VIBRATION TESTING COURSE

SCIENTIFIC ATLANTA SPECTRAL DYNAMICS

INTRODUCTION TO RANDOM VIBRATION TESTING

What Is Random Vibration Testing?

To understand random vibration testing, it is necessary to fully understand random waves, how they are controlled, shaped, measured, and analyzed. It is also necessary to know why random signals are used, how specifications are derived, and what methods are used in random vibration testing.

What Are Random Waves?

A random wave is one which never repeats itself; the same set of instantaneous amplitude values will never occur again. A random wave is generally thought of as being a broad spectrum process such as thermal noise or the output of a random noise generator. However, most random processes are very narrow bandwidth, usually due to filtering. When observing on an oscilloscope, we see a typical scrambled, jumbled picture. If we were to observe a random wave using a narrowband filter, such as 2 Hz or 3 Hz, we would see a wave undulating in amplitude at a 2 to 3 Hz rate. The wave appears to be sinusoidal with an amplitude modulation. Nevertheless, it is not a sinusoid. No matter how narrow we make the filter, a random wave will never become sinusoidal, even though we may have difficulty in distinguishing such a narrowband random wave from a sinusoid.

How Are Random Waves Measured?

When we observe random waves on an oscilloscope, we are looking at the time domain. Ordinarily, we begin with a time function from which we wish to produce the corresponding spectrum. This process of obtaining a spectrum from a time function is called analysis. The complementary process of obtaining a time function from a spectrum is called synthesis.

Since random waves are neither periodic nor stationary, we resort to describing the wave characteristic in terms of a "law of averages". In fact, one way in which a random wave can be described is by its average or mean value. The instantaneous values of a random wave will fluctuate about the average value. This is called dispersion. To measure the average, we can use the common full-wave rectifier type voltmeter.

While the mean value of a random wave is analogous to its "center of gravity" and is called the first moment of the distribution, the second moment is also a

valuable statistic. Higher moments also exist mathematically, but are not in common use. The second moment of the signal's distribution is the "mean square". Voltmeters are available to measure this function and are sometimes called "true rms meters" since some of them measure rms as well as mean square. By employing a full-wave electronic rectifier circuit, the bottom half or negative portion of the wave is folded up into the top or positive half and squared using a squaring circuit.

Another function often required in analysis work is the root mean square or rms. The rms is the positive square root of the mean square. From this definition, we note that, in order to compute the rms, we must first compute the mean square. This is the classical way of obtaining rms: first square, then add the squares, divide by the number of squares and finally extract the square root. In analog instruments, the process is continuous.

When a mechanical structure is excited into vibration, the frequencies are, in general, contained in the exciting force and the natural frequencies of resonance of the structure. The method used for this analysis is called PSD Analysis or power spectral density. The purpose of power spectral density is to plot these vibrations in terms of the mean of the square of the amplitude against frequency. The reason the square of amplitude is used instead of the linear amplitude is that we wish to know the frequency distribution of the power or intensity dissipated or the real part of the vibration, and we are not concerned with the reactive component of the vibration. We also require the differential power rather than the power at discrete frequencies because there are no discrete frequencies in a random process as there are in the case of periodic waves. Hence, we think not of the power at a certain specific frequency, but of the power in a band of frequencies.

What are the Requirements for Random Vibration Systems?

The random vibration system (consisting of a shaker or exciter head, power amplifier, vibration sensor or pick-up, charge amplifier, equalizer/analyzer, true rms meter, X-Y plotter or recorder and noise generator/source) shall be capable of generating a random vibration for which the magnitude has Gaussian (normal) amplitude distribution, except that the acceleration magnitudes of the peak values may be limited to a minimum of three times the rms (three-sigma limits).

The system shall be capable of being equalized so that the magnitude of its spectral-sensity curve will be between specified limits when the test item, or a substitute equivalent mass, is appropriately secured to the shaker or exciter machine. The equalization of an electro-dynamic vibration machine system is

the adjustment of the electrical amplifier and control system so that the ratio of the output-vibration amplitude to the input-signal amplitude is of a constant value (or given values) throughout the required frequency spectrum.

Why Use Random Testing?

Many parts or components, when exposed to vibration, may have more than one resonant frequency that could be detrimental to the unit's life or its capability to function. This test is conducted for the purpose of determining the ability of component parts to withstand the dynamic stress exerted by random vibration applied between upper and lower frequency limits to simulate the vibration experienced in various service-field environments. Random vibration is more characteristic of modern-field environments produced by missiles, high-thrust jets, and rocket engines. In these types of environments, the random vibration provides a more realistic test.

By exposing component parts to random vibration, we can:

(1) Determine if the unit will meet the requirements that it was designed for;

(2) Determine if a change in mounting or re-design is required;

(3) Predict the life expectancy or failures that may occur; and,

(4) Guarantee that all units of the same design will be acceptable.

PROPERTIES OF RANDOM WAVES

Wide and Narrow Band Random Waves

A random wave is one which never repeats itself; the same set of instantaneous amplitude values will never occur again. We generally conceive of a random wave as being a broad spectrum process such as Brownian motion, thermal noise or the output of a random noise generator. However, many random processes are of very narrow bandwidth, usually due to filtering. When we observe a wide-band random wave on a cathode ray oscilloscope, we are accustomed to seeing a typical scrambled, jumbled picture. When we observe a narrow-band random wave, particularly when the bandwidth is as narrow as 2 or 3 Hz, we observe a wave undulating in amplitude at a 2 to 3 Hz rate, most probably in accordance with a Gaussian amplitude distribution and most probably having a uniform rectangular phase distribution from $-\pi$ to $+\pi$. Actually, on a cathode ray scope, we will not see much phase variation because the synchronizing CRO signal tends to hold the picture steady on the X axis. The wave appears to be a sinusoid with an amplitude modulation. Nevertheless, it is <u>not</u> a sinusoid. No matter how narrow we make the filter, a random wave will never become sinusoidal, even though we may have difficulty in distinguishing such a narrow-band random wave from a sinusoid.

Stationarity

Aside from the statistical functions which characterize a random wave, such as autocorrelation, PSD, etc., there are some other properties which we will now discuss. The first of these is stationarity. A random wave is said to be weakly stationary when its primary statistics are invariant with time. The primary statistics are the mean, mean square and the autocorrelation function (or PSD). If these functions do not vary outside of statistical limits with time, the wave is said to be weakly stationary. To be strongly stationary, all the higher moments of the mean and autocorrelation functions must also be invariant. In general, if the mean and mean square are reasonably steady, we can

ssume stationarity for practical purposes. The property of stationarity is esirable. With stationarity, it does not matter where a sample is taken in ıe sequence of an experiment.

Aside from practical aspects such as this, stationarity means a reat deal to the theoretical engineer. It means that he is relieved of the reat labor associated with the study of non-stationary events. This is so mportant to him that he will invoke stationarity whenever he can and he will olerate some nonstationarity rather than deal with a much more difficult problem, and this is rightfully so.

Actually, very few physical processes are stationary in a full mathematical sense. A practical example is the vibration of a liquid fuel rocket ngine. As the fuel is gradually depleted, a change in the vibration characteristics results. However, the process may be considered approximately stationary for short periods of time but not from the beginning to the end of the ight. On the other hand, the output of a random noise generator is stationary s long as it is turned on. In the rocket engine case, we may have to analyze series of five second samples over a period of sixty seconds. In the noise enerator case, we can make an on-line analysis.

If we were to make an on-line analysis of non-stationary data, we ould get the average of all the different vibrational conditions from the beinning to the end of the experiment. Little useful purpose would be achieved / such an experiment. In some cases, it may be desirable to test the entire xperiment for stationarity before we analyze a sample. We observe the data om the output of an averaging type meter and also a true mean squared eter. In each case, there will be a fluctuation of the meter readings about e mean value. If the meter movement is fairly steady and to a good first pproximation, does not swing too much about its average reading, we can say e process is stationary. This statement leads to criticism since it is purely ıalitative, but be assured that a little experience will enable a good decision be made. Although the definition of "weakly stationary" includes the requirent of a stationary autocorrelation function in addition to the mean and mean quare, it is often not necessary to make this measurement to prove stationarity.

Since the spectrum does not often change during a trial, PSD may be found to be more revealing in such cases than the autocorrelation function as a spectrum stationarity guidepost despite the fact that both contain the same basic information, being Fourier transforms of each other. An exception might be a solid fuel rocket engine which will have a spectral shift as the combustion cavity grows in volume as the fuel is depleted. In this case, the spectrum shifts downward from some high frequency as the burning proceeds. This is readily detected by listening to the sample. Accordingly, mean, mean square and listening tests are usually enough to establish a practical level of stationarity. We should, however, use our intuitive knowledge of the experiment and particularly our eyes and ears as they are excellent analyzers in their own right.

