UNDAMPED OSCILLATION OF THE SAMPLE AUTOCOVARIANCE FUNCTION AND THE EFFECT OF PREWHITENING OPERATION*)

BY HIROTUGU AKAIKE

(Received Oct. 7, 1961)

1. Introduction and summary

In the spectral analysis of stationary time series it is well known that the periodgram is an inconsistent estimate of the spectral density function. As is well known and will be seen in the following sections of the present paper, this inconsistency is another expression of the "undamped oscillation" of the sample autocovariance function. The "undamped oscillation" of the sample autocovariance function was well investigated by many statisticians, especially by Bartlett [2]. Nevertheless, it seems to the present author that in the recent investigations of the method of spectral analysis, such as those reported in the recent issue of the Technometrics [Vol. 3, No. 2 (1961)], the practical importance of the results obtained earlier is not fully recognized.

In this paper we shall give a heuristic exposition of the statistical property of the sample autocovariance function to see that for a sufficiently large number of lags the sequence of the sample autocovariances may, under appropriate conditions, be considered to be a stationary stochastic process which has a spectral density function approximately equal to the square of that of the original process. Taking into account of this fact we shall give a warning as to a possible misinterpretation of considering the observed sample autocovariance function to suggest the existence of a "beat phenomenon", where the true power spectrum has a unimodal density.

We can show, further, that the ratio of the root mean square of the amplitude of the undamped oscillation of the sample autocovariance function to the power or variance of the original process is minimum for white noise. This suggests a way how to see the effect of the prewhitening operation advocated by Tukey [3]. Indeed, in this way, we can get

*) A part of the results of this paper was announced in February 8, 1961, at the weekly meeting of the Institute of Statistical Mathematics.

much information concerning the truncation point in the computation of the sample autocovariances to get an estimate of the power spectral density function. With the aid of numerical examples we shall show that the prewhitening operation has, as was frequently pointed out by Tukey, far more importance in the practical application of the spectral analysis than the choice of the spectral windows.

It will also be briefly mentioned that the prewhitening operation brings more important effect to the estimation of the frequency response function by using the cross-spectral density.

2. Undamped oscillation of the sample autocovariance function

In this section we shall give a simple heuristic exposition of the "undamped oscillation" of sample autocovariance function. Extensions of the result to more general cases will be straightforward.

Now, consider the stationary stochastic process $\{X_n\}$ which is defined by

$$X_n = a_0\varepsilon_n + a_1\varepsilon_{n-1} + \cdots a_k\varepsilon_{n-k} \tag{1}$$

where $a_0 \cdots a_k$ is a given set of constants and $\{\varepsilon_n\}$ is a purely white noise of which ε_n has zero-mean, unit variance and finite fourth order moment.*) Such a process is known as a process of moving summation and if we define $R(h) = EX_{n+h}X_n$ the spectral density function of the process is given by $p_x(f) = \sum_{k=-\infty}^{\infty} e^{-2\pi ifh}R(h) = |a_0 + a_1e^{-2\pi if} + \cdots + a_ke^{-2\pi ikf}|^2$ $(\frac{1}{2} \geqq f \geqq -\frac{1}{2})$. Now we define the sample autocovariance function $C(j)$ by

$$C(j) = \frac{1}{N}\sum_{n=1}^{N} X_{n+j}X_n \quad {}^{**)}$$

where N is a given positive integer.

Then, from the defining relation (1) of the process we have

$$\frac{1}{N}\sum_{n=1}^{N} X_{n+j}X_n = a_0\frac{1}{N}\sum_{n=1}^{N}\varepsilon_{n+j}X_n + a_1\frac{1}{N}\sum_{n=1}^{N}\varepsilon_{n+j-1}X_n + \cdots + a_k\frac{1}{N}\sum_{n=1}^{N}\varepsilon_{n+j-k}X_n\,.$$

If we define

$$D(j) = \frac{1}{N}\sum_{n=1}^{N}\varepsilon_{n+j}X_n$$

*) We shall here call the process $\{\varepsilon_n\}$ a purely white noise when ε_n's are mutually independent and indentically distributed.

**) This definition of sample autocovariance function is somewhat different from the ordinary definition. But the difference will be small when $j/N \ll 1$.

we get the important representation

$$C(j)=a_0D(j)+a_1D(j-1)+\cdots+a_kD(j-k)\ . \tag{2}$$

From this representation of $C(j)$ we can see at once that the sample autocovariance function $C(j)$ may be considered to be the response of the filter, which has generated $\{X_n\}$ from the purely white noise $\{\varepsilon_m\}$, to the input process $\{D(j)\}$. Now, let us consider the statistical property of the input $\{D(j)\}$. Taking into account the pure whiteness of the process $\{\varepsilon_n\}$ we can get

$$ED(j)=0 \qquad \text{for } j>0$$

and

$$\begin{aligned} ED(j+h)D(j)&=\frac{N-|h|}{N}\frac{R(h)}{N} && \text{for } j, j+h>0 \text{ and } |h|<N \\ &=0 && \text{for } j, j+h>0 \text{ and } |h|\geqq N\,. \end{aligned}$$

Thus we can see that the process $\{C(j); j\geqq k\}$ is obtained by a linear transformation from a weakly stationary process $\{D(j); j\geqq 0\}$ and that for large N the covariance function of the process $\{D(j); j\geqq 0\}$ is, except for a constant factor N^{-1}, very nearly equal to that of the process $\{X_n\}$.

The following is the well known equation concerning the linear time invariant transformation of the stationary process;

$$p_{\text{out}}(f)=|G(f)|^2p_{\text{in}}(f)$$

where $p_{\text{out}}(f)$ is the spectral density function of the output, $p_{\text{in}}(f)$ is that of the input and $G(f)$ is the frequency response function of the system. Taking into account of this relation we can at once see that the sequence of the sample autocovariances $C(j)$ $(j\geqq k)$ may be considered to be a realization of a stationary process with a spectral density approximately equal, except for the constant factor N^{-1}, to the square of that of the original process $\{X_n\}$.

