LIBRARY REFERENCE ONLY

DSIR AND FOC
FIRE RESEARCH
ORGANIZATION
REFERENCE LIBRARY
A99FR.N193

F.R. Note No. 193

DEPARTMENT OF SCIENTIFIC AND INDUSTRIAL RESEARCH AND FIRE OFFICES' COMMITTEE
JOINT FIRE RESEARCH ORGANIZATION

This report has not been published and should be regarded as confidential advance information. No reference should be made to it in any publication without the written consent of the Director of Fire Research. (Telephone: ELStree 1341 and 1797).

CORRELATION FUNCTIONS AND STATISTICAL ANALYSIS

by

D. I. Lawson

June, 1955.

Fire Research Station,
Boreham Wood,
Herts.

© BRE Trust (UK) Permission is granted for personal noncommercial research use. Citation of the work is allowed and encouraged.

CORRELATION FUNCTIONS AND STATISTICAL ANALYSIS

by

D. I. Lawson

During recent years attention has been turned to the limitations of communication channels in respect of the information which can be passed through them in a given time. In order to transmit information, it must be coded in some way so that the information appears as a series of electrical impulses which are passed through the channel and subsequently decoded at the receiver. In ordinary radio transmitters the coding and decoding are accomplished by means of a microphone and loudspeaker. In other systems the coding may be carried out by a teleprinter or by a morse sender. In any case the rate at which information can be handled will be determined by the bandwidth of the channel as this will determine the rate at which signals can be transmitted.

In any practical channel there will be a chance of errors occurring during the transmission because noise will always be present due to the random motion of electrons in the transmission channel and these interfering signals will give rise to uncertainty as to the character of the signals at the receiver. The noise due to the thermal agitation of the electrons has a root mean square value proportional to the square root of the bandwidth of the channel so that as the channel is designed to handle information at a greater rate the greater will be the probability of errors due to noise. This can be offset by increasing the power of the transmitted signal and the following relation has been derived by Shannon (1) between the information handling capacity of a channel (the bandwidth) and the power of the transmission.

$$I\ max = WT \log (1 + P/N)$$

where I max is the maximum amount of information which may be transmitted in a time T. P is the mean signal power received. W is the bandwidth and N is the mean noise power.

This interest in the ultimate limits of information transmission has led to the development of devices for recovering weak signals which may be badly inferfered with by random noise. It seems remarkable that so far no attempt appears to have been made to apply the techniques to statistical analysis for the problems are essentially the same. Statistics are collected with two ends in view; to be able to predict, from an examination of the past, probable future trends, and to examine the correlation between two sets of data in order to see if a casual relation exists between them. All statistical data by definition are subject to random variations but within any data there may be a periodic trend (the signal) and it is one of the functions of the duty of the statistician to discover this. This is, of course, the same problem that faces the communication engineer of separating the signal from the noise. In order to do this use is made of the fact that the signal is coherent while the noise is random. A way of doing this is to compare the signal at any instant of time with its value after a given time delay. A convenient method of making this comparison is by using an auto-correlation function ϕ_τ which is defined by

$$\phi_\tau = \frac{1}{2T}\int_{-T}^{T} y(t).y(t-\tau)\,dt$$

Where $y(t)$ is the displacement of the signal at time t

τ is the time delay

T is the time for which the examination is carried out.

The integral $\int_{-T}^{T} y(t)\,y(t-\tau)\,dt$ is not of course a true correlation as its value will vary with $y(t)$ the amplitude of the signal. This may be normalised by dividing the signal values by their R.M.S. values before performing the integration. If it is required to keep the correlation within the usual limits of ± 1 and the normalised auto-correlation function becomes

$$\phi_{TN} = \frac{1}{2T}\int_{-T}^{T} \frac{y(t)}{\sqrt{\bar{y}^2}}\,\frac{y(t-\tau)}{\sqrt{\bar{y}^2}}\,dt$$

$$= \frac{1}{2T\bar{y}^2}\int_{-T}^{T} y(t)\,y(t-\tau)\,dt$$

If $\phi(\tau)$ is plotted as a function of τ the resulting graph is called an autocorrelogram, this has several important properties which will be stated without proof.

(1) The autocorrelation function is an even function i.e.

$$\phi(\tau) = \phi(-\tau)$$

(2) Any periodic function $y(t)$ will give rise to a periodic autocorrelogram having the same period as the function.

(3) Any noise signal generated in a channel of bandwidth W will have a form $\frac{\sin 2\pi Wt}{2\pi Wt}$ i.e. the form of a damped wave train. This means that if the bandwidth of the channel is considerably greater than the frequency of the signal being transmitted, the autocorrelogram will reveal the signal, as shown in Figure 1. Putting this in statistical language, if the statistical fluctuations occur at a higher rate than the trands which are being sought, the autocorrelogram will reveal the trends.