Ergodicity

In addition to being stationary, we also desire the process we are analyzing to be ergodic. To explain this property, suppose we have a large number of independent experiments on the same problem. Let us say that each of these experiments is one minute long. We know that, if the statistics of the test are steady throughout the one-minute interval of each experiment, we can call the process stationary. Now, in addition to this, if we average "across the ensemble" of experiments at some time, t_1, all the values at time, t_1, for each member of the ensemble being added and then divide by the number of the members of the ensemble, we will obtain an "ensemble average". This would apply to all such statistics as the mean, mean square, etc. If, now, the average in time over the period of one minute is the same as the average over the ensemble, the process is said to be ergodic. This is a very powerful attribute since it means that we can base all of our calculations on a single record; we will not require an analysis of each individual record. Ergodicity is said to be weak if the primary ensemble statistics of mean, mean square and autocorrelation are steady; strong if all higher moments of these functions are included. We shall have more to say later about the moments of a function. Ergodicity is difficult to prove quantitatively. We

do not usually have a large number of identical tests to form an ensemble. Ordinarily, we must depend on a qualitative intuition which tells us that if a certain experiment were exactly repeated, it would lead to the same results as the first.

Refer to Figure 3-1. Here we see some random processes. $y_1(t)$, $y_2(t)$, ... , $y_k(t)$. The time averages, shown for $y_1(t)$ and $y_k(t)$ only are indicated by a bar over the average quantity as $\overline{y}_1$ or $\overline{y}_k$. The ensemble averages are indicated by arrow shaped brackets, $\langle y(t_1)\rangle$, $\langle y(t_2)\rangle$. If the process is stationary, $\langle y(t_1)\rangle = \langle y(t_2)\rangle$. That is, the average remains statistically constant no matter where in time we average across the ensemble. It is ergodic if $\langle y(t)\rangle = \overline{y}_i$ for i = 1, 2, 3, ...

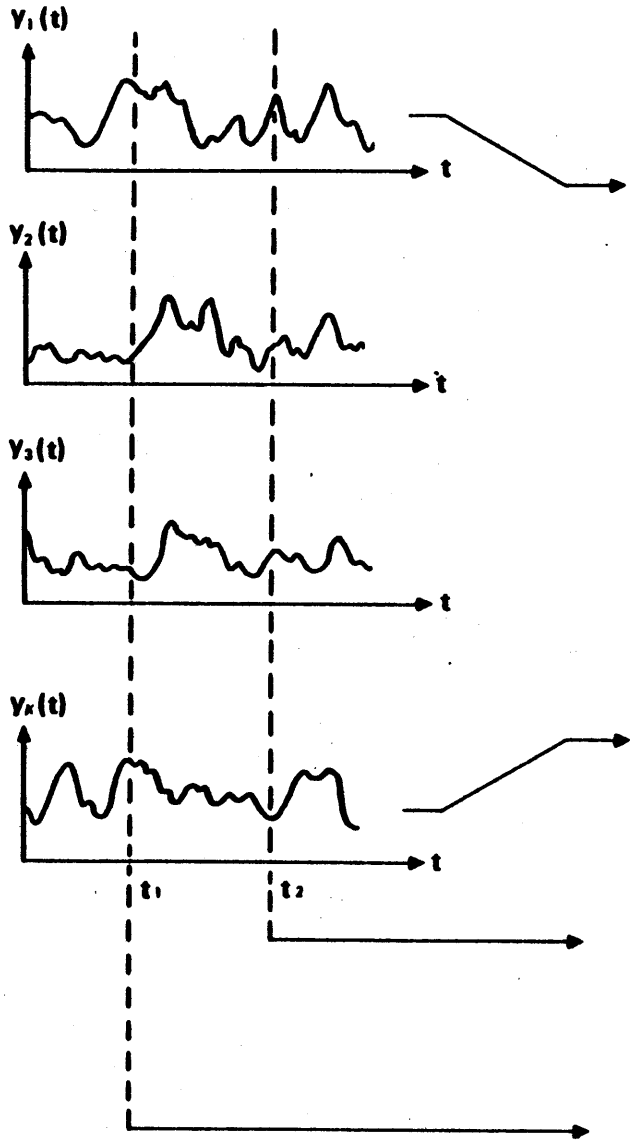
Figure 3-2 shows this condition for mean squared values and Figure 3-3 for the autocorrelation function.

If the mean or another statistic for two members of an ensemble are plotted, the plots are not to be expected to overlay. However, they will have the same mean or average value, they will both have the same amount of dispersion about the mean and these conditions will hold true within the given interval of statistical confidence.

Normalcy

The terms "normal" and "Gaussian" are identical when applied to the amplitude distribution of a random wave. It is generally argued that the Gaussian distribution arises from the implications of the "Central Limit Theorem". This theorem states that if there are a large number of independent wavelets, such as there are in thermal noise where a very large number of molecular and electronic collisions each cause a minute wavelet of electrical current and each wavelet has its own amplitude distribution different from any other, the overall amplitude distribution of the process tends to become Gaussian. The mathematical expression for a Gaussian distribution is:

$$p(x) = \left(\sqrt{2\pi\sigma^2}\right)^{-1} \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right] \qquad (3\text{-}1)$$



$\langle\ \rangle$ Indicates ensemble average

— Indicates time average

$$\bar{y}_i = \lim_{T \to \infty} \frac{1}{T} \int_0^T y_i(t)\, dt$$

= time average for record $y_1(t)$

$$\bar{y}_K = \lim_{T \to \infty} \frac{1}{T} \int_0^T y_K(t)\, dt$$

= time average for record $y_K(t)$

$$\langle y(t_2) \rangle = \lim_{k \to \infty} \frac{1}{k} \sum_{i=1}^{k} y_i(t_2)$$

= ensemble average for time $t = t_2$

$$\langle y(t_1) \rangle = \lim_{k \to \infty} \frac{1}{k} \sum_{i=1}^{k} y_i(t_1)$$

= ensemble average for time $t = t_1$

If the random process is stationary,

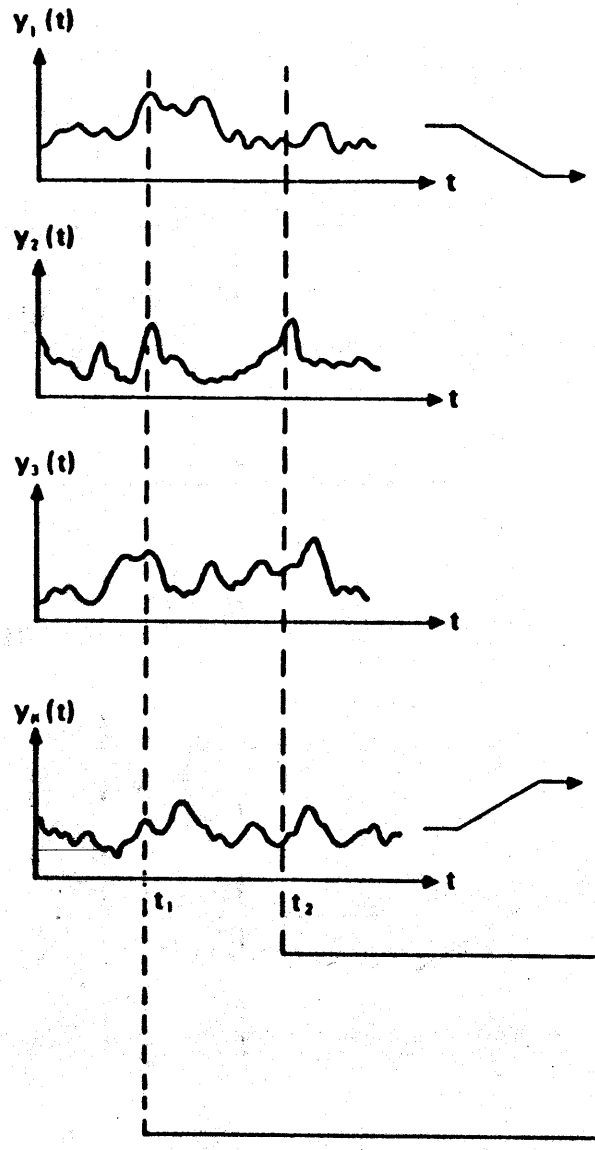
$$\langle y(t_1) \rangle = \langle y(t_2) \rangle$$

If the stationary random process is ergodic,

$$\langle y(t) \rangle = \bar{y}_i\ ;\ i = 1, 2, 3, \ldots$$

COMPUTATION OF MEAN VALUES

FIGURE 3-1



$$\overline{y_i^2} = \lim_{T \to \infty} \frac{1}{T} \int_0^T y_i^2(t)\, dt$$

= mean square time average for record $y_i(t)$

$$\overline{y_K^2} = \lim_{T \to \infty} \frac{1}{T} \int_0^T y_K^2(t)\, dt$$

= mean square ensemble average for record $y_K(t)$

$$\langle y^2(t_2) \rangle = \lim_{k \to \infty} \frac{1}{k} \sum_{i=1}^{k} y_i^2(t_2)$$

= mean square ensemble average for time $t = t_2$

$$\langle y^2(t_1) \rangle = \lim_{k \to \infty} \frac{1}{k} \sum_{i=1}^{k} y_i^2(t_1)$$