To extend the present result to more general cases, recall that if a stationary process X_n has an absolutely continuous spectral distribution then it can always be represented [5. Chap. X §8] as an infinite moving summation of white noise with coefficients satisfying $\sum_{-\infty}^{\infty}|a_\nu|^2<+\infty$

$$X_n = \sum_{-\infty}^{\infty} a_\nu \varepsilon_{n-\nu} \,. \quad ^{*)}$$

If in the above representation of X_n we can assume the boundedness of the fourth order moments of $\varepsilon_\nu s'$ and further a condition which assures the validity of the relation

$$E(\varepsilon_{\nu_1}\varepsilon_{\nu_2}\varepsilon_{\nu_3}\varepsilon_{\nu_4})=0 \quad \text{when } \nu_1 \neq \nu_i \ (i=2, 3, 4) \tag{3}$$

then the process X_n can be approximated arbitrarily closely and uniformly in n (in the sense of fourth order moment) by a process of finite moving summation for which the present result applies and then the process $C(j)$ of X_n can be approximated arbitrarily closely (in the sense of mean square) by the corresponding sample autocovariance function of this process of finite moving summation, and in this sense we can extend our result to this more general case. Thus we can see that if the process X_n is Gaussian and with an absolutely continuous spectral distribution, i.e., if the spectrum of X_n can be represented by using a density function, then our present result applies. Further, if we can assume the finiteness of the fourth order moment of ε_n in the definition of the generalized autoregressive process X_n which is defined by

$$A_0X_n + A_1X_{n-1} + \cdots + A_hX_{n-h} = B_0\varepsilon_n + B_1\varepsilon_{n-1} + \cdots + B_k\varepsilon_{n-k}$$

where A's are satisfying the necessary stability condition and $\{\varepsilon_n\}$ is a purely white noise, then our result applies to this case, too.

3. A warning to the misinterpretation of the sample autocovariance function

The analysis of section 2 shows that when the process X_n has a unimodal spectral density function its sample autocovariance function shows "undamped oscillation" with the same central frequency as that of X_n and with the band width narrower than that of X_n. For the Gaussian process with a unimodal spectral density, when the band width of the spectrum is relatively narrow, the realization of the process shows an oscillation which resembles an amplitude-modulated sine wave of which frequency is equal to the midband or central frequency of the process [10, p. 75, 4. p. 87]. Thus we can guess for the present case that the sample autocovariance function will often show a shape which makes

*) We shall call the weakly stationary process $\{\varepsilon_n\}$ a white noise when the process has a flat spectral density, i.e., ε_n's are mutually orthogonal.

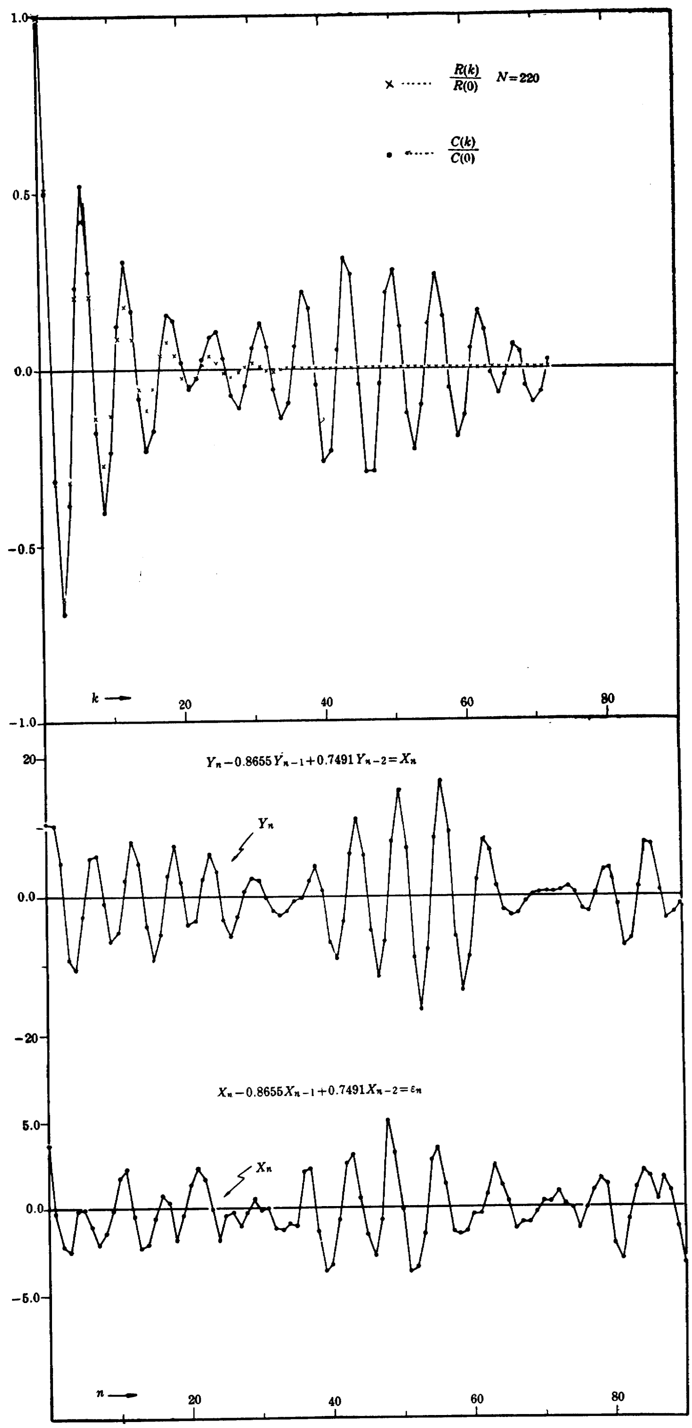


Fig. 1

"Beat phenomenon" of a sample autocovariance function.

