TECHNICAL REPORT NO. 158
May 1968
THE VARIANCE OF
SPECTRUM ESTIMATES

By
Henry R. Neave
(University of Nottingham
and
University of Wisconsin).

This research was supported in part by the University of Wisconsin Computing Center, the National Science Foundation, and the Science Research Council of Great Britain.
SUMMARY

Since the asymptotic expression for the variance of estimates of the spectrum of a stationary time series was derived, it has often been used as an approximation to the variance of estimates using finite samples. Little attempt seems to have been made however to justify this or to investigate the nature of the convergence to the asymptotic form. In this paper, an exact expression for the variance is derived on the additional assumption that the time series is a normal process. This expression is then used to study estimates of four different spectra, using fairly small sample sizes, the Tukey and Parzen lag window generators, and three different truncation points in each case. Graphs of the standard deviation and also of the bias and root mean square error are given for the complete survey of one of the spectra investigated; tables are also included showing the behaviour of these quantities for all the four spectra.
ACKNOWLEDGEMENTS

This is the last in a series of seven Technical Reports based mainly on research carried out when preparing my Ph.D. thesis at Nottingham University. I wish to express my sincere gratitude to Professor Clive W. J. Granger for his guidance and encouragement throughout this work.

I must also thank Miss Anne Louise Brady for her skill and patience in writing the computer programs to prepare the many graphs on the Calcomp plotter at the University of Wisconsin Computing Center.
§1. Background

Let \( \{X_t; t = \ldots -2, -1, 0, 1, 2, \ldots \} \) be a real-valued, weakly stationary, discrete stochastic process (or time series) with zero mean and covariance function:

\[
R(v) = E[X_tX_{t+v}] = R(-v). \tag{1.1}
\]

The power spectrum or spectral density \( f(.) \) of the process is the Fourier transform of the covariance function, i.e.

\[
f(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} R(v) \cos v\omega \, dv \tag{1.2}
\]

\[
R(v) = \int_{-\pi}^{\pi} f(\omega) \cos v\omega \, d\omega. \tag{1.3}
\]

The process itself has a spectral representation:

\[
X_t = \int_{-\pi}^{\pi} e^{it\omega} \, dZ(\omega) \tag{1.4}
\]

where \( Z(.) \) is a complex-valued stochastic function with uncorrelated increments, and

\[
E[|dZ(\omega)|^2] = f(\omega) \, d\omega \tag{1.5}
\]

(Cramer (1940)). This expresses \( X_t \) as a superposition of harmonics with random amplitudes, the expectation of whose square
is proportional to the spectral density. The fundamental im-
portance of the spectrum in the analysis of a stationary stochastic
process is therefore that the tendency of the process to oscil-
late with a frequency \( \omega \) is reflected directly in the value of
\( f(\omega) \).

The apparently obvious way to estimate the spectrum using
a finite sample \( \{X_t; t = 1, 2, \ldots, T\} \) is to form estimates \( \hat{R}(v) \)
of the covariances and, following (1.2), use

\[
\hat{f}(\omega) = \frac{1}{2\pi} \sum_{v=-T+1}^{T-1} \hat{R}(v) \cos v\omega.
\]

(1.6)

The limits of summation here have been written \( \pm(T-1) \) because
products \( X_iX_j \) can be formed only for \( |i-j| \leq T-1 \), and therefore
meaningful estimates of \( R(v) \) can be formed only for \( |v| \leq T-1 \).
Unbiased estimates for \( R(v) \) are

\[
\hat{R}_T(v) = \frac{1}{T-|v|} \sum_{t=1}^{T-|v|} X_tX_{t+|v|}, \quad |v| < T
\]

(1.7)

but it is more convenient computationally to replace the divisor
\( T-|v| \) by \( T \); this is also justifiable through mean square error
and other considerations (Schaerf (1964)). The resulting
estimate is the periodogram:

\[
f_T(\omega) = \frac{1}{2\pi} \sum_{v=-T+1}^{T-1} \hat{R}_T(v) \cos v\omega
\]

(1.8)
where $R_T(v)$ is the sample covariance function:

$$
R_T(v) = \frac{1}{T} \sum_{t=1}^{T-|v|} X_t X_{t+|v|}, \quad |v| < T.
$$

(1.9)

However $f_T(\omega)$ is unsatisfactory as an estimate because of excessive variance; in fact it is not consistent, i.e. its variance does not tend to zero as $T$ tends to infinity.

The almost universally adopted form of spectrum estimate is

$$
f_T^*(\omega) = \frac{1}{2\pi} \sum_{v=-T+1}^{T-1} k_T^*(v) R_T(v) \cos v\omega, \quad (-\pi \leq \omega \leq \pi)
$$

(1.10)

where $k_T^*(.)$ is a weighting function called the covariance averaging kernel (Parzen (1961)) or the lag window (Blackman and Tukey (1958)). The lag window is usually formed from

$$
k_T^*(v) = k^*(\frac{v}{M_T})
$$

(1.11)

where $k^*(.)$ is a lag window generator, which is a bounded function having the properties:

(i) $k^*(0) = 1$
(ii) $k^*(\theta) = k^*(-\theta)$
(iii) $k^*(\theta) = 0$, \quad $|\theta| > 1$.
and $M_T$ is a function of $T$ satisfying

$$M_T \to \infty, \quad \frac{M_T}{T} \to 0 \quad \text{as} \quad T \to \infty. \quad (1.13)$$

It is well-known (Parzen (1957), Grenander and Rosenblatt (1957)) that, under these conditions, the variance of the estimate $f_T^*(\cdot)$ is given by

$$\lim_{T \to \infty} \frac{T}{M_T} \text{var}(f_T^*(\omega)) = 2(1+\delta_0,\omega, +\delta_{\pi},\omega)\hat{f}(\omega)^2 \int_0^1 k^*(\Theta)^2 d\Theta \quad (1.14)$$

where $\delta_{a,b}$ is the Kronecker delta function, equal to 1 if $a = b$ and 0 otherwise.

In deriving (1.14), terms of the order of $M_T/T$ are neglected, and so the use of (1.14) as an approximation to the finite case, where $M_T/T$ cannot be zero and on occasions may be as big as, say, .3 or even .5, becomes rather inappropriate. It is shown in Neave (1968a, 1968b) that if terms of the order of $M_T/T$ are retained, then the above result is amended to

$$\lim_{T \to \infty} \text{var}(f_T^*(\omega)) = 2\gamma(1+\delta_0,\omega, +\delta_{\pi},\omega)\hat{f}(\omega)^2 \int_0^1 k^*(\Theta)^2 (1-\gamma\Theta) d\Theta, \quad T \to \infty \quad (1.15)$$

where, instead of $M_T$ being $o(T)$, we assume that

$$\frac{M_T}{T} \to \gamma, \quad (1.16)$$
γ being some constant between 0 and 1. Such an assumption results in the spectrum estimate being inconsistent, but nevertheless various numerical examples in Neave (1968a) show that (1.15) provides a much better approximation to finite cases than does (1.14).

