Gallilei and Huygens recognized that the oscillations of the pendulum were due to principles expressed by momentum, inertia and gravity, which per se say nothing about oscillations.

Let us consider the situation in the case of an oscillating spring balance.



The weight L has mass m. The deviation $u(t)$ = distance between position at time t and equilibrium position. The coefficient of friction is f. We have three structural relations.

Inertial force $= m\,u''(t) = m\frac{d}{dt}\,u(t)$

Frictional force $= f\,u'(t)$

Restoring force $= k\,u(t)$ (Hooke's law).

These are the inner forces of the system. In addition the weight is exposed to positive and negative shocks from the outside. Let the shock at time t be $K(t)$. We find,

$$K(t) = m\,u''(t) + f\,u'(t) + k\,u(t).$$

The shock acts against inertia, friction and the stress of the compressed or lengthened spring. This differential equation has the solution

$$u(t) = \int_{-\infty}^{t} \frac{K(\tau)}{m\ \ \lambda}\, e^{-\frac{f}{2m}(t-\tau)} \sin \lambda(t-\tau)\,d\tau$$

provided $\lambda = \sqrt{\frac{k}{m} - (\frac{f}{2m})^2}$ is real and $> 0$. It is seen that the sequence of past and present shocks has been transformed to a composition of damped oscillations, all of them with the same period $p = 2\pi/\lambda$. The transformation has the form of a filter

$$u(t) = \int_{-\infty}^{t} K(\tau)\psi(t-\tau)d\tau = \int_{0}^{\infty} \psi(\tau)K(t-\tau)d\tau$$

where the form of the transformation (filter) is determined by $\psi$, i.e. by the structural parameters m,f,k alone.

The essence of the situation just described is that a model is used, which gives relationships between position, velocity and acceleration, i.e. between position *at a certain point of time and at two points of time just past. There is a kind of sluggishness in the system,* and that is the explanation of many oscillatory phenomenon.

sometimes such an ascertainment might seem trivial. However, to make use of it may not be trivial. It will be the main theme in this paper, where we shall have in mind situations were shocks or impulses will be non-observed random variables generating a sequence of observable $u(t)$. Contrary to the situation in the example, we shall consider discrete processes $u(t)$ where the $u(t)$ are observed for integer values of $t$. Replacing $u'(t)$ and $u''(t)$ by $u(t)-u(t-1)$ and $u(t)-2u(t-1)+u(t-2)$ respectively in the differential equation above, we obtain

$$u(t) = \phi_1 u(t-1) + \phi_2 u(t-2) + K(t).$$

We shall be concerned with models described by such linear difference equation, where more generally, $p$ instead of 2 lags may be involved.

But first we shall say something about mathematical tools used for describing, estimating and analyzing time series.

## B. Search for periodicities. Spectrum and autocovariance.

First some words about periodic functions. $f(t)$ has period $p$ if $f(t) = f(t+p)$ for all $t$. Considering $f(t)$ as a value arising at time $t$, then $\nu = 1/p$ is the number of periods per time unit. A prominent periodic function with period $p$ is $f(t) = A\sin(\omega t+\phi) = A\sin(\frac{2\pi}{p}t+\phi)$. $\phi$ is called the phase and $\omega = 2\pi/p = 2\pi\nu$ is the angular velocities. (Projecting the graph of $A\sin(\omega t+\phi)$ on the circle with center in origin and radius $A$, it is seen that $\nu$ is the number of revolutions per unit time and $\omega$ the number of radians per unit time.)

Consider now a time series of observations

$$Y_1, Y_2, \ldots Y_T \tag{1}$$

To investigate if $p = 2\pi/\omega$ could roughly be considered as a period, the covariance between $Y(t)$ and $f(t) = \sin(\omega t+\phi)$, i.e.

$$\frac{1}{T}\sum_{t=1}^{T}(Y_t-\bar{Y})(f(t)-\overline{f(t)}) = \frac{1}{T}\sum_{t=1}^{T}(Y_t-\bar{Y})\sin(\omega t+\phi) = \frac{1}{T}K(\phi) \qquad (2)$$

is studied, where the averages $\bar{Y}$ and $\overline{f(t)}$ are over all the T value of t.

Let us vary $\phi$ i.e. shift the graph of the sine function back and forth horizontally to make the oovariance as large as possible. Obviously $\hat{\phi}$ = maximizing $\phi$


must be such that

$$K'(\hat{\phi}) = \sum(Y_t-\bar{Y})\cos(\omega t+\hat{\phi}) = 0$$

or

$$\cos\hat{\phi}\sum(Y_t-\bar{Y})\cos\omega t - \sin\hat{\phi}\sum(Y_t-\bar{Y})\sin\omega t = 0.$$

On the other hand from (2)

$$\cos\hat{\phi}\sum(Y_t-\bar{Y})\sin\omega t + \sin\hat{\phi}\sum(Y_t-\bar{Y})\cos\omega t = 0.$$

Multiplying the first equation by $\cos\hat{\phi}$ and the second by $\sin\hat{\omega}$ and adding we get

$$K(\hat{\phi})\sin\hat{\phi} = \sum(Y_t-\bar{Y})\cos\omega t = A(\omega) \qquad (3)$$

Similarly

$$K(\hat{\phi}) \cos \hat{\phi} = \sum(Y_t-\bar{Y}) \sin \omega t = B(\omega).$$

Hence

$$K(\hat{\phi}) = \sqrt{[\sum(Y_t-\bar{Y})\cos\omega t]^2 + [\sum(Y_t-\bar{Y})\sin\omega t]^2} = \tag{4}$$

$$= \sqrt{A(\omega)^2 + B(\omega)^2}$$

It is seen from (3), that with $Z(\omega) = A(\omega) + i\,B(\omega)$ we may also write

$$R^2(\omega) = |Z(\omega)|^2 = \frac{4}{T^2}|\sum(Y_t-\bar{Y})e^{i\omega t}|^2 \tag{5}$$

Obviously to search for periodicities $p$, we may try out different values of $p$, i.e. different values of $\omega$, to find the $p$ for which an adjusted sine function is highly correlated with $Y_t$. It is convenient to make a graph of $R(\omega) = \frac{2}{T}K(\hat{\phi})$ as a function of $\omega$, or $p$, or $\nu$. (The reason for the factor 2 will be clear later.) $R(\omega)$ as a function of $p$ is the *periodogram*, as a function of $\nu$ or $\omega$, the *spectrum*. However, today the words are used interchangebly, whatever is the argument. Spectrum seems to be preferred. Summing up what has been said above. The *spectral value* for a given frequency *is the phase adjusted covariance between the sine-function and the time series.*

In the socalled periodogram analysis only periods $p$ which are factors in $T$ are chosen, hence $p = T/n$, where $n$ is an integer. A cyclical trend

$$\rho_0 + \sum_{j=1}^{r} \rho_j \sin(\frac{2\pi}{T}n_j t+\phi_j) = \tag{6}$$

$$= \rho_0 + \sum_{j=1}^{r} \alpha_j \cos\frac{2\pi}{T} n_j t + \sum_{j=1}^{r} \beta_j \sin\frac{2\pi}{T} n_j t$$

is fitted to the observations $Y_t$ by least square method. Here the $n_j$ are integers and

$$\alpha_j = \rho_j \sin\phi_j \quad , \quad \beta_j = \rho_j \cos\phi_j \quad , \quad \rho_j^2 = \alpha_j^2 + \beta_j^2 \tag{7}$$

We get for the least square estimates of $\rho_j$, $\alpha_j$, $\beta_j$

$$\hat{\rho}_0 = \bar{Y} \quad , \quad \hat{\alpha}_j = \frac{2}{T}\sum_t Y_t \cos\frac{2\pi}{T} n_j t \quad , \quad \hat{\beta}_j = \frac{2}{T}\sum_t Y_t \sin\frac{2\pi}{T} n_j t \tag{8}$$

$\hat{\rho}_j^2 = \hat{\alpha}_j^2 + \hat{\beta}_j^2$ .

Note that $\hat{\alpha}_j = \frac{2}{T} A(\frac{2\pi}{T}n_j)$ , $\hat{\beta}_j = \frac{2}{T} B(\frac{2\pi}{T}n_j)$ (see (3), note also that $\bar{Y}$ drops out because of the special choice of $\omega$). We see that $\hat{\rho}_j^2 = R(\frac{2\pi}{T}n_j)$. *The spectrum gives the square of the amplitude of the different periods.*

Obviously if $Y_t$ is roughly periodic with period $p$ then $Y_t$ is high whenever $Y_{t+p}$ is high and $Y_t$ is low whenever $Y_{t+p}$ is low. Hence, for given $p$, there must be high correlation between $\{Y_t\}$ and $\{Y_{t+p}\}$.

This leads us to study the *autocovariance function*

$$c_p = \frac{1}{T-p}\sum_{t=1}^{T-p}(Y_t-\bar{Y})(Y_{t+p}-\bar{Y}) \tag{9}$$

as a function of $p$. For convenience we define $c_p$ also for negative $p$, by $c_p = c_{-p}$. It is seen that

$$c_{-p} = \frac{1}{T-p}\sum_{p+1}^{T}(Y_t-\bar{Y})(Y_{t-p}-\bar{Y}) \tag{10}$$

Now, let us consider the relationship between spectrum and covariance.

Let $Z_t = Y_t - \bar{Y}$. We get

$$\frac{T^2}{4}R^2 = \sum_{t,s} Z_t Z_s \cos\omega t \sin\omega s + \sum_{t,s} Z_t Z_s \sin\omega t \sin\omega s =$$

$$= \sum_{t,s} Z_t Z_s \cos\omega(t-s) = \sum_{h=-(T-1)}^{T-1}(\sum_t Z_t Z_{t+h})\cos\omega h =$$

$$= \sum_{-(T-1)}^{-1} \cos\omega h \sum_{t=-h+1}^{T} Z_t Z_{t+h} + \sum_{0}^{T-1} \cos\omega h \sum_{t=1}^{T-h} Z_t Z_{t+h} =$$

$$= \sum_{-(T-1)}^{T-1} (T-|h|) c_h \cos\omega h$$

having made use of (10). Hence

$$R^2(\omega) = \frac{4}{T} \sum_{-(T-1)}^{T-1} \left(1 - \frac{|h|}{T}\right) c_h \cos\omega h \tag{11}$$

which expresses the spectrum by means of the covariance function.

Let us now multiply (11) by $\cos\omega k$ and integrate over $\omega$ from $-\pi$ to $\pi$. We get ($|k| < T$)

$$\frac{8\pi}{T}\left(1 - \frac{|k|}{T}\right) c_k = \int_{-\pi}^{\pi} R^2(\omega) \cos\omega k \, d\omega .$$

Hence there is a one-to-one correspondance between $R(\omega)$ and $c_h$. Often the spectrum is defined as $I_T(\omega) = \frac{T}{8\pi} R^2(\omega)$. Then

$$\left(1 - \frac{|k|}{T}\right) c_k = \int_{-\pi}^{\pi} I_T(\omega) \cos\omega k \, d\omega \tag{12}$$

and by (11),

$$I_T(\omega) = \frac{1}{2\pi} \sum_{-(T-1)}^{T-1} \left(1 - \frac{|h|}{T}\right) c_h \cos\omega h \tag{13}$$

The above definitions of spectrum and covariance functions are relative to an empirical material $Y_1, \dots, Y_T$. We shall now define the same concepts relatively to a well defined stochastic process.

Consider an arbitrary stochastic process $Y(t)$ ; $t = \dots -1, 0, 1, \dots$ Then the joint distribution of $Y(t_1), \dots, Y(t_n)$ exists for any $n, t_1, \dots, t_n$. If $E\,Y(t)^2 < \infty$ for any $t$ then the means $\eta(t) = E\,Y(t)$ and the covariances $\sigma(t,s) = E(Y(t)-\eta(t))(Y(s)-\eta(s))$ exist. An arbitrary process is said to be strictly stationary if

$Y(t_1),\ldots,Y(t_n)$ has the same distribution as $Y(t_1+\tau),\ldots,Y(t_n+\tau)$ for any $n,t_1,\ldots,t_n,\tau$. It follows in particular that if $E\,Y(t)^2 < \infty$ then since $(Y(t),Y(s))$ and $(Y(t+\tau),Y(s+\tau))$ has the same distribution then $\eta(t) = \eta(t+\tau)$ and $\sigma(t,s) = \sigma(t+\tau,s+\tau)$, i.e. $\eta(t)$ is a constant and (setting $\tau = -s$) $\sigma(t,s) = \sigma(t-s,o)$ depends on $t,s$ only through $t-s$. We write $\eta(t) = \eta$ and $\sigma(t,s) = \sigma(t-s)$. Regardless of whether $Y(t)$ is strictly stationary or not, we shall say that $Y(t)$ is *second order stationary* if $\eta(t) = \eta$ is a constant and $\sigma(t,s) = \sigma(t-s)$ depends only on $t-s$. $\sigma(h) = \sigma(t,t+h)$ is called the autocovariance function. Of course $\sigma(o) = \operatorname{var} Y(t)$ and $\sigma(h)/\sigma(o)$ (if $\sigma(o) > 0$) is the correlation between $Y(t)$ and $Y(t+h)$, the socalled serial correlation. The autocovariance $\sigma(h)$ is obviously quite analogous to the empirical autocovariance discussed above. Referring to equation (2) (and neglecting the small term $|h|/T$ for large T) it might seem natural to define the cumulative spectrum $F(\omega)$ by means of

$$\sigma(h) = \int_{-\pi}^{\pi} \cos \omega h \, d F(\omega) \tag{14}$$

We shall take the integration to be over the closed interval $[-\pi,\pi]$.

Theorem 1. Let $\sigma(h)$ be the autocovariance of a second order stationary process. There exists a unique non-negative non-decreasing function $F(\omega)$ defined over the closed interval $[-\pi,\pi]$, (i) satisfying (14), (ii) having $F(\pi) = \sigma(o)$, (iii) being continous from the right, (iv) having symmetric increments i.e. $F(\omega_1) - F(\omega_2) = F(-\omega_2) - F(-\omega_1)$ for any continuity points $\omega_1,\omega_2,-\omega_1,-\omega_2$. Furthermore (v)

$$F(\omega) = \frac{\sigma(o)}{2\pi}(\pi+\omega) + \frac{1}{\pi} \sum_{h=1}^{\infty} \frac{\sigma(h)}{h} \sin \omega h \tag{15}$$

for any continuity point $\omega$ of $F$||.

Obviously (15) indicates that the spectral density may be given by

$$f(\omega) = \frac{\sigma(o)}{2\pi} + \frac{1}{\pi} \sum_{h=1}^{\infty} \sigma(h) \cos \omega h \tag{16}$$

This is true if $\Sigma|\sigma(h)| < \infty$ , because in that case we may multiply (16) by $\cos \omega h$ and integrate to obtain

$$\sigma(h) = \int_{+\infty}^{-\infty} \cos k\omega \; f(\omega) d\omega$$

Thus an $F(\omega)$ exists in this case and $F'(\omega) = f(\omega)$.

<u>Proof of the theorem:</u> Let us first prove that if a non-decreasing $F$ exists, satisfying (i),(ii),(iii) and (iv), then $F$ is unique and given by (15) in the continuity points. Since $F(\omega)$ is non-decreasing, we can expand it in a Fourier series in all continuity points (See Appendix II). The sine Fourier coefficients for $F$ over $[-\pi,\pi]$ are

$$b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} F(\omega) \sin \omega k \, d\omega = \frac{1}{\pi} \int_{-\pi}^{\pi} \sin \omega k \int_{-\pi}^{\pi} I_{(\nu \leqq \omega)}(\nu,\omega) dF(\nu) d\omega$$

$h=1,2,\ldots$, where $I_S(\nu,\omega)$ denotes the indicator function for a set $S$. We get

$$\begin{aligned} b_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \sin \omega k \; I_{(\nu \leqq \omega)}(\nu,\omega) d\omega \, dF(\nu) = \\ &= \frac{1}{\pi} \int_{-\pi}^{\pi} \int_{\nu}^{\pi} \sin \omega k \, d\omega \, dF(\nu) = \\ &= -\frac{1}{\pi k} \int_{-\pi}^{\pi} ((-1)^k - \cos \nu k) \, dF(\nu) = \frac{1}{\pi k}(\sigma(k) - (-1)^k \sigma(o)) \end{aligned} \tag{17}$$

using (14). For the cosine Fourier coefficients we get in the same manner for $h \geqq 0$

$$a_h = \frac{1}{\pi} \int_{-\pi}^{\pi} F(\omega) \cos \omega h \, d\omega = -\frac{1}{\pi k} \int_{-\pi}^{\pi} \sin \nu k \; dF(\nu) = 0$$

making use of (iv). Furthermore by (iv)

$$a_o = \frac{1}{\pi} \int_{-\pi}^{\pi} F(\omega) d\,\omega = 0$$

Hence

$$F(\omega) = \sum_{h=1}^{\infty} b_k \sin \omega\, k \tag{18}$$

where $b_k$ is given by (17). Writing down this expansion in the particular case when $F(\omega) = \frac{\sigma(o)}{2\pi}(\pi+\omega)$ noting that then $\sigma(h) = 0$ (by (14)), and subtracting from (18) we get (15). Now $F(\omega)$ is given for an arbitrary $\omega$ by $F(\omega) = F(\omega+o)$.

It remains to prove the existence of $F$.

For that purpose we introduce

$$c_h = \frac{1}{T-h} \sum_{t=1}^{T-h} Y(t)Y(t+h) \tag{19}$$

which it is now natural to use instead of the $c_h$ in (10) since $E\,Y(t) = 0$. Furthermore also in the definitions of $R^2(\omega)$ and $I_T(\omega)$ we now use $Y(t)$ in place of $Y_t - \bar{Y}$, obtaining

$$\begin{aligned} I_T(\omega) &= \frac{T}{8\pi} R^2(\omega) = \frac{T}{8\pi} \left[ (\Sigma Y(t) \cos \omega\, t)^2 + (\Sigma Y(t) \sin \omega\, t)^2 \right] \\ &= \frac{1}{2\pi T} \left| \Sigma Y(t) e^{i\omega t} \right|^2 \end{aligned} \tag{20}$$

(see (5) and (4)). Then (13) is still valid, it is only in the development from (9) to (11) to replace $Z_t$ by $Y(t)$.