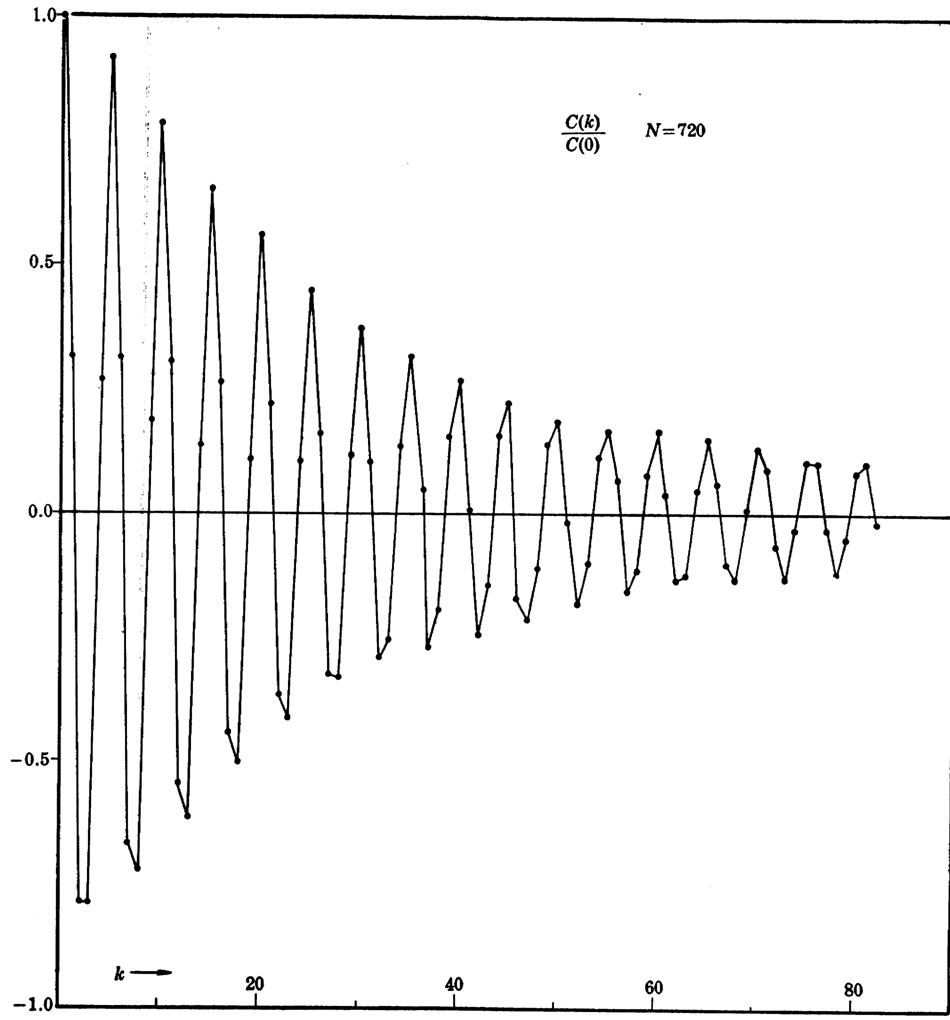


Fig. 2

Sample autocovariance function of oscillation of a ship model.

one to suppose the existence of a beat phenomenon in the original X_n.

Our numerical example in fig. 1 illustrates a typical one of this phenomenon.*) In most of all our numerical computations of sample autocovariance functions hitherto performed we encountered with this apparent "beat phenomena". The present example was computed from an artificially constructed time series which simulates the sampling con-

*) The sample autocovariance function for fig. 1 was computed following the definition of section 2, but others in this section were computed following the definition

$$C(j)=\frac{1}{N}\sum_{n=1}^{N-j} X_{n+j}X_n \quad (j=0, 1, 2, \cdots).$$

dition of the experiment reported in a book of statistical analysis of brain waves [6, chap. 4 and chap. 10]. In this book a correlogram which is very much like that in fig. 1 of the present paper is classified as a type showing a beat in the brain wave while the present artificial series X_n was constructed by using the relation

$$X_n - 0.8655X_{n-1} + 0.7491X_{n-2} = \varepsilon_n$$

where $\{\varepsilon_n\}$ is a purely white noise of which ε_n's follow one and the same Gaussian distribution with zero-mean and unit variance. Thus our present numerical example strongly suggests the inappropriateness of the classification, and it would be more pertinent to regard that sample correlogram as showing the existence of rather lightly damped oscillating mechanism with one degree of freedom. Even if there existed really a beating mechanism we should have to analyse much longer record of the brain wave to attain the desired resolvability against the two dominant frequencies situated very near to each other. In fig. 1 are also illustrated a part of original X_n and its corresponding part of Y_n which was generated by the relation $Y_n - 0.8655Y_{n-1} + 0.7491Y_{n-2} = X_n$. The "undamped oscillation" of the sample autocovariance function has a strong resemblance to the oscillation of Y_n, as was expected by our argument in section 2.

In the case where the band width of the spectrum of X_n is much narrower than the present example, the undamped oscillation of the sample autocovariance function will have a sharp line-like spectrum, and the sample autocovariance function continues a movement which is very much like a sine wave with nearly a constant amplitude. Such an example is illustrated in fig. 2. Thus in the practical applications of correlogram analysis, it will be better first to split the record into two parts of equal length and to compute the $C(k)$'s for each two parts and then to make differences between the corresponding $C(k)$'s. The values of these differences contain only the noise or the effect of sampling fluctuations, and the signal or the true value $\{R(k)\}$ is completely suppressed. In most practical cases, the variances of these differences will be approximately the twice of those of $C(k)$'s, and the differences will give insight into the magnitude of sampling fluctuation of $C(k)$.