It will be convenient in most of the subsequent work to use the unbiased sample covariance function \( \tilde{R}_T(.) \) in our spectrum estimates. Accordingly we define

\[
\tilde{f}_T(\omega) = \frac{1}{2\pi} \sum_{v=-T+1}^{T-1} k_T(v) \tilde{R}_T(v) \cos v\omega
\]  

(1.17)

which differs in form from (1.10) only in the use of \( \tilde{R}_T(.) \) instead of \( R_T(.) \). It is apparent that

\[
f^*_T(.) = \tilde{f}_T(.)
\]

(1.18)

if and only if the weighting functions \( k_T^*(.) \) and \( k_T(.) \) satisfy the relation

\[
k_T(v) \equiv (1-\gamma |v|)k_T^*(v)
\]

(1.19)

where \( \gamma = M_T/T \). If \( k_T^*(.) \) is defined by (1.11) and (1.12), \( k_T(.) \) may be defined similarly by using the corresponding modified lag window generator:

\[
k(\theta) = (1-\gamma |\theta|)k^*(\theta).
\]

(1.20)
Assuming consistency of the estimate, which requires that $\gamma \to 0$ as $T \to \infty$, the two generators $k^*(.)$ and $k(.)$ are asymptotically identical.

§2. Summary of this paper

Since the asymptotic expression for the variance of spectrum estimates was derived, it has often been used (in the form (1.14), not (1.15)) as an approximation to the variance of estimates using finite samples. With the exception of Scheinok (1960), little attempt seems to have been made to justify this or to investigate the nature of the convergence to the asymptotic form. In this paper, an exact expression for the variance is derived on the additional assumption that $\{X_t\}$ is a normal process. This expression is then used to study estimates of four different spectra, using fairly small sample sizes and three truncation points in each case. Graphs of the standard deviation and also of the bias and root mean square error are given for the complete survey of one of the spectra investigated; there are also tables showing the behaviour of these quantities for all the four spectra.

§3. The exact expression for the variance

We shall be using the expression (1.17) for the spectrum estimate. Since the form (1.10) is more frequently used in practice, we shall assume that (1.18), and therefore (1.20), holds; the
numerical examples given in §6,7 are computed in this way and therefore apply to the usual practical case.

We therefore have

\[
\tilde{f}_T^*(\omega) = \tilde{f}_T(\omega) = \frac{1}{2\pi} \sum_{|v| < M_T} k_T(v) \tilde{R}_T(v) \cos v\omega
\]  

(3.1)

where

\[
\tilde{R}_T(v) = \frac{1}{T-|v|} \sum_{t=1}^{T-|v|} x_t x_{t+|v|}, \quad |v| < T
\]

and therefore

\[
E[\tilde{R}_T(v)] = R(v), \quad |v| < T.
\]  

(3.2)

It is preferable to write (3.1) and subsequent expressions in terms of non-negative summation variables. Thus

\[
\tilde{f}_T(\omega) = \frac{1}{\pi} \left\{ k_T(0) \frac{\tilde{R}_T(0)}{2} + \sum_{v=1}^{M_T} k_T(v) \tilde{R}_T(v) \cos v\omega \right\},
\]  

(3.3)

\[
E[\tilde{f}_T(\omega)] = \frac{1}{\pi} \left\{ k_T(0) \frac{R(0)}{2} + \sum_{v=1}^{M_T} k_T(v) R(v) \cos v\omega \right\},
\]  

(3.4)
\[
E[\tilde{f}_T(\omega)^2] = \frac{1}{\pi^2} \left\{ k_T(0)^2 \frac{E[\tilde{R}_T(0)^2]}{4} \right\} + \sum_{v=1}^{M_T} k_T(0)k_T(v)E[\tilde{R}_T(0)\tilde{R}_T(v)] \cos v\omega
\]
\[
+ \sum_{u,v=1}^{M_T} k_T(v)k_T(u)E[\tilde{R}_T(v)\tilde{R}_T(u)] \cos v\omega \cos u\omega \right\}. \quad (3.5)
\]

Equation (3.5) may be rearranged to produce
\[
E[\tilde{f}_T(\omega)^2] = \frac{1}{\pi^2} \left\{ k_T(0)^2 \frac{E[\tilde{R}_T(0)^2]}{4} \right\} + \sum_{v=1}^{M_T} k_T(0)k_T(v)E[\tilde{R}_T(0)\tilde{R}_T(v)] \cos v\omega
\]
\[
+ \sum_{v=1}^{M_T} k_T(v)^2 E[\tilde{R}_T(v)^2]
\]
\[
+ 2 \sum_{1 \leq v < u \leq M_T} k_T(v)k_T(u) E[\tilde{R}_T(v)\tilde{R}_T(u)] \cos v\omega \cos u\omega \right\}. \quad (3.6)
\]

Writing
\[
\text{var} (\tilde{f}_T(\omega)) = E[\tilde{f}_T(\omega)^2] - \{E[\tilde{f}_T(\omega)]\}^2,
\]
we therefore have, from (3.4) and (3.6),
\[ \text{var} \left( \tilde{f}_T(\omega) \right) = \]

\[ \frac{1}{\pi^2} \left( k_T(0)^2 \frac{\text{var} \left( \tilde{R}_T(0) \right)}{4} \right) + \sum_{v=1}^{M_T} k_T(0)k_T(v) \text{cov} \left( \tilde{R}_T(0), \tilde{R}_T(v) \right) \cos vw \]

\[ + \sum_{v=1}^{M_T} k_T(v)^2 \text{var} \left( \tilde{R}_T(v) \right) \]

\[ + 2 \sum_{1 \leq v < u \leq M_T} k_T(v)k_T(u) \text{cov} \left( \tilde{R}_T(v), \tilde{R}_T(u) \right) \cos vw \cos uw. \]  

(3.7)

It is therefore necessary to calculate expressions for the covariances

\[ \text{cov} \left( \tilde{R}_T(v), \tilde{R}_T(u) \right), \quad 0 \leq v \leq u \leq M_T. \]

Now,

\[ E[\tilde{R}_T(v)\tilde{R}_T(u)] = E\left[ \frac{1}{(T-v)(T-u)} \sum_{t=1}^{T-v} \sum_{s=1}^{T-u} X_tX_{t+v}X_sX_{s+u} \right] \]

\[ = \frac{1}{(T-v)(T-u)} \sum_{t=1}^{T-v} \sum_{s=1}^{T-u} E[X_tX_{t+v}X_sX_{s+u}]. \]  

(3.8)

It is proved in Iserlis (1918) that for normally distributed variates A, B, C, D,

Using this formula, we obtain

\[ E[X_t X_{t+v} X_s X_{s+u}] \]

\[ = R(v)R(u) + R(t-s)R(t-s+v-u) + R(t-s-u)R(t-s+v). \] (3.10)

It then follows from (3.8) and (3.10) that

\[ \text{cov} \left( \tilde{R}_T(v), \tilde{R}_T(u) \right) \]
\[ = E[\tilde{R}_T(v)\tilde{R}_T(u)] - R(v)R(u) \]
\[ = \frac{1}{(T-v)(T-u)} \sum_{t=1}^{T-v} \sum_{s=1}^{T-u} \{ R(t-s)R(t-s+v-u) + R(t-s-u)R(t-s+v) \}. \] (3.11)

It is clear that the summand in (3.11) is not dependent on \( t \) and \( s \) explicitly, but only on their difference \( x = t-s \), and the evaluation of (3.7) is carried out by effecting this change of variable.