= mean square ensemble average for time $t = t_1$

$\langle\ \rangle$ Indicates ensemble average

— Indicates time average

If the random process is stationary,

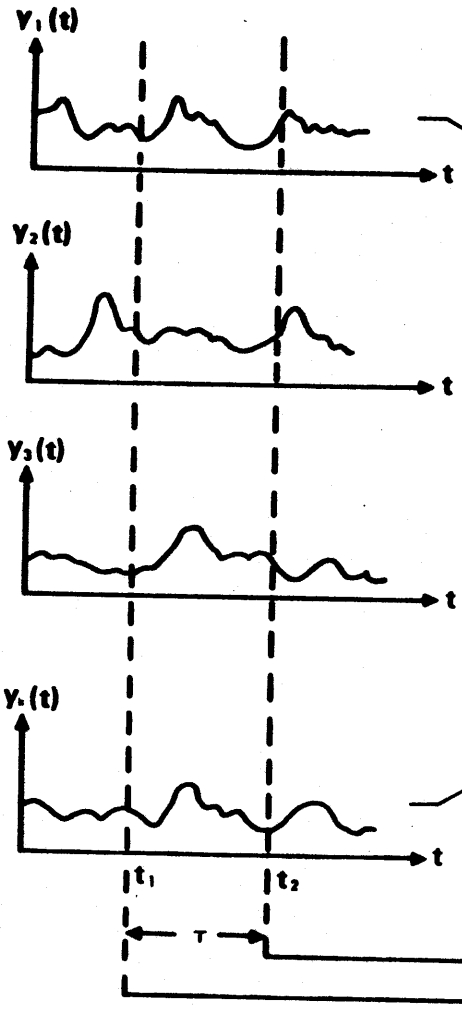
$$\langle y^2(t_1) \rangle = \langle y^2(t_2) \rangle$$

If the stationary random process is ergodic,

$$\langle y^2(t) \rangle = \overline{y_i^2};\ i = 1, 2, 3, \ldots$$

COMPUTATION OF MEAN SQUARE VALUES

FIGURE 3-2



$$\overline{y_1(t)y_1(t+\tau)} = \lim_{T\to\infty} \frac{1}{T} \int_0^T y_1(t)y_1(t+\tau)\,dt$$

= time correlation for record $y_1(t)$

$$\overline{y_K(t)y_K(t+\tau)} = \lim_{T\to\infty} \frac{1}{T} \int_0^T y_K(t)y_K(t+\tau)\,dt$$

= time correlation for record $y_K(t)$

$$\langle y(t_1)y(t_2)\rangle = \lim_{k\to\infty} \frac{1}{k} \sum_{i=1}^{k} y_i(t_1)y_i(t_2)$$

= ensemble correlation for times $t = t_1$ and t_2

$\langle\ \rangle$ Indicates ensemble average

— Indicates time average

If the random process is stationary,

$$\langle y(t_1)y(t_2)\rangle = \langle y(t)y(t+\tau)\rangle$$

If the stationary random process is ergodic,

$$\langle y(t)y(t+\tau)\rangle = \overline{y(t)\ y(t+\tau)} = R_y(\tau)$$

COMPUTATION OF AUTOCORRELATION VALUES

FIGURE 3-3

In words, this says that the probability density p (x) is equal to the product of a term containing the variance σ^2 and an exponential with a complicated negative exponent. See 3-1. The exponent has the term $-(x - m)^2$ which is the mean value, m, of the function subtracted from x, all squared. This is divided by two times the variance. It is not necessary to compute this equation since it is tabulated in many texts.

Before we go to Figure 3-4, we should discuss the quantity, σ. σ^2 is the variance. The variance is the dispersion or scatter of the squared values of x about the Y axis or centroid. It is closely related to the mean square value since it is equal to the mean square value when the square of the average or "dc" is zero, thus mean square - mean^2 = variance = σ^2.

The positive square root of the variance is called the standard deviation, σ. The standard deviation is the positive square root of the mean square and also equal to the rms when there is no "d-c" component. Since there is seldom a "d-c" component, mean = zero, in vibration work because the instrumentation renders the mean to zero, the variance can usually be taken as the mean square and the standard deviation, σ, taken as the rms. This makes life easy since we can read the standard deviation and the variance on a true rms meter.

Refer now to Figure 3-4a. The abscissa is normalized in terms of the σ and plotted in terms of instantaneous values of the function, x. Thus, when x = 0, the probable instantaneous value is about 40%. When the instantaneous value is equal to σ, the probable value of the instantaneous value is about 25%. Put in another way, if a function has a Gaussian distribution, the chance that the instantaneous value is zero is about 40%. The chance that the instantaneous value is equal to the rms ($\sigma = 1$) is about 25%. The chance that it is three times the rms ($\sigma = 3$) is about 1/2 of 1%. This distribution is tabulated in many books. σ is both positive and negative since an instantaneous peak can occur in either the positive or negative region. Figure 3-4b shows the amplitude profile.

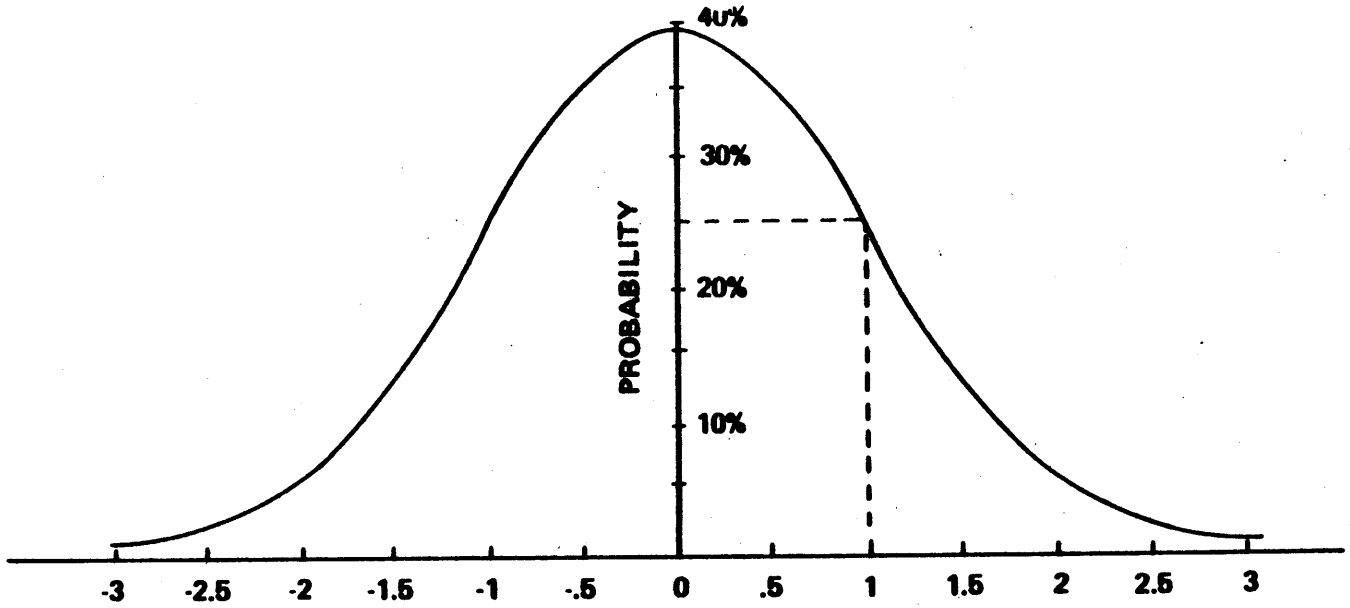
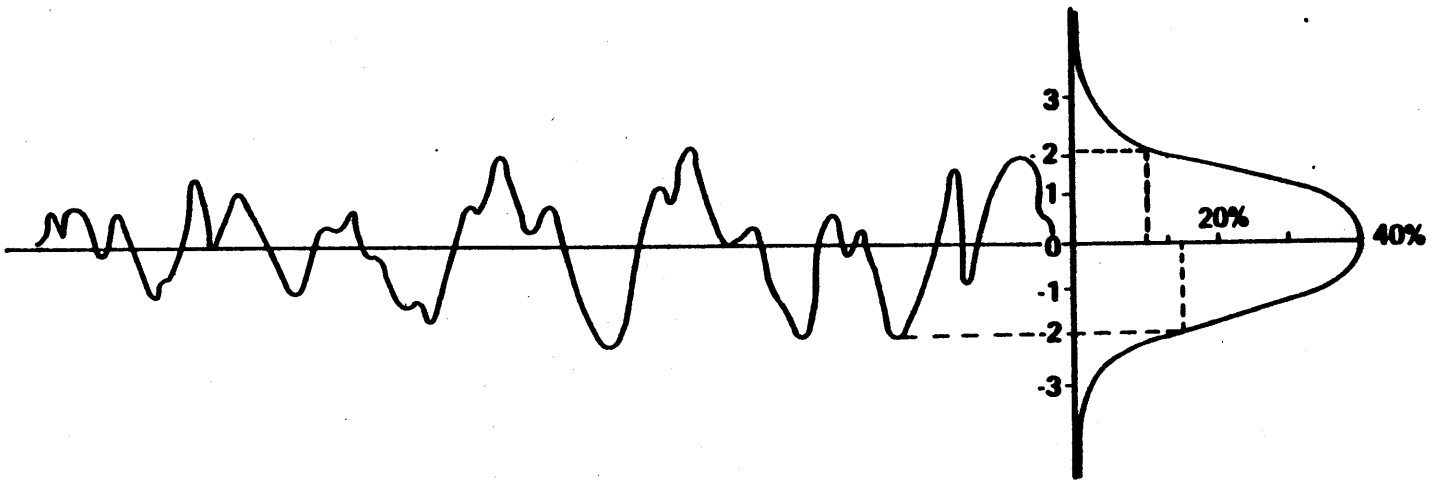


FIGURE 3-4a



AMPLITUDE DISTRIBUTION OF A RANDOM WAVE

FIGURE 3-4b

Skew and Kurtosis

The property of a Gaussian distribution is invaluable to the theoretical engineer because it simplifies his mathematics since it possesses an analytical definition. What happens then, when the distribution is not Gaussian? As we might have known, the Greeks have a word for it -- kurtosis -- which means bent or curved. Imagine a bell shaped curve of Figure 3-4a to become narrower and taller. The probability of three-sigma peaks becomes less and the probability at $\sigma = 0$ becomes greater. Such a shift in the amplitude probability distribution sometimes occurs when a structure approaches its elastic limit. The amplitude distribution of the vibration changes from Gaussian to one which has positive kurtosis distortion. Negative kurtosis is the opposite; the probability of high amplitude peaks becomes greater, and of sigma equaling zero becomes less. Sometimes random noise generators have a small amount of negative kurtosis when the magnetic cap on the noise generating gas tube is improperly adjusted. Kurtosis is sometimes called "excess".