By way of discussing the sampling fluctuation of sample autocovariance function, we shall make a slight degression to see that for the estimation of the spectral density it is the signal-to-noise ratio of the sample autocovariance function that matters. The sample covariance $\tilde{C}(k)$ smoothed by the real smoothing kernel $\{d_\nu;\ \nu=0, \pm 1, \pm 2, \cdots\}$ is defined by

$$\tilde{C}(k)=d_k C(k) \quad k=0, \pm 1, \pm 2, \cdots ,$$

where it is assumed that $d_{-k}=d_k$.

Now we want to select a smoothing kernel which will make $\{\tilde{C}(k)\}$ a good estimate of $\{R(k)\}$. For this purpose we seek the d_k's which minimize the expected total mean square error

$$\sum_k E|\tilde{C}(k)-R(k)|^2 .$$

It is assumed in this section that some necessary conditions are satisfied to assure the convergence of infinite sums and integrals. Under the condition $\sum_k E|\tilde{C}(k)-R(k)|^2<\infty$ we have $\sum_k E|\tilde{C}(k)-R(k)|^2=E(\sum_k|\tilde{C}(k)-R(k)|^2)$ and we have

$$\sum_k E|\tilde{C}(k)-R(k)|^2=E\int_{-1/2}^{1/2}|\tilde{p}(f)-p(f)|^2\ \mathrm{df}$$

where $\tilde{p}(f)=\sum_k e^{-2i f k}\tilde{C}(k)$.

Therefore we get

$$E\int|\tilde{p}(f)-p(f)|^2\ \mathrm{df}$$
$$=\sum_k [D^2(C(k))+R^2(k)]\left[d_k-\frac{R^2(k)}{D^2(C(k))+R^2(k)}\right]^2$$
$$+\sum_k \frac{D^2(C(k))R^2(k)}{D^2(C(k))+R^2(k)} \qquad (4)$$

and we can see that the d_k which minimizes the total mean square is given by

$$d_k=\left[1+\left(\frac{D(C(k))}{R(k)}\right)^2\right]^{-1} .$$

This is a result already given by Lomnicki and Zaremba [9], and it shows clearly that for the estimation of the spectral density function it is the signal-to-noise ratio or, assuming the unbiasedness of $C(k)$, the coefficient of variation of the sample autocovariance that matters, and for the estimation of the spectral density the use of the sample autocovariances corresponding to the true covariances which are nearly equal to zero merely increases the sampling variability of the estimate. The result (4) is an expression of the uncertainty in the estimation of the power spectral density, and shows that if we adopt d_k such as

$$\begin{aligned} d_k &= 1 \quad \text{for } |k| \leqq \rho N \quad (\rho < 1) \\ &= 0 \quad \text{otherwise} \end{aligned}$$

and take into account of the fact that for large k it holds that

$$D^2(C(k)) \doteqdot \frac{1}{N}\sigma^2 \qquad (\sigma^2\text{: constant}),$$

then for increasing N the expected total mean square remains nearly constant. This means that the "undamped oscillation" of the sample autocovariance function causes the inconsistency of the periodgram.

Observations made in this and former sections will be of some use to those who will try to estimate the spectral density of some stationary process through the correlogram or the sample autocovariance function. The results show that to protect the unsophisticated analyst from the misinterpretation of the sample autocovariance function it will be most strongly recommended to tell him not to compute sample auto-covariances with lags more than a given number which we tentatively suppose, as was often recommended by Tukey, to be the 10% of the total number of observations used for the computation. Without remembering the results of the present observations, computing the sample autocovariances with lags up to 30% of the total number of observations, which was suggested by Jenkins [7, p. 159], will often be a cause of the trouble in the course of interpretation of the numerical result.

We shall here summarize the content of this and former sections:

If a fairly stable estimate of the power spectral density of a stationary time series is desired, we should use the record of which length is at least ten times longer than that of the lag after which the true auto-covariances will become negligibly small.

In the next section more informations will be obtained to make the above statement practically useful.

4. Effect of prewhitening

In this section, by using the results obtained in section 2, we shall discuss the effect of prewhitening operation on the undamped oscillation of the sample autocovariance function. *)

Assuming the conditions in section 2, we have

$$E|C(k)|^2 \approx N^{-1}\int_{-1/2}^{1/2} p^2(f)df$$

$$E|C(0)| = D^2(X) = \int_{-1/2}^{1/2} p(f)df$$

where k is assumed to be in the range of lags where the true autocovariances are negligibly small compared with the standard deviations of $C(k)$'s and the sign $\approx$ is used to indicate that the difference of the both side members tends to be negligibly small compared with their own magnitude as N increases infinitely. From these relations we have

$$\frac{E|C(k)|^2}{[EC(0)]^2} \approx \frac{1}{N}\frac{\int_{-1/2}^{1/2} p^2(f)df}{\left[\int_{-1/2}^{1/2} p(f)df\right]^2}$$

$$= \left(\int_{-1/2}^{1/2}\left[p(f) - \int_{-1/2}^{1/2} p(f')df'\right]^2 df + \left[\int_{-1/2}^{1/2} p(f)df\right]^2\right)(ND^4(X_n))^{-1}$$

$$= N^{-1}\left(\int_{-1/2}^{1/2}\left[\frac{p(f)}{D^2(X_n)} - 1\right]^2 df + 1\right).$$