Let us therefore consider a general expression of the form

\[ \sum_{t=1}^{a} \sum_{s=1}^{b} h(t-s). \] (3.12)
This can clearly be written as

\[
\sum_{x=-b}^{a} U(x,a,b)h(x).
\]  \hspace{1cm} (3.13)

The definition of \(U(.)\) depends on which is the larger of \(a\) and \(b\);
if \(b > a\) then

\[
U(x,a,b) = b-|x|, \quad x \leq -(b-a)
\]
\[= a, \quad -(b-a) \leq x \leq 0\]
\[= a-x, \quad 0 \leq x \leq a, \] \hspace{1cm} (3.14)

and if \(a > b\) then

\[
U(x,a,b) = b-|x|, \quad -b \leq x \leq 0
\]
\[= b, \quad 0 \leq x \leq a-b\]
\[= a-x, \quad a-b \leq x \leq a. \] \hspace{1cm} (3.15)

Assuming the latter case, \(a \geq b\), we therefore have from (3.12),
(3.13) and (3.15) that

\[
\frac{1}{ab} \sum_{t=1}^{a} \sum_{s=1}^{b} h(t-s) = \frac{1}{a} \sum_{x=0}^{a-b} h(x) + \frac{1}{ab} \sum_{x=a-b+1}^{a} (a-x)\{h(x)+h(a-b-x)\}. \]
\hspace{1cm} (3.16)
Now setting

\[ h(x) = R(x)R(x+v-u) + R(x-u)R(x+v), \]

\[ a = T-v, \quad b = T-u, \quad u \geq v, \]

(3.16) becomes

\[
\text{cov} \left( \tilde{R}_T(v), \tilde{R}_T(u) \right) \\
= \frac{1}{T-v} \sum_{x=0}^{u-v} \{ R(x)R(x+v-u) + R(x-u)R(x+v) \} \\
+ \frac{1}{(T-v)(T-u)} \sum_{x=u-v+1}^{T-v} (T-v-x) \{ R(x)R(x+v-u) + R(x-u)R(x+v) \\
+ R(u-v-x)R(-x) + R(-x-v)R(u-x) \}.
\]

Since \( R(.) \) is an even function, this is

\[
\frac{1}{T-v} \sum_{x=0}^{u-v} \{ R(x)R(x+v-u) + R(u-x)R(x+v) \} \\
+ \frac{2}{(T-v)(T-u)} \sum_{x=u-v+1}^{T-v} (T-v-x) \{ R(x)R(x+v-u) + R(x-u)R(x+v) \}.
\]
This can be written with non-negative arguments throughout as

\[
\frac{1}{T-v} \sum_{x=0}^{u-v} \{ R(x)R(u-v-x) + R(u-x)R(x+v) \}
\]

\[
+ \frac{2}{(T-v)(T-u)} \left\{ \sum_{x=u-v+1}^{T-v} (T-v-x)R(x)R(x+v-u) + \sum_{x=u-v+1}^{u} (T-v-x)R(u-x)R(x+v) \right. \\
\left. + \sum_{x=u+1}^{T-v} (T-v-x)R(x-u)R(x+v) \right\}
\]

(3.17)

if \( u < T/2 \), or by replacing \( x \) by \( x-v+u \) in the last three sums,

\[
\frac{1}{T-v} \sum_{x=0}^{u-v} \{ R(x)R(u-v-x) + R(u-x)R(x+v) \}
\]

\[
+ \frac{2}{(T-v)(T-u)} \left\{ \sum_{x=1}^{T-u} (T-u-x)R(x)R(x+u-v) + \sum_{x=1}^{V} (T-u-x)R(u+x)R(v-x) \right. \\
\left. + \sum_{x=v+1}^{T-u} (T-u-x)R(x+u)R(x-v) \right\}. \]

(3.18)

This expression for \( \text{cov}(\tilde{R}_T(v), \tilde{R}_T(u)) \) is then valid for

\[ 0 \leq v \leq u < T/2 \]
and the last inequality here is clearly satisfied if \( \frac{M_T}{T} < \frac{1}{2} \).

It follows in particular from (3.18) that

\[
\begin{align*}
(a) \quad \text{var} \, (\widetilde{R}_T(0)) &= E[\widetilde{R}_T(0)^2] - R(0)^2 \\
&= \frac{2}{T^2} \{ TR(0)^2 + 2 \sum_{x=1}^{T} (T-x)R(x)^2 \}, \quad (3.19) \\
(b) \quad \text{cov} \, (\widetilde{R}_T(0), \widetilde{R}_T(v)) &= E[\widetilde{R}_T(0)\widetilde{R}_T(v)] - R(0)R(v) \\
&= \frac{2}{T} \sum_{x=0}^{V} R(x)R(v-x) + \frac{4}{T(T-v)} \sum_{x=1}^{T-v} (T-v-x)R(x)R(x+v), \quad 0 < v < T/2 \quad (3.20)
\end{align*}
\]

and

\[
\begin{align*}
(c) \quad \text{var} \, (\widetilde{R}_T(v)) &= E[\widetilde{R}_T(v)^2] - R(v)^2 \\
&= \frac{1}{T-v} \{ R(0)^2 + R(v)^2 \} + \frac{2}{(T-v)^2} \sum_{x=1}^{T-v} (T-v-x)R(x)^2 \\
&+ \sum_{x=1}^{V} (T-v-x)R(v+x)R(v-x) + \sum_{x=v+1}^{T-v} (T-v-x)R(x+v)R(x-v), \quad 0 < v < T/2. \quad (3.21)
\end{align*}
\]
From (3.7) and (3.18) to (3.21) we may now state the final result:

**Theorem**

If \( \{X_t\} \) is a normal process and \( M_T/T < \frac{1}{2} \), then the variance of \( \tilde{f}_T(\omega) \), as defined by (1.17), is given exactly by

\[
\frac{k_T(0)^2}{2\pi^2 T^2} \left\{ TR(0)^2 + 2 \sum_{x=1}^{T} (T-x)R(x)^2 \right\} + \frac{2k_T(0)}{\pi^2} \sum_{v=1}^{M_T} k_T(v) \cos \omega v \left[ \sum_{x=0}^{v} R(x)R(v-x) \right]
\]

\[
+ \frac{2}{T-v} \sum_{x=1}^{T-v} (T-v-x)R(x)R(x+v) \]

\[
+ \frac{1}{\pi^2} \sum_{v=1}^{M_T} k_T(v)^2 \cos^2 \omega v \left\{ \frac{1}{T-v}(R(0)^2+R(v)^2) + \frac{2}{(T-v)^2} \sum_{x=1}^{T-v} (T-v-x)R(x)^2 \right\}
\]

\[
+ \sum_{x=1}^{v} (T-v-x)R(x+v)R(v-x) + \sum_{x=v+1}^{T-v} (T-v-x)R(x+v)R(x-v) \]

\[
+ \frac{2}{\pi^2} \sum_{0<v<u<\frac{M_T}{T}} k_T(v)k_T(u) \cos \omega u \cos \omega v \left\{ \frac{1}{T-v} \sum_{x=0}^{u-v} \left[ R(x)R(u-v-x)+R(u-x)R(x+v) \right] \right\}
\]
\[
\begin{align*}
\frac{2}{(T-u)(T-v)} & \sum_{x=1}^{T-u} (T-u-x)R(x)R(x+u-v) \\
+ \sum_{x=1}^{T-u} (T-u-x)R(x+u)R(v-x) & + \sum_{x=v+1}^{T-u} (T-u-x)R(x+u)R(x-v)\}. \quad (3.22)
\end{align*}
\]