Taking the expected value in the new (13) we get

$$f_T(\omega) = E\, I_T(\omega) = \frac{1}{2\pi} \sum_{-(T-1)}^{T-1} \left(1 - \frac{|h|}{T}\right) \sigma(h) \cos \omega\, h \tag{21}$$

since by (19) $E\,c_h = \sigma(h)$. Multiplying (21) by $\cos \omega\, k$ and integrating, we get (see (12))

$$\int_{-\pi}^{\pi} f_T(\omega) \cos \omega\, h\, d\omega = \begin{cases} \sigma(h)\left(1 - \frac{|k|}{T}\right) & \text{if } |k| < T \\ 0 & \text{if } |k| \geq T \end{cases}$$

We introduce

$$F_T(\omega) = \begin{cases} \sigma(o) & \text{if} \quad \omega \geqq \pi \\ \int_{-\pi}^{\omega} f_T(\lambda) d\lambda & \text{if} \quad -\pi \leqq \omega \leqq \pi \\ 0 & \text{if} \quad \omega \leqq -\pi \end{cases} \tag{22}$$

and have

$$\int_{-\pi}^{\pi} \cos \omega k \, d\, F_T(\omega) = \begin{cases} \sigma(k)(1 - \frac{|k|}{T}) & \text{if } |k| < T \\ 0 & \text{if } |k| \geqq T \end{cases} \tag{23}$$

Obviously $\frac{1}{\sigma(o)} F_T(\omega)$ is a cumulative distribution function. We make use of two famous limit theorems in probability theory. First we know that there exist a non-decreasing function $\frac{1}{\sigma(o)} F(\omega)$, continuous from the right and a sequence $T_1, T_2, \ldots$ such that $\lim_{j\to\infty} \frac{1}{\sigma(o)} F_{T_j}(\omega) = \frac{1}{\sigma(o)} F(\omega)$ in all continuity points of F. It is seen from (22) that $\frac{1}{\sigma(o)}$ F($\omega$) is a cumulative distribution function equal to 1 if $\omega \geqq \pi$. Then we also know that for any bounded continuous $\phi(\omega)$ we have (dropping $1/\sigma(o)$)

$$\lim \int_{-\infty}^{+\infty} \phi(\omega)\, d\, F_{T_j}(\omega) = \int_{-\infty}^{+\infty} \phi(\omega)\, d\, F(\omega).$$

In particular this is true with $\phi(\omega) = \cos \omega k$. Hence, replacing T by $T_j$ in (23) and going to the limit we get

$$\sigma(k) = \int_{-\pi}^{\pi} \cos \omega k \, d\, F(\omega)$$

Hence we have found an F($\omega$) satisfying (14). It has symmetric increments since $F_T(\omega)$ has symmetric increments by (22) and (21). Therefore everything is proved.

It should be noted that it follows from the development above that

$$\lim_{T\to\infty} E \int_{-\pi}^{\omega} I_T(\lambda) d\lambda = F(\omega)$$

given by (14) or (15). (We use the same argument as in the probability theory to show that $T = T_1, T_2, \ldots$ could be replaced by $T = 1, 2, \ldots$ .) Hence if search for periodicities in Y(1),...,Y(T) motivates studying $I_T(\omega)$, then search for periodicites in the process $\{Y(t)\}$ motivates studying $F(\omega)$, or $f(\omega) = F'(\omega)$ if existing.

We refer to T.W. Anderson (1971) p.374-379, 383-384 for examples of autocovarians functions $\sigma(h)$ and spectra $F(\omega)$.

## II. ALGEBRAIC SOLUTION OF THE AUTOREGRESSIVE EQUATION

Theorem 1.

For given $a_t; t > k$ ; $Z_k, Z_{k-1}, \ldots, Z_{k-p+1}$ there is a unique solution $Z_t; t > k$ of the difference equation

$$Z_t = \sum_{j=1}^{p} \phi_j Z_{t-j} + a_t \tag{1}$$

To find this solution proceed as follows.

1. Let the characteristic roots, i.e. the roots of the characteristic algebraic equation

$$\phi_p B^p + \phi_{p-1} B^{p-1} + \ldots + \phi_1 B = 1 \tag{2}$$

be $H_j^{-1} e^{\pm\alpha_j i} (0 \leq \alpha_j < 2\pi)$; $j=1,2,\ldots,q$, with multiplicities $m_1, m_2, \ldots, m_q$ respectively. Then set

$$c_t(A) = \sum_{j=1}^{q} \left[ Q_{m_j-1}(t) \cos \alpha_j t + R_{m_j-1} \sin \alpha_j t \right] H_j^t \tag{3}$$

where $Q_\nu$, $R_\nu$ are polynomials of degree $\nu$ and $A$ denotes the vector having the $p$ coefficients of the polynomials as vectors. (The sum of the $m_j$ - each of which is counted twice if $\alpha_j \neq 0$ - should be $p$.)

2. Let $A_0$ be the solution of the system of linear equations in $A$,

$$c_{r-k}(A) = Z_r \; ; \; r=k, k-1, \ldots, k-p+1 \tag{4}$$

3. Find $\psi_1,\ldots,\psi_{p-1}$ recursively from $\psi_0 = 1$,

$$\psi_r = \sum_{j=1}^{r} \phi_j \psi_{r-j} \; ; \; r=1,2,\ldots p-1 \tag{5}$$

4. Let $D_0$ be the solution of the system of linear equations in $D$,

$$c_j(D) = \psi_j \; ; \; j=1,2,\ldots p-1 \tag{6}$$

and set

$$\psi_j = c_j(D_0) \; ; \; j=p,p+1,\ldots \tag{7}$$

Alternatively $\psi_j$ is given by the recursion

$$\psi_r = \sum_{j=1}^{p} \phi_j \psi_{r-j} \; ; \; r=p,p+1,\ldots \tag{8}$$

or by the formal identity in $B$,

$$\psi(B) = \sum_{j=o}^{\infty} \psi_j B^j = \frac{1}{\phi(B)} \quad \text{or} \quad \phi(B)\psi(B) = 1 \tag{9}$$

where

$$\phi(B) = 1 - \sum_{j=1}^{p} \phi_j B^j \tag{10}$$

5. Then the unique solution is

$$Z_t = c_{t-k}(A_0) + \sum_{i=k+1}^{t} a_i \psi_{t-i} = c_{t-k}(A_0) + V_{tk} \tag{11}$$

(Note that the difference equation for determining $c_t$ is the same as for determining $\psi_j$; only the initial conditions are different.)

6. $V_{th}$ satisfies the difference equation for every $h$ and corresponds to the case when

$$Z_{k-p+1} = \ldots = Z_{k-1} = Z_k = 0$$

Proof: The uniqueness follows recursively from (1).

By a "general" solution of (1) we mean a solution containing arbitrary constants so that it can be adjusted uniquely to any value of $Z_k, Z_{k-1}, \ldots, Z_{k-p+1}$. Now, if $C_t$ is the general solution of

$$C_t = \sum_{j=1}^{p} \phi_j C_{t-j} \tag{12}$$

and $V_t$ is any special solution of (1), then it is seen that $Z_t = C_t + V_t$ is the general solution of (1). Hence we shall first find a $V_t$. We shall verify that

$$V_t = 0 \; ; \; t \leqq k \; ; \quad V_t = \sum_{i=k+1}^{t} a_i \psi_{t-i} \; ; \; t \geqq h+1 \tag{13}$$

where

$$\psi_r = \sum_{j=1}^{\min(p,r)} \phi_j \psi_{r-j} \; ; \; \psi_0 = 1 \tag{14}$$

is a special solution of (1).

For that purpose we write $\psi_j = 0 \; ; \; j < 0 \; ;$ and $\phi_j = 0 \; ;$ $j > p$. Then (1), (13) and (14) may be written

$$Z_t = \sum_{j=1}^{\infty} \phi_j Z_{t-j} + a_t \tag{1)'}$$

$$V_t = \sum_{i=k+1}^{\infty} a_i \psi_{t-i} \tag{13)'}$$

$$\psi_r = \sum_{j=1}^{\infty} \phi_j \psi_{r-j} \tag{14)'}$$

Substituting (13)' in the right hand side of (1)' we obtain

$$\sum_{j=1}^{\infty} \phi_j \sum_{k+1}^{\infty} a_i \psi_{t-j-i} + a_t = \sum_{k+1}^{\infty} a_i \sum_{j=1}^{\infty} \phi_j \psi_{t-i-j} + a_t$$

However, if $i \geqq t$ then $\psi_{t-i-j}$ equals 0, hence we obtain

$$\sum_{i=k+1}^{t-1} a_i \sum_{j=1}^{\infty} \phi_j \psi_{t-i-j} + a_t = \sum_{i=k+1}^{t-1} a_i \psi_{t-i} + a_t$$

making use of (14)'. But by (13) this equals $V_t$ and the verification is completed.

Let us now consider the general solution of (12). It is easily verified that

$$C_t = (-t)^{(\nu)}G^t \ ; \ \nu=0,1,\ldots,n \tag{15}$$

where $g = G^{-1}$ is a (complex or real) root of $\phi(B) = 0$ of multiplicity $m$, are solutions of (12). ($x^{(\nu)}$ denotes $x(x-1)\ldots(x-\nu+1)$, $x^{(o)} = 1$). Because we have for the right hand side of (12), substituting (15)

$$\sum_{j=1}^{p} \phi_j (j-t)^{(\nu)}G^{t-j} = g^{\nu} \sum_{j=1}^{p} \phi_j (j-t)^{(\nu)} g^{j-t-\nu} =$$

$$= g^{\nu}\frac{d^{\nu}}{dg^{\nu}}[g^{-t}(1-\phi(g))] = (-t)^{(\nu)}G^t = C_t$$

(using the multiplication rule for differentiation of the product in the bracket). Obviously then also

$$C_t = t^{\nu}G^t \ ; \ \nu=0,1,2,\ldots,n \tag{16}$$

are solutions of (12), since they are linear combinations of the functions (15).

Thus if $G_1^{-1},G_2^{-1},\ldots,G_q^{-1}$ are the roots of $\phi(B) = 1$ of orders $n_1,\ldots,n_q$, respectively, then

$$C_{t-k} = \sum_{j=1}^{q} P_{n_j-1}(t-k)G_j^{\,t-k} \tag{17}$$

is a solution of (12), where $P_{n_j-1}$ ; $j=1,2,\ldots,q$ ; denote arbitrary polynomials of degrees $n_j-1$ ; $j=1,2.,,,.q$.

We shall *now show that*

$$Z_t = C_{t-k} + V_t$$

*where* $C_{t-k}$ and $V_t$ *are given by (17) and (13) is the general solution of (1).* Thus we have to show that the coefficient of the polynomials are determined uniquely for any $Z_{h-p+1},\dots,Z_h$ by $C_{t-k}+V_t = Z_t$ ; $t=k-p+1,\dots,k$ ; i.e. by

$$\sum_{j=1}^{q} P_j(t-k)G_j^{t-k} = Z_t \;;\; t=h-p+1,\dots,k \tag{18}$$

since $V_t = 0$ ; $t \leq k$. Here we write $P_{n_j-1} = P_j$ for convenience. However, it is known from the theory of linear equations that (18) has a unique solution if and only if the unique solution of

$$\sum_{j=1}^{q} P_j(t-h)G_j^{t-k} = 0 \;;\; t=k-p+1,\dots,k \tag{19}$$

is that all coefficients of the polynomials are 0, hence $P_j = 0$. We write (19)

$$\sum_{j=1}^{q} P_j(v)G_j^{v} = 0 \;;\; v=-p+1,\dots,-1,0 \tag{20}$$

To prove that such is the case, assume that there exists a non-zero solution

$$\sum_{j=1}^{r} P_j(v)G_j^{v} = 0 \;;\; v=-p+1,\dots,-1,0 \tag{21}$$

where the degrees of the $P_j$ are $N_j-1 \leq n_j-1$ ; $j=1,\dots,r$ ; and where we leave out polynomials which are 0 ; $r \leq q$. (Below we follow a type of proof in Edouard Goursat (1933),15.ed. Tome II,p.456-7.It is much simpler than the customary proof using van der Monde determinants.) We get from (21)

$$P_1(v) + \sum_{j=2}^{r}\left(\frac{G_j}{G_1}\right)^{v} P_j(v) = 0 \;;\; j=-p+1,\dots,-1,0 \tag{22}$$

Replacing $v$ by $v+1$ and subtracting we get

$$\Delta P_1(v) + \sum_{j=2}^{r}\left(\frac{G_j}{G_1}\right)^{v}\left[P_j(v+1)\frac{G_j}{G_1} - P_j(v)\right] = 0 \tag{23}$$

or

$$\Delta P_1(v) + \sum_{j=2}^{r} g_j^{v}Q_j(v) = 0 \;;\; j=-p+1,\dots,-1 \tag{24}$$

where the $Q_j$ are polynomials of degrees $N_j-1$ ; $j=2,\dots,r$ ; but $\Delta P_1(v)$ has degree $N_1-2$. (We have used the notation $\Delta f(v) = f(v+1)-f(v)$). Note that now $v$ goes to $-1$ (not 0) because of the argument $v+1$ in (23).

The operation above on (22) to obtain (24) is now repeated recursively to obtain

$$\Delta^{N_1} P_1(v) + \sum_{j=2}^{r} g_j^v R_j(v) = 0$$

or

$$\sum_{j=2}^{r} g_j^v R_j(v) = 0 \quad ; \quad v=-p+1,\dots,-N_1. \tag{25}$$

since $\Delta^{N_1} P_1(v) = 0$. Note that $v$ goes to $-N_1$.

We now repeat recursively r-1 times on (25) the whole procedure used on (21), to get

$$h^r S(v) = 0 \quad ; \quad v=-p+1,\dots,-\sum_{j=1}^{r-1} N_j \tag{26}$$

where $h \neq 0$ and $S(v)$ is of degree $N_r-1$. However, this polynomial is by (26) 0 for

$$p - \sum_{j=1}^{r-1} N_j = N_r + p - \sum_{j=1}^{r} N_j \geq N_r + p - \sum_{j=1}^{q} n_j = N_r$$

values, *which is a contradiction*. Hence everything is proved. Now all statements in the theorem follow easily. This completes the proof.

Note that by (4) the coefficients in the polynomials in (3) are independent of k.

Example 1. Let

$$Z_t = Z_{t-1} - Z_{t-2} + a_t \quad ; \quad t=1,2,\dots$$

with given values of $Z_o, Z_{-1}$ and $a_t$. Recursively we find

$$Z_1 = Z_o - Z_{-1} + a_1$$

$$Z_2 = -Z_{-1} + a_1 + a_2$$

$$Z_3 = -Z_o + a_2 + a_3$$

$$Z_4 = -Z_o + Z_{-1} - a_1 + a_2 + a_3$$

$$Z_5 = Z_{-1} - a_1 - a_2 + a_4 + a_5$$

It seems hard to find any general pattern even if we write down many $Z_t$.

Let us then make use of Theorem 1 with $k = 0$. The solution of the characteristic equation $-B^2 + B = 1$ gives the roots

$$\tfrac{1}{2}(1 \pm i\sqrt{3}) = e^{\pm \frac{\pi}{3} i}$$

Hence by (3)

$$C(A_1, A_2) = A_1 \cos \frac{\pi}{3} t + B \sin \frac{\pi}{3} t$$

From (4) we get

$$A_1 = Z_o \quad , \quad A_2 = -\frac{1}{\sqrt{3}}(2Z_{-1} - Z_o)$$

$$C_t = Z_o \cos \frac{\pi}{3} t - \frac{1}{\sqrt{3}}(2Z_{-1} - Z_o) \sin \frac{\pi}{3} t$$

From (5) we get $\psi_o = 1$, $\psi_1 = 1$ and from (7)

$$\psi_r = D_1 \cos \frac{\pi}{3} r + D_2 \sin \frac{\pi}{3}$$

(satisfying (8), i.e. $\psi_r = \psi_{r-1} - \psi_{r-2}$). From $\psi_o = \psi_1 = 1$ we then get $D_1 = 1$, $D_2 = \frac{1}{\sqrt{3}}$

$$\psi_r = \cos \frac{\pi}{3} r + \frac{1}{\sqrt{3}} \sin \frac{\pi}{3} r$$

Hence we have the solution (by (11))

$$Z_t = Z_o\left( \cos \frac{\pi}{3} t + \frac{1}{\sqrt{3}} \sin \frac{\pi}{3} t\right) - Z_{-1} \frac{2}{\sqrt{3}} \sin \frac{\pi}{3} t + \sum_{j=1}^{t} a_j \left[ \cos \frac{\pi}{3}(t-j) + \frac{1}{\sqrt{3}} \sin \frac{\pi}{3}(t-j)\right]$$

Note that the last sum from $j = 1$ to $j = t$ equals

$$\sum_{m=o} (a_{t-6m}+a_{t-6m-1}-a_{t-6m-3}-a_{t-6m-4})$$

$$= (a_t+a_{t-1}-a_{t-3}-a_{t-4}) + (a_{t-6}+a_{t-7}-a_{t-9}-a_{t-10})$$

$$+ \ldots$$

until the last term $a_{t-j}$ for which the subscript is $\geq 1$. The two first terms have one of the forms $Z_o-Z_{-1}$, $Z_{-1}$, $-Z_o$, $-Z_o+Z_{-1}$, $Z_{-1}$, $Z_o$. Thus the general pattern has been discovered.

Example 2. Let

$$Z_t = Z_{t-1} - Z_{t-2} + a_t \quad ; \quad t=1,2,\ldots$$

with given values of $Z_o$, $Z_{-1}$ and $a_t$. Using the equation for $Z_t$ recursively it is hard to find any general pattern for $Z_t$ expressed by means of $Z_o$, $Z_{-1}$, $a_1 a_2, \ldots$ .

We denote the characteristic roots by $G_1^{-1}$, $G_2^{-1}$. They are the solutions of $B^2 + B = 1$, hence

$$G_1 = \tfrac{1}{2}(\sqrt{5}+1) \quad , \quad G_2 = \tfrac{1}{2}(-\sqrt{5}+1)$$

We have

$$C_t = A_1 G_1^t + A_2 G_2^t$$

Hence by (4)

$$A_1 = \frac{1}{\sqrt{5}} (Z_{-1}+G_1 Z_o) \quad , \quad A_2 = \frac{1}{\sqrt{5}}(-Z_{-1}-G_2 Z_o)$$

$$C_t = \frac{1}{\sqrt{5}}\left[(G_1^t-G_2^t)\, Z_{-1} + (G_1^{t+1}-G_2^{t+1})\, Z_o\right]$$

We find $\psi_r$ from $\psi_o = \psi_1 = 1$ and

$$\psi_r = D_1 G_1^r + D_2 G_2^r$$

(see (6) and (7). This gives

$$D_1 = G_1/\sqrt{5} \quad , \quad D_2 = - G_2/\sqrt{5} \quad , \quad \psi_r = G_1^{r+1}-G_2^{r+1})/\sqrt{5}$$

Making use of

$$\frac{1}{\sqrt{5}}(G_1^n - G_2^n) = (\tfrac{1}{2})^{n-1} \sum_{i=o}^{n} \binom{n}{2i+1} 5^i$$

We obtain the solution

$$Z_t = Z_o(\tfrac{1}{2})^t \sum_{i=o}^{t+1} \binom{t+1}{2i+1} 5^i + Z_{-1}(\tfrac{1}{2})^{t-1} \sum_{i=o}^{t+1} \binom{t}{2i+1} 5^i +$$

$$+ \sum_{j=1}^{t} a_j \psi_{t-j}$$

where

$$\psi_r = (\tfrac{1}{2})^r \sum_{i=o}^{r+1} \binom{r+1}{2i+1} 5^i$$

## III. THE STATIONARY AUTOREGRESSIVE ANTEIMPULSE-GENERATED PROCESS

### A. The general AR-process

Consider a (second order) stationary process $Y_t$ ; $t=\ldots-2,-1,0,1,2,\ldots$ ; generated by another stationary process $a_t$ ; $t=\ldots-2,-1,0,1,2,\ldots$ by means of the difference equation

$$Y_t - \eta = \sum_{j=1}^{p} \phi_j (Y_{t-j} - \eta) + a_t \qquad (1)$$

$t=\ldots-2,-1,o,1,2,\ldots$ where $E\,Y_t = \eta$. We shall have in mind situations where the $a_t$ are unobserved impulses (hence $E\,a_t = 0$), the past and present values $(a_j\ ;\ j \leq t)$ of which influence the observed time series $Y_t$. The process is then said to be ante-impulse-generated. Of course it is only in that case that the process is meaningful as a timeseries. Since we are not going to treat the inference problems, we may concentrate our attention on $Z_t = Y_t - n$. We shall take care to make an assumption ((ii) in Theorem 1 below) which secures that the process is really anteimpulse-generated. In the later chapters we shall investigate to which extent this assumption is also necessary.