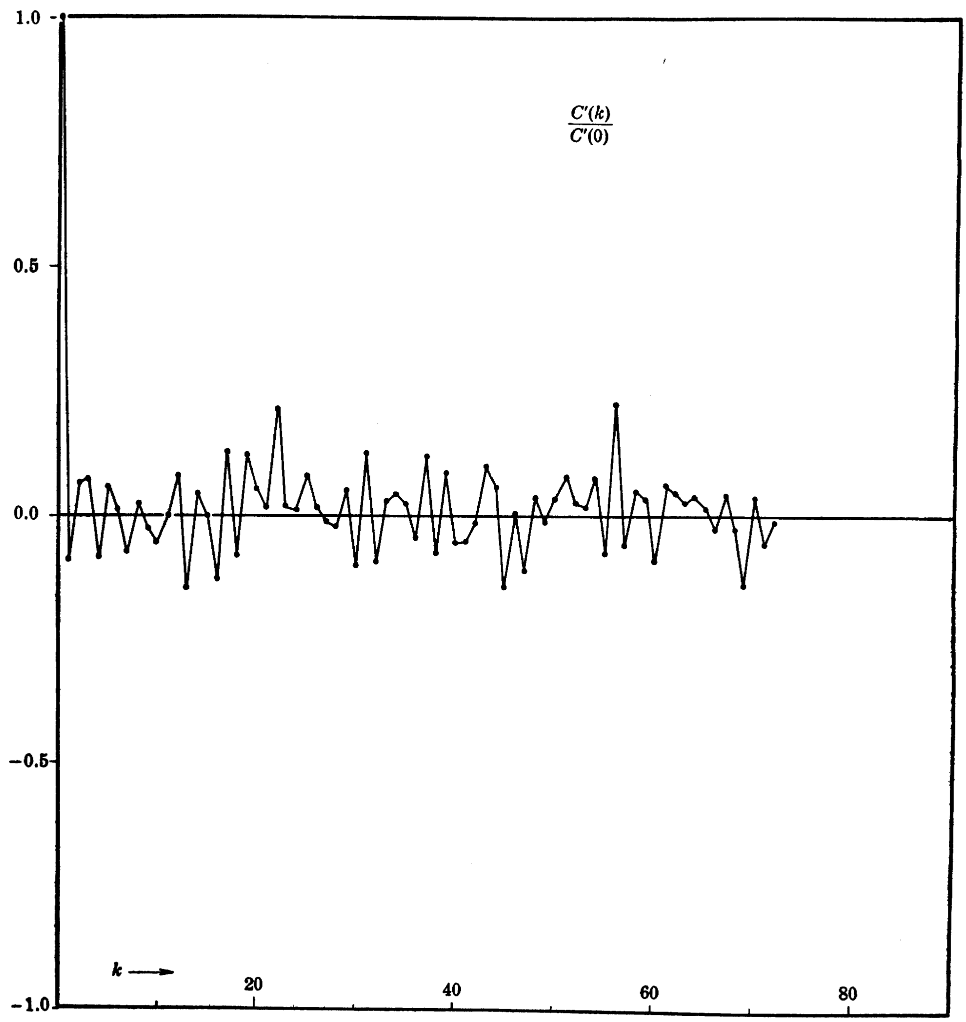
Thus, considering the sample autocovariance function to be composed of the signal $R(k)$ and the noise $C(k)-R(k)$, we can see that the ratio of the mean power $E|C(k)|^2$ of the stationary noise to the square of the maximum signal level $EC(0)$ attains its smallest possible value when $p(f)=D^2(X)$ holds, i.e., when X_n is a white noise. Obviously the above stated ratio takes larger value for $p(f)$ with larger variation, and we can see that if, for a process X_n with highly peaked power spectral density function $p(f)$, we can design a proper numerical filter which will fairly whiten the X_n, then we can markedly improve the signal-to-noise ratio of the sample autocovariance function to avoid the misinterpretation discussed in the preceding section.

We shall see the practical meaning of this signal-to-noise ratio by some numerical examples. The first example is concerned with the already mentioned sample autocovariance function of the artificial series simulating a record of brain wave experiment. The numerical filter for this case was designed by the following procedure:

*) As to the details of the prewhitening after the data have been obtained, see the paper by Blackman and Tukey [3, §15].

Fig. 3.

Effect of prewhitening operation applied to the $\{C(k)\}$ of fiig. 1.