Due to the complicated structure of this expression, the only way to study it in general is by using a computer, but one case where a machine is unnecessary is when \(\{X_t\}\) is a \text{WHITE NOISE}, i.e. completely uncorrelated, process, satisfying

\[
\text{cov} (X_t, X_s) = 0 \quad , \quad \text{all } t \neq s, \quad (3.23)
\]

that is,

\[
R(v) = 0 \quad , \quad v \neq 0 .
\]

Therefore in this case

\[
\text{var} (f_T(\omega)) = \frac{1}{\pi^2} R(0)^2 \left( \frac{1}{2T} + \sum_{v=1}^{M_T} \frac{1}{T-v} k^2 \left( \frac{v}{M_T} \right)^2 \cos^2 v\omega \right). \quad (3.24)
\]

If we now wish to calculate the variance of the usual form of estimate \(f_T^*(\omega)\), using the lag window generator \(k^*(\cdot)\), we may make use of the fact that
\[ f^*_T(\omega) \equiv f_T(\omega) \]

if the modified lag window generator (1.20):

\[ k(\theta) = (1 - |\theta|)k^*(\theta) \]

is used in the definition of \( f_T(\cdot) \). It therefore follows from (3.24) that

\[
\text{var} (f^*_T(\omega)) = \frac{1}{\pi^2} R(0)^2 \left\{ \frac{1}{2T} + \sum_{v=1}^{M_T} \frac{1}{T-v} k\\left( \frac{v}{M_T} \right)^2 (1 - \frac{v}{T})^2 \cos^2 v\omega \right\} \\
= \frac{1}{\pi^2} R(0)^2 \left\{ \frac{1}{2T} + \frac{1}{T} \sum_{v=1}^{M_T} k\\left( \frac{v}{M_T} \right)^2 (1 - \frac{v}{T}) \cos^2 v\omega \right\}.
\]

Then as \( T \to \infty \) with \( \frac{M_T}{T} \to \gamma \),

\[
\text{var} (f^*_T(\omega)) \sim \frac{1}{4\pi^2} R(0)^2 \frac{M_T}{T} \int_{-1}^{1} k(\theta)^2 (1 - \gamma \theta) d\theta
\]

for \( \omega \neq 0 \) or \( \pi \), and double this at these exceptional points. [Use has been made here of the fact that the 'average value' of \( \cos^2 v\omega \), i.e.

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{v=0}^{N-1} \cos^2 v\omega
\]
is given by

\[ \frac{1}{2} (1 + \delta_\omega, 0 + \delta_\omega, \pi) \] 

Since in the white noise case

\[ f(\omega) = \frac{1}{2\pi} R(0), \]

the form of (1.15) is obtained.

§4. The Bias and Mean Square Error

From (1.17) it is immediate that

\[ E[\tilde{f}_T(\omega)] = \frac{1}{2\pi} \sum_{|v| \leq M_T} k_T(v) R(v) \cos v\omega \]

so that the bias, \( b_T(\omega) \), is given by

\[ -\frac{1}{2\pi} \sum_{|v| \leq M_T} (1 - k_T(v)) R(v) \cos v\omega + \sum_{|v| > M_T} R(v) \cos v\omega. \]

The mean square error of \( \tilde{f}_T(\omega) \) is defined as

\[ E[(\tilde{f}_T(\omega) - f(\omega))^2] = \text{var} (f_T(\omega)) + b_T(\omega)^2, \]
and the "root mean square (RMS) error" is the square root of this quantity.

From (4.1) it can be seen that, like \( \tilde{f}_T(\omega) \) itself, the expectation of \( \tilde{f}_T(\omega) \) is a trigonometric polynomial of degree \( M_T \). This is one of the facts that must be borne in mind when choosing a spectrum estimate, particularly with a small truncation point. Also, although

\[
\frac{1}{2\pi} \sum_{|v| < M_T} R(v) \cos v\omega
\]

(4.3)
is the optimum approximating polynomial of degree \( M_T \) from the criterion of least squares (see e.g. Sneddon (1961), Theorem 8), the weightings \( k_T(v) \) in (4.1) leave us with only a smoothed version of this optimum. Consequently when considering the amount of precision required in the estimate, one must allow \( M_T \) to exceed, by some fairly sizable factor, the degree of polynomial for which (4.3) would be satisfactory. Unfortunately from (1.14) or (1.15), it is apparent that the variance increases with \( M_T \), and one of the basic problems in spectrum estimation is the forecast of which value of \( M_T \) results in the best balance between bias and variance, given some particular lag window generator. Some authors have advocated the minimisation of the mean square error as the suitable criterion, but this is not universally accepted: in any case any results obtained depend on the spectrum being estimated. Such considerations are nevertheless interesting, and a contribution to this theory is presented in Neave (1968c). More detailed
discussion of the problems outlined above follows later in this paper and in Neave (1968d).

One of the results proved in Parzen (1957) was effectively that in fairly general conditions, the bias approximately satisfies

\[ b_T(\omega) = \frac{k_r}{M_T^2} f^{(r)}(\omega) \]

where \( r \) is the characteristic exponent of the lag window generator (which takes the value 2 for most generators in use), \( k_r \), the characteristic coefficient, is the value of

\[ \lim_{z \to 0} \frac{1-k(z)}{|z|^r} \]  \hspace{1cm} (4.4)

and \( f^{(r)}(.) \) is the \( r \)th derivative of the spectral density, assumed continuous. Putting \( r = 2 \),

\[ b_T(\omega) = \frac{k_2}{M_T^2} f''(\omega). \]  \hspace{1cm} (4.5)

Thus asymptotically the bias is proportional to the second derivative of the spectral density. With \( f''(.) \) being continuous, it is easily shown that (4.5) is approximately

\[ \frac{k_2}{M_T^2} \cdot \frac{1}{\delta^2} \{ f(\omega+\delta) - 2f(\omega) + f(\omega-\delta) \} \]  \hspace{1cm} (4.6)
\begin{equation}
R(0) = 2\pi c
\end{equation}

\begin{equation}
R(v) = 0, \quad v \neq 0.
\end{equation} \tag{5.1}

This is the 'white noise' process, referred to in §3.

B. Autoregressive (first order)

This form has the covariance function:

\begin{equation}
R(v) = R(0)\alpha |v|, \quad 0 < \alpha < 1.
\end{equation} \tag{5.2}

Its importance lies in its representation by the generating process:

\begin{equation}
X_t = \alpha X_{t-1} + \varepsilon_t \tag{5.3}
\end{equation}

where \{\varepsilon_t\} is a white noise process. Then
\[ f(\omega) = \frac{1}{2\pi} R(0) + \frac{1}{\pi} \sum_{\nu=1}^{\infty} R(\nu) \cos \nu \omega \]

\[ = \frac{R(0)}{2\pi} \frac{1-\alpha^2}{1+\alpha^2 - 2\alpha \cos \omega} . \]

In the numerical examples studied, \( R(0) \) will be taken as 1.