We shall investigate the existence and uniqueness of $Z_t$ defined by means of (1) and we shall study how to express $Z_t$

explicitly by means of $a_\tau$ ; $\tau \leq t$.

We shall call such a process a general AR (p)-process and we shall later use the same name even if the process is not anteimpulse-generated). AR indicates that the process is autoregressive (i.e. given by (1)).

Two special cases are important in practical application. The first one is the case when the $a_t$ are uncorrelated (or independent). In that case the process is usually referred to as an AR(p) process. The second case is the situation when $a_t$ is given by

$$a_t = a_t' - \sum_{i=1}^{q} \theta_i a_{t-i}'$$

where $\{a_t'\}$ is an uncorrelated process ($E\, a_t' = 0$). This process is referred to as an ARMA (p,q), since the $a_t$ are moving averages (MA) of the $a_t'$. We shall include the case when $q = -\infty$.

The concept of convergence in quadratic mean is used extensively below. For the properties of this concept the reader is referred to Appendix I.

Unless something else is stated, we shall assume everywhere below that "stationary" means "second order stationary" and "convergence" means convergence in quadratic means.

<u>Theorem 1</u>. If

(i) $a_t$ ; $t = \ldots -1, 0, 1, \ldots$ is *second order stationary*, $E\, a_t = 0$,

(ii) the equation

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^D = 0 \qquad (2)$$

has ***all roots outside the unit circle***, then the only stationary process $Z_t$ satisfying

$$Z_t = \sum_{j=1}^{p} \phi_j Z_{t-j} + a_t \;;\; t = \ldots -1, 0, 1, \ldots \qquad (1)'$$

(with $E\, Z_t = o$) is

$$Z_t = \sum_{j=o}^{\infty} a_{t-j}\psi_j = \sum_{i=-\infty}^{t} a_i\psi_{t-i}, \tag{3}$$

where the convergence of the series is in quadratic mean and the $\psi_j$ are defined recursively by

$$\psi_o = 1, \ \psi_r = \sum_{j=1}^{\min(p,r)} \phi_j \psi_{r-j} \ ; \ j=1,2,... \tag{4}$$

or more conveniently from the power series identity

$$\phi(B)\psi(B) = 1 \quad \text{where} \quad \psi(B) = \sum_{o}^{\infty} \psi_j B^j \tag{5}$$

If in addition the $a_t$ are *independent* then $Z_t$ is given by (3) with convergence with probability 1.

Note that the theorem really gives a property of an *arbitrary stationary* process $Z_t$. If for such a process $a_t$ is defined by (1)', then it is stationary and $Z_t$ is given uniquely by (3) and (4).

Proof: We assume (i) and (ii) and introduce

$$V_{tk} = \sum_{k+1}^{t} a_i\psi_{t-i} \quad , \quad V_t = \sum_{-\infty}^{t} a_i\psi_{t-i} \tag{6}$$

All the poles of

$$\psi(B) = \frac{1}{\phi(B)} = \sum_{i=o}^{\infty} \psi_i B^i \tag{7}$$

are outside the unit circle (by (ii)). Hence (7) converges for $|B| \leq$ minimal absolute value of the roots of $\phi(B) = 0$, hence for $|B| = 1$, hence $\Sigma|\psi_i|$ converges. By Lemma 13 in Appendix I, the series in (6) defining $V_t$ converges in quadratic mean and is stationary. Now by the theorem of Chapter II, $V_{tk}$ satisfies (1)' and hence

$$V_{tk} = \sum_{j=1}^{p} \phi_j V_{t-jk} + a_t \tag{8}$$

Letting $k \to \infty$ we obtain that $V_t$ satisfies (1)'.

To prove that $V_t$ is the only stationary solution we have to prove that if $Z_t$ is stationary and satisfies (1)', then $Z_t = V_t$.

By the theorem of Chapter II we know that

$$Z_t = c_{t-k}(A_o) + V_{tk} \tag{9}$$

where $c_{t-k}$ is given by II·(3) and $A_o = (A_{o1}, A_{o2}, \dots)$ are the vector of the coefficients in the polynomials in II.(3) determined such that

$$c_j(A_o) = Z_{k+j} \ ; \quad j=0,-1,\dots,-p+1 \tag{10}$$

(see II·(4)). The left hand side of (10) is linear in $A_o$ and does not contain $k$. The right hand side have joint first and second order moments independent of $k$, since $Z_t$ is stationary. Hence so have the components of $A_o$. $c_{t-k}$ may be written as a sum of terms of the form

$$(t-k)^m \ A_{oi} \ H_j^{t-k} \ K_j(t-k) \tag{11}$$

where $K_j(t-k) = \cos\alpha(t-k), \sin\alpha(t-k)$ and $|H_j| < 1$ (by (ii)).

Hence

$$c_{t-h}(A_o) = \sum_{j=1}^{p} \alpha_{jk} \ A_{oj} \tag{12}$$

where $\alpha_{jk} \to 0$ ; $j=1,2,\dots,p$ ; as $k \to \infty$. By Lemma 11 in the Appendix I,

$$\operatorname{var} c_{t-k}(A_o) \leq \left( \sum_{j=1}^{p} |\alpha_{jk}| \right)^2 \ \max_i \operatorname{var} A_{oi} \to 0 .$$

Therefore $c_{t-k}(A_o) \xrightarrow{2} 0$ and $Z_t \xrightarrow{2} V_t$. However $Z_t$ is independent of $k$, hence $Z_t = V_t$ and we have proved the first statement in the theorem. The second statement follows from lemma 10 in the appendix.

B. <u>The ARMA process</u>

We shall apply Theorem 1 to the special case when $a_t$ is a moving average process with infinite many terms in the average. We denote the impulse by $A_t$, which has the form

$$A_t = a_t - \sum_{i=1}^{\infty} \theta_i a_{t-i} \tag{13}$$

In the moving average process the $a_t$ should be uncorrelated and stationary. However, we need only to assume that $a_t$ is a stationary process.

Theorem 2. If

(i) $a_t$ ; $t=\ldots-1,0,1,\ldots$ ; is second order stationary,

(ii) the equation

$$\phi(B) = 1-\phi_1 B-\cdots-\phi_p B^p = 0$$

has all roots outside the unit circle,

(iii) $$\sum_{i=1}^{\infty} |\theta_i| < \infty \tag{13}$$

then the only stationary solution satisfying

$$Z_t = \sum_{j=1}^{p} \phi_j Z_{t-j} + a_t - \sum_{i=1}^{\infty} \theta_i a_{t-i} \tag{14}$$

is

$$Z_t = \sum_{j=o}^{\infty} a_{t-j}\psi_j \tag{15}$$

where the convergence is in quadratic mean and the $\psi_j$ are given by the power series identity

$$\psi(B)\phi(B) = \theta(B) \tag{16}$$

where

$$\psi(B) = \sum_{j=o}^{\infty} \psi_j B^j \quad , \quad \theta(B) = 1-\sum_{j=1}^{\infty} \theta_j B^j \tag{17}$$

Proof: Note that if $\tilde{\psi}_j$ is defined by

$$\phi(B)\tilde{\psi}(B) = 1 \quad , \text{ where } \tilde{\psi}(B) = \sum_{j=o}^{\infty} \tilde{\psi}_j B^j \tag{18}$$

then

$$\psi(B) = \frac{\theta(B)}{\theta(B)} = \theta(B)\tilde{\psi}(B)$$

and hence

$$\psi_o = \tilde{\psi}_o \quad , \quad \psi_j = -\sum_{i=o}^{j} \theta_{j-i}\tilde{\psi}_i \tag{19}$$

Obviously $A_t$ given by (13) is a wide sense stationary process by Lemma 13 of Appendix I. We apply Theorem 1 to

$$Z_t = \sum_{j=1}^{p} \phi_j Z_{t-j} + A_t \tag{20}$$

We obtain

$$Z_t = \sum_{i=o}^{\infty} A_{t-i}\tilde{\psi}_i \tag{21}$$

This may be written

$$Z_t = -\sum_{i=o}^{\infty} \tilde{\psi}_i \sum_{j=o}^{\infty} \theta_j a_{t-i-j} = -\sum_{i=o}^{\infty} \sum_{j=-\infty}^{t-i} \tilde{\psi}_i \theta_{t-i-j} a_i \tag{22}$$

We now apply Lemma 14 in the Appendix to this $Z_t \equiv Z$ where $\psi_{ij}$ in the Lemma is given by

$$\begin{aligned} \psi_{ij} &= \tilde{\psi}_i \theta_{t-j-i} \; ; \; i \leq t \quad , \quad j \leq t-i \\ \psi_{ij} &= 0 \qquad \text{otherwise} \end{aligned} \tag{23}$$

to obtain

$$Z_t = -\sum_{j=-\infty}^{\infty} a_j \sum_{i=-\infty}^{\infty} \psi_{ij} = -\sum_{j=-\infty}^{t} a_j \sum_{i=o}^{t-j} \tilde{\psi}_i \theta_{t-j-i} \tag{24}$$

We only have to check the assumptions 2), 3) and 4) in Lemma 14. Assumption 2) is obviously true by Assumption (iii) of the Theorem. Assumption 3) of the Lemma is true because

$$\sum_{j=o}^{\infty} \left| \sum_{j=-\infty}^{t-i} \tilde{\psi}_i \theta_{t-i-j} \right| \leq \sum_{j=o}^{\infty} |\tilde{\psi}_i| \sum_{j=-\infty}^{t-i} |\theta_{t-i-j}| =$$

$$= \sum_{j=o}^{\infty} |\tilde{\psi}_i| \sum_{k=o}^{\infty} |\theta_k| < \infty$$

since $\Sigma|\tilde{\psi}_i|$ converges by the proof of Theorem 1 (all roots of $\phi(B) = 0$ are outside the unit circle). Assumption 4) is true by (22) with

$$Y_i = -\sum_{j=-\infty}^{t-i} \tilde{\psi}_i \theta_{t-i-j} a_i$$

Combining (24) and (19) we obtain (15) and Theorem 2 is proved.

Note that the solution (15) of (14) can be constructed by the following "formal" procedure. Let $B$ denote the backward shift operator $B Z_t = Z_{t-1}$. Then $B^j Z_t = Z_{t-j}$ and (14) may be written

$$\phi(B) Z_t = \theta(B)\, a_t \qquad (14)'$$

where $\theta(B)$ is a power series given by (17). Now operating on (14)' as if $B$ was a number, we get

$$Z_t = \frac{\theta(B)}{\phi(B)}\, a_t = (1 + \psi_1 B + \psi_2 B^2 + \ldots)\, a_t$$

$$= a_t + \psi_1 B\, a_t + \psi_2 B^2 a_t + \ldots$$

having developed the function $\theta(B)/\phi(B)$ in a power series. Now returning to the original interpretation of $B$ as an operator we get $B^j a_t = a_{t-j}$ from which (15) is obtained.

C. Autocovariance and spectrum of second of stationary ante-impuls-generated AR and ARMA processes.

We now assume that the $a_t$ are uncorrelated with $E\, a_t = 0$ and $\operatorname{var} a_t = \sigma^2$.

Let us first consider any stationary anteimpulse-generated process

$$Z_t = \sum_{j=o}^{\infty} \psi_j a_{t-j} \quad , \quad \psi_o = 1 \quad , \quad \Sigma |\psi_j| < \infty \qquad (25)$$

where the convergence is in quadratic mean,

$$Z_{tm} = \sum_{j=1}^{m} \psi_j a_{t-j} \xrightarrow{2} Z_t \qquad (26)$$

as $m \to \infty$.

We refer below to the lemmas in Appendix I. We have from Lemma 3:

$$E\, Z_t = 0 \quad , \operatorname{var} Z_t = E\, Z_t^2 = \sigma_a^2 \sum_{j=o}^{\infty} \psi_j^2 \qquad (27)$$

Applying Lemma 3 to $E\, Z_{tm} a_{t+k}$ we get

$$E\,Z_t a_{t+k} = 0 \quad \text{for} \quad k > 0 \;, \quad E\,Z_t a_{t+k} = \psi_k \sigma_a^2 \text{ for } k \leq 0 \qquad (28)$$

Applying Lemma 3 to $E\,Z_{tm} Z_{t-k}$ we get from (28) and (26) for the *autocovariance function (autocorrelogram)*

$$\sigma(k) = E\,Z_t Z_{t-k} = \sum_{j=o}^{\infty} \psi_j \psi_{j-k} \sigma_a^2 \qquad (29)$$

(setting $\psi_j = 0$ for $j < 0$).

Let us now consider the stationary $Z_t$ satisfying the autoregressive equation

$$Z_t = \phi_1 Z_{t-1} + \cdots + \phi_p Z_{t-p} + a_t \qquad (30)$$

where all roots of $\phi(B) = 0$ are outside the unit curcle. If we multiply (30) by $Z_{t-h}$ and make use of (28) we get for $h > 0$ and $h = 0$,

$$\begin{aligned} \sigma(h) &= \phi_1 \sigma(h-1) + \cdots + \phi_p \sigma(h-p) \quad ; \; h > 0 \\ \sigma(0) &= \phi_1 \sigma(-1) + \cdots + \phi_p \sigma(-p) + \sigma_a^2 \end{aligned} \qquad (31)$$

With $h = 1,2,\ldots,p$, these are the Yule-Walker equations. They can be solved by using the theorem of Chapter II. We get

$$\sigma(h) = \sum_{j=1}^{q} \left[ Q_{m_j-1}(t) \cos \alpha_j t + R_{m_j-1}(t) \sin \alpha_j t \right] H_j^t \qquad (32)$$

where the $p$ coefficients in the polynomials $Q$ and $R$ are determined by setting $h = 0,1,2,\ldots,p-1$ and making use of $\sigma(h) = \sigma(-h)$. Hence $\sigma(h)$ is expressed by means of $\sigma(0)$. Inserting in the second equation (31) we can express $\sigma(0)$ by means of $\sigma_a^2$. Hence we have obtained the autocovariance function in terms of the structural parameters $\phi_1,\ldots,\phi_p$ and $\sigma_a$.

Example: $\quad Z_t = \phi Z_{t-1} + a_t \;; \; |\phi| < 1$

We get

$$\sigma(h) = \phi\,\sigma(h-1) \quad \text{for} \quad h > 0$$

hence

$$\sigma(h) = \phi^h \sigma(0)$$

Furthermore

$$\sigma(0) = \phi\,\sigma(-1) + \sigma_a^2$$

Since $\sigma(-1) = \sigma(1)$ we get

$$\sigma(0) = \phi\,\phi\,\sigma(0) + \sigma_a^2$$

Hence $\sigma(0) = \frac{1}{1-\phi^2}\,\sigma_a^2$ and

$$\sigma(h) = \phi^h \sigma_a^2/(1-\phi^2)$$

for $h \geqq 0$.

The Yule-Walker equations (31) are often used to find estimates of $\phi_1,\ldots,\phi_p,\sigma_a^2$. The estimates

$$c(h) = \frac{1}{T-h} \sum_{t=1}^{T-h} Z_t Z_{t+h}$$

$$c(0) = \frac{1}{T} \sum_{t=1}^{T} Z_t^2 \qquad (33)$$

are used for $\sigma(h)$ and $\sigma_a^2$ in (31) and (31) is solved for $\phi_1,\ldots,\phi_p,\sigma_a$. (If $E\,Z_t \neq 0$, then $Z_t$ is replaced by $Z_t - \bar{Z}$ in (33), where $\bar{Z} = \frac{1}{T}\sum_{t=1}^{T} Z_t$).

We now consider the case of an ARMA $(p,\infty)$ process (14). It follows from (15) and (28) that $Z_t$ and $a_{t+h}$, $h > 0$, are still uncorrelated. We introduce the cross-autocovariance function

$$\gamma(h) = E\,Z_t a_{t-h} \qquad (34)$$

which is 0 for $h < 0$. Multiplying (14) by $a_{t-h}$ and taking expected value we get

$$\gamma(h) = \phi_1\gamma(h-1) + \phi_2\gamma(h-2) + \ldots + \phi_p\gamma(h-p) + \theta_h\sigma_a^2\ ;\ h > 0 \qquad (35)$$

$$\gamma(0) = \sigma_a^2$$

From these equations $\gamma(h)$ can be expressed recursively by means of $\phi_i$, $\theta_i$, $\sigma_a^2$. It is seen that $\gamma(h)$ is "stationary", i.e. independent of $t$.

Let us now multiply (14) by $Z_{t-h}$ and take expected value. We then get

$$\sigma(h) = \phi_1\sigma(h-1)+\cdots+\phi_p\sigma(h-p) - \theta_1\gamma(-h+1) + \theta_2\gamma(-h+2)+\ldots \qquad h>0 \tag{36}$$

$$\sigma(0) = \phi_1\sigma(-1)+\cdots+\phi_p\sigma(-p) + \sigma_a^2 + \theta_1\gamma(1) + \theta_2\gamma(2)+\cdots \tag{37}$$

With $\gamma(h)$ determined from (35),we can obtain $\sigma(h)$ from (36) and (37), making use of $\sigma(h) = \sigma(-h)$ ; $h=1,2,\ldots,p-1$.