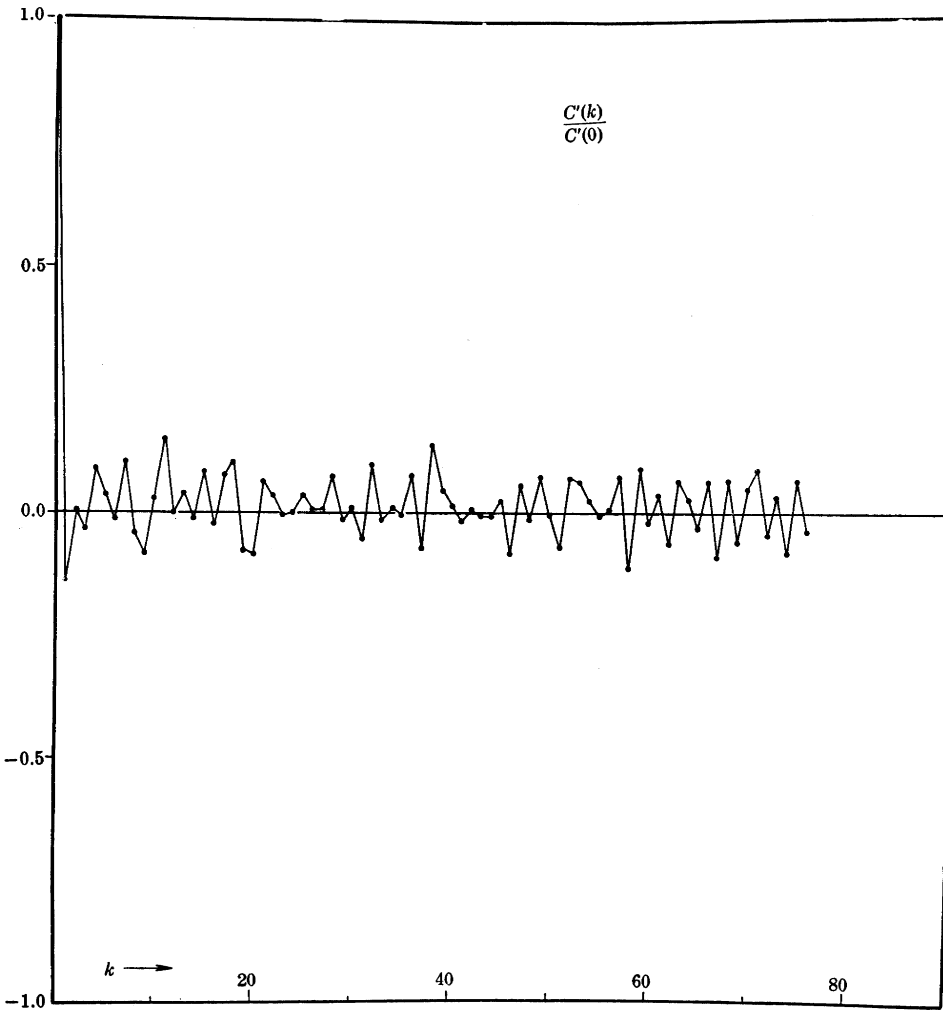
i) Taking into account the shape of the sample autocovariance function we have decided to solve the following simultaneous equations for the unknowns α and β

$$C(1)+\alpha C(0)+\beta C(-1)=0$$
$$C(2)+\alpha C(1)+\beta C(0) \quad =0\ ,$$

ii) then we have calculated the prewhitened sample autocovariance function $C'(k)$, which is the sample autocovariance function of the prewhitened process $X_n'=X_n+\alpha X_{n-1}+\beta X_{n-2}$, by the relation

Fig. 4

Effect of prewhitening operation applied to the $\{C(k)\}$ of fig. 2.



$$C'(k)=\beta C(k+2)+\alpha(1+\beta)C(k+1)+(1+\alpha^2+\beta^2)C(k)$$
$$+\alpha(1+\beta)C(k-1)+\beta C(k-2)\ .$$

In this example, as was already mentioned in the preceding section, we can observe a significant "beat phenomenon" in the original sample autocovariance function, whitse inı[prewhitened form illustrated in fig. 3 we can hardly rəcognize any trace of the regular undamped oscillation. Fig. 4 shows the results of successive applications of the prewhitening operation of the type stated above to the sample autocovariance function

Fig. 5
Effect of prewhitening operation applied to the $\{C(k)\}$ of oscillation of the front axle of an automobile.