C. Half peak cosine wave

This is defined by the spectrum

\[ f(\omega) = \frac{1}{2} A (1 + \cos \frac{\pi \omega}{a}) , \quad 0 \leq \omega \leq a \]

\[ = 0 \quad , \quad a \leq \omega \leq \pi \]

\[ 0 < a < \pi \quad , \quad A > 0 \]

and is thus a half-period cosine peak with maximum height \( A \) at \( \omega = 0 \). We have

\[ R(\nu) = 2 \frac{1}{2} A \int_0^a (1 + \cos \frac{\pi \omega}{a}) \cos \nu \omega \, d\omega \]

\[ = A \int_0^a \cos \nu \omega + \frac{1}{2} \{ \cos(\nu + \frac{\pi}{a}) \omega + \cos(\nu - \frac{\pi}{a}) \omega \} \, d\omega \]

\[ = A \left[ \frac{\sin \nu \omega}{\nu} + \frac{1}{2} \left( \frac{\sin(\nu + \frac{\pi}{a}) \omega}{\nu + \frac{\pi}{a}} + \frac{\sin(\nu - \frac{\pi}{a}) \omega}{\nu - \frac{\pi}{a}} \right) \right]_0^a \]
\[ \int_{\omega_o - a}^{\omega_o + a} \cos v\omega \, d\omega = \frac{1}{v} \left[ \sin v\omega \right]_{\omega_o - a}^{\omega_o + a}, \quad v \neq 0 \]

\[ = \frac{1}{v} \{ \sin v(\omega_o + a) - \sin v(\omega_o - a) \} \]

\[ = 2 \frac{1}{v} \cos v\omega_o \sin va. \]

Also

\[ \int_{\omega_o - a}^{\omega_o + a} \cos v\omega \cos \frac{\pi}{a} (\omega - \omega_o) \, d\omega \]

\[ = \frac{1}{2} \int_{\omega_o - a}^{\omega_o + a} \cos((v+\pi/a)\omega - \frac{\pi\omega_o}{a}) + \cos((v-\pi/a)\omega + \frac{\pi\omega_o}{a}) \, d\omega \]

\[ = \frac{1}{2} \left[ \frac{\sin((v+\pi/a)\omega - \frac{\pi\omega_o}{a})}{v+\pi/a} + \frac{\sin((v-\pi/a)\omega + \frac{\pi\omega_o}{a})}{v-\pi/a} \right]_{\omega_o - a}^{\omega_o + a} \]
if \(|v| \neq \pi/a\),

\[
= \frac{1}{2} \left\{ \sin((v+\frac{\pi}{a})(\omega_o+a) - \frac{\pi \omega_o}{a}) - \sin((v+\frac{\pi}{a})(\omega_o-a) - \frac{\pi \omega_o}{a}) \right\}_{v+\pi/a} \\
+ \frac{\sin((v-\frac{\pi}{a})(\omega_o+a) + \frac{\pi \omega_o}{a}) - \sin((v-\frac{\pi}{a})(\omega_o-a) + \frac{\pi \omega_o}{a}) \right\}_{v-\pi/a}
\]

\[
= \frac{1}{2} \left\{ \sin(v+\omega_o+a) - \sin(-v+\omega_o-a) \right\}_{v+\pi/a} \\
+ \frac{\sin(-v+\omega_o+a) - \sin(v+\omega_o-a) \right\}_{v-\pi/a}
\]

\[
= \frac{1}{2} \left\{ \sin(v_0-a) - \sin(v_0+a) \right\} \left\{ \frac{1}{v+\pi/a} - \frac{1}{v-\pi/a} \right\}
\]

\[
= \frac{2A\pi^2}{v(\pi^2 - v^2a^2)} \cos(v\omega_o) \sin va , \quad |v| \neq 0, \pi/a.
\]

At the exceptional values,

\[R(0) = 2Aa,\]

\[R(\pi/a) = Aa \cos(\frac{\pi}{a}\omega_o), \quad \text{if } a|\pi .\]

For \(f(\omega)\) to exist, it follows from (1.2) that \(R(v) \to 0\) as \(v \to \infty\). The manner in which \(R(v) \to 0\) for these spectrum elements is
of interest. In the constant case there is no problem is

\[ R(v) = 0, \quad v \neq 0. \]

For the autoregressive form,

\[ R(v) = R(0) \alpha^{-|v|} \]

and it is well-known that this tends to zero quicker than any negative power of \( v \), i.e.

\[ v^n \alpha^n \rightarrow 0 \]

for any \( n \). In the half peak and whole peak cases, \( R(v) \) damps down quicker than

\[ \frac{1}{(v - \pi/a)^2} \]

and is eventually of the same order as \( v^{-3} \). This latter observation is then true of all spectra compounded from these four spectrum elements.

§6. A computer investigation of the variance and bias of spectrum estimates.

The spectra studied in this investigation are illustrated
in Figure 1(a-d). They are formed by superposition of the following spectrum elements:

**Figure 1(a)**

(a) Autoregressive, \( a = 0.75 \).
(b) Whole peak, \( A = 1.0, a = 11\pi/90, \omega_o = 8\pi/45 \).
(c) Whole peak, \( A = 0.8, a = 11\pi/90, \omega_o = 19\pi/45 \).
(d) Whole peak, \( a = 0.6, a = 11\pi/90, \omega_o = 2\pi/3 \).

**Figure 1(b)**

(a) Constant, \( c = 1 \).
(b) Half peak, \( A = 5, a = \pi/2 \).
(c) Whole peak, \( A = 5, a = \pi/6, \omega_o = \pi/6 \).
(d) Whole peak, \( a = 5, a = \pi/6, \omega_o = \pi/2 \).

**Figure 1(c)**

Autoregressive, \( a = 0.8 \).

**Figure 1(d)**

(a) Constant, \( c = 1 \).
(b) Half peak, \( A = 5, a = \pi \).
(c) Whole peak, \( A = 2, a = \pi/6, \omega_o = \pi/6 \).
(d) Whole peak, \( A = 2, a = \pi/18, \omega_o = \pi/4 \).
(e) Whole peak, \( A = 2, a = \pi/18, \omega_o = 3\pi/4 \).
(f) Whole peak, \( A = 2, a = \pi/6, \omega_o = 5\pi/6 \).
An ALGOL program was written to calculate

(1) the variance

(2) the expected value and the bias

(3) the mean square error

for each of these four spectra, given sample lengths $T$:

$30, 60, 90, 120,$

at the points

$$\omega = \frac{\pi j}{180}, \quad j = 0, 1, 2, \ldots, 180$$

and using the lag window generators

(1) Parzen: \[ k_p(\theta) = 1 - 6\theta^2 + 6\theta^3, \quad 0 \leq \theta \leq \frac{1}{2} \]

$$= 2(1-\theta)^3, \quad \frac{1}{2} \leq \theta \leq 1 \quad (6.1)$$

(2) Tukey: \[ k_T(\theta) = \frac{1}{2}(1+\cos \pi \theta), \quad 0 \leq \theta \leq 1 \quad (6.2)\]

with ratios of truncation point to sample size:

$$\frac{M_T}{T} = .1, .2, \frac{1}{3}$$

The sample sizes used were smaller than the generally accepted order of size for satisfactory estimates. This was because of the large amount of computation involved in the calculation of the
covariances (3.18) and of the variances (3.22). For each spectrum, the program took about 13 minutes on a Burroughs B5500 computer.