Example: Consider the ARMA (1,1) process

$$Z_t = \phi Z_{t-1} + a_t - \theta a_{t-1}$$

We obtain by (35)

$$\gamma(0) = \sigma_a^2 \quad , \quad \gamma(1) = \phi\gamma(0) + \theta\sigma_a^2 \quad , \quad \gamma(h) = \phi\,\gamma(h-1)$$

if $h>1$. Hence

$$\gamma(h) = \phi^{h-1}(\phi+\theta)\sigma_a^2$$

From (36) and (37) we get

$$\sigma(0) = \phi_1\sigma(-1) + \sigma_a^2 + \theta\,\gamma(1)$$

$$\sigma(1) = \phi\,\sigma(0) + \theta\,\sigma_a^2$$

Since $\sigma(-1) = \sigma(1)$ we get

$$\sigma(0) = \frac{1+\theta^2-2\phi\theta}{1-\phi^2}\,\sigma_a^2$$

From (36) we have

$$\sigma(h) = \phi\sigma(h-1) \quad ; \quad h>1$$

hence $\sigma(h) = \phi^{h-1}\sigma(1)$ and

$$\sigma(h) = \phi^{h-1}\,\frac{(1-\phi\theta)(\phi-\theta)}{1-\phi^2}\,\sigma_a^2 \quad ; \quad h>0$$

Let us now introduce the *generating function* for the autocovariance function

$$\sum(B) = \sum_{-\infty}^{+\infty} \sigma(h)\, B^h \tag{38}$$

Introducing (29) into (38) and rearranging the terms we get

$$\sum(B) = \sigma_a^2 \psi(B)\, \psi(B^{-1}) \tag{39}$$

In the case of an ARMA process (14) we have $\psi(B) = \theta(B)/\phi(B)$ and hence

$$\sum(B) = \sigma_a^2 \frac{\theta(B)\theta(B^{-1})}{\phi(B)\phi(B^{-1})} \tag{40}$$

So by expanding $\sum(B)$ given by (40) in a series (38), $\sigma(h)$ can be expessed by means of $\sigma_a^2$, $\phi_j$, $\theta_j$. (It is easy to see that (38) converges in a proper annular region if $\sum(B)$ is given by (40)).

An important general result follows from (40). We have for an arbitrary second order process for which $\sum|\sigma(h)| < \infty$, that the spectral density is given by

$$f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{+\infty} \sigma(h) e^{i\lambda h} \tag{41}$$

Combining (40) and (41) *we get for the spectrum of the ARMA (p,q) process*

$$f(\lambda) = \frac{1}{2\pi}\, \sigma_a^2\, \frac{\theta(e^{\lambda i})\theta(e^{-\lambda i})}{\phi(e^{\lambda i})\phi(e^{-\lambda i})} \tag{42}$$

Hence the *spectrum of an ARMA process (14) can immediately be written down.* Thus if $Z_t$ is ARMA (1,1)

$$Z_t = \phi\, Z_{t-1} + a_t - \theta\, a_{t-1}$$

we get

$$f(\lambda) = \frac{\sigma_a^2}{2\pi}\, \frac{1+\theta^2-2\theta \cos \lambda}{1+\phi^2+2\theta \cos \lambda}$$

We refer to Box and Jenkins (1970) page 67-84 where the autocovariance and spectrum for some important AR and ARMA processes are given.

## D. The Gaussian ARMA process.

Let $Z_t$ be stationary and defined as in Theorem 2. Assume in addition that the $a_t$ are independent and Gaussian. We make use of (15)

$$Z_t = \sum_{j=o}^{\infty} a_{t-j}\,\psi_j$$

to prove that then $Z_t$ is Gaussian. By (26) $Z_t - Z_{tm}$ converges in probability to 0. On the other hand

$$Z_{tm} = \sum_{j=1}^{m} \psi_j\, a_{t-j} \;;\; t=t_o,t_o+1,\ldots,t_1 \tag{43}$$

has a multivariate normal distribution with covariance matrix given by

$$\sigma_m(k) = E\,Z_{tm}Z_{t-hm} = \sum_{j=0}^{m} \psi_j\psi_{j-h}$$

Since $\sum|\psi_j| < \infty$, then $\sum\psi_j^2$ and $\sum\psi_j\psi_{j-h}$ converges. Hence it follows that as $m \to \infty$ then the variables (43) converge in distribution (e.g. by means of characteristic functions). Then this also is true for the variables $Z_t$ ; $t=t_o,\ldots,t_1$. However, these variables are independent of $m$. Hence they *are* multinormally distributed.

## E. Prediction in AR processes.

At time $t$ we want to predict $Z_{t+m}$ ; $m > 0$ (disregarding sampling errors in estimating the $\phi_i$). It is natural to use as a predictor a function $\phi(Z_t,Z_{t-1},\ldots,)$ of past and present observations which minimizes

$$E(\phi(Z_t,Z_{t-1},\ldots) - Z_{t+m})^2$$

i.e. the predictor

$$\hat{Z}_t(m) = E(Z_{t+m}|Z_t,Z_{t-1},\ldots) \tag{44}$$

In the case of an AR-process with independent impulses we get from Theorem 1 of Chapter II

$$Z_{t+m} = c_t(A_o) + V_{t+m\,m} \tag{45}$$

where $c_t(A_o)$ and $V_{t+m\,m}$ are given by equations (3) and (11) of Chapter II, and the vector $A_o$ is a function of the random variables $Z_t, Z_{t-1},\ldots,Z_{t-p}$ given by

$$c_{r-t}(A) = Z_t \; ; \quad r=t,t-1,\ldots,t-p+1$$

We have from (45)

$$\hat{Z}_t(m) = c_t(A_o) = \sum_{j=1}^{q}\left[Q_{m_j-1}(t)\cos\alpha_j t + R_{m_j-1}(t)\sin\alpha_j t\right]H_j^t \qquad (46)$$

Also from (1), replacing $t$ by $t+m$,

$$\begin{aligned}\hat{Z}_t(m) &= \phi_1\hat{Z}_t(m-1)+\ldots+\phi_{m-1}\hat{Z}_t(1) + \phi_m Z_t+\ldots+\phi_p Z_{t+m-p} \; ; \; m < p \\ \hat{Z}_t(m) &= \phi_1\hat{Z}_t(m-1)+\ldots+\phi_p\hat{Z}_t(m-p) \; ; \; m > p\end{aligned} \qquad (47)$$

Except in very simple situations the recursive formulae (47) is preferable for numericable purposes, but (46) gives an explicit analytical expression and is useful in qualitative discussions of the prediction formula.

A third expression follows from the second expression (3)

$$\hat{Z}_t(m) = \sum_{i=-\infty}^{t} a_i\psi_{t+m-i} = \sum_{j=o}^{\infty} a_{t-j}\psi_{j+m} \qquad (48)$$

where the $a_i$ can be computed from (1). (48) reflects the fact that $a_i$ is unpredictable for $i > t$.

# CHAPTER IV. THE STATIONARY AUTOREGRESSIVE PROCESS WHICH IS NOT ANTEIMPULSE-GENERATED

## A. Sufficiency of the assumption about the characteristic roots

Theorem 1: If

(i) $a_t$ ; $t = \cdots -1,0,1,\cdots$ is second order stationary, $Ea_t = 0$,

(ii) the equation

$$\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \cdots - \varphi_p B^p = 0 \tag{1}$$

has no roots on the unit circle, then the only stationary process satisfying

$$Z_t = \sum_{j=1}^{p} \varphi_j Z_{t-j} + a_t \text{ ; } t = \cdots -1,0,1,\cdots \tag{2}$$

is

$$Z_t = \sum_{-\infty}^{+\infty} a_{t-j}\psi_j \tag{3}$$

where the $\psi_j$ are determined in the following manner. Write $\varphi(B) = \varphi_1(B)\varphi_2(B)$ where $\varphi_1$ and $\varphi_2$ are polynomials and all the p-q roots of $\varphi_1(B)=0$ are outside the unit circle, whereas all the q roots of $\varphi_2(B)=0$ are inside the unit circle. Then expand $\varphi_1(B)$ according to powers of B and $\varphi_2(B)$ according to powers of $\frac{1}{B}$ and multiply the two power series to obtain the identity

$$\frac{1}{\varphi_1(B)\varphi_2(B)} = \sum_{i=-\infty}^{\infty} \psi_i B^i \tag{4}$$

defining the $\psi_i$.

Proof: ($\alpha$) We shall first consider the case when $q = p$ and hence $\varphi_1(B) = 1$, $\varphi_2(B) = \varphi(B)$, i.e. all the roots of $\varphi(B) = 0$ are inside the unit circle and the $\psi_{-j}$ are found from

$$\frac{1}{\varphi(B)} = -\left(\frac{1}{B}\right)^p \Big/ \sum_{j=0}^{\infty} \varphi_{p-j}\left(\frac{1}{B}\right)^j = \sum_{j=p}^{\infty} \psi_{-j} B^{-j} \tag{5}$$

(where $\varphi_0 = -1$). Hence in this case $\psi_j = 0$ for $j > -p$, and we have to prove that

$$Z_t = \sum_{j=p}^{\infty} \psi_{-j}\, a_{t+j} \tag{6}$$

We shall reduce this case to the case dealt with in Theorem 1 of Chapter III by introducing $\tilde{Z}_t = Z_{-t}$, i.e. reversing the process. We find from (2)

$$\tilde{Z}_{-t} = \sum_{j=1}^{p} \varphi_j\, \tilde{Z}_{-t+j} + a_t$$

We replace $t$ by $t-p$ and then $p-j$ by $j$ under the summation sign and get

$$\tilde{Z}_{t-p} = \sum_{j=0}^{p-1} \varphi_{p-j}\, \tilde{Z}_{t-j} + a_{p-t}$$

Solving for $Z_t$ we get

$$\tilde{Z}_t = \sum_{j=1}^{p} \tilde{\varphi}_j\, \tilde{Z}_{t-j} + \tilde{a}_t \tag{7}$$

where

$$\begin{aligned} \tilde{\varphi}_j &= -\varphi_{p-j}/\varphi_p \ ; \quad j = 1,2,\ldots,p-1 \ , \\ \tilde{\varphi}_p &= 1/\varphi_p \ , \quad \tilde{a}_t = -a_{p-t}/\varphi_p \end{aligned} \tag{8}$$

The characteristic polynomial for (7) is now seen to be

$$\tilde{\varphi}(B) = 1 - \sum_{j=1}^{p} \tilde{\varphi}_j\, B^j = -\varphi\left(\frac{1}{B}\right) B^p/\varphi_p \tag{9}$$

Thus $\tilde{\varphi}(B) = 0$ has all roots outside the unit circle and we can apply Theorem 1 of Chapter III to obtain

$$\tilde{Z}_t = \sum_{j=0}^{\infty} \tilde{\psi}_j\, \tilde{a}_{t-j} \tag{10}$$

where the $\tilde{\psi}_j$ are given by

$$1/\tilde{\varphi}(B) = \sum_{j=0}^{\infty} \tilde{\psi}_j B^j \tag{11}$$

We shall now rewrite (10) and (11) in terms of $\varphi_j$ and $\psi_j$. From (9) and (11) we find the relationship between $\psi_j$ and $\tilde{\psi}_j$.

$$-\varphi_p/B^p\varphi(\tfrac{1}{B}) = \sum_{j=0}^{\infty} \tilde{\psi}_j B^j$$

i.e.

$$\varphi_p/\varphi(B) = -\sum_{j=0}^{\infty} \tilde{\psi}_j B^{-j-p}$$

Hence by comparison with (5)

$$\tilde{\psi}_j = -\varphi_p \psi_{-j-p} \tag{12}$$

Hence from (10) and (8)

$$Z_t = \tilde{Z}_{-t} = \sum_{j=0}^{\infty} \tilde{\psi}_j \tilde{a}_{-t-j} = \sum_{j=0}^{\infty} (-\varphi_p \psi_{-j-p})(-a_{p+t+j}/\varphi_p) =$$

$$= \sum_{j=p}^{\infty} \psi_{-j} a_{t+j}$$

which proves (6) and the theorem in the case when $q = p$.

(β). We now consider the case $0 < q < p$. The equation (2), i.e. $\varphi(B)Z_t = a_t$ may be written

$$\varphi_1(B)A_t = a_t \tag{13}$$

where

$$A_t = \varphi_2(B)Z_t \tag{14}$$

$A_t$ is stationary by (14) and by Theorem 1 of Chapter III we have from (13)

$$A_t = \sum_{j=0}^{\infty} \lambda_j a_{t-j} \tag{15}$$

By (14) and ($\alpha$) above we have

$$Z_t = \sum_{j=q}^{\infty} \kappa_j A_{t+j} \tag{16}$$

We have from (4)

$$\psi_i = \sum_{j=\max(-i,q)}^{\infty} \lambda_{i+j} \kappa_j = \sum_{j=-\infty}^{+\infty} \lambda_{t+j} \kappa_j \tag{17}$$

defining $\lambda_s = 0$, $\kappa_r = 0$ if $s < 0$, $r < q$.

Now, from Theorem 1 of Chapter III, ($\alpha$) above and (4) we have

$$1/\varphi_1(B) = \sum_0^{\infty} \lambda_j B^j , \quad 1/\varphi_2(B) = \sum_q^{\infty} \kappa_j B^{-j} ,$$
$$\psi(B) = 1/\varphi_1(B)\,\varphi_2(B) = \sum_{i=-\infty}^{\infty} \psi_i B^i \tag{18}$$

By well known theorems about analytic functions the three series (18) converge absolutely in the regions $|B| < R_1$, $|B| > R_2$, $R_2 < |B| < R_1$, respecively, where $R_1$ is the minimum modulus of roots of $\varphi_1(B) = 0$ and $R_2$ is the maximum modulus of the root of $\varphi_2(B) = 0$. For $B = 1$ this implies

$$\sum|\lambda_i| < \infty , \quad \sum|\kappa_i| < \infty , \quad \sum|\psi_i| < \infty \tag{19}$$

We now apply Lemma 14 in Appendix I with $Z = Z_t$ and

$$Y_i = \kappa_i A_{t+i} = \sum_{j=0}^{\infty} \kappa_i \lambda_{t+i-j}$$

to obtain

$$Z_t = \sum_{j=-\infty}^{+\infty} a_{t-j} \sum_{i=-\infty}^{\infty} \kappa_i \lambda_{t+j}$$

which by (17) equals (3). We only have to check the assumptions in Lemma 14. Assumption 3 is true because

$$\sum_j \left|\sum_i \kappa_i \lambda_{t+i-j}\right| \le \sum_i |\kappa_i| \sum|\lambda_{t+i-j}| = \sum|\kappa_i| \sum|\lambda_j| < \infty$$

by (19). The other assumptions are trivially true.

B. The necessity of the assumptions about the characteristic roots

We shall now show that the requirement about the characteristic roots not being on the unit circle, is a necessary condition for the AR-equation to have a stationary solution. For convenience we make the assumption that the spectrum of the impulse process has a density.

Theorem 2. Suppose that $\{a_t\}$ is a second order stationary process, the spectrum of which has continuous positive density. Furthermore $Ea_t = 0$, $\operatorname{var} a_t > 0$.

A necessary and sufficient condition for AR-equation (1) to have a stationary solution is that the characteristic algebraic equation (2) has no roots on the unit circle. If this condition is satisfied, then the stationary solution is given by (3).

We shall give two proofs. The first proof assumes that $\{a_t\}$ is an uncorrelated process. It is intuitively appealing and is based only on very elementary properties about second order mean convergence. The second proof assumes no extra restriction and is somewhat briefer; it is based on the fundamental relationship between the autocovariance function and the spectrum (see Chapter I, Theorem 1). The impatient reader should turn to the second proof on page 40.

Proof (i). We assume now that $\{a_t\}$ is uncorrelated. We use proof by contradiction.

Assume first that $p = 1$ and that the root is equal to $1$. Then [following R.W. Andersen (1971) page 171],

$$Z_t = Z_{t-1} + a_t$$

and hence

$$Z_t - Z_{t-1} = a_t + \cdots + a_{t-s+1}$$

$\text{var}(Z_t - Z_{t-1}) = EZ_t^2 + EZ_{t-s}^2 - 2EZ_t\,Z_{t-s} = s\sigma_a^2$ where $\sigma_a^2 = \text{var}\,a_t$ . Hence if $Z_t$ is stationary $-2EZ_t^2 \leq 2EZ_t\,Z_{t-s} = 2EZ_t^2 - s\sigma_a^2$ . The left inequality follows from Schwartz inequality. Hence $4EZ_t^2 \geq s\sigma_a^2$ which is impossible since $s$ may be chosen arbitrarily large. [However, if $\sigma_a^2 = 0$, $a_t = 0$, then $Z_t = Z_{t-s}$ and $Z_t = Z_0$ is a trivial stationary solution].

Assume now that $p > 1$ and the characteristic equation (1) has a root $B = 1$. Then we may write $\varphi(B) = (1-B)(1-B\tilde{\varphi}_1 - \cdots - B^{p-1}\tilde{\varphi}_{p-1})$ and $V_t$ defined by

$$V_t = Z_t - \tilde{\varphi}_1 Z_{t-1} - \cdots - \tilde{\varphi}_{p-1}\,Z_{t-p+1} \tag{20}$$

satisfies

$$(1-B)V_t = V_t - V_{t-1} = a_t \tag{21}$$

Now, if $Z_t$ is stationary, then by (20) $V_t$ is stationary. This violates (21) by what has just been proved. [However if $\sigma_a^2 = 0$, then $V_t = V$ and by (3) if no other roots are on the unit circle

$$Z_t = \sum_{j=-\infty}^{+\infty} \psi_j\,V = \lambda V = Z\,]$$

Consider now the case of two conjugate complex roots of (1), $e^{\pm\alpha i}$ ; $0 < \alpha < \pi$ ; on the unit circle. Take first the case when $p = 2$ .

$$Z_t = \varphi_1\,Z_{t-1} + \varphi Z_{t-2} + a_t \tag{22}$$

Then the roots of $\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2$ are $e^{\pm i\alpha}$ if and only if $\varphi_1 = 2\cos\alpha$ , $\varphi_2 = -1$ , hence

$$\varphi(B) = 1 - (2\cos\alpha)B + B^2 = (1 - B\,e^{i\alpha})(1 - B\,e^{-i\alpha})$$

We define

$$V_t = Z_t - e^{-i\alpha} Z_{t-1} \tag{23}$$

and have from (22)

$$V_t - e^{i\alpha} V_{t-1} = a_t \tag{24}$$

from which we obtain

$$V_t = e^{si\alpha} V_t + a_t e^{i\alpha} a_{t-1} + \cdots + e^{(s-1)i\alpha} a_{t-s+1} \tag{25}$$

Hence

$$\mu = E|V_t - e^{si\alpha} V_{t-s}|^2 = E|\sum_{k=1}^{s-1} a_{t-k} e^{i\alpha k}|^2 =$$

$$= E \sum a_{t-k} e^{i\alpha k} \sum a_{t-k} e^{-i\alpha k} = s\sigma_a^2 \tag{26}$$

On the other hand

$$\mu = E(V_t - e^{si\alpha} V_{t-s})(V_t^* - e^{-si\alpha} \bar{V}_{t-s}) =$$

$$= E V_t V_t^* - e^{si\alpha} E V_t V_{t-s}^* - e^{si\alpha} E V_t^* V_{t-s} + E V_{t-s} V_{t-s}^* \tag{27}$$

(where $V^*$ denotes the complex conjugate of $V$).