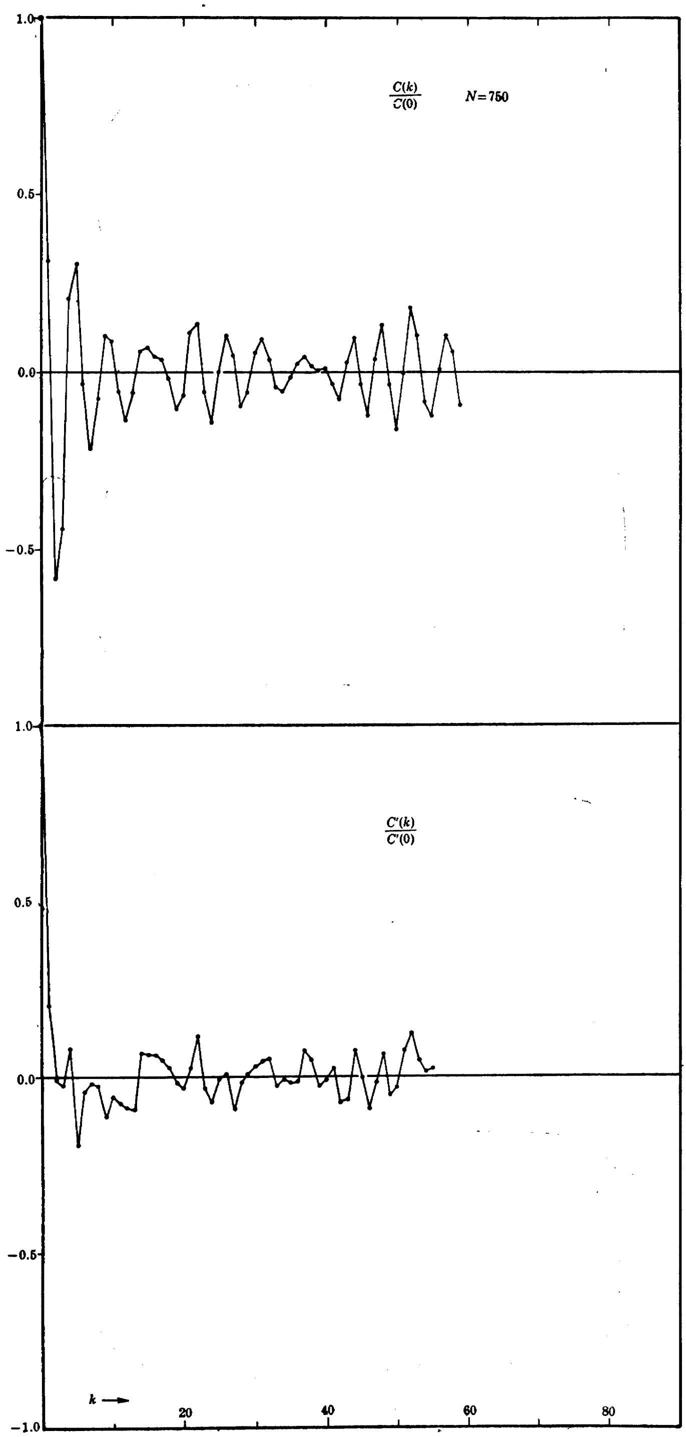
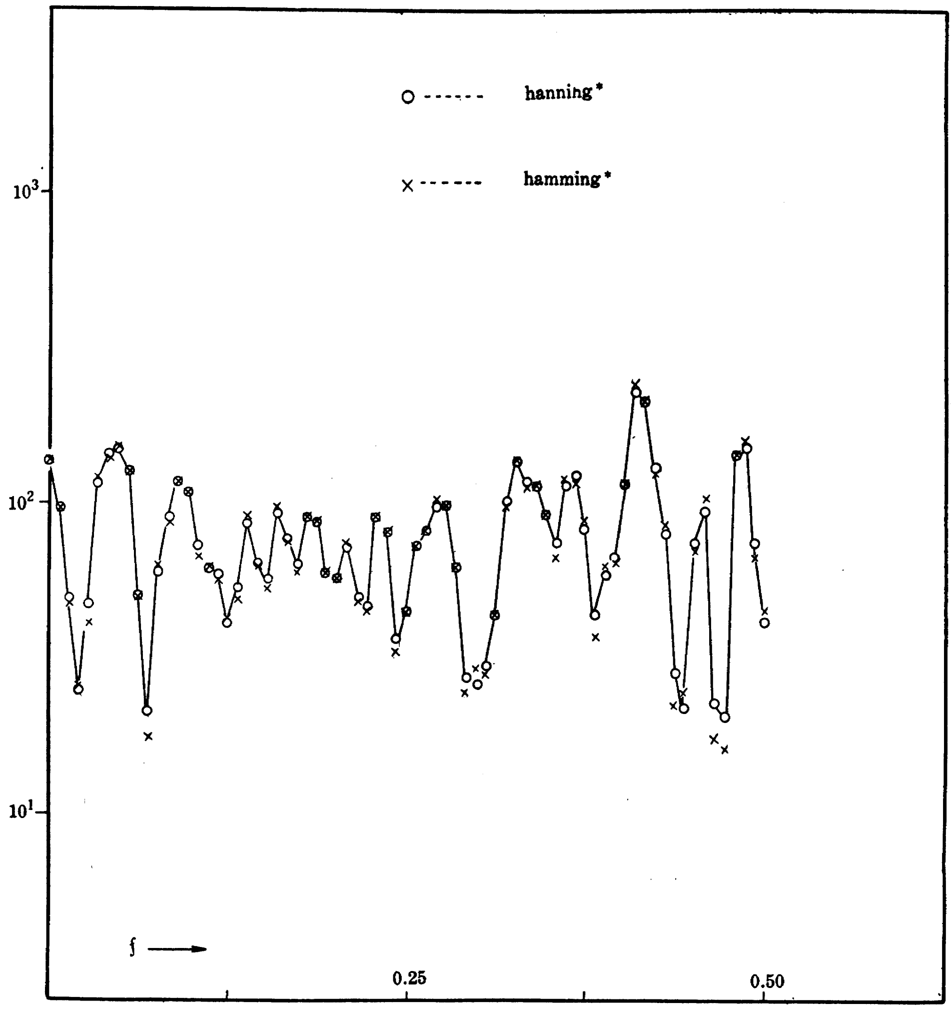


Fig. 6

Comparison of the spectral windows with respect to the $\{C'(k)\}$ of fig. 3.