If the center of gravity is shifted away from the Y axis making the bell-shaped curve of Figure 3-4a lopsided, we call the condition skew. Skew may be either positive or negative. It often occurs when the structural vibration encounters more resistance in the compression mode than it does in the tensile mode, or vice versa.

To understand the measures of skew and kurtosis, we shall set down some rules and definitions. When we know the mean value of a random variable, we often need to know how widely the values of the variable are spread on either side of the mean. This is called dispersion and is indicated by the variance, σ^2. When $\xi(x)$ is such a variable with a mean value of m and a variance of σ^2, we may wish to consider a standardized variable

$$\frac{\xi - m}{\sigma} \qquad (3\text{-}2)$$

which is the deviation of ξ from its mean, m, expressed in terms of the standard deviation, σ, (positive square root of the variance). The distribution

function of this standardized variable, when the number of events comprising ξ is large, is approximately equal to a Gaussian distribution which we will call $\Phi(z)$. When $\Phi(z)$ is differentiated, giving $\phi(z)$, the function can be expanded in a series due to Edgeworth.

$$f(x) = \phi(x) - \frac{1}{3!}\frac{\mu_3}{\sigma^3}\phi^{(3)}(x) + \frac{1}{4!}\left(\frac{\mu_4}{\sigma^4} - 3\right)\phi^{(4)}(x) + \cdots \quad (3\text{-}3)$$

In the abbreviated form of Equation 3-3, we show only the first terms of which the second is

$$-\frac{1}{3!}\frac{\mu_3}{\sigma^3}\phi^{(3)}(x) \quad (3\text{-}4)$$

and the third is

$$+\frac{1}{4!}\left(\frac{\mu_4}{\sigma^4} - 3\right)\phi^{(4)}(x) \quad (3\text{-}5)$$

The second term is skewness and can be seen to be proportional to μ_3 which is the third moment of the mean. When μ_3 is zero, skewness is zero. The third is kurtosis and is seen to be proportional to the fourth moment of the mean. $\phi^{(3)}(x)$ and $\phi^{(4)}(x)$ are the third and fourth derivatives of the Gaussian mathematical expression. As can be seen from the expression for the third term, when $\frac{\mu_4}{\sigma^4}$ is greater than 3, the term is positive and representative of positive kurtosis; the opposite is true when $\frac{\mu_4}{\sigma^4}$ is smaller than 3. Plots of skew and kurtosis versus the instantaneous amplitude may be made indicating the departure of the amplitude distribution of a Gaussian function for all values of x.

Section 3

PROPERTIES OF RANDOM WAVES

If we pass a random wave through an extremely narrowband filter, does it become periodic? Why?

What are the conditions for a weakly stationary random process?

Define ergodicity.

What is meant by a Gaussian amplitude distribution?

What is the phase distribution of a Gaussian amplitude distributed random wave?

What is skew?

What is kurtosis?

FUNCTIONS OF RANDOM WAVES

Averages

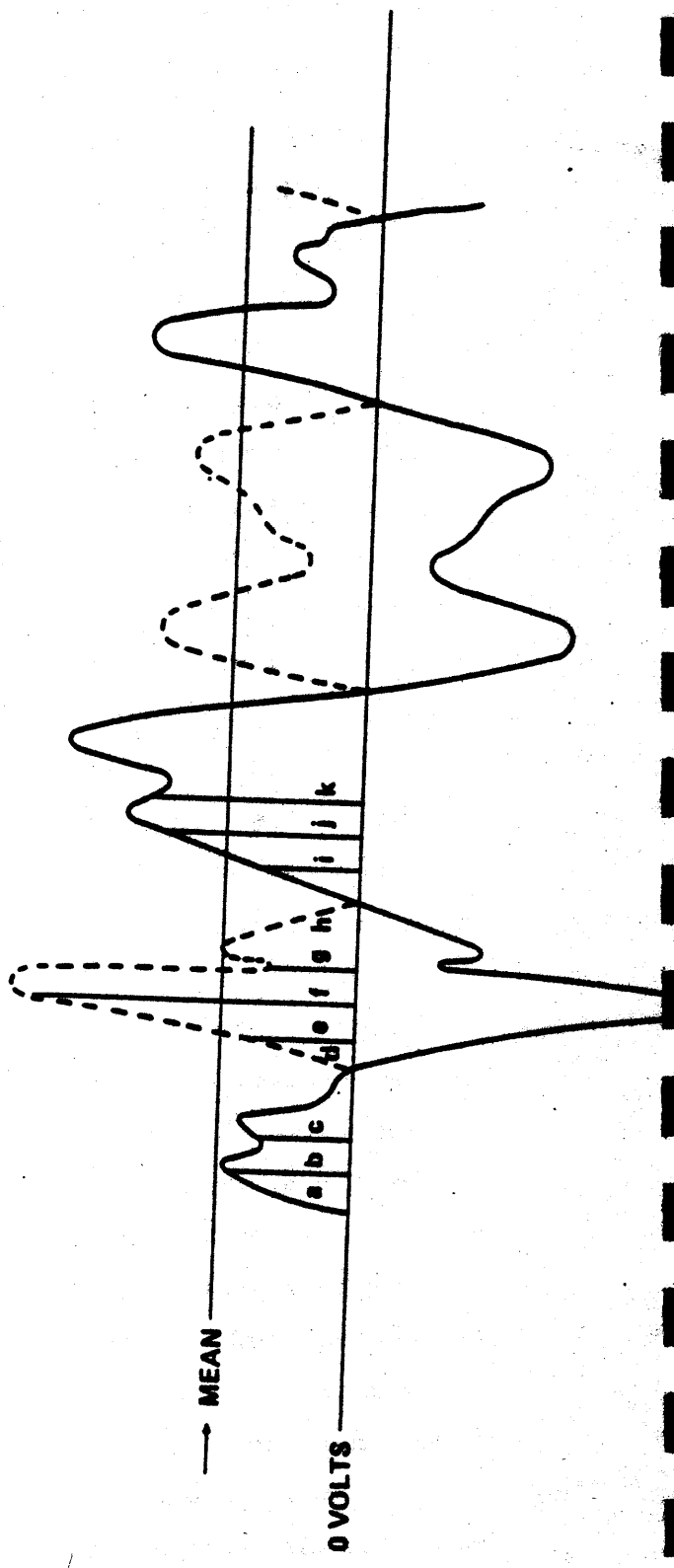
The definition of a random wave is intuitive. A random wave never repeats itself; for example, an identical set of amplitudes will never occur again. Obviously, it would be impossible to write an equation describing such a wave as we can do in the case of periodic and other analytic waves. Random waves are sometimes called stochastic, coming from a Greek root word meaning "to guess".

Since we cannot write an analytic expression for a random wave, we resort to describing the wave characteristics in terms of "the law of averages". In fact, one way in which a random wave can be described is by its average or mean value. The instantaneous values of a random wave will fluctuate about the average value. This is called dispersion. To measure the average, we can use the common full-wave rectifier type voltmeter. Such a measurement is useful when establishing stationarity.

While the mean value of a random wave is analogous to its "center of gravity" and is called the first moment of the mean, the second moment of the mean is also a valuable statistic. Higher moments of the mean exist mathematically but are not in common use. The second moment of the mean is the "mean square". Voltmeters are available to measure this function and are sometimes called "true rms meters" since some of them measure rms as well as mean square. An illustration might be helpful. Figure 4-1 shows a random wave. By employing a full-wave electronic rectifier circuit, the bottom half or negative portion of the wave is folded up into the top or positive half as shown by the dotted lines. Each interval along the X axis, as shown by the perpendiculars (a, b, c, etc.) is measured, added to the others and divided by the total number of perpendiculars. In analog meters, the number of perpendiculars is infinite and the averaging of the ordinate is done by an RC circuit. Accordingly, the process is continuous instead of discrete as is shown in Figure 4-1. In both continuous and discrete examples, however, the average or mean is computed. In the example, the mean might be about as shown. We can think of it as a

The nth moment is:

It is intere
correlation function
4-5 is the variance.
density" is introduc
the filter in the ana
Lastly, the area un
there is no dc in f (
functions have high
are important math
made their computa



horizontal axis about which the full-wave rectified signal can freely rotate; in other words, the linear center of gravity or first moment of the mean.