However, from (23) we have,

$$E V_t V_{t-s}^* = 2E Z_t Z_{t-s} - e^{i\alpha} E Z_t Z_{t-s-1} - e^{-i\alpha} E Z_{t-1} Z_{t-s-1}$$

which is independent of $t$. Hence (27) reduces to

$$\mu = 2E|V_t|^2 - e^{si\alpha} E V_t V_{t-s}^* - e^{-si\alpha} E V_t^* V_{t-s}$$

Hence

$$|\mu| \leqq 2E|V_t|^2 + 2|E V_t V_{t-s}^*| \leqq 2E|V_t|^2 + 2E|V_t|^2 \tag{28}$$

by Schwartz inquality. Combining that with (26) we get again $4E|V_t|^2 \geqq s\,\sigma_a^2$, which is a contradiction if $\sigma_a^2 > 0$. Hence there is no stationary solution of (22).

[However, if $\sigma_a^2 = 0$, then by Chapter II,(22) has a solution

$$Z_t = A\cos\alpha t + B\sin\alpha t$$

Hence we get with $\sigma^2 = \operatorname{var} A$, $\tau^2 = \operatorname{var} B$ and $\gamma = \operatorname{cov}(A,B)$,

$$\operatorname{cov}(Z_t, Z_{t-s}) = \sigma^2 \cos\ (t-s) + (\tau^2-\sigma^2)\sin\alpha t \sin\alpha s + \gamma \sin(s+t)\alpha .$$

With $s = 0$ we have $\frac{d}{dt}\operatorname{var} Z_t = 2\alpha\gamma$, hence $\gamma = 0$ if $Z_t$ is stationary. Furthermore $\sigma^2 = \tau^2$ and from $EZ_t = 0$ we have $EA = EB = 0$. Thus

$$EA = EB = 0 \ , \quad \operatorname{var} A = \operatorname{var} B \ , \quad \operatorname{cov}(A,B) = 0$$

Then we verify that $Z_t$ is stationary.]

Consider now the case of at least one pair of complex conjugate roots $e^{\pm i\alpha}$ on the unit circle and $p > 2$. Then we write

$$\varphi(B) = (1-\varphi B+B^2)(1-\tilde{\varphi}_1 B-\cdots-\tilde{\varphi}_{p-2} B^{p-2})$$

where $\varphi = 2\cos\alpha$. Then

$$V_t = Z_t - \tilde{\varphi}_1 Z_{t-1} - \cdots - \tilde{\varphi}_{p-2} Z_{t-p+2} \tag{29}$$

satisfies

$$V_t = \varphi V_{t-1} - V_{t-2} + a_t \tag{30}$$

Thus, as before, if $Z_t$ is stationary, then $V_t$ is stationary by (29) and hence from what we have proved above about (22), $V_t$ can not satisfy (30), which is a contradiction. Then everything is proved.

<u>Proof (ii)</u>. Let us assume that $Z_t$ is stationary and satisfies (2). We want to investigate if this is consistent with $\varphi(B) = 0$ having roots on the unit circle.

For that purpose we first find a relationship between

$$\sigma_a(h) = E\, a_t\, a_{t-h} \ , \quad \sigma_Z(h) = E\, Z_t\, Z_{t-h} \tag{31}$$

We have from (2), with $\varphi_0 = -1$

$$\sigma_a(h) = E \sum_{j=0}^{p} \varphi_j Z_{t-j} \sum_{j=0}^{p} \varphi_j Z_{t-j-h} = \sum \varphi_j \varphi_k \sigma_Z(h-j+k) \tag{32}$$

For the cumulative spectra $F_Z$ and $F_a$ for $\{Z_t\}$ and $\{a_t\}$, respectively, we have by I B(14)

$$\sigma_Z(h) = \int_{-\pi}^{\pi} e^{i\lambda h} dF_Z(\lambda) \quad , \quad \sigma_a(h) = \int_{-\pi}^{\pi} e^{i\lambda h} dF_a(\lambda) \tag{33}$$

recalling that $F_Z$ and $F_a$ have symmetric increments. Introducing (33) in (32) we get

$$\int_{-\pi}^{\pi} e^{i\lambda h} dF_a(\lambda) = \sum_{j,k} \varphi_j \varphi_k \int_{-\pi}^{\pi} e^{i\lambda(h-j+k)} dF_Z(\lambda) =$$

$$= \int_{-\pi}^{\pi} e^{i\lambda h} \sum \varphi_j e^{-i\lambda j} \sum \varphi_k e^{i\lambda k} dF_Z(\lambda)$$

Hence

$$\int_{-\pi}^{\pi} e^{i\lambda h} dF_a(\lambda) = \int_{-\pi}^{\pi} e^{i\lambda h} |\varphi(e^{-i\lambda})|^2 dF_Z(\lambda)$$

By the uniqueness property in Theorem 1 of Chapter I it then follows that

$$dF_a(\lambda) = |\varphi(e^{-i\lambda})|^2 dF_Z(\lambda) \tag{35}$$

Hence we get, integrating and using the mean value theorem,

$$|\varphi(e^{-i\lambda})|^2 (F_Z(\mu+\delta) - F_Z(\mu) = F_a(\mu+\delta) - F_a(\mu) \tag{36}$$

where $\min(\mu+\delta,\mu) \leqq \lambda \leqq \max(\mu+\delta,\mu)$.

There are at most $p$ roots of $\varphi(B) = 0$ on the unit circle. Hence $|\varphi(e^{-i\lambda})|^2 = 0$ only for (say) $\lambda = \lambda_1,\dots,\lambda_q$ ; $q \leqq p$. Now since $F_a$ is continuous by the assumption we get from (36) $F_Z(\mu-0) = F_Z(\mu)$ if $\mu \neq \lambda_1,\dots,\lambda_q$. But

$$f_0(\lambda_j) = F_Z(\lambda_j) - F_Z(\lambda_j-0)$$

may be $> 0$.

If $\mu \neq \lambda_1,\dots,\lambda_q$ we also have from (36)

$$f_Z(\mu) = \frac{d}{d\mu} F_Z(\mu) = f_a(\mu)/|\varphi(e^{-i\mu})|^2 \tag{37}$$

since $f_a(\mu) = \frac{d}{d\mu} F_a(\mu)$ exists by assumption. Hence $f_Z(\mu)$ exists for $\mu \neq \lambda_1,\dots,\lambda_q$ and (see Theorem 1 of Chapter I)

$$\int_{-\pi}^{\pi} f_Z(\mu)d\mu + \sum_{j=1}^{q} f_0(\lambda_j) = \text{var } Z_t < \infty \tag{38}$$

It follows from (38) and (37) that

$$I = \int_{-\pi}^{\pi} \frac{f_a(\mu)}{|\varphi(e^{-i\mu})|^2} d\mu < \infty \tag{39}$$

We have, if $\lambda$ is one of the roots $\lambda_j \neq 0$ and $\pi$ with multiplicity $r$,

$$\varphi(e^{-i\mu}) = (e^{-i\mu}-e^{-i\lambda})^r (e^{-i\mu}-e^{i\lambda})^r \tilde{\varphi}(e^{-i\mu})$$

Writing

$$e^{-i\mu} - e^{-i\lambda} = e^{-i\frac{1}{2}(\lambda+\mu)}\, 2i \sin \tfrac{1}{2}(\mu-\lambda)$$

we get

$$I = \int_{-\pi}^{\pi} \frac{h(\mu)\, d\mu}{4 \sin^{2r} \frac{1}{2}(\mu-\lambda) \sin^{2r} \frac{1}{2}(\mu+\lambda)}$$

where

$$h(\mu) = f_a(\mu)/|\tilde{\varphi}(e^{-i\mu})|^2$$

However, since

$$\sin^{2r} \tfrac{1}{2}(\mu-\lambda) \approx [\tfrac{1}{2}(\mu-\lambda)]^{2r}$$

asymptotically, when $\mu$ is close to $\lambda$, we get that $I$ diverges, contradicting (39). The case $\lambda=\lambda_j=0$ or $\pi$ is treated similarly. Hence $\varphi(e^{-i\mu}) = 0$ has no real roots and the necessity of the assumption in Theorem 2 is again proved.

The sufficience follows from Theorem 1.

The theorem has as an immediate consequence

<u>Theorem 3</u>. Let $\{a_t\}$ be an uncorrelated second order stationary process, $Ea_t = 0$, $\operatorname{var} a_t > 0$. Assume also that $1-\theta_1 B-\cdots-\theta_q B^q = 0$ has no roots on the unit circle.

A necessary and sufficient condition for the ARMA(p,q) equation

$$Z_t - \varphi_1 Z_{t-1} - \cdots - \varphi_p Z_{t-p} = a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q} \tag{40}$$

too have a stationary solution is that the characteristic algebraic equation $\varphi(B) = 0$ has no roots on the unit circle. In the case when all roots are outside the unit circle the solution is given by Theorem 2 of Chapter III.B.

<u>Proof</u>: Obviously $A_t = a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q}$ is stationary, and the covariance-function is given by

$$\sigma_A(h) = E\, A_t\, A_{t-h} = \sum_{j,k=0}^{q} \theta_j\, \theta_k\, \sigma_a(h-j+k) \tag{41}$$

where $\theta_0 = 1$. Hence by (33)

$$\sigma_A(h) = \sum_{j,k} \theta_j\, \theta_k \int_{-\pi}^{\pi} e^{i\lambda(h-j+k)} d\, F_a(\lambda) =$$

$$= \int_{-\pi}^{\pi} |1 - \Sigma\, \theta_j\, e^{-i\lambda j}|^2\, d\, F_a(\lambda)\,.$$

Hence the spectrum of $\{A_t\}$ is by Chapter I eq. (14) given by

$$d F_A(\lambda) = |1 - \theta_j\, e^{-i\lambda j}|^2\, d F_a(\lambda)$$

But by Chapter I eq. (15), $F_a(\lambda) = \frac{\sigma(0)}{2\pi}(\pi+\lambda)$. Thus the density of the spectrum of $\{A_t\}$ exists and is given by

$$f_A(\lambda) = |1 - \theta_j\, e^{-i\lambda j}|^2\, \sigma(0)/2\pi \tag{42}$$

Hence $f_A$ is a positive and continuous function of $\lambda$.
This proves Theorem 3.

C. Another proof of sufficiency (Theorem 1 in section A)

We shall give an alternative proof of the theorem in IV.A, which naturally links up with the proof (ii) of necessity given in the foregoing section. The proof is based on a famous result about the Fourier representation of a discrete stationary process by means of a process with uncorrelated increments. It states that to a second order stationary process $Z_t$ can be assigned a process $\zeta_Z(\lambda)$ with orthogonal increments, i.e.

$$E\,\zeta_Z(\lambda') - \zeta_Z(\lambda''))(\zeta_Z(\lambda''') - \zeta_Z(\lambda^{iv})) = 0 \, ,$$

$$\lambda' > \lambda'' \geq \lambda''' > \lambda^{iv} \tag{43}$$

such that for each t

$$Z_t = \int_{-\pi}^{\pi} e^{i\lambda t} \, d\,\zeta_Z(\lambda) \tag{44}$$

the integral being defined as a limit in quadratic mean. $\zeta_Z(\lambda) - \zeta_Z(\lambda')$ is uniquely given by Z , for any $(\lambda, \lambda')$. Hence we may fix $\zeta_Z(-\pi) = 0$ . Then

$$E|\zeta(\lambda)|^2 = F(\lambda) \tag{45}$$

where $F(\lambda)$ is the spectrum of the process given in section I.B.

Then let us consider the stationary $Z_t$ defined by (2), i.e.

$$Z_t = \sum_{j=1}^{p} \varphi_j Z_{t-j} + a_t \tag{46}$$

where $a_t$ is stationary. We shall prove that provided

$$\varphi(B) = 1 - \varphi_1 B - \cdots - \varphi_p B^p$$

has no roots on the unit circle, then

$$Z_t = \sum_{j=-\infty}^{+\infty} \psi_j Z_{t-j} \tag{47}$$

where the $\psi_j$ are given by (4).

For that purpose we introduce (44) and

$$a_t = \int_{-\pi}^{\pi} e^{i\lambda t} \, d\zeta_a(\lambda) \tag{48}$$

in (46). We obtain

$$\int_{-\pi}^{\pi} e^{i\lambda t} \varphi(e^{-i\lambda}) \, d\zeta_Z(\lambda) = \int_{-\pi}^{\pi} e^{ik\lambda} \, d\zeta_a(\lambda)$$

However since $\zeta_a(\lambda)$ is unique, we have

$$\int_{-\pi}^{\mu} \varphi(e^{-i\lambda}) \, d\zeta_Z(\lambda) = \zeta_a(\mu) \tag{49}$$

or

$$d\zeta_Z(\lambda) = \frac{1}{\varphi(e^{-i\lambda})} \, d\zeta_a(\lambda)$$

where $\varphi(e^{-i\lambda}) \neq 0$ for $\lambda$ real by our assumption. Hence from (44), we have

$$Z_t = \int_{-\pi}^{\pi} e^{i\lambda t} \frac{1}{\varphi(e^{-i\lambda})} \, d\zeta_a(\lambda) \tag{50}$$

as the unique solution of (46).

Introducing (4) in (50) we get (47), since

$$\int_{-\pi}^{\pi} e^{i\lambda t} \psi_j (e^{-i\lambda})^j \, d\zeta_a(\lambda) = \psi_j \, a_{t-j}$$

by (48).

## V THE UNCONDITIONAL LIKELIHOOD BY AR PROCESSES

We shall derive the likelihood of the observations $Z_1, Z_2, \ldots, Z_n$ from a stationary AR process defined by

$$Z_t - \phi_1 Z_{t-1} - \cdots - \phi_p Z_{t-p} = a_t \tag{1}$$

where $E Z_t = 0$, and the $a_t$ are independent and normal with variance $\sigma_a^2$. The roots of $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p = 0$ are outside the unit circle. Thus $\{Z_t\}$ is Gaussian and the observations $Z^{(n)} = (Z_1, \ldots, Z_n)'$ have density

$$f_n(z) = (2\pi)^{-\frac{n}{2}} \sigma_a^{-n} |m_n^{(p)}|^{\frac{1}{2}} e^{-\frac{1}{2\sigma_a^2} z' m_n^{(p)} z} \tag{2}$$

We must determine $m_n^{(p)}$.

We have for the conditional density of $Z_{p+1}, \ldots, Z_n$, given $Z^{(p)} = (Z_1, \ldots, Z_p)'$

$$f(z_{p+1}, \ldots, z_n | z^{(p)}) = (2\pi)^{-\frac{n-p}{2}} \sigma_a^{-n+p} e^{-\frac{1}{2\sigma_a^2} \sum_{t=p+1}^{n} (z_t - \sum_{j=1}^{p} \phi_j z_{t-j})^2} \tag{3}$$

From

$$f_n(z^{(n)}) = f_p(z^{(p)}) f(z_{p+1}, \ldots, z_n | z^{(p)}), \tag{4}$$

(3) and (2) we then get

$$z^{(n)'} m_n^{(p)} z^{(n)} = z^{(p)'} m_p^{(p)} z_p^{(p)} + \sum_{t=p+1}^{n} \left( \sum_{j=1}^{p} \phi_j z_{t-j} \right)^2 \tag{5}$$

where we write $\phi_o = -1$. Hence we can concentrate on finding

$$m_p^{(p)} = \left( m_{ijp}^{(p)} \right)_{i,j} = (m_{ij})_{ij}$$

writing briefly $m_{ijp}^{(p)} = m_{ij}$.

Now we have from Chapter III, B (15) that

$$Z_t = \sum_{j=o}^{\infty} \psi_j a_{t-j}$$

$$Z_{c-t} = \sum_{j=o}^{\infty} \psi_j \, a_{c-t-j}$$

However, for any c, the process $\{a_t\}$ has precisely the same distribution as $\{a_{c-t}\}$. Hence $\{Z_t\}$ and $\{Z_{c-t}\}$ have the same distribution and we then have (with $c=n+1$) that $(Z_1,\ldots,Z_n)$ and $Z_n,\ldots,Z_1)$ are identically distributed. Hence

$$m_{ijn}^{(p)} = m_{n+1-j,n+1-i,n}^{(p)} \tag{6}$$

From (5) we get

$$m_{p+1}^{(p)} = \begin{bmatrix} & \vdots & 0 \\ m_p^{(p)} & \vdots & \vdots \\ & \vdots & 0 \\ \cdots\cdots & \cdots & \cdots \\ 0 \ldots\ldots 0 & \ldots & 0 \end{bmatrix} + \begin{bmatrix} \phi_p^2, \phi_p\phi_{p-1}, \ldots\ldots, \phi_p\phi_o \\ \phi_{p-1}\phi_p, \phi_{p-1}^2, \ldots, \\ \vdots \\ \phi_o\phi_p, \ldots\ldots\ldots\ldots, \phi^2 \end{bmatrix} \tag{7}$$

Hence

$$m_{ijp+1}^{(p)} = m_{ij} + \phi_{p-i+1}\phi_{p-j+1} \quad ; \quad i,j \le p$$

$$m_{p+1,j,p+1}^{(p)} = \phi_o\phi_{p-j+1} \tag{8}$$

On the left hand side of (8) we make use of (6) with $n = p+1$ and of (8). Then (8) reduces to

$$m_{p+2-i,p+2-i} + \phi_i\phi_j = m_{ij} + \phi_{p-i+1,p-j+1} \quad ; \; i,j \ge 2 \tag{9}$$

$$\phi_o\phi_{j-1} = m_{1j} + \phi_p\phi_{p-j+1} \tag{10}$$

We now make use of (6) with $n = p$, to obtain the recursion formula

$$m_{ij} = m_{i-1,j-1} + \phi_{i-1}\phi_{j-1} - \phi_{p-i+1}\phi_{p-j+1} \tag{11}$$

to be combined with (10). We get for $j \ge i$,

$$m_{ij\,p}^{(p)} = m_{ij} = \sum_{s=1}^{i} \phi_{i-s}\phi_{j-s} - \sum_{j=1}^{i} \phi_{p+s-i}\phi_{p+s-j} \tag{12}$$

*Hence the likelihood of $(Z_1,..,Z_n)$ is explicitly given by (2) with n=p, (3)(4)(12).*
Obviously we may also write

$$f_n(z) = (2\pi)^{-\frac{n}{2}}\sigma_a^{-n}\left|m_p^{(p)}\right| e^{-\frac{1}{2\sigma_a^2}S_\phi(z)} \tag{13}$$

where

$$S_\phi(z) = \sum_{s,t=o}^{p} D_{st}^{(n)}\phi_s\phi_t$$

and where the $D_{st}^{(n)}$ are quadratic forms in $Z_1,...,Z_n$. They are the sufficient statistics relatively to our observations. We shall derive explice expressions for them.