The other function of statistical analysis is to examine relationships between sets of data and in order to do this use may be made of the cross-correlation function. Using the language of communications for the moment, if two signals $x(t)$ and $y(t)$ are to be compared, then the cross correlation function will be

$$\phi(\tau) = \frac{1}{2T}\int_{-T}^{T} x(t)\,y(t-\tau)\,dt$$

As with the autocorrelation function the existence of a periodic cross-correlation function will indicate a periodic relation between x and y.

An automatic method of auto and cross-correlation

Clearly the value of auto and cross-correlograms as a method of statistical analysis will depend on the rapidity with which an examination of sets of data can be made. Fortunately, a simple method has been devised by Revesz [2] for computing auto correlograms. The signals, i.e. the data is recorded after suitable modulation on to a magnetic tape as it arrives and after a sufficient quantity has been collected the tape is played back through two pick-up heads as shown in Figure 2 and the delay between the heads can be varied to provide various values of . The two signals after amplification and demodulation are multiplied together by feeding one to the stator and the other to the moving coil of an electricity meter. In order to integrate the product it is only necessary to count the revolutions of the rotor and this is done photoelectrically. The autocorrelogram is constructed by plotting the integrated product for any fixed time period as a function of the delay τ between the pick-up heads.

Cross correlograms may be prepared by recording the two sets of statistical data on two synchronised tape recorders or by recording two tracks on a single tape. The two recordings are then reproduced each by a separate pick-up head arranged so that a variable time delay can be introduced and the resulting signals are handled as before.