The second moment of the mean, or mean square, is calculated as follows: Each of the perpendiculars is squared before being added and averaged. This results in a different statistic from the mean because the higher amplitudes after being squared give more weight to the peaks; in other words, the dispersion is greater. The mean square, or second moment of the mean, is useful because it is proportional to power or energy. In mechanics, it is the "moment of inertia".

Another function often required in analysis work is the root mean square or rms. The rms is the positive square root of the mean square. From this definition, we note that, in order to compute the rms, we must first compute the mean square. This is the classical way of obtaining rms: first square, then add the squares, divide by the number of squares and finally extract the square root. Of course, in analog instruments, the process is continuous. Some meters produce a quantity which is proportional to rms provided the random signal is Gaussian. Such instruments are therefore sensitive to waveform and should be used with caution. Other instruments rectify the random signal and then pass it through a segmented diode weighting network which gives the proper weight to the instantaneous values as they change. Such instruments produce the mean square voltage and the square rooting is accomplished by a meter scale calibration. One instrument computes the logarithm of the average value of the random wave and divides the result by two to obtain a square root. The result is the square root of the log of the average, not rms. Still another very acceptable way to compute rms is by means of a thermocouple or other such thermal device. In such rms meters, the output current is proportional to heat, which is energy. We repeat an important previous statement: in most practical cases, the rms and the standard deviation, σ, are equivalent; the mean square and the variance, σ^2, are likewise equivalent.

Autocorrelation

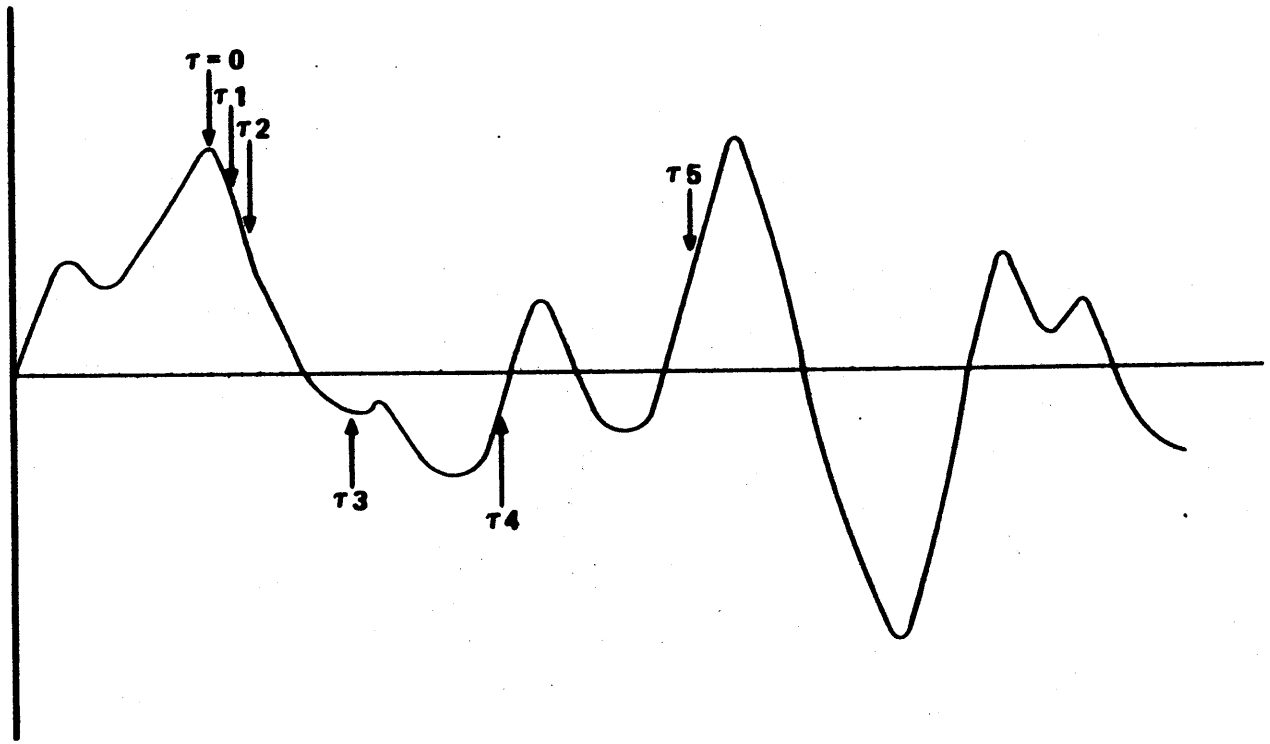
Another function in the time domain which is often used to characterize a random wave is the autocorrelation function. This function describes the dependence of the instantaneous values at one time with the instantaneous values at another time. To illustrate this, refer to Figure 4-2. We observe that the instantaneous amplitude at $\tau = 0$, will obviously not have much influence on the amplitude, τ_5, at a later time. τ_5 is said to be uncorrelated with τ_0. The same can be said of τ_3 and τ_4. However, we can say that there appears to be a degree of dependence of τ_2 on τ_0 and still more of τ_1 on τ_0. τ_1 and τ_2 are then said to be correlated with τ_0; τ_1 more so than τ_2. It is evident in our example that the amount of correlation dies out quickly as time increases-- a characteristic of wide band random waves. A correlogram is shown in Figure 4-3 which plots the conditions shown in Figure 4-2.

We see that at τ_0 the wave is completely correlated with itself and the correlation is normalized to unity at this point. Also, negative values are shown along the X axis. As shown in Figure 4-4, this wide band random noise situation becomes somewhat different in the case of narrow band random noise in that the narrow band noise correlation persists over a greater time interval. Figure 4-4 should be compared with Figure 4-3. In passing, it may be of interest to note that the autocorrelation function of a sinewave is a cosine wave.

The autocorrelation function can be computed graphically. It is a very laborious process but a simple example will be given to illustrate the method. The mathematical expression for autocorrelation is given by Equation 4-1.

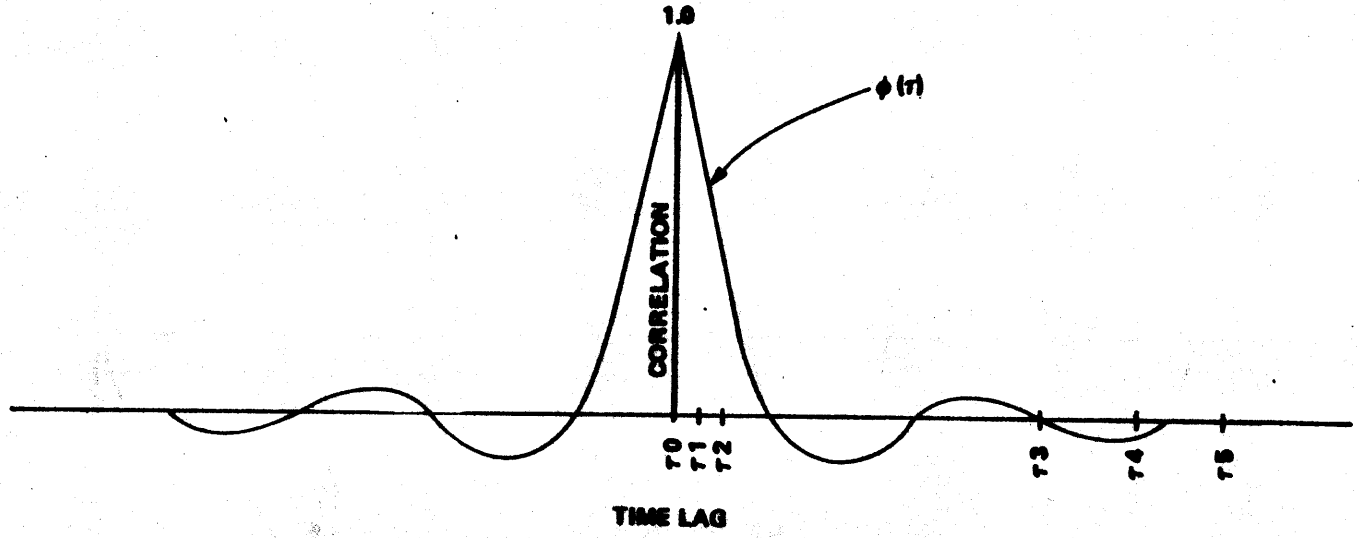
$$\phi(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{0}^{t} x(t)\, x(t+\tau)\, dt \qquad (4\text{-}1)$$

This says that the function x (t) is multiplied by itself lagged by a time, τ; that is x (t) times x (t + τ). This is integrated over the period of time, T, and divided by T and the function, ϕ (T), approaches exactness as the time of observation approaches infinity, $T \to \infty$. Equation 4-1 has to be computed over the time, T, for each value of τ which results, as we shall see, in a tremendous amount of labor. Note that the mean power in Equation 4-1 is obtained when τ is set to zero.