We then rewrite (12)

$$m_{i\ i+d} = \sum_{s=o}^{i-1}\phi_s\phi_{s+d} - \sum_{s=p-d-i}^{p-d}\phi_s\phi_{s+d} \tag{14}$$

and get for the first term on the right hand side of (4)

$$z'\,m\,z = \sum_d\sum_{i=1}^{p-d} m_{i,i+d}\, z_i z_{i+d} \tag{15}$$

The inner sum in (15) can, by (14), be written

$$\sum_{i=1}^{p-d}\sum_{s=o}^{i-1}\phi_s\phi_{s+d}\, z_i z_{i+d} - \sum_{i=1}^{p-d}\sum_{s=p-d-i+1}^{p-d}\phi_s\phi_{s+d} z_i z_{i+d} = Q_1 - Q_2 \tag{16}$$

We have for $Q_2$, substituting $p-d-i+1 = i'$ and rearranging the summation by $\sum_{i=1}^{p-d}\sum_{s=i}^{p-d} = \sum_{s=1}^{p-d}\sum_{i=1}^{s}$,

$$Q_2 = \sum_{s=1}^{p-d}\phi_s\phi_{s+d}\sum_{i=1}^{s} z_{p-d-i+1} z_{p-i+1}$$

Hence, substituting again $p-d-i+1 = i'$,

$$Q_2 = \sum_{s=1}^{p-d}\phi_s\phi_{s+d}\sum_{i=p-d-s+1}^{p-d} z_i z_{i+d} \tag{17}$$

For $Q_1$ we have

$$Q_1 = \sum_{i=1}^{p-d} \sum_{s=1}^{i} \phi_{s-1}\phi_{s-1+d} z_i z_{i+d} = \sum_{s=1}^{p-d} \sum_{i=s}^{p-d} \phi_{s-1}\phi_{s-1+d} z_i z_{i+d}$$

or

$$Q_1 = \sum_{s=o}^{p-d-1} \phi_s \phi_{s+d} \sum_{i=s+1}^{p-d} z_i z_{i+d} \tag{18}$$

Comparing the lower limits in the inner sums in (17) and (18) we observe that the expression for $Q_1 - Q_2$ must be written differently, according as $2s < p-d$, $= p-d$, $> p-d$. We get for the coefficients of $\phi_s \phi_{s+d}$ in (15), in the three cases respectively

$$D^{(p)}_{r,s+d} = \sum_{i=s+1}^{p-d-s} z_i z_{i+d}, \quad 0, \quad -\sum_{p-d-s+1}^{s} z_i z_{i+d}$$

or

$$D^{(p)}_{st} = \sum_{i=s+1}^{p-t} z_i z_{i+t-s} \quad , \quad 0, \quad -\sum_{p-t+1}^{s} z_i z_{i+t-s} \tag{19}$$

according as $s < p-t$, $= p-t$, $> p-t$. Hence we may write for the first term in (4)

$$z' m^{(p)}_p z = \sum_{s,t}^{p} D^{(p)}_{st} \phi_s \phi_t \tag{20}$$

where $D^{(p)}_{s,t}$ is given by (19).

Now let us find $D^{(n)}_{s,t}$ in (13) by adding the last term in (5) which may be written

$$\sum_{t,s} \phi_t \phi_s \sum_{j=p+1}^{n} z_{j-t} z_{j-s} = \sum_{t,s} \phi_t \phi_s \sum_{i=p+1-t}^{n-t} z_i z_{i+t-s} \tag{21}$$

Combining (21), (20), (19) and (5) we see that we get the same expression for $D^{(n)}_{st}$ as for $D^{(p)}_{st}$, only replace $p$ by $n$.

Hence we have

<u>Theorem</u>. The likelihood function of the observations $Z_1, Z_2, \ldots, Z_n$ in a p-th order stationary anteimpulse-generated AR-process $\{Z_t\}$ with independent and normal $(0,\sigma)$ impulses is given

by (13) where

$$D_{st}^{(n)} = D_{st}(z) = \begin{cases} \sum_{i=s+1}^{n-t} z_i z_{i+t-s} & ,\text{if} \quad s+t < n \\ 0 & ,\text{if} \quad s+t = n \\ -\sum_{i=n-t+1}^{s} z_i z_{i+t-s} & ,\text{if} \quad s+t > n \end{cases} \qquad (22)$$

and $m_p^{(p)}$ is a symmetric $p \times p$ matrix with elements given by (12). The $D_{st}(Z)$ constitutes a sufficient set of statistics for $Z_1, \ldots, Z_n$.

Note that $m_p^{(p)}$ does not depend on $n$, hence the numerical work connected with computing $m_p^{(p)}$ will usually be trifling. Note also that since $s,t \leq p$, only the first sum will be of interest if $2p < n$.

In the case when $\mu = E\,Z_t$ is unknown, i.e. $Z_t - \mu$ is given by (1), the likelihood is easily written down. It is only to replace $D_{st}(z)$ in (22) by

$$D_{st}(z-\mu) = \sum_{i=s+1}^{n-t} z_i z_{i+t-s} - (n-t-s)(\bar{z}_{st} - \mu)(\bar{z}_{ts} - \mu)$$

where $\bar{z}_{st}$ is the average over $z_i$ ; $i = s+1, \ldots, n-t$.

## VI. THE DISTRIBUTION OF THE EMPIRICAL SPECTRUM

We shall in this chapter consider a process $Z_t$ given by

$$Z_t = \zeta + \sum_{s=-\infty}^{\infty} \gamma_s V_{t-s} \tag{1}$$

where $\{V_t\}$ is a process of independent variables which is second order stationary with mean 0 and variance $\sigma^2$. Furthermore

$$\sum \gamma_s^2 < \infty \tag{2}$$

We then have for the covariance function

$$\sigma(h) = \operatorname{cov} Z_t Z_{t-h} = \sigma^2 \sum \gamma_s \gamma_{s-h} \tag{3}$$

and the spectral density of $\{Z_t\}$ exists and is given by (see Chapter I, B, eq.16)

$$f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{+\infty} e^{i\lambda h} \sigma(h) = \frac{\sigma^2}{2\pi} \left| \sum_{-\infty}^{+\infty} \gamma_s e^{i\lambda s} \right|^2 \tag{4}$$

Obviously, the assumptions are satisfied in the case when $\{Z_t\}$ is an ARMA process (see Chapter II,B, Eq.(15)).

We shall derive the asymptotic distributions of the empirical spectrum given in Chapter I. B. More precisely we shall be interested in the joint asymptotic distribution of

$$\begin{aligned} A(\lambda) &= \frac{2}{T} \sum_{t=1}^{T} (Z_t - \xi) \cos \lambda t \\ B(\lambda) &= \frac{2}{T} \sum_{t=1}^{T} (Z_t - \xi) \sin \lambda t \end{aligned} \qquad -\pi \leq \lambda \leq \pi \tag{5}$$

(see Chapter I, B, eq.(4)) for different values of $\lambda$ as $T \to \infty$.

The general idea of the proof of the theorem as given below, is that presented in T.W. Anderson (1970). However, I have found it necessary to expand upon and deviate from some details in Anderson's proof.

We obviously have $E\,A(\lambda) = E\,B(\lambda) = 0$. The exact expressions for the covariances for the pairs $(A(\lambda),(A\lambda'))$, $(B(\lambda),B(\lambda'))$, $(A(\lambda),B(\lambda')$ can be found in Anderson (1970) p. 457, see eq.'s (73), (79), (80), (81). It is also proved in Anderson (see p.477), that

$$\lim_{T\to\infty} T \operatorname{var} A(\lambda) = 4\pi f(\lambda), \quad \lim T \operatorname{var} B(\lambda) = 4\pi(\lambda) \tag{6}$$

when $\lambda \neq 0, \pm\pi$ and

$$\lim T \operatorname{cov}(A(\lambda),A(\lambda')) = 0 \quad ; \quad \lambda \neq \pm\lambda' \tag{7}$$

Also the pairs $(B(\lambda),B(\lambda')$ and $(A(\lambda),B(\lambda'))$ are uncorrelated.

In the course of the derivation we shall need two theorems. First the well known

<u>Lindeberg's Theorem</u>: For each $n=1,2,\ldots$ let $X_{1n},\ldots,X_{nn}$ be independent $E\,X_{jn} = 0$,

$$\sum_{j=1}^{n} \operatorname{var} X_{jn} = 1$$

and for any $\delta > 0$,

$$\lim \sum_{j=1}^{n} \int_{|x|>\delta} x^2\, dF_{jn} = 0$$

where $F_{jn}(x) = \Pr(X_{jn} \leq x)$. Then $\sum_{j=1}^{n} X_{jn}$ converges in distribution to the normal (0,1) as $n \to \infty$. Then the useful

<u>T.W. Anderson's Theorem</u>: Let

$$S_T = Z_{mt} + X_{mt}$$

where $\operatorname{plim}_{m\to\infty} X_{mt} = 0$, uniformly in $T$,

$$\lim_{T\to\infty} \Pr(Z_{mT} \leq z) = F_m(z)$$

for all $m$ and $z$, where

$$\lim F_m(z) = F(z)$$

for all continuity point of the probability distribution function $F(z)$. Then

$$\lim_{T\to\infty} Pr(S_T \leq z) = F(z)$$

for all continuity points of $F(z)$.

Note that this result is similar to Cramér's Theorem about $S_T = Z_T + X_T$ (where $\operatorname{plim} X_T = 0$ and $Z_t$ converges in distribution). However, here $Z_{mT}$ and $X_{mT}$ depend on an m, which cancel out in the sum $S_T$.

We also need some simple lemmas.

Lemma 1.
$$\left(\sum_{s=-\infty}^{\infty} \gamma_s \cos\lambda s\right)^2 + \left(\sum_{s=-\infty}^{\infty} \gamma_s \sin\lambda s\right)^2 = \frac{2\pi f(\lambda)}{\sigma^2} \tag{8}$$

Proof by substituting $e^{i\lambda s} = \cos\lambda s + i\sin\lambda s$ in (4).

Lemma 2.
$$\lim_{N\to\infty} \frac{1}{N}\sum_{r=1}^{N} \cos(\lambda r+\phi) = 0 \quad ; \quad 0<\lambda<2\pi \tag{9}$$

which is also true with "cos" replaced by "sin".

Proof by using

$$\frac{1}{N}\sum_{r=1}^{N}\cos(\lambda r+\phi) = \frac{1}{2N}\left[e^{i(\lambda+\phi)}(1-e^{i\lambda N}) + e^{-i(\lambda+\phi)}(1-e^{-i\lambda N}\right]\Big/(1-e^{-i\lambda})$$

Lemma 3.
$$\lim \frac{1}{N}\sum_{r=1}^{N} \sin^2(\lambda r+\phi) = \frac{1}{2} \quad ; \quad 0<\lambda<2\pi \tag{10}$$

Proof: By
$$\sin^2(\lambda r+\phi) = \tfrac{1}{2}(1-\cos(2\lambda r+2\phi))$$

and Lemma 2.

Lemma 4. Let $0<\lambda_j<\pi$ ; $\lambda_j \neq \lambda_i$ ; $i \neq j$ ; $H_j \neq 0$ and

$$\alpha_r(n) = \sum_{j=1}^{n} H_j \sin(\lambda_j r+\phi_j) \tag{11}$$

Then

$$\beta(n) = \lim \frac{1}{N} \sum_{r=1}^{N} (\alpha_r(n))^2 = \frac{1}{2} \sum_{j=1}^{n} H_j^2$$

Proof is by induction on n, starting from (10). We have

$$(\alpha_r(n+1))^2 = (\alpha_r(n))^2 + H_{n+1}^2 \sin^2(\lambda_{n+1}r+\phi_{n+1}) +$$

$$+ 2\alpha_r(n)\, H_{n+1} \sin(\lambda_{n+1}r+\phi_{n+1})$$

Hence by (10)

$$\beta(n+1) = \beta(n) + \tfrac{1}{2}H_{n+1}^2 + 2H_{n+1} \sum_{j=1}^{n} \lim \frac{1}{N} \sum_{r=1}^{n} \sin(\lambda_j r+\phi_j)\sin(\lambda_{n+1}r+\phi_{n+1})$$

But the product of sines in the last term can be written

$$\tfrac{1}{2}\left[\cos((\lambda_j-\lambda_{n+1})r+\phi_j-\phi_{m+1}) - \cos((\lambda_j+\lambda_{n+1})\phi_j+\phi_j+\phi_{n+1})\right]$$

Hence the last term goes to 0 by Lemma 2.

Theorem: If $\{Z_t\}$ is defined by (1) and

$$\lim_{c\to\infty} \sup_{t=1,2,\ldots} \int_{|v|>c} v^2\, d\, F_t(v) = 0 \qquad (12)$$

where $F_t(v) = \Pr(V_t \leq v)$, then

$$\sqrt{T}(A(\lambda_1), B(\lambda_1), \ldots, A(\lambda_n), B(\lambda_n)) \qquad (13)$$

$0<\lambda_1<\lambda_2<\ldots<\lambda_n < \pi$, converges in distribution to a multinormal variable with 0 mean, uncorrelated components and variances

$$4\pi f(\lambda_1), 4\pi f(\lambda_1, \ldots, 4\pi f(\lambda_n), 4\pi f(\lambda_n) \qquad (14)$$

respectively.

Proof: We may suppose $\zeta = 0$. Consider first the limit distribution of $\sqrt{T}\, A(\lambda)$. We shall apply T.W. Anderson's Theorem to

$$\sqrt{T}\, A(\lambda) = a_{mT}(\lambda) + (\sqrt{T}\, A(\lambda) - a_{mT}(\lambda)) \qquad (15)$$

where $a_{mT}(\lambda) = \frac{2}{\sqrt{T}} \sum_{t=1}^{T} Z_{tm} \cos\lambda T$ ; and $Z_{tm} = \sum_{s=-m}^{m} \gamma_s V_{t-s}$

We have

$$E\left[a_{mT}(\lambda) - \sqrt{T}\,A(\lambda)\right]^2 = \frac{4}{T}\,E\left(\sum_{t=1}^{T} \cos\lambda t \sum_{|s|>m} \gamma_s V_{s-t}\right)^2 =$$

$$= \frac{4}{T} \sum_{t,t'=1}^{T} \cos\lambda t \cos\lambda t' \sum_{|s|,|s'|>m} \gamma_s \gamma_{s'}\, E\, V_{t-s} V_{t'-s'} =$$

$$= \frac{4\sigma^2}{T} \sum_{|s|,|s'|>m} \gamma_s \gamma_{s'} \sum_{t=1}^{T} \cos\lambda t \cos\lambda(t-s+s') \leq \qquad (16)$$

$$\leq 4\sigma^2 \left(\sum_{|s|>m} |\gamma_s|^2\right)$$

having made use of the fact that $E\,V_{t-s}V_{t'-s'} = \sigma^2$ or 0 according as $t' =$ or $\neq t-s+s'$. The interchange of E and $\sum_{|s|,|s'|}$ is permitted since the general term is dominated by $|\gamma_s \gamma_{s'}|\,|V_{t-s}|\,|V_{t-s'}|$, the expectation of which is $\leq$ a general term $|\gamma_s|\,|\gamma_{s'}|\sigma^2$ of a convergent series.

We get from (16)

$$\operatorname*{plim}_{m\to\infty} |\sqrt{T}\,A(\lambda) - a_{mT}(\lambda)| = 0 \qquad (17)$$

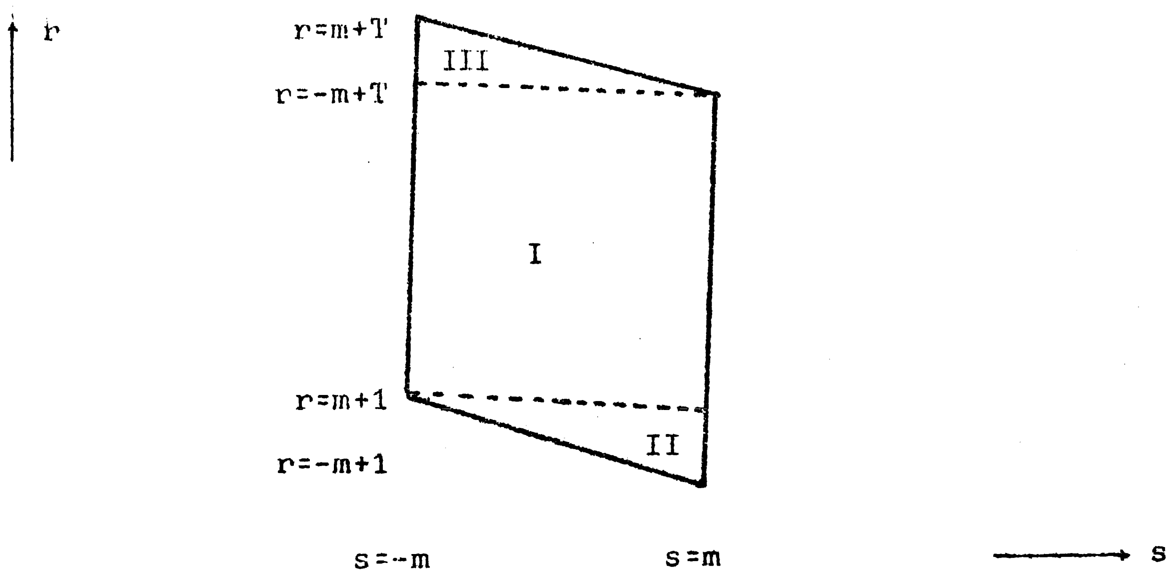
uniformly in T.