The two lag windows (6.1, 6.2) are those most commonly employed in spectrum estimation at the present time. The Tukey generator is the best of those produced in the early stages of spectrum estimation theory (Tukey (1949)). In recent years, many researchers have however adopted the Parzen generator, mainly because, unlike Tukey's, it produces estimates which are necessarily non-negative. Parzen's generator originated in a paper presented to the I.M.S. meeting in 1957. It is adapted from a function known in the theory of approximation as the Jackson de la Vallée Poussin kernel - see Achieser (1956), p. 119.

To investigate the corresponding asymptotic expressions (1.15) for the variance, the following definite integrals are needed:

\[
2 \int_{0}^{1} k_p(\theta)^2 \, d\theta = 0.5393
\]

\[
2 \int_{0}^{1} k_p(\theta)^2 (1-\theta) \, d\theta = 0.5301
\]

\[
2 \int_{0}^{1} k_p(\theta)^2 (1-2\theta) \, d\theta = 0.5209
\]

\[
2 \int_{0}^{1} k_p(\theta)^2 (1-\frac{1}{3}\theta) \, d\theta = 0.5086
\]
\[ \frac{1}{2} \int_{0}^{1} k_T(\theta)^2 d\theta = 0.7500 \]
\[ 2 \int_{0}^{1} k_T(\theta)^2 (1 - \theta) d\theta = 0.7328 \]
\[ 2 \int_{0}^{1} k_T(\theta)^2 (1 - 2\theta) d\theta = 0.7155 \]
\[ 2 \int_{0}^{1} k_T(\theta)^2 (1 - \frac{1}{3}\theta) d\theta = 0.6925 . \]

As regards the bias, the characteristic coefficients (4.4) of the Parzen and Tukey generators are 6 and \( \frac{1}{4} \pi^2 = 2.467 \) respectively. Hence asymptotically the bias of the Parzen estimate is nearly two and a half times greater than that of the Tukey estimate, while the variance of the latter exceeds Parzen's by a factor of about 40%. Not surprisingly though, the expressions for the bias (4.5, 4.6) are infrequently used as approximations to the finite case, since for example their only dependence on the lag window generator is its behaviour on approaching \( \theta = 0 \). Their use here is in any case not entirely relevant as the trigonometrical spectrum elements do not have continuous second derivatives.

For economy of space, the graphs for only one of the four spectra are presented here - Figure 2(a-x). The spectrum chosen was that with the most irregular shape, Figure 1(a). The results from all four are however summarised in Table 1(a-d), where the averages of the square of the bias, the variance and the mean square error for all the cases considered are given. Throughout,
the results have been scaled as percentages relative to

\[
\frac{1}{180} \left\{ \frac{1}{2} f(0)^2 + \sum_{j=1}^{179} f\left(\frac{\pi j}{180}\right) + \frac{1}{2} f(\pi)^2 \right\}
\]

(6.3)
as 100\%, this being the best estimate from the frequencies considered of

\[
\frac{1}{\pi} \int_{0}^{\pi} f(\omega)^2 d\omega .
\]

This base was chosen because, according to (1.14) or (1.15), the average variance of the spectrum estimate is asymptotically proportional to it.
| SAMPLE SIZE | $\frac{M_T}{T}$ = | PARZEN | | TUKEY |
|-------------|------------------|--------|--------|
|             | .1               | .2     | $\frac{1}{3}$ | .1 | .2 | $\frac{1}{3}$ |
| 30          | Square Bias      | 22.91  | 17.25  | 15.14 | 19.41 | 16.09  | 12.99  |
|             | Variance         | 8.66   | 14.04  | 20.19 | 10.89 | 17.76  | 25.98  |
|             | M.S.E.           | 31.57  | 31.30  | 35.33 | 30.31 | 33.85  | 38.97  |
| 60          | Square Bias      | 17.09  | 13.30  | 6.46  | 15.96 | 9.60   | 3.21   |
|             | Variance         | 7.29   | 12.24  | 18.78 | 9.26  | 16.12  | 24.89  |
|             | M.S.E.           | 24.38  | 25.54  | 25.24 | 25.22 | 25.72  | 28.10  |
| 90          | Square Bias      | 15.48  | 7.53   | 2.40  | 13.80 | 3.70   | 1.02   |
|             | Variance         | 6.62   | 11.74  | 18.32 | 8.60  | 15.68  | 24.42  |
|             | M.S.E.           | 22.10  | 19.27  | 20.72 | 22.40 | 19.39  | 25.43  |
| 120         | Square Bias      | 13.12  | 3.90   | 1.06  | 9.23  | 1.60   | 0.43   |
|             | Variance         | 6.29   | 11.49  | 18.04 | 8.34  | 15.41  | 24.13  |
|             | M.S.E.           | 19.40  | 15.39  | 19.11 | 17.57 | 17.01  | 24.56  |

**TABLE 1(a)**
<table>
<thead>
<tr>
<th>SAMPLE SIZE</th>
<th>$\frac{M_T}{T} = $</th>
<th>PARZEN</th>
<th>TUKEY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>.1</td>
<td>.2</td>
</tr>
<tr>
<td>30</td>
<td>Square Bias</td>
<td>15.58</td>
<td>10.15</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>8.89</td>
<td>14.40</td>
</tr>
<tr>
<td></td>
<td>M.S.E.</td>
<td>24.47</td>
<td>24.55</td>
</tr>
<tr>
<td>60</td>
<td>Square Bias</td>
<td>10.00</td>
<td>5.13</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>7.38</td>
<td>12.09</td>
</tr>
<tr>
<td></td>
<td>M.S.E.</td>
<td>17.39</td>
<td>17.22</td>
</tr>
<tr>
<td>90</td>
<td>Square Bias</td>
<td>7.65</td>
<td>1.96</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>6.59</td>
<td>11.38</td>
</tr>
<tr>
<td></td>
<td>M.S.E.</td>
<td>14.24</td>
<td>13.34</td>
</tr>
<tr>
<td>120</td>
<td>Square Bias</td>
<td>4.93</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>6.17</td>
<td>11.08</td>
</tr>
<tr>
<td></td>
<td>M.S.E.</td>
<td>11.09</td>
<td>11.94</td>
</tr>
</tbody>
</table>

**TABLE 1(b)**
<table>
<thead>
<tr>
<th>SAMPLE SIZE</th>
<th>$\frac{M_T}{T}$</th>
<th>PARZEN</th>
<th>T UKEY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>.1</td>
<td>.2</td>
</tr>
<tr>
<td>30</td>
<td>Square Bias</td>
<td>53.57</td>
<td>31.34</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>9.76</td>
<td>18.35</td>
</tr>
<tr>
<td></td>
<td>M.S.E.</td>
<td>63.34</td>
<td>49.69</td>
</tr>
<tr>
<td>60</td>
<td>Square Bias</td>
<td>30.60</td>
<td>11.90</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>9.72</td>
<td>17.63</td>
</tr>
<tr>
<td></td>
<td>M.S.E.</td>
<td>40.32</td>
<td>29.53</td>
</tr>
<tr>
<td>90</td>
<td>Square Bias</td>
<td>18.25</td>
<td>5.24</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>9.46</td>
<td>16.62</td>
</tr>
<tr>
<td></td>
<td>M.S.E.</td>
<td>27.71</td>
<td>21.85</td>
</tr>
<tr>
<td>120</td>
<td>Square Bias</td>
<td>11.33</td>
<td>2.60</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>9.15</td>
<td>15.74</td>
</tr>
<tr>
<td></td>
<td>M.S.E.</td>
<td>20.48</td>
<td>18.34</td>
</tr>
</tbody>
</table>