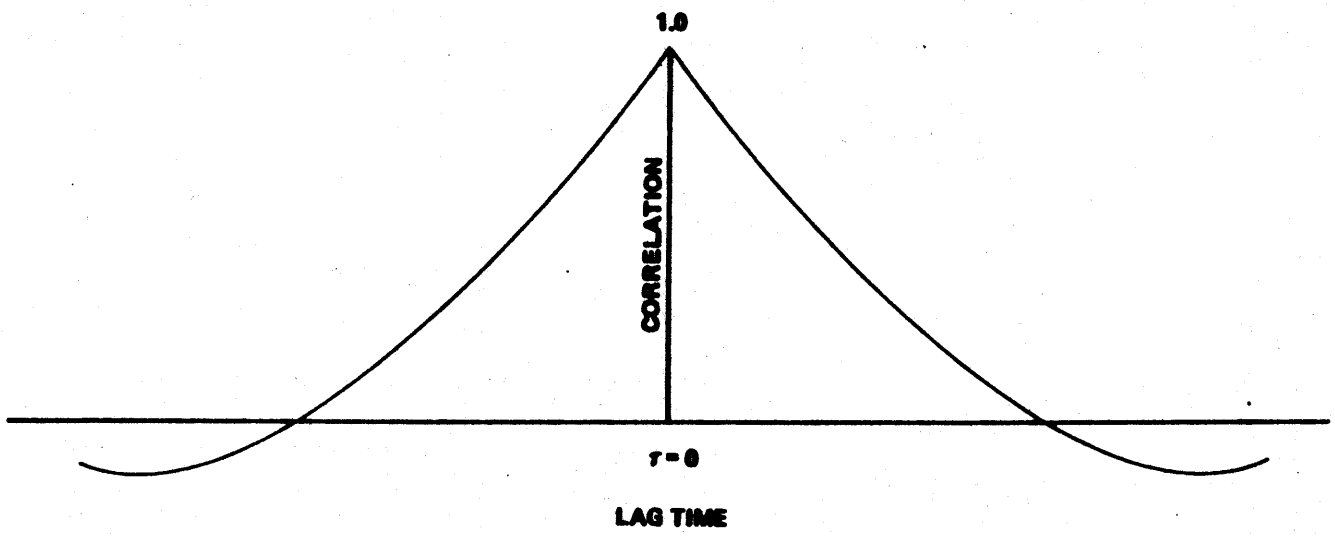
AUTOCORRELATION OF A RANDOM VALUE

FIGURE 4-2



AUTOCORRELATION FUNCTIONS OF WIDE BAND RANDOM NOISE

FIGURE 4-3



AUTOCORRELATION FUNCTION OF NARROW BAND RANDOM NOISE

FIGURE 4-4

Figure 4-5a shows a wave extending from -T to T. The same wave is shown lagging by an amount, τ, and extending from $(-T - \tau)$ to $T - \tau$. The first plot is x (t) and the lagged plot is x $(t - \tau)$. At each point on the time axis, t, the function, x (t), is multiplied by x $(t - \tau)$ and the result plotted in Figure 4-5b. Enough points must be taken on the time axis so that the wave is clearly defined. After all multiplications are complete and plotted in Figure 4-5b, the graph must be integrated or averaged. When this is done, a single point on the autocorrelation curve is established, as shown for $\tau = 1$ in Figure 4-6. A large number of such points must be calculated and plotted to produce an adequate correlation function.

Next, the lagged curve is "slid over" by an amount equal to τ, now making the lag 2τ. See Figure 4-5a. The above process is repeated for 2τ resulting in the next point on the correlation curve shown in Figure 4-6. It is evident that, for accurate results, when T becomes large and there are a large number of lags, the graphical process is very laborious.

Refer to Figure 4-7. Here we see a typical wide band random shot noise in 4-7 a. This is in the time domain. In 4-7 b we see its autocorrelation function. This shows that the autocorrelation function is both positive and negative, although often the absolute value is plotted. This is also in the time domain. If a sinusoid is combined with the shot noise and an autocorrelation function then produced, a function such as shown in 4-7 c results. As time passes, the correlation of the random shot noise approaches zero but a periodic signal attributable to the sinusoid remains. This is a cosine wave which is the autocorrelation function of a sine wave and it remains as long as the sinusoid does.

When the Fourier transform of the shot noise function is taken, the power spectrum results. This is shown in Figure 4-7 d. Note that negative frequencies are implied. Such frequencies do not really exist but are the result of expressing the Fourier transform in exponential terms instead of trigonometric terms. Easier mathematical handling results but we must be careful to take this into account when mathematical and experimental results are compared. In the mathematical case, we have a minor image of the positive side reproduced in the

negative frequency regions. In the experimental case, no negative frequencies exist and the results will be as though the negative frequencies are folded over to the positive side along the zero frequency ordinate. This results in doubling the amplitude of the mathematical model at each point on the positive abscissa.

Power Spectral Density

Loosely we can define power spectral density as the distribution of mean power with frequency. The same information that the autocorrelation function in the time domain contains is also possessed by the power spectral density function in the frequency domain. The autocorrelation function and the power spectral density function are Fourier transforms of each other and, although they contain the same information, one is a time history and the other is a spectrum. This remarkable mathematical relationship was discovered by Professor Norbert Wiener of MIT near the beginning of World War II. This far reaching development was a landmark in the theory of statistical communication since it wedded the disciplines of statistics and communication. It is called the Wiener-Khintchine theorem. Because the central idea in the statistical theory of communication is that messages and noise are both considered random phenomena, control problems are seen in a better light, statistical filtering can be formulated and the theory of generalized harmonic analysis has developed.

GRAPHICAL SOLUTION OF AUTO CORRELATION

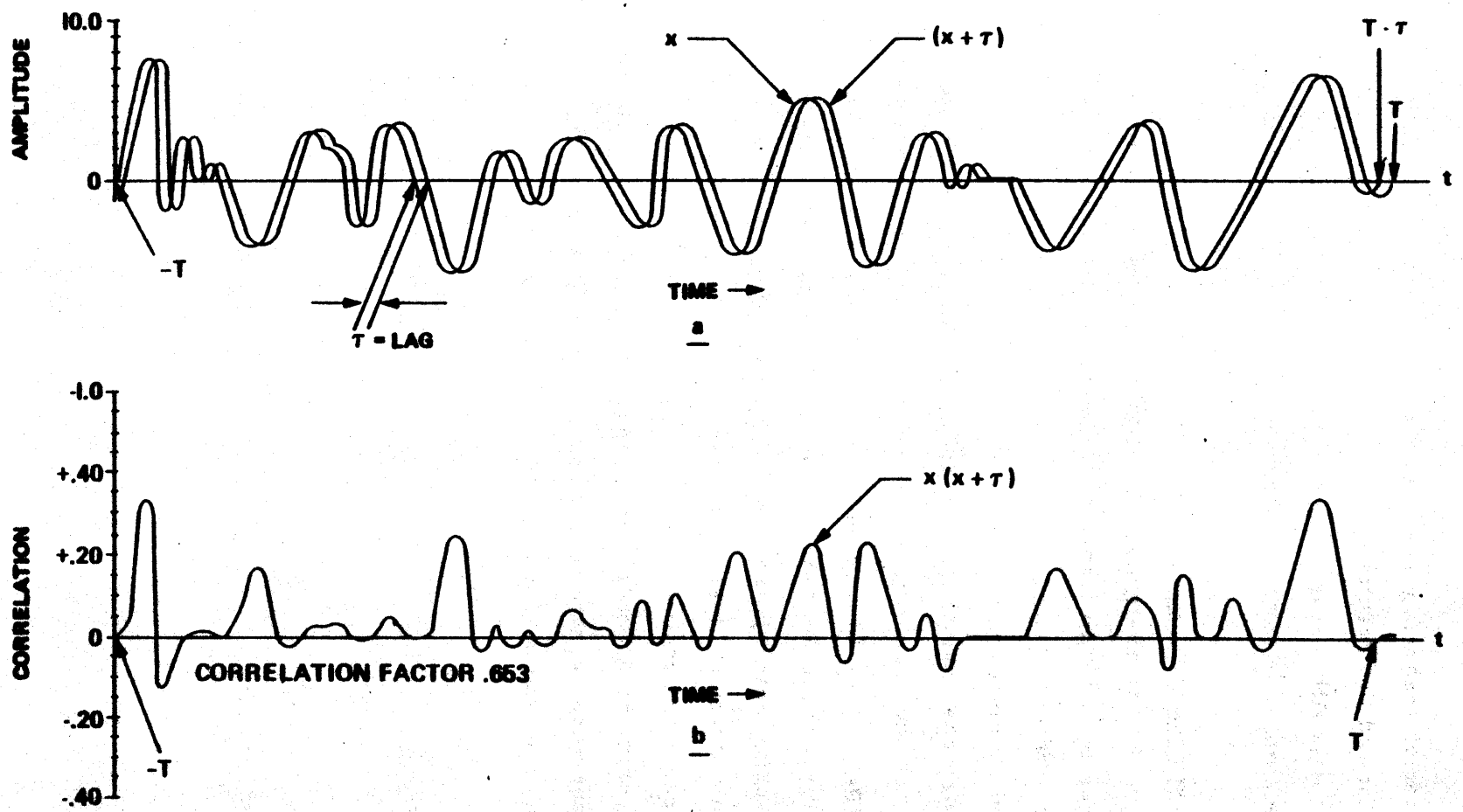
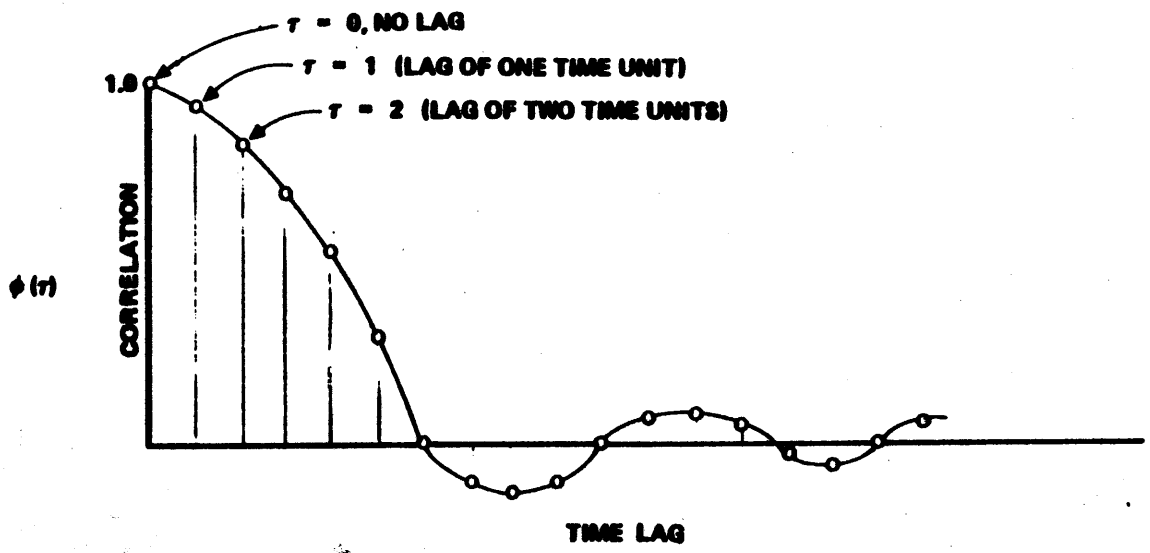