We now consider the first term in (15), which we write

$$a_{mT}(\lambda) = \frac{2}{\sqrt{T}} \sum_{t=1}^{T} \sum_{s=-m}^{m} \gamma_s V_{t-s} \cos\lambda t =$$

$$= \frac{2}{\sqrt{T}} \sum_{s=-m}^{m} \sum_{r=-s+1}^{-s+T} \gamma_s V_r \cos\lambda(s+r) \qquad (18)$$

We divide the region of summation into I + II + III




The regions II and III have constant number of points. Hence the variances of the sums over these regions (including the factor $2/\sqrt{T}$) go to 0 as $T \to \infty$. Thus the limit distribution of $a_{mT}$ is the same as the limit distribution of

$$\bar{a}_{mT}(\lambda) = \frac{2}{\sqrt{T}} \sum_{r=m+1}^{T-m} \alpha_r V_r \tag{19}$$

where

$$\alpha_r = \sum_{-m}^{m} \gamma_s \cos \lambda(r+s) \tag{20}$$

We introduce

$$\alpha^T = \left(\sum_{m+1}^{T-m} \alpha_r^2\right)^{\frac{1}{2}} \quad , \quad W_r^T = \alpha_r V_r / \sigma \alpha^T \tag{21}$$

Then

$$\bar{a}_{mT} = 2\,\sigma \frac{\alpha^T}{\sqrt{T}} \sum_{r=m+1}^{T-m} W_r^T \tag{22}$$

Now obviously from (20)

$$|\alpha_r| \le \sum_{s=-\infty}^{\infty} |\gamma_s| = \alpha \quad \text{(say)} \tag{23}$$

$$\alpha_r = \cos\lambda r \sum_{s=-m}^{m} \gamma_s \cos\lambda s - \sin\lambda r \sum_{s=-m}^{m} \gamma_s \sin\lambda s =$$

$$= K_m \sin(\lambda r + \phi_m) \tag{24}$$

defining $K_M$ and $\phi_m$ by the two sums $\sum_{-m}^{m}$ . Then by Lemma 1,

$$\lim_{m\to\infty} K_m = 2\pi f(\lambda)/\sigma^2 \tag{25}$$

From (24) and Lemma 3 we get

$$\lim_{T\to\infty} \frac{1}{\sqrt{T}} \alpha^T = K_m/\sqrt{2} \quad , \quad \lim_{T\to\infty} \alpha^T = \infty \tag{26}$$

We now make use of Lindeberg's Theorem on $\sum_{r=m+1}^{T-m} W_r^T$. We verify that

$$\sum_{r=-m+1}^{T-m} \operatorname{var} W_r^T = 1 \tag{27}$$

Let $F_r$ and $F_r^T$ be the cumulative distribution functions of $V_r$ and $W_r^T$ , respectively. We have for the Lindeberg criterion

$$\sum_{r=m+1}^{T-m} \int_{|w|>\delta} w^2 dF_r^T(w) = \sum_{r=m+1}^{T-m} \frac{\alpha_r^2}{(\alpha^T)^2\sigma^2} \int_{|v|>\sigma\delta\alpha^T/\alpha_r} v^2 dF_r(v) \le$$

$$\le \frac{1}{\sigma^2} \sup_r \int_{|v|\ge\sigma\delta\alpha^T/\alpha} v^2 dF_r(v) \tag{28}$$

making first use of (23) in the region of integration, then replacing the integral by its supremum and finally using (21). Now by the second relation (26) and the assumption (12) of our theorem, the last term of (28) goes to 0. It follows from Lindeberg's Theorem that $T\to\infty$, $\sum_{m+1}^{T-m} W_r^T$ converges in distribution to norm (0,1). By (22) and the first equation (26) it then follows that $\bar{a}_{mT}(\lambda)$

converges in distribution to a variable $\bar{a}_m(\lambda)$, where $\bar{a}_m(\lambda)$ is norm$(0,\sqrt{2}\,\sigma K_m)$. Now by (25), $\bar{a}_m(\lambda)$ converges in distribution to norm $(0,\sqrt{4\pi f(\lambda)}$ as $m \to \infty$. The assumptions of T. W. Anderson's Theorem are satisfied and we have proved that $\sqrt{T}\,A(\lambda)$ (see eq.(15)) converges in distribution to norm $(0,\sqrt{4\pi f(\lambda)}$, which is a special case of our theorem.

In the general case we make use of the fact that if a random vector X is such that for any sure vector g, g'X converges in distribution norm$(0,\kappa)$, where $\kappa^2 = g'\sigma g$, then X converges in distribution to the multinormal with expectation 0 and covariance matrix $\sigma$.

In our case we choose $g = (g_1,h_1,\ldots,g_n,h_n)$ and consider

$$\sqrt{T}\sum_{j=1}^{n}\left[g_j A(\lambda_j) + h_j B(\lambda_j)\right] =$$

$$\sqrt{\frac{2}{T}}\sum_{t=1}^{T}\sum_{j=1}^{n}(g_j \cos\lambda_j t + h_j \sin\lambda_j t)Z_t = \sqrt{T}\,A_T \text{ (say)} \tag{29}$$

As above we now introduce $a_{mT}$ by replacing $Z_t$ by $Z_{mt}$ in $\sqrt{T}\,A_T$ where $Z_{mt}$ is given by (15). Then as in (15) we write

$$\sqrt{T}\,A_T = a_{mT} + (\sqrt{T}\,A_T - a_{mT}) \tag{30}$$

and prove that the last term converges in probability to 0 uniformly in T as $m \to \infty$. (See eq. (17)). Thus $\sqrt{T}\,A_T$ has the same limit in distribution as $a_{mT}$, which may be written

$$a_{mT} = \sqrt{\frac{2}{T}}\sum_{s=-m}^{m}\sum_{r=-s+1}^{-s+T}\gamma_s V_r \sum_{j=1}^{m}(g_j \cos\lambda_j(s+r) + h_j \sin\lambda_j(s+r)) \tag{31}$$

(see eq. (18)). We deduce as above that $a_{mT}$ has the same limit in distribution as

$$\bar{a}_{mT} = \sqrt{\frac{2}{T}}\sum_{s=m+1}^{T-m} V_r \alpha_r(n) \tag{32}$$

where

$$\alpha_r(n) = \sum_{s=-m}^{m} \gamma_s \sum_{j=1}^{n} (g_j \cos \lambda_j(r+s) + h_j \sin \lambda_j(r+s)) \tag{33}$$

(see eq. (19) and (20)).

We have

$$|\alpha_r(n)| \le \sum_{\infty-}^{\infty} |\gamma_s| \sum_{j=1}^{n} (|g_j|+|h_j|) = \alpha(n) \quad \text{(say)} \tag{34}$$

(see eq. (23)).

(33) may be written

$$\alpha_r(n) = \sum_{s=-m}^{m} \gamma_s \sum_{j=1}^{n} H_j \cos [\lambda_j(r+s)+\theta_j] \tag{35}$$

where

$$H_j^2 = \sum (g_j^2 + h_j^2) \tag{36}$$

As in eq. (24) we may use the addition rule for cosine on the two terms $\lambda_j r + \theta_j$ and $\lambda_j s$ and get

$$\alpha_r(n) = \sum_{s=-m}^{n} H_j \; K_{jm} \sin (\lambda_j r + \theta_j + \phi_j m) \tag{37}$$

where $K_{jm}$ and $\phi_{jm}$ are defined as $K_m$ and $\phi_m$ in connection with eq. (24).

By Lemma 4 we get

$$\lim_{T\to\infty} \frac{1}{T} \sum_{r=m+1}^{T-m} (\alpha_r(n))^2 = \tfrac{1}{2} \sum_{j=1}^{n} K_{jm}^2 (g_j^2 + h_j{}^2) \tag{38}$$

We can now proceed as before, making use of Lemma 1 to obtain that (38) goes to

$$\pi \sum f(\lambda_j)(g_j^2 + h_j^2)/\sigma^2 \tag{39}$$

as $m \to \infty$ and we obtain that $\sqrt{T}\, A_T$ converges in distribution to norm $(0,\kappa)$, where

$$\kappa^2 = 4\pi \sum_{j=1}^{n} f(\lambda_j)(g_j^2 + h_j^2) \tag{40}$$

but $\kappa^2$ may be written $g' \sigma g$ where $g = (g_1, h_1, \ldots, g_n, h_n)$ and $\sigma$ is a diagonal matrix with diagonal given by (14). Hence by the general result cited before eq. (29), the theorem is proved.

The asymptotic distribution of the spectrum

$$I_T(\lambda) = \frac{T}{8\pi}(A(\lambda)^2 + B(\lambda)^2) \tag{41}$$

(see Chapter I,B (12)) now follows immediately.

$$2\, I_T(\lambda_j)/f(\lambda_j) \quad ; \quad 0 < \lambda_1 < \pi \quad ; \quad j=1,2,\ldots \tag{42}$$

where $f(\lambda)$ is given by (4), are independent and chi-square distributed with two degrees of freedom.

We easily derive that this result and the result in the theorem is true if $\zeta = E\, Z_t$ is replaced by the empirical mean

$$\bar{Z}_T = \frac{1}{T} \sum_{t=1}^{T} Z_t$$

in $A(\lambda)$ and $B(\lambda)$.

## APPENDIX I: SECOND ORDER MEAN CONVERGENCE

We shall in general assume the random variables to be complex with finite second order moments. The sequence $X_n$ ; $n=1,2,\ldots,$ is said to converge to $X$ in quadratic mean; $X_n \overset{2}{\to} X$; if $\lim_{n\to\infty} E|X_n-X|^2 = 0$.

Lemma 1.

$$\sqrt{E|X|^2} - \sqrt{E|Y|^2} \leq \sqrt{E|X+Y|^2} \leq \sqrt{E|X|^2} + \sqrt{E|Y|^2}.$$

Proof: by squaring and using Schwartz inequality $|E\,X\,Y^*| \leq \sqrt{E|X|^2 E|Y|^2}$.

Note that $\sqrt{E|X|^2}$ has the property of a norm and $\sqrt{E|X-Y|^2}$ the property of a metric.

Lemma 2.

$$X_n \overset{2}{\to} X, Y_n \overset{2}{\to} Y \text{ imply } aX_n+bY_n \overset{2}{\to} aX+bY.$$

Proof: Use lemma 1 on $E|aX_n+bY_n-aX-bY|^2 = E|a(X_n-X)+b(Y_n-Y)|^2$.

Lemma 3.

$X_n \overset{2}{\to} X$ and $Y_n \overset{2}{\to} Y$ imply

a) $E|X_n|^2 \to E|X|^2$ , b) $EX_nY_n \to E\,X\,Y$ , c) $EX_n \to E\,X$.

Proof: a) follows from the left inequality in lemma 1 with $X = X_n$ and $Y = -X$. b) follows from $|EX_nY_n - E\,X\,Y| \leq$ $E|X_n-X|\,|Y_n| + E|X|\,|Y_n-Y| \leq \sqrt{E|Y_n|^2 E|X_n-X|^2} + \sqrt{E|X|^2 E|Y_n-Y|^2} \to 0$ since $E|Y_n|^2 \to E|Y|^2$ by a). c) follows from b) with $Y_n = Y = 1$. (Note that it does not follow from $X_n \overset{2}{\to} X$, $Y_n \overset{2}{\to} Y$ that $X_nY_n \overset{2}{\to} XY$.)

Lemma 4. $X_n \overset{2}{\to} X$ implies $E|X_n-X| \to 0$.

Proof: Use $0 \leq \operatorname{var}|X_n-X| = E|X_n-X|^2 - (E|X_n-X|)^2$.

Lemma 5. $X_n \overset{2}{\to} X \Rightarrow \operatorname{plim} X_n = X$.

Proof: By Markov's inequality (that for any $Z$ such that $\Pr(Z \geq 0) = 1$ we have $\Pr(Z \geq \varepsilon) \leq EZ/\varepsilon$ with $\varepsilon > 0$. Take $Z = |X_n - X|^2$).

Lemma 6.

$\operatorname{plim}_{m,n\to\infty}(X_m - X_n) = 0$ implies that there exist a subsequence $X_{n_i}$ ; $i=1,2,\ldots$ and an $X$ such that $\lim X_{n_i} = X$ with probability 1.

Proof: For any $\delta$, $\varepsilon > 0$ there is an $N$ such that $m > n \geq N$ implies $\Pr(|X_m - X_n| > \delta) < \varepsilon$. Let $\varepsilon_i$, $\delta_i > 0$ and $\sum \delta_i, \sum \varepsilon_i$ convergent. We can then define $n_i$ ; $i=1,2,\ldots$ such that

$n_{i+1} > n_i$ and

$$\Pr(|X_m - X_{n_i}| > \delta_i) < \varepsilon, \text{ for all } m > n_i.$$

In particular

$$\Pr(|X_{n_{i+1}} - X_{n_i}| > \delta_i) < \varepsilon_i.$$

Hence

$$\sum_i \Pr(|X_{n_{i+1}} - X_{n_i}| > \delta_i)$$

converges and by Borel-Cantelli

$$\Pr(|X_{n_{i+1}} - X_{n_i}| > \delta_i \text{ for infinitely many } i) = 0.$$

Thus with probability 1, $|X_{n_{i+1}} - X_{n_i}| \leq \delta_i$ for all but a finite number of $i$. Hence $\sum |X_{n_{i+1}} - X_{n_i}|$ converges almost certainly and

$$X_{n_j} = X_{n_1} + \sum_{i=1}^{j} (X_{n_{i+1}} - X_{n_i})$$

has a limit $X$ with probability 1.

Lemma 7.

$|X_m - X_n| \overset{2}{\to} 0$ as $m,n \to \infty$ implies that there exists an $X$ such that $X_n \overset{2}{\to} X$.

Proof: By $E|X_m - X_n|^2 \to 0$ and Markov's inequality we have $\operatorname{plim}|X_m - X_n| = 0$. Hence by Lemma 6 there exist an $X$ and a subsequence $X_{n_i}$ such that $X_{n_i} \to X$ a.e. Hence by Fatou's lemma

$$E|X_m - X|^2 \leq \liminf_i E|X_m - X_{n_i}|^2.$$

However, by the assumption the expectation on the right hand side can be made $< \varepsilon$ for $m$ and $n_i$ sufficiently large. Hence the liminf of it can be made $\leq \varepsilon$ for $m$ sufficiently large, which proves the assertion.

(Fatou's lemma: $f_n \geq 0$ implies $\liminf \int f_n d\mu \geq \int \liminf f_n d\mu$).

Lemma 8. (Loeve)

$X_n$ ; $n=1,2,\ldots$ converges in quadratic mean if and only if

$$E X_m X_n^* \to \text{finite limit } C$$

as $m,n \to \infty$.

Proof: The "only if" part follows from Lemma 3. The "if" part follows because $E|X_m - X_n|^2 = EX_m X_m^* + EX_n X_n^* - EX_m X_n^* - EX_n X_m^*$.

Below we shall assume, partly for convenience, that the random variables in question are real.

The following two lemmas are fundamental in probability theory and are given here without proofs.

Lemma 9.

If $X_1, X_2, \ldots$ are independent and $\sum_{j=1} X_j$ converges in probability (or in quadratic mean) then $\sum X_j$ converges with probability 1.

Lemma 10.

If the $V_t$ ; $t=1,2,\ldots$ are independent identically distributed with $E V_t = 0$, $\operatorname{var} V_t = \sigma^2 < \infty$ and $\sum \alpha_i^2$ converges, then $\sum_{i=1}^{\infty} \alpha_i V_i$ converges with probability 1.

For the proofs of Lemmas 9-10 see M. Loéve: Probability Theory I (4.ed. 1977) Section 17.3 and 18.2.

<u>Lemma 11</u>.

For any set of random variables $V_1, V_2, \ldots, V_n$

$$\operatorname{var} \sum_{i=1}^{n} \alpha_i V_i \le \left(\sum_{i=1}^{n} |\alpha_i|\right)^2 \max \operatorname{var} V_i .$$

<u>Proof</u>: It follows from $|\operatorname{cov}(V_i, V_j)| \le \sqrt{\operatorname{var} V_i \operatorname{var} V_j}$ .

<u>Lemma 12</u>.

If $\sum_{i=1}^{\infty} |\alpha_i| < \infty$ and the process $\{V_t\}_{t=1,2,\ldots}$ has uniformly bounded variance, $EV_t = 0$, then $\sum_{1}^{\infty} \alpha_i V_i$ converges in quadratic mean.

<u>Proof</u>: We have by Lemma 11

$$E\left(\sum_{m}^{n} \alpha_i V_i\right)^2 \le K\left(\sum_{m}^{n} |\alpha_i|\right)^2$$

and use Lemma 7.

<u>Lemma 13</u>.

If $\sum_{i=0}^{\infty} |\psi_i| < \infty$ and $\{a_t\}$ is stationary in the wide sense, then

$$V_t = \sum_{i=0}^{\infty} a_{t-i} \psi_i = \sum_{i=-\infty}^{t} a_i \psi_{t-i}$$

converges in quadratic mean and is stationary in the wide sense.

<u>Proof</u>: The first assertion follows from Theorem 12. To prove the second statement, introduce

$$V_{tm} = \sum_{i=0}^{m} a_{t-i} \psi_i .$$

Then $V_{tm} \xrightarrow{2} V_t$. Hence by Lemma 3

$$EV_{tm} V_{t+hm} \to EV_t V_{t+h} .$$

However,

$$EV_{tm} V_{t+hm} = \sum_{i,i'=0}^{m} \psi_i \psi_{i'} E a_{t-i} a_{t-i'}$$

which is independent of $t$, since $a_t$ is stationary. Hence, also the limit $EV_tV_{t+h}$ is independent of $t$. ||

Lemma 14.

If

1) $a_t$ is stationary in the wide sense

2) $\sum_{j=-\infty}^{+\infty} |\psi_{ij}| < \infty$ for all $i$

3) $\sum_{j=-\infty}^{\infty} |\sum_{i=-\infty}^{\infty} \psi_{ij}| < \infty$

4) $Z = \sum_{-\infty}^{\infty} Y_i$ exists in quadratic mean, where

$Y_i = \sum_{j=-\infty}^{\infty} \psi_{ij} a_j$ in quadratic mean.

then $Z = \sum_{j=-\infty}^{\infty} a_j \sum_{i=-\infty}^{\infty} \psi_{ij}$ in quadratic mean.

Proof: The last double sum converges in quadratic mean by assumption 3) and Lemma 12. Hence we just have to prove that it equals $Z$. We have

$$\sum_{-n}^{m} Y_i = \sum_{-n}^{m} \sum_{-\infty}^{+\infty} \psi_{ij} a_j = \sum_{j=-\infty}^{\infty} a_j \sum_{i=-n}^{m} \psi_{ij}$$

by Lemma 2 and assumption 2). Hence

$$Q = \left[E(Z - \sum_{j=-\infty}^{\infty} a_j \sum_{i=-\infty}^{\infty} \psi_{ij})^2\right]^{\frac{1}{2}} \leq$$

$$\leq \left[E(Z - \sum_{i=-n}^{m} Y_i)^2\right]^{\frac{1}{2}} + \left\{E\left[\sum_{j=-\infty}^{\infty} a_j (\sum_{i=m+1}^{\infty} \psi_{ij} + \sum_{i=-\infty}^{-n-1} \psi_{ij})\right]^2\right\}^{\frac{1}{2}}$$

by Lemma 1. The first term on the right hand side goes to 0 by assumption 4). The last term squared is

$$\text{var} \sum_{-\infty}^{+\infty} a_j \Big( \sum_{i=m+1}^{\infty} \psi_{ij} + \sum_{i=-\infty}^{-n-1} \psi_{ij} \Big) \leq$$

$$\leq \Big( \sum_{j=-\infty}^{+\infty} \Big| \sum_{i=m+1}^{\infty} \psi_{ij} + \sum_{i=-\infty}^{-n-1} \psi_{ij} \Big| \Big)^2 \text{ var } a$$

by Lemma 11. Hence by assumption 3) the last term goes to 0.