**TABLE 1(c)**
<table>
<thead>
<tr>
<th>SAMPLE SIZE</th>
<th>$\frac{M_T}{T}$</th>
<th>PARZEN</th>
<th></th>
<th>TUKEY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>.1</td>
<td>.2</td>
<td>$\frac{1}{3}$</td>
</tr>
<tr>
<td>30</td>
<td>Square Bias</td>
<td>6.98</td>
<td>3.74</td>
<td>2.26</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>8.93</td>
<td>14.07</td>
<td>20.14</td>
</tr>
<tr>
<td></td>
<td>M.S.E.</td>
<td>15.91</td>
<td>17.81</td>
<td>22.40</td>
</tr>
<tr>
<td>60</td>
<td>Square Bias</td>
<td>3.65</td>
<td>1.66</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>7.16</td>
<td>11.96</td>
<td>18.32</td>
</tr>
<tr>
<td></td>
<td>M.S.E.</td>
<td>10.81</td>
<td>13.63</td>
<td>19.14</td>
</tr>
<tr>
<td>90</td>
<td>Square Bias</td>
<td>2.40</td>
<td>0.90</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>6.42</td>
<td>11.36</td>
<td>17.82</td>
</tr>
<tr>
<td></td>
<td>M.S.E.</td>
<td>8.82</td>
<td>12.26</td>
<td>18.24</td>
</tr>
<tr>
<td>120</td>
<td>Square Bias</td>
<td>1.60</td>
<td>0.57</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>6.08</td>
<td>11.10</td>
<td>17.59</td>
</tr>
<tr>
<td></td>
<td>M.S.E.</td>
<td>7.68</td>
<td>11.67</td>
<td>17.82</td>
</tr>
</tbody>
</table>

TABLE 1(d)
FIRST SPECTRUM USED IN COMPUTER STUDY
SECOND SPECTRUM USED IN COMPUTER STUDY
THIRD SPECTRUM USED IN COMPUTER STUDY

Figure 2 (C)

0 30 60 90 120 150 180
0 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0

05/12/68
FOURTH SPECTRUM USED IN COMPUTER STUDY
§7. Comments on the Computer Study

The mean value curve of a spectrum estimate may be termed 'satisfactory' if it tells fairly accurately where the important peaks are without necessarily showing too well how high or how wide they are; for the illustrated spectrum, Figure 1(a), the corresponding truncation points are about 20 for the Tukey estimate and 25 for the Parzen. A reasonably accurate mean value curve is obtained with $M_T$ about 30 for the Tukey estimate and about 40 for the Parzen. From Table 1(a), it would appear that these descriptions correspond to an average square bias of about 3% and 1% of the average of $f(.)^2$ respectively. The aptness of these figures as judgments of the above rough descriptions is borne out by the other cases, but as the other spectra are smoother, the corresponding truncation points are then somewhat smaller. For given $M_T$, $T$ and $k(.)$, the lowest figures for bias in Table 1 are for 1(d), then 1(b), then 1(c) and worst, of course, is 1(a). 1(d) came out best because the spectrum mainly consists of

$$f(\omega) = \frac{1}{2}(1 + \cos \omega)$$

i.e. it is well approximated by the first two terms of its Fourier series, which for any but the smallest truncation points have practically unit weight in the spectrum estimate even where
(as is essentially the case here) \( R_T(.) \) and not \( \hat{R}_T(.) \) is being used. The very narrow peaks centred on \( \omega = \pi/4, 3\pi/4 \) are not really visible however except perhaps with the Tukey estimate with \( M_T = 40 \).

It has been observed in Neave (1968d) that if there is to be any hope of obtaining a useful spectrum estimate then it must be approximately true that

\[
E[f^*_T(\omega)] = f(\omega)
\]

and that therefore the mean value curve is a smoothed version of the true spectrum, which means that peaks are not so high as they should be and troughs not so deep. This is a well-known phenomenon. What is not well-known, but is apparent from this study, is that, away from 0 and \( \pi \), the same is true of the standard deviation curve, after the application of a scale factor, which from (1.15) is asymptotically

\[
\left(2 \frac{M_T}{T} \int_0^\frac{1}{2} k(\theta)^2 (1 - \frac{M_T}{T} \theta) d\theta\right)^{\frac{1}{2}}.
\]

The rise at the end points indicated by (1.14), (1.15) is visible in several cases, though not by as much as the factor of \( \sqrt{2} \) following from the asymptotic result. However, instead of being restricted to the end points, this rise is also seen to contribute extensively to the standard deviation at neighbouring frequencies, an effect
of the overall smoothing just mentioned: this is also a fact which is probably not widely recognised. In fact the "contamination" or "leakage" due to smoothing of the standard deviation function often seems to affect wider frequency bands than that of the mean value function. Its effect is shown impressively in the case of \( l(c) \), the autoregressive spectrum. From Table 1, it can be seen that the variance in this case is proportionally much greater than in the other cases. This is almost entirely due to the peak in the variance at \( \omega = 0 \) being reinforced by the factor of about 2, and the variance being slow to throw off the effects of this factor as \( \omega \) increases. The high figures in Table 1(c) cannot just be due to the high value at \( \omega = 0 \), as the type of average calculated is similar to (6.3), i.e. only giving half the weight to this point (and \( \pi \)) as to any of the others. The discrepancy between the variance figures for the autoregressive case and the others is greatest at \( M_T/T = .1 \), and to illustrate this effect, Figure 3(a) shows the standard deviation diagram for the Parzen estimate with \( M_T = 12 \) and \( T = 120 \), and the asymptotic standard deviation for \( M_T/T = .1 \) (both approximations (1.14) and (1.15) are drawn, but with \( M_T/T \) so small, they are almost coincident); it is apparent for example that in the frequency range \((0, \pi/4)\) the standard deviation is much greater than the approximation would indicate (except of course at \( \omega = 0 \)).

Due to the normalisation with respect to the average of
\( f(.)^2 \), the figures for the variance in Table 1 should tend to limits almost equal to

\[
\frac{2M_T}{T} \int_0^1 k(\theta)^2 \left(1 - \frac{M_T}{T} \theta\right) d\theta \times 100\% ,
\]

and these limits are shown in Table 2. For reference, the classical approximations

\[
\frac{2M_T}{T} \int_0^1 k(\theta)^2 d\theta \times 100\%
\]

are also given. It is seen that invariably the actual variance exceeds the approximation, but for fixed \( M_T/T \) it decreases steadily as \( T \) increases. It is worth noticing however that the classical approximations, (1.14), which are larger than the more relevant (1.15), actually overestimate the figures for the higher values of \( T \). It is also interesting to observe that, conversely to what might be expected, the larger the value of \( M_T/T \), the quicker is the convergence to the limit. This is because several of the terms dispensed with during the proof of (1.15) are of the order of \( \frac{1}{M_T} \). In other words, the smaller the value of \( M_T/T \) (for fixed \( T \)), the worse is the asymptotic formula as an approximation.
\begin{table}
\begin{tabular}{|c|c|c|c|}
\hline
 & $\gamma$ & $\frac{M_T}{2T} \int_0^1 k(\theta)^2 (1-\gamma \theta) d\theta$ & $\frac{M_T}{2T} \int_0^1 k(\theta)^2 d\theta$ \\
\hline
\hline
\text{PARZEN} & .1 & 5.301% & 5.393% \\
 & .2 & 10.418% & 10.786% \\
 & $\frac{1}{3}$ & 16.953% & 17.977% \\
\hline
\text{TUKEY} & .1 & 7.328% & 7.500% \\
 & .2 & 14.310% & 15.000% \\
 & $\frac{1}{3}$ & 23.083% & 25.000% \\
\hline
\end{tabular}
\caption{Table 2}
\end{table}

As previously mentioned, the approximations to the bias mentioned in §4 are not altogether appropriate here as the second derivative of the spectrum is not continuous. However for reference, the shapes of these approximations for the bias of the illustrated spectrum are shown in Figure 3(b).