FIGURE 4-5



AUTOCORRELATION CURVE OBTAINED BY GRAPHICAL METHOD

FIGURE 4-6

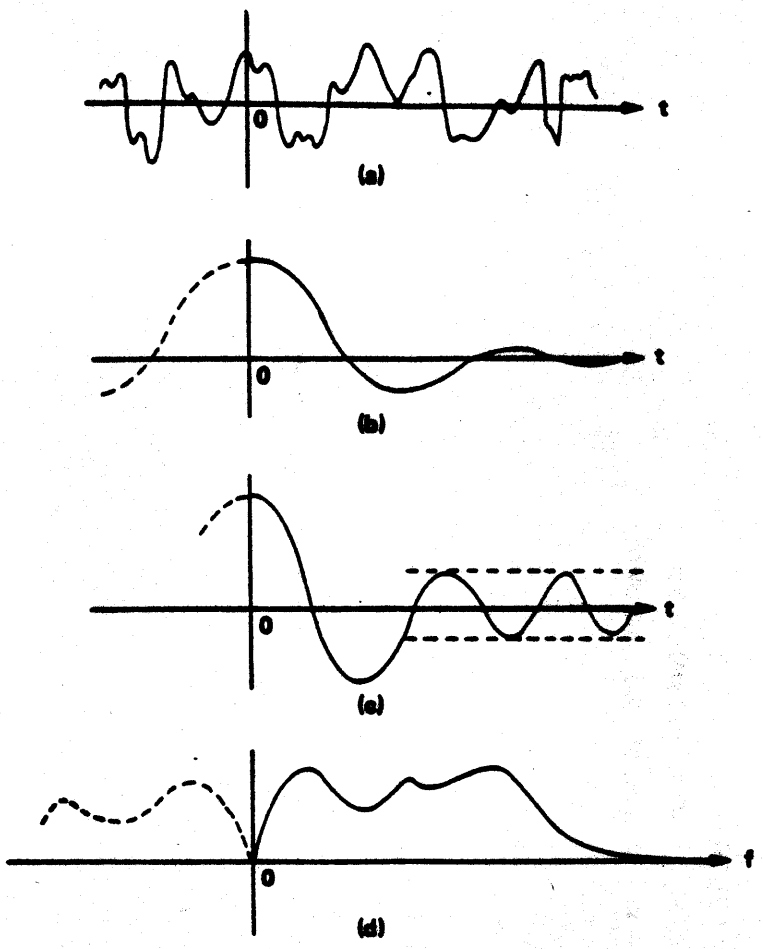


FIGURE 4-7

(a) TYPICAL WIDE-BAND SHOT NOISE
(b) ITS AUTO CORRELATION FUNCTION
(c) ITS AUTO CORRELATION FUNCTION PLUS A SINUSOID
(d) ITS POWER SPECTRUM

We will not attempt, as some writers do, to define power spectral density in terms of the autocorrelation function; it will be defined on its own merits. Since in mechanical systems the vibration of a structure concerns us, we shall begin from that point of view. When a mechanical structure is excited into vibration, the frequencies are, in general contained in the exciting force and the natural frequencies of resonance of the structure. The purpose of power spectral density is to plot these vibrations in terms of the mean of the square of the amplitude against frequency. The reason the square of amplitude is used instead of the linear amplitude is that we wish to know the frequency distribution of the <u>power</u> or intensity dissipated or real part of the vibration, and we are not concerned with the reactive component of the vibration. We also require the <u>differential</u> power rather than the power at discrete frequencies because there are no discrete frequencies in a random process as there are in the case of periodic waves. Hence, we think not of the power at a certain specific frequency, but of the power in a <u>band</u> of frequencies. To obtain the differential of power, we reduce the width of this band until it approaches zero. This is standard differential calculus practice. However, in real life, we cannot do this. A practical filter must have a finite bandwidth, even if it is as narrow as 1 to 2 Hz. We define power spectral density as "the limiting mean square acceleration (velocity, displacement, etc.) per unit of bandwidth, divided by the bandwidth as the bandwidth approaches zero". The expression for power spectral density becomes exact as the period of observation approaches infinity. We will later see what the effect is when the bandwidth of the filter is not infinitesimally narrow and when the period of observation is relatively short.

As stated previously, the autocorrelation function and power spectral density are Fourier transforms of one another. The transformation is:

$$\Psi(f) = \int_{-\infty}^{\infty} \phi(\tau)\, \epsilon^{-j2\pi f\tau}\, d\tau \qquad (4\text{-}2)$$

where $\Psi(f)$ is the power spectral density function and $\phi(\tau)$ is the autocorrelation function. The inverse transform is:

$$\phi(\tau) = \int_{-\infty}^{\infty} \Psi(f)\, \epsilon^{j2\pi f\tau}\, df \qquad (4\text{-}3)$$

In digital computation techniques, the classical approach to power spectral density is to first compute the autocorrelation function, Equation 4-1, and then take the Fourier transform to obtain PSD from Equation 4-2. In analog computation practice, it is usual to compute PSD directly. When the proper procedures are used, both analog and digital methods are equivalent. Figure 4-8 shows the PSD curve of a random wave exciting a filter.