Bibliography

1. Shannon, C.E. A mathematical theory of communication bell system Technical Journal 1949, 27 p.379 and 623.

2. Revesz, G. An autocorrelogram computer. Journal of Scientific Instruments vol 31, No. 11 1954, p.p. 406-410.

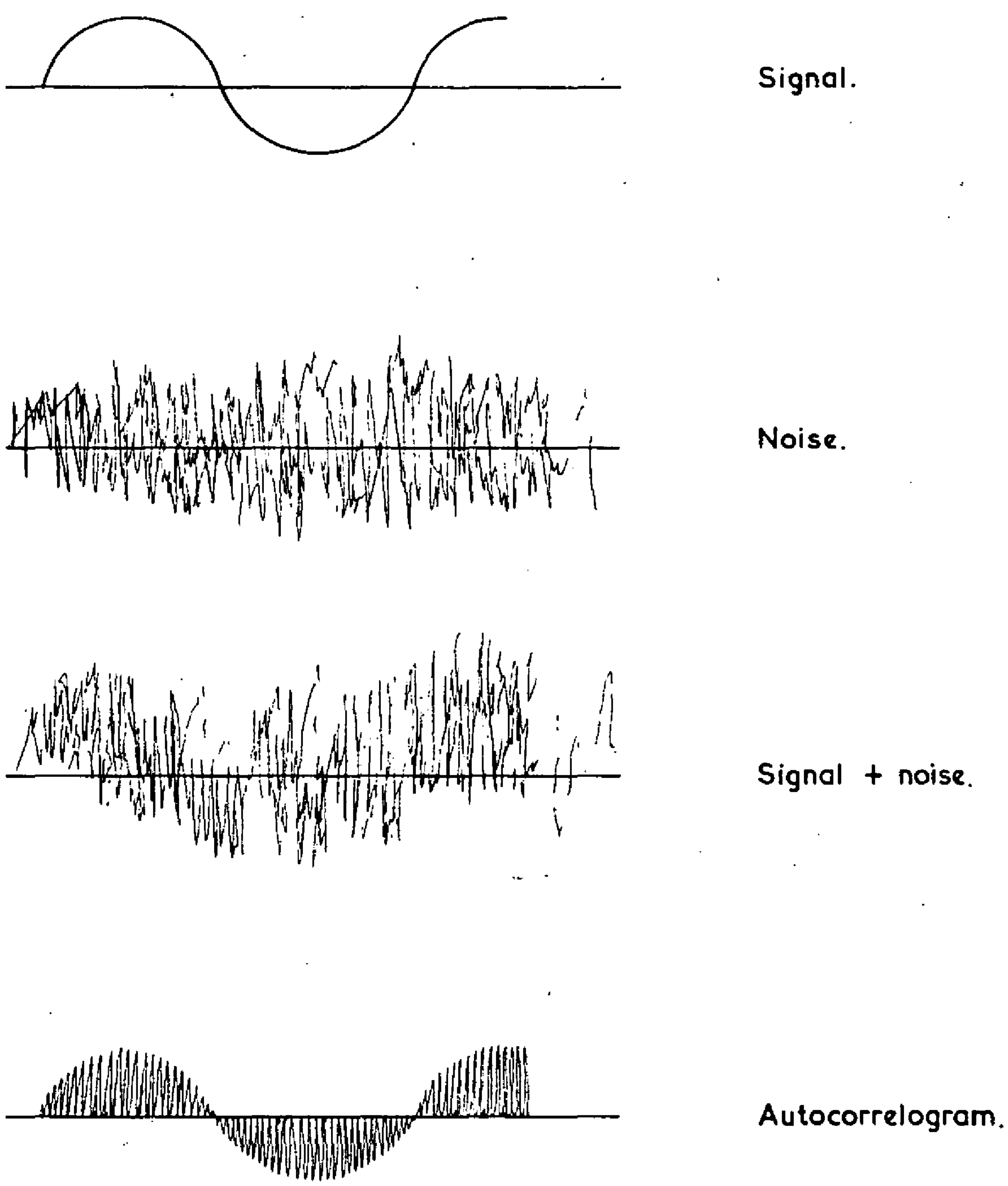

FIG. I. A SIGNAL AND ITS AUTOCORRELOGRAM.

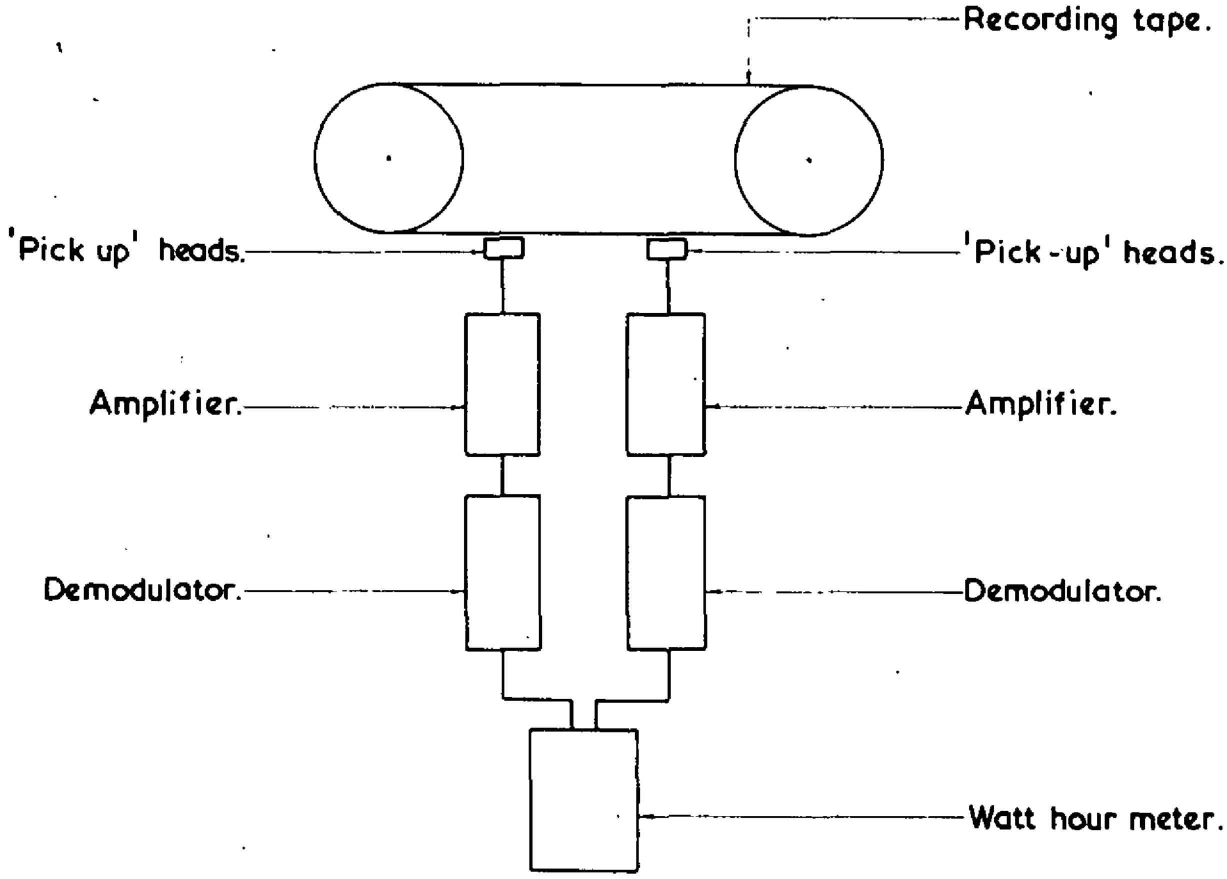

FIG. 2. A SCHEMATIC OF A AUTOCORRELOGRAM COMPUTER.