Hence

$$E(Z - \sum_{j=-\infty}^{\infty} a_j \sum_{i=-\infty}^{\infty} \psi_{ij})^2 = 0$$

and the result follows.

(Above we have used Schwartz inequality $|EXY^*| \leq \sqrt{E|X|^2 E|Y|^2}$ in the case of complex variables. This is proved as follows. For any complex number $\lambda$, we have that

$E|\lambda X-Y|^2 = \lambda\lambda^* E|X|^2 - \lambda EXY^* - \lambda^* EYX^* + E|Y|^2$ is $\geq 0$. We write $EXY^* = re^{i\alpha}$, hence $EX^*Y = re^{-i\alpha}$, and substitute $\lambda = te^{-i\alpha}$, to obtain

$$t^2 E|X|^2 - 2tr + E|Y|^2 \geq 0$$

But the polynomial in $t$ cannot be $\geq 0$ for all $t$ if it has two distinct real zeros. Hence $r^2 \leq E|X|^2 E|Y|^2$.)

## APPENDIX II : FOURIER SERIES

We shall prove the following (see Apostol (1957))

Theorem. Let f be a real function of a real variable with period $2\pi$ (i.e. $f(x) = f(x+2\pi)$ for all $x$). If

(i) f is Riemann integrable over $[-\pi,\pi]$

(ii) for a given x there is an interval $x-\delta, x+\delta$ , $\delta > 0$, where f is of bounded variation; then

$$\tfrac{1}{2}[f(x+)+f(x-)] = \tfrac{1}{2} a_0 + \sum_{j=1}^{\infty} (a_j \cos j x + b_j \sin j x) \tag{1}$$

where

$$a_n = \frac{1}{\pi} \int_{-\pi}^{+\pi} f(t) \cos nt \, dt \quad , \quad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin nt \, dt \tag{2}$$

Note that if f is of bounded variation over $[-\pi,\pi]$, then (i) and (ii) are fulfilled, since any function of bounded variation is Riemann integrable.

In order to prove the theorem, we shall first prove some lemmas.

Lemma 1. (Riemann-Lebesgue). If f is Riemann-integrable over $[a,b]$, then for any $\beta$ , we have

$$\lim_{\alpha\to\infty} \int_a^b f(t) \sin(\alpha t+\beta) \, dt = 0 \tag{3}$$

Proof. For $f(t) = 1$ if $t \in [c,d] \subset [a,b]$, $f(t) = 0$ otherwise, we get for the integral in (3); $(\cos(c\alpha+\beta) - \cos(d\alpha+\beta))/\alpha$ , which is $\le 2/\alpha$ in absolute value. Hence (3) is true for this f , hence for any step function $m(t)$

$$\int_a^b m(t) \sin(\alpha t+\beta) \, dt \to 0 \tag{4}$$

as $\alpha \to \infty$. However, from the theory of the Riemann integral we know that there exist two step functions $m(t)$ and $M(t)$ such that $m(t) \le f(t) \le M(t)$ and $\int_0^b |M(t)-m(t)|dt$ is arbitrarily small. Hence

$$\left|\int_a^b (f(t)-m(t)) \sin(\alpha t+\beta)\, dt\right| \le \int_a^b |M(t)-m(t)|\, dt \tag{5}$$

Combining (4) and (5) we get the lemma.

<u>Lemma 2</u>. (Jordan). If $g$ is of bounded variation over $[0,\delta]$, $\delta > 0$, then

$$\lim_{\alpha\to\infty} \frac{2}{\pi} \int_0^\delta g(t) \frac{\sin\alpha t}{t} dt = g(0+) \tag{6}$$

<u>Proof</u>. We may assume $g$ to be increasing. We then have for $0 < h < \delta$,

$$\int_0^\delta g(t) \frac{\sin\alpha t}{t} dt = \int_0^h [g(t)-g(0+)] \frac{\sin\alpha t}{t} dt + g(0+)\int_0^h \frac{\sin\alpha t}{t} dt + \int_h^\delta g(t) \frac{\sin\alpha t}{t} dt \tag{7}$$

Now it is seen after substituting $\alpha t = u$ under the sign of integration, that the second integral on the right hand side goes to $\frac{\pi}{2} g(0+)$. The third integral goes to 0 by lemma 1. Thus it suffices to prove that the first integral can be made arbitrarily small. We have with

$$g(t) - g(0+) = G(t) \quad , \quad F(t) = \int_t^h \frac{\sin\alpha v}{v} dv \tag{8}$$

that this integral may be written

$$I = -\int_0^h G(t)\, d\,F(t) = \int_0^h F(t)\, d\,G(t) = F(\varepsilon)\, G(h) \tag{9}$$

for some $\varepsilon \in [0,h]$ (having used partial integration and the mean value theorem). Introducing (8) in (9) we get

$$I = [g(h)-g(0+)] \int_{\varepsilon\alpha}^{h\alpha} \frac{\sin t}{t} dt$$

However, $\int_p^q \frac{\sin t}{t} dt$ is bounded under variation of $p$ and $q$ since $\int_0^\infty \frac{\sin t}{t} dt$ converges. Hence $I$ can be made arbitrarily small by choosing $h$ small (we may set $h = 1/\alpha$ ). Hence the lemma is proved.

Lemma 3. The partial sum in (1)

$$S_n(x) = \frac{a_0}{2} + \sum_{j=1}^{n} (a_j \cos jx + b_j \sin jx) \tag{10}$$

may be written

$$S_n(x) = \frac{2}{\pi} \int_0^{\pi/2} g(t) \frac{\sin(2n+1)t}{\sin t} dt \tag{11}$$

where

$$g(t) = \tfrac{1}{2}[f(x+2t)+f(x-2t)] \tag{12}$$

Proof. Introducing (2) in (1) we get

$$S_n(x) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t)\, D_n(t-x)\, dx \tag{13}$$

where

$$D_n(x) = \tfrac{1}{2} + \sum_{j=1}^{n} \cos jt = \begin{cases} \sin (n+\tfrac{1}{2})t/\sin \frac{t}{2} & \text{if } t \neq m\pi \\ n + \tfrac{1}{2} & \text{if } t = m\pi \end{cases} \tag{14}$$

for any integer $m$.

Since $D_n(x)$ and $f(x)$ are periodic and $D_n(x) = D_n(-x)$

$$S_n(x) = \frac{1}{\pi} \int_{x-\pi}^{x+\pi} f(t)\, D_n(t-x)\, dt = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x+t)\, D_n(t)\, dt =$$
$$= \frac{2}{\pi} \int_0^{\pi/2} g(t)\, D_n(2t)\, dt \tag{15}$$

where $g$ is given by (12). Combining (15) and (14) we get (11).

We can now prove the theorem by proving that $S_n(x)$ given by (11) goes to $g(x)$ as $n \to \infty$. We first note that $S_n(x)$ has the same limit as

$$s_n(x) = \frac{2}{\pi} \int_0^{\pi/2} g(t) \frac{\sin(2n+1)t}{t} \, dt \qquad (16)$$

since we may apply lemma 1 to

$$\int_0^{\pi/2} \left(\frac{1}{t} - \frac{1}{\sin t}\right) g(t) \sin(2n+1)t \, dt$$

with $\alpha = 2n+1$ (note that the first factor in the integrand can be defined to be continuous for $t = 0$). Now we may divide the integral in (16) in two parts

$$\frac{2}{\pi} \int_0^{\pi/2} = \frac{2}{\pi} \int_0^{\delta} + \frac{2}{\pi} \int_{\delta}^{\pi/2}$$

where $\delta$ is defined in the theorem (and assumed $< \frac{\pi}{2}$). The first integral on the right hand side goes to $g(0+)$ by lemma 2 and the second integral goes to $0$ by lemma 1. (The case $\delta \geq \frac{\pi}{2}$ is easier.) This proves the theorem.

## APPENDIX III : PARTIAL AUTOCORRELATION

With a given random set of variables $X_1, X_2, \ldots, X_n$, let $\xi_1(X_3,\ldots,X_n) = E(X_1|X_3,\ldots,X_n)$ and $\xi_2(X_3,\ldots,X_n) = E(X_2|X_3,\ldots,X_n)$. Then the partial correlation coefficient between $X_1$ and $X_2$, relatively to $X_3,\ldots,X_n$ is defined as

$$\rho_{12\cdot 3\cdots n} = \frac{E[X_1-\xi_1(X_3,\ldots,X_n)][(X_2-\xi_2(X_3,\ldots,X_n)]}{\{E[X_1-\xi_1(X_3,\ldots,X_n)]^2 E[X_2-\xi_1(X_3,\ldots,X_n)^2]\}^{\frac{1}{2}}} \tag{1}$$

The coefficient may be interpreted in the following manner. Suppose that there is negative correlation between $X_1$ = the bulk of crop and $X_2$ = temperature during growth period. Then this may be "explained" by a negative correlation between temperature $X_2$ and $X_3$ = average amount of rain in the growth period. The partial correlation between $X_1$ = crop and $X_2$ = temperature, relatively to $X_3$ = amount of rain, may be positive. Thus calculating the partial correlation coefficient amounts to "correcting" for the amount of rain.

We shall be concerned with the least square partial correlation coefficient which is defined in a slightly different manner. $\varphi_3,\ldots,\varphi_n$ and $\kappa_3,\ldots,\kappa_n$ are defined by minimizing

$$E(X_1-\sum_{j=3}^{n}\varphi_j X_j)^2 \quad , \quad E(X_2-\sum_{j=3}^{n}\kappa_j X_j)^2 \tag{2}$$

with respect to $\varphi$ and $\kappa$. Then $\varphi$ and $\kappa$ are given by

$$E X_i(X_1-\sum_{j=3}^{n}\varphi_j X_j) = 0 \quad , \quad E X_i(X_2-\sum_{j=3}^{n}\kappa_j X_j) = 0 \tag{3}$$

For convenience we assume $EX_i = 0$ and we write $\mathrm{cov}(X_i,X_j) = \gamma_{ij}$. Then from (3), $\sum_{j=3}^{n}\gamma_{ij}\varphi_j = \gamma_{1j}$ , $\sum_{j=3}^{n}\gamma_{ij}\kappa_j = \gamma_{2j}$ ; $j=3,\ldots,n$ (4)

Then we use $\xi_1(X_3,\ldots,X_n) = \sum_{j=3}^{n}\varphi_j X_j$, $\xi_2(X_3,\ldots,X_n) = \sum_{j=3}^{n}\kappa_j X_j$ in (1).

Using (3),the numerator in (1) may be written

$$E(X_1 - \sum_3^n \varphi_j X_j) X_2 = \gamma_{12} - \sum_3^n \gamma_{2j} \varphi_j \tag{5}$$

or, alternatively

$$E X_1 (X_2 - \sum_2^n \kappa_j X_j) = \gamma_{12} - \sum_3^n \gamma_{1j} \kappa_j \tag{6}$$

Furthermore

$$E(X_1 - \xi_1)^2 = E(X_1 - \xi_1) X_1 = \gamma_{11} - \sum_3^n \gamma_{1j} \varphi_j \tag{7}$$

and

$$E(X_2 - \xi_2)^2 = \gamma_{22} - \sum_3^n \gamma_{2j} \kappa_j \tag{8}$$

Hence we have

$$\rho_{12\cdot 3\cdots n} = \frac{\gamma_{12} - \sum_{j=3}^n \gamma_{2j} \varphi_j}{\{(\gamma_{11} - \sum_3^n \gamma_{1j} \varphi_j)(\gamma_{22} - \sum_3^n \gamma_{2j} \kappa_j)\}^{\frac{1}{2}}} \tag{9}$$

where the $\varphi_j$ and $\kappa_j$ are given by (4).

---

Now let $\{Z_t\}_{t=\cdots-1,0,1,\cdots}$ be a stationary process

$$E Z_t = 0 , \quad \gamma_i = E Z_t Z_{t-i} , \quad \rho_i = \gamma_i / \gamma_0 \tag{10}$$

We define the k-th order partial autocorrelation coefficient $\pi_k$ as the partial correlation coefficient between $Z_t$ and $Z_{t-k}$ relatively to $Z_{t-1},\ldots,Z_{t-k}$. $\pi_k$ has the following interpretation. Having "explained" $Z_t$ by the last k-1 observations $Z_{t-1},\ldots$ $\ldots,Z_{t-k+1}$, is there any reason to involve $Z_{t-k}$ also? $\pi_k$ measures the importance of involving $Z_{t-k}$ in order to explain $Z_t$.

<u>Theorem.</u> Let $\varphi_{k,j}$ ; $j = 1,2,\ldots,k$ ; $k = 1,2,\ldots,$ be defined by

$$\sum_{j=1}^{k} \rho_{i-j}\,\varphi_{kj} = 0 \; ; \quad i = 1,2,\ldots,k \tag{11}$$

then

$$\pi_k = \varphi_{kk}$$

$\varphi_{kj}$ may also be found from the recursion formulae

$$\varphi_{mm} = (\rho_m - \sum_{j=1}^{m-1} \varphi_{m-1j}\,\rho_{m-j})/(1 - \sum_{j=1}^{m-1} \varphi_{m-1\,m-j}\,\rho_{m-j}) \,, \tag{12}$$

$$\varphi_{mj} = \varphi_{m-1j} - \varphi_{mm}\,\varphi_{m-1\,m-j} \,, \tag{13}$$

starting with

$$\varphi_{22} = (\rho_2 - \rho_1^2)/(1-\rho_1^2) \,, \quad \varphi_{21} = \rho_1(1-\rho_2)/1-\rho_1^2 \tag{14}$$

<u>Proof</u>: We get from (9) and (4)

$$\pi_k = \frac{\rho_k - \sum_{j=1}^{k-1} \varphi_j\,\rho_{k-j}}{[(1-\sum_{j=1}^{k-1} \varphi_j\,\rho_j)(1-\sum_{j=1}^{k-1} \kappa_j\,\rho_{k-j})]^{\frac{1}{2}}} \tag{15}$$

$$\sum_{j=1}^{k-1} \varphi_j\,\rho_{i-j} = \rho_i \quad , \quad \sum_{j=1}^{k-1} \kappa_j\,\rho_{i-j} = \rho_{k-i} \tag{16}$$

Now we have,

$$\sum_{j=1}^{k-1} \kappa_j\,\rho_{k-i} = \sum_{j=1}^{k-1} \kappa_j \sum_{m=1}^{k-1} \varphi_m\,\rho_{k-j-m} = \sum_{m=1}^{k-1} \varphi_m \sum_{j=1}^{k-1} \kappa_j\,\rho_{k-m-j} = \sum_{m=1}^{m-1} \varphi_m\,\rho_m$$

having succesively made use of the first and the second equation (16). Hence (15) may be written

$$\pi_k = (\rho_k - \sum_{j=1}^{k-1} \varphi_j\,\rho_{k-j})/(1-\sum_{j=1}^{k-1} \varphi_j\,\rho_j) \tag{17}$$

where $\varphi_1, \ldots, \varphi_k$ are defined by the first equation (16). To emphazise that $\varphi_j$ depends on k, we now write $\varphi_j = \varphi_{k-1 j}$. We have by (16),

$$\sum_{j=1}^{m} \varphi_{mj}\, \rho_{i-j} = \rho_i \; ; \quad i = 1,2,\ldots,m \tag{18}$$

which is a set of linear sets of equations determining $\varphi_{mj}$ ; $j = 1,2,\ldots,m$ ; for each $m = 2,3,\ldots$ . (We may define $\varphi_{11} = \rho_1$ .) The matrix of (18) is

$$R_m = \begin{pmatrix} 1 , & \rho_1 , & \rho_2 , & \cdots , & \rho_{m-1} \\ \rho_1 , & 1 , & \rho_1 , & \cdots , & \rho_{m-2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \rho_{m-1} , & \cdot & \cdot & \cdot & 1 \end{pmatrix}$$

Now (18) may be written

$$\rho_i - \varphi_{mm}\, \rho_{m-i} = \sum_{1}^{m-1} \varphi_{mj}\, \rho_{i-j} \; ; \quad i = 1,2,\ldots,m-1$$

Hence

$$\begin{pmatrix} \varphi_{m\,1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \varphi_{m\,m-1} \end{pmatrix} = R_{m-1}^{-1} \begin{pmatrix} \rho_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \rho_{m-1} \end{pmatrix} - \varphi_{mm}\, R_{m-1}^{-1} \begin{pmatrix} \rho_{m-1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \rho_1 \end{pmatrix} \tag{20}$$

On the other hand from (18) with m replaced by m-1 ,

$$R_{m-1}^{-1} \begin{pmatrix} \rho_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \rho_{m-1} \end{pmatrix} = \begin{pmatrix} \varphi_{m-1\,1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \varphi_{m-1\,m-1} \end{pmatrix} \tag{21}$$

Now in (18) with m replaced by m-1 , we reverse both the equations and the terms in the sum, i.e. we replace $i$ by $m-i$ and $j$ by $m-j$. We then get

$$\rho_{m-i} = \sum_{j=1}^{m-1} \varphi_{m-1\, m-j}\, \rho_{i-j} ; \quad i = 1,2,\dots,m-1$$

Hence

$$\begin{bmatrix} \varphi_{m-1\, m-1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \varphi_{m-1\, 1} \end{bmatrix} = R_{m-1}^{-1} \begin{bmatrix} \rho_{m-1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \rho_1 \end{bmatrix} \tag{22}$$

Combining (20), (21), (22) we get

$$\begin{bmatrix} \varphi_{m\, 1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \varphi_{m\, m-1} \end{bmatrix} = \begin{bmatrix} \varphi_{m-1\, 1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \varphi_{m-1\, m-1} \end{bmatrix} - \varphi_{mm} \begin{bmatrix} \varphi_{m-1\, m-1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \varphi_{m-1\, 1} \end{bmatrix} \tag{23}$$

which proves (13). We now insert (13) into (18) with $i = m$ and get (12). However the right hand side of (12) is equal to the right hand side of (17) with $k = m$. Hence $\pi_k = \varphi_{kk}$, and everything is proved.

## References

T.W. Anderson (1971). The Statistical Analysis of Time Series, John Wiley & Sons, Inc., New York.

Tom M. Apostol (1957). Mathematical Analysis, Addison-Wesley Publ. Co. Inc., Reading, Mass., U.S.A.

George E.P. Box and Gwilym M. Jenkins (1970). Time Series Analysis, Holden-Day, San Francisco.

Harald Cramér and M.R. Leadbetter (1967). Stationary and Related Stochastic Processes, John Wiley & Sons, Inc., New York.

Edouard Goursat (1933). Cours D'Analyse Mathématique, Tome II, Gauthiers-Villars et Cie, Paris.

I.I. Gihmann and A.V. Skorohod (1974). The Theory of Stochastic Processes I, Springer Verlag, Berlin.

Ulf Grenander and Murray Rosenblatt (1956). Statistical Analysis of Stationary Time Series, Almquist & Wiksell, Stockholm.