From this brief computer experiment, a few rough conclusions may be made concerning the choice of spectrum estimates, especially if T is not large. Firstly the choice of $M_T$ should not be some arbitrary function of T which increases in some vague way with T. One should attempt to estimate the lowest value of $M_T$ which should give a satisfactory mean value function for the estimate (see also Neave (1968d)). For spectra of the type studied here, the value would seem to be in the range 30-40 using the Parzen generator,
and slightly lower using the Tukey generator. The order of the variance can then be estimated from the asymptotic formula, remembering the points raised concerning smoothing and the increases in possibly wide neighbourhoods of the end points. If this variance seems sufficiently large to destroy the degree of precision required, then it is unlikely that any estimate will produce satisfactory results. To lower $M_T$ would result in a mean value curve of insufficient precision and to increase $M_T$ would result in even higher variance but with little improvement to the bias - this argument follows since, after obtaining a fairly accurate mean value curve, a big increase in $M_T$ is then usually required to bring out much finer detail, which would probably then be useless because of the increased variance.

It would thus appear that for spectra of the class of those in Figure 1, the choice of $M_T$ could be almost independent of $T$. Published remarks on choosing $M_T$ as about $T/6$ or $T/4$ or $T/3$ etc. seem to express the wrong emphasis, since as $M_T$ is increased beyond the 'satisfactory' region, much more is lost by high variance than is gained by decreased bias, unless $T$ is extremely large.
FINITE PROPERTIES OF SPECTRUM ESTIMATES

---
true spectrum
mean value
standard deviation
root mean square error
bias

PARZEN

T = 30
M_T = 3
TUKEY

$T = 30$

$M_T = 3$
PARZEN

T = 30

M_T = 6
TUKEY

T = 30

M_T = 6
PARZEN

T = 30
M_T = 10

FIGURE 2 (a)

-54-
TUKEY

T = 30
M_T = 10
PARZEN

T = 60

M_T = 6
TUKEY

$T = 60$

$M_T = 6$
PARZEN

\[ T = 60 \]

\[ M_T = 12 \]
TUKEY

T = 60

M_T = 12
PARZEN

\[ T = 60 \]

\[ M_T = 20 \]
PARZEN

\[ T = 120 \]
\[ M_T = 12 \]
TUKEY

\[ T = 120 \]

\[ M_T = 12 \]
PARZEN

T = 120

M_T = 24
TUKEY

$T = 120$

$M_T = 24$
TUKEY

T = 120
M_T = 40
FIGURE 3(a): STANDARD DEVIATION WITH $T=120$, $M_T=12$ FOR THIRD SPECTRUM.
FIGURE 3(b): ASYMPTOTIC FORMS OF THE BIAS CURVES.
§8 The 'Typical Shape' of Economic Spectra

Of course, some spectra met in practice are much more irregular than those considered here. Granger (1966) produces evidence to the effect that the logarithms of many economic spectra have mainly an autoregressive shape, i.e. the ratios of the spectrum ordinates at low frequencies to those at high frequencies are very large indeed. The obvious thought is that these series are subject to complicated trend contributions which have not been efficiently removed. It is pertinent from considerations of the very complex system of influences affecting economic processes that it must be very difficult to remove trend and general non-stationary factors. While admitting this, Granger denies that this is the real reason for the occurrence of his 'typical shape'.

He also illustrates an interesting example of spectrum estimation for a sample of United States bank clearing data (monthly, 1875-1958), from which an exponential linear 'trend' has been estimated and removed. Using the Parzen generator and truncation points of between 60 and 100, the estimate of \( f(0) \) varies between about 6 and 10. The curves then decrease rapidly to about \( .01 \) at \( \pi/10 \), and, apart from seasonal peaks rising to about \( .04 \), decreases from the rest of the frequency range to around \( .001 \). The enormous peak centred on zero frequency must still give some cause for doubt as to whether this is indeed evidence of true low frequency power or if it is mainly the effect of leakage from a deterministic
component at \( \omega = 0 \). After all, economic data is always particularly subject to phenomena such as inflation, stop-go, and wars and other irregular catastrophes, and from these considerations it is very difficult to see how a true economic spectrum could really be without some deterministic effect at zero frequency. However the seasonal peaks are neither excessively high or narrow, so that away from zero frequency, the spectrum could probably be said to belong to the class covered by the computer study.

It is also particularly encouraging that even with such a sharp low frequency disturbance, the main features of the spectrum are so obvious, though certainly little dependence can be placed on the actual magnitude of the low frequency peak. It would therefore appear that when moving into the field of such complicated spectral shapes, the figures concerning 'good' ranges of the truncation points relative to satisfactory mean value curves might not have to be increased to any alarming extent.

§9 The importance of pre-whitening

In general, the conclusions of this study seem to be in favour of Blackman and Tukey's (1958) 'prewhitening' technique, although Granger (1964, p. 45) reports that the general consensus of opinion seems to be against it. This technique is to perform a pilot analysis to discover the location of the main peaks, and then reduce their effect by applying a filter to the data (this effectively means that instead of dealing with raw data, one uses linear
combinations of the data). The importance of the technique stems
from the double leakage from peaks in both mean value and variance
inferred by the empirical results above, and in that the variance
leakage may even affect wider frequency ranges than mean value
leakage. If the peaks are mainly removed, the small values of \( M_T \)
mentioned above became appropriate, and with samples of the size of
the bank-clearing data illustrated by Granger, the variance is then
sufficiently small.

§10. Hypothesis on the behaviour of the variance

From the results in §7, it appears that the exact standard
deivation curve is a smoothed version of the asymptotic form and
therefore might be closely approximated by an expression of the form

\[
A_T(\omega) B_T\left(\frac{M_T}{T}, k\right) C_T(\omega, f, K)
\]

where \( A_T(0) \) increases to \( \sqrt{2} \) as \( T \to \infty \), \( A_T(\pi/2) = 1 \), with \( A_T(\cdot) \)
strictly decreasing in \( (0, \pi/2) \) and symmetrical about \( \pi/2 \),

\[
B\left(\frac{M_T}{T}, k\right) = \left\{ 2 \frac{M_T}{T} \int_0^1 k(\theta)^2 (1 - \frac{M_T}{T} \theta) \theta \, d\theta \right\}^{1/2}
\]

the usual constant in the asymptotic expression, and

\[
C_T(\omega, f, K) = \int_{-\pi}^{\pi} L_K(\lambda - \omega) f(\lambda) \, d\lambda
\]
where $L_{K_T}(.)$ is a function similar to $K_T(.)$ and possibly being definable in terms of $K(.)$, $M_T$ and $T$. 
REFERENCES


ISSERLIS, L. (1918). On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. Biometrika, 12, 134.