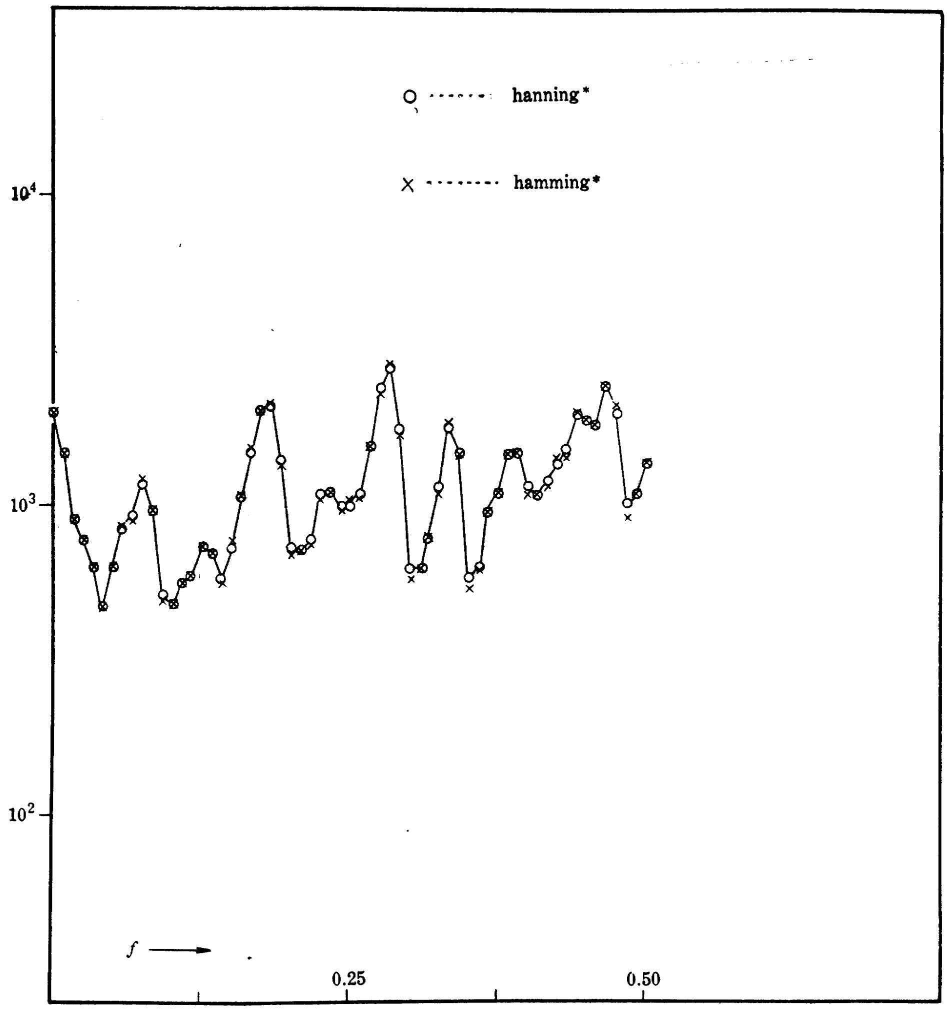


of a record of oscillation of a ship model, which was illustrated in fig. 2.*) The second example given in fig. 5 is concerned with a record of oscillation of the front axle of an automobile running over a paved road.**) These examples show that we should always prepare against the "undamped oscillation" in practical applications of correlogram analysis, especially when the length N of the record used in the calculation of $C(k)$ is not

*) This sample autocovariance function was made available to the present author by courtesy of Mr. Y. Yamanouchi of the Transportation Technical Institute of the Ministry of Transportation.

**) This record was made available to the present author by countesy of Mr. I. Kanesige of the Isuzu Motor Company. It is one of the experimental results treated by Kanesige [7].

Fig 7.
Comparison of the spectral windows with respect to the $\{C'(k)\}$ of fig. 4.



long enough, and that the application of some kind of prewhitening operation to the sample autocovariance function will always be of great help to avoid the misinterpretation of the numerical result.

In figs. 6 and 7 are illustrated the estimates of spectral density which were obtained, after prewhitening, by using different smoothing kernels. From these examples we can obviously see that the effect of prewhitening is so drastic that when prewhitening is performed properly the choice of the spectral window practically matters very little for the estimate of the spectral density function. This point was often stressed

* As to the definition of the window, see Blackman and Tukey [3].

by Tukey in many occasions but without numerical examples, and some of the theoretically minded statisticians are, lacking in the experience of numerical computation, still giving too much weight to the choice of the spectral windows.

It seems to the present author that there remains much to be theoretically investigated in the nature of the prewhitening operation when it is applied after the data are obtained. Some of them are 1) how to design the practically most efficient numerical filter and 2) to study how is the statistical variability of the estimate obtained after prewhitening. In this paper we shall only content ourselves with pointing out the necessity of further theoretical study of the prewhitening operation.

Here we want to mention very briefly the importance of prewhitening operation in the estimation procedure of the frequency response function by using the cross-spectral density. In the statistical estimation of the frequency response function by Fourier-method, we have always to use some sort of smoothing operation not only to reduce the sampling fluctuation but also to avoid the bias due to the truncation of the record, and it becomes essential to keep the input process as white as possible to avoid the bias in the estimation of the frequency response function [1]. In this case the input record may be considered to be a deterministic one and, contrary to the case of estimation of the spectral density, the prewhitening operation applied to the input record does not cause any statistical difficulty. More precise discussion of this estimation procedure will be presented in the forthcoming paper.

Acknowledgement

The author expresses his thanks to Mr. Y. Yamanouchi and to Mr. I. Kanesige for their kindness of providing the practical numerical data. Thanks are also due to Miss Y. Saigusa and Mrs. T. Isii for their preparations of the numerical results presented in this paper.

THE INSTITUTE OF STATISTICAL MATHEMATICS

REFERENCES

[1] H. Akaike, "Note on power spectra," *Bulletin of the Japan Statistical Society*, (1960), pp. 59-60 (in Japanese).
[2] M. S. Bartlett, "On theoretical specification and sampling properties of autocorrelated time-series," *J. R. Statist. Soc. Suppl.*, Vol. 8 (1946), pp. 27-41.

[3] R. B. Blackman and J. W. Tukey, "The measurement of power spectra from the point of view of communications engineering," *Bell System Technical Journal*, Vol. 37 (1958), pp, 185-282, pp. 485-569 (also published seperately by Dover, (1958)).

[4] S. H. Crandall, "Statistical properties of response to random vibration," Chapter 4 in *Random Vibration*, S. H. Crandall editor, Technology Press of M. I. T. and John Wiley and Sons, Inc. New York, (1958), pp. 77-90.

[5] J. L. Doob, *Stochastic Process*, John Wiley and Sons, Inc. New York, (1953).

[6] B. Fujimori and T. Wakabayashi editors, *Analysis of EEG and It's Applications*, Igaku Shoin Ltd. Tokyo, (1957) (in Japanese).

[7] G. M. Jenkins, "General considerations in the analysis of spectra," *Technometrics*, Vol. 3 (1961), pp. 133-166.

[8] I. Kanesige, "Measurement of power spectra of vehicle vibration and vehicle road roughness," *Proceedings of the* 10*th Japan National Congress for Appl. Mech.*, (1960), pp. 371-374.

[9] Z. A. Lomnicki and S. K. Zaremba, "On estimating the spectral density function of a stochastic process," *J. R. Statist. Soc. B.*, Vol. 19 (1957), pp. 13-37.

[10] S. O. Rice, "Mathematical analysis of random noise," *Bell System Technical Journal*, Vol. 24 (1945), pp. 46-156 (also in *Selected Papers on Noise and Stochastic Processes*, N. Wax editor, Dover Publ., New York (1954)).