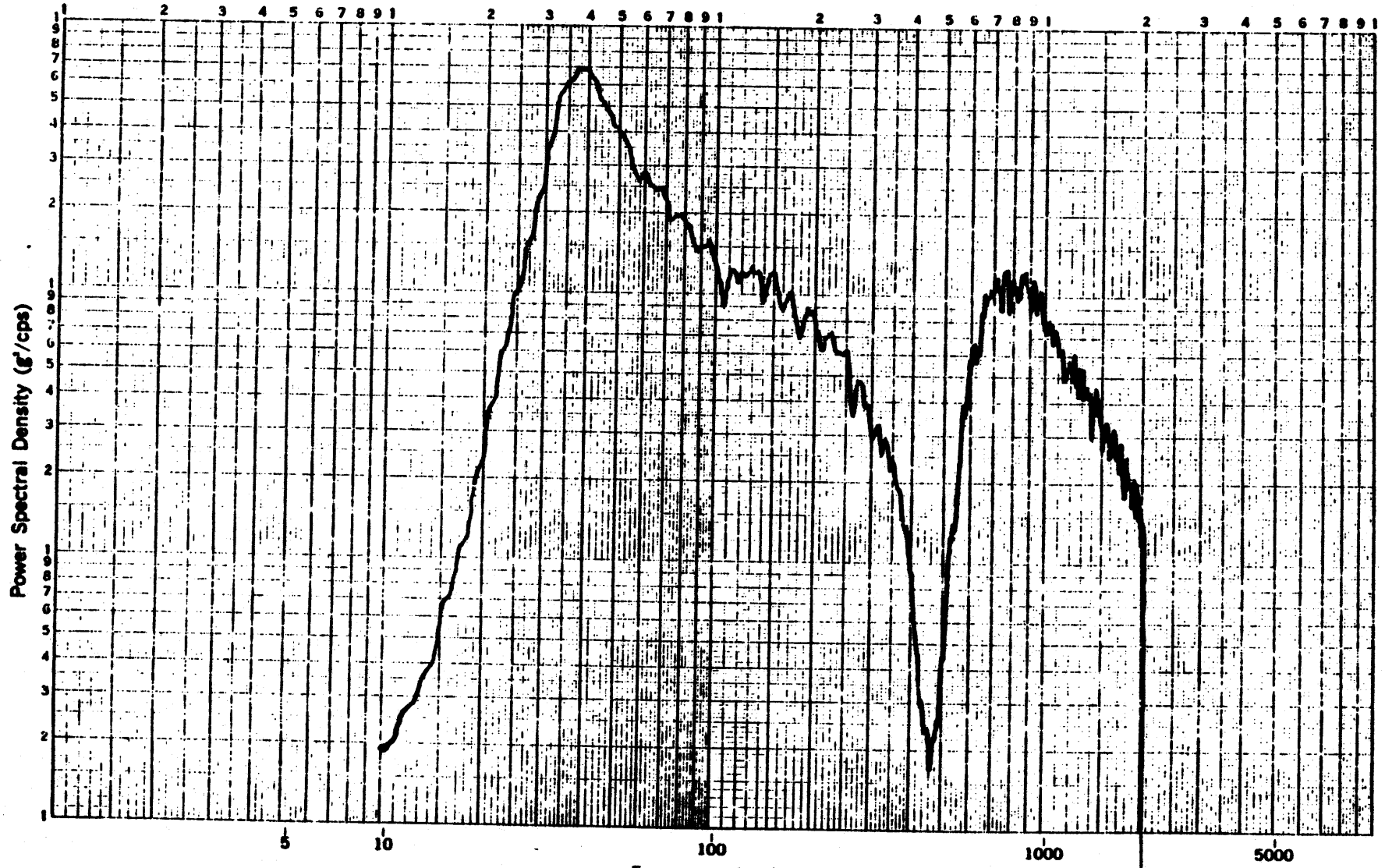
Cross Functions

In autocorrelation and power spectral density, we have dealt with single valued functions in which we have mathematically operated on the function itself. Similar functions are possible when two or more variables are involved. These are called the cross-correlation functions and cross spectral density functions. The usefulness of these functions arises, for example, when it is desired to see what dependence the vibration on one part of a structure has on the vibration of another part of the same structure. For instance, we may want to see how much of the vibration of an engine is affecting a gyro mount. The normalized cross-correlation function will give us the percentage amount of vibration on the gyro mount and also the time of propagation from the engine to the mount. The frequency distribution of that amount of vibration getting to the gyro mount will be derived from the cross spectral density function. From both these data, remedial measures can usually be determined and taken if necessary. As would be expected, the cross-correlation function and the cross spectral density functions are related by Fourier transforms similar to those in Equations 4-2 and 4-3.

There are some differences between the auto- and the cross-correlation functions. The amplitude of the autocorrelation function can never be greater than it is at $\tau = 0$. In fact, at $\tau = 0$ the value of the autocorrelation function is equal to the mean square of the process being analyzed. In the case of the cross-correlation function, the maximum amplitude can occur anywhere but ordinarily never reaches a value as great as the autocorrelation of either parent function, which it cannot exceed. Calibration of the autocorrelation function is rather simple. We set in the lag value to $\tau = 0$ and adjust the

J17 BRACKET RESPONSE

FIGURE 4-7



Job No. 6-249 Item 06 WARHEAD Serial No. X-23 Date 6/26/64 Time

Axis & Condition Pickup S/N & Location 1765-J17 Pickup Sensitivity mv rms

Overall mv rms = g rms Analyzer Calibration g²/cps Sweep Speed g rms Decades/min. 0.4 cps/sec

Analyzer Filter cps BW Multiplier Range Cap. uf Operator

S-019

amplitude control for a normalized value of unity. We cannot do this with the cross-correlation function so usually we correlate one of the two functions with itself at $\tau = 0$ and use this normalized unity value for calibration purposes.

The cross spectral density is a complex variable having both phase and amplitude, whereas the power spectral density has no phase information since power has no phase. Cross spectral density can be plotted in several ways. The usual presentation is in an orthogonal rectilinear coordinate system displaying the real part of the cross spectrum plotted against frequency on one graph and the imaginary part of the cross spectrum plotted against frequency on the other graph. The first is called the coincident spectrum or co-spectrum; the second is called the quadrature spectrum or quad-spectrum. The principal advantage of this method of display is that it is the natural computational process to follow in analog computing. In addition, coordinates are not crowded and resonances are clearly shown. Also used are the Nyquist plot and the Bode plot. The Nyquist plot clearly shows regions of conditional stability but it has a polar coordinate form from which it is somewhat difficult to tell the exact frequency since it is sometimes crowded on the scale. The Bode plot shows magnitude on one graph and phase on the other. This is generally the form in which mechanical engineers like to see their cross spectral density data but it has the drawback of requiring further analog computations, thus diminishing the accuracy.

Higher Moments

All the above described functions have higher moments which, in general, are not used in practice. To illustrate the idea of moments, the first moment of the mean for random processes is:

$$\mu_1 = \lim_{T \to \infty} \frac{1}{T} \int_0^T f(t)\, dt \qquad (4\text{-}4)$$

The second moment of the mean or mean square is:

$$\mu_2 = \lim_{T \to \infty} \frac{1}{T} \int_0^{T_2} f^2(t)\, dt \qquad (4\text{-}5)$$

The nth moment is:

$$\mu_n = \lim_{T \to \infty} \frac{1}{T} \int_0^T f^n(t)\, dt \qquad (4\text{-}6)$$

It is interesting to note that Equation 4-5 is the value of the auto-correlation function when $\tau = 0$. When f (t) has no d-c component, Equation 4-5 is the variance. When the idea of a differential frequency or "spectral density" is introduced into Equation 4-5 and we divide by the bandwidth of the filter in the analyzer, we have the power spectral density function. Lastly, the area under a power spectral density curve is Equation 4-5 when there is no dc in f (t). The correlation functions and spectral density functions have higher moments, as well as does the mean. Although they are important mathematically, the state of the practicing art has not yet made their computation necessary.

Section 4

FUNCTIONS OF RANDOM WAVES

What is the difference between RMS and the standard deviation?

Define autocorrelation.

Define power spectral density.

What is the amplitude of an autocorrelation function at tau equals zero?

What does a cross spectral density plot tell us?

Define the first moment of the mean.

What is the difference between the first moment of the mean and its higher moments?

SOME STATISTICS

Estimators

Our insight tells us that certain functions such as autocorrelation and power spectral density become exact when the number of observations approaches infinity. In real life, there is no such thing as an infinite number of observations so we become interested in how closely we can approach exactitude when we have only a comparatively few observations. We would hope to attain a reasonable estimate so we give attention to the quantity used to obtain this reasonable estimate, which is called an "estimator". Some estimators are good and some are better. Three principal factors influence the "goodness" of an estimator.

1. The estimators should have, if possible, an "expected value" equal to that of the parameter being established. The "expected value" is the mean or average value. That is to say, the estimator should be equal to the mean value of the parameter under investigation. If this is true, the estimator is said to be "unbiased".

2. It is desirable that the mean square error of the estimator be smaller than that for other possible estimators. That is

$$\left(\hat{\phi}_a - \phi\right)^2 \leq \left(\hat{\phi}_b - \phi\right)^2 \qquad (5\text{-}1)$$

where $\hat{\phi}_a$ is the estimator (designated by the little hat $\hat{}$) under consideration, $\hat{\phi}_b$ is any other possible estimator and ϕ is the true value of the parameter we seek to estimate. If this condition is true, the estimator is called "efficient".

3. It is desirable that the estimator approach the parameter being estimated with a probability of unity when the sample size becomes large. If this is true, the estimator is said to be "consistent".

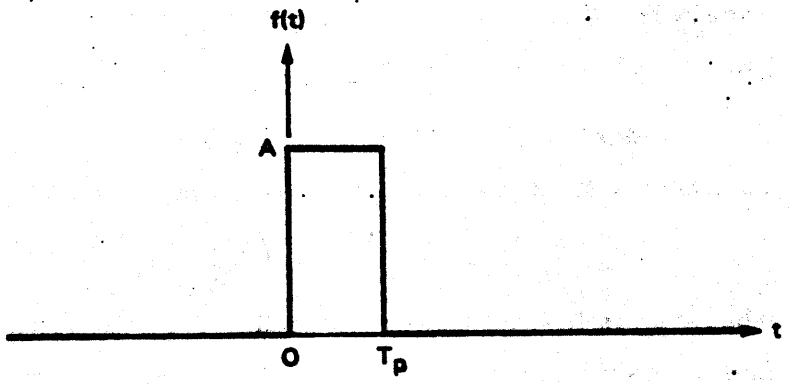
For the functions of interest to us, competent mathematical statisticians have investigated the goodness of certain estimators and we are accordingly advised to use the Chi-Square probability distribution and the normalized

standard error function. These will be discussed in some detail after a little sampling theory which is desirable to our understanding of the derivation of these distributions.

Sampling Theory

If a continuous wave is sampled enough times per second, it becomes obvious that every detail of the wave shape can be reconstructed exactly. If the continuous wave is a periodic or random wave having no components higher than some frequency, f, sampling theory tells us that only two samples per wavelength of this highest frequency are necessary to completely identify the wave. This is not so obvious and requires further explanation.

To synthesize a wave, we employ a rectangular sampling pulse. This pulse is in the time domain and must be transferred into the frequency domain. This is done by a Fourier transformation.



A SINGLE SAMPLING PULSE OF HEIGHT A AND DURATION T_p

FIGURE 5-1.

The pulse can be described as:

$$f(t) = \begin{cases} A \text{ from } 0 \text{ to } T_p \\ 0 \text{ everywhere else} \end{cases}$$

Then

$$G(j\omega) = \int_0^{T_p} A\, e^{-j\omega t} dt \qquad (5\text{-}2)$$

integrating

$$G(j\omega) = \frac{A}{j\omega}\left(1 - e^{-j\omega T_p}\right) \tag{5-3}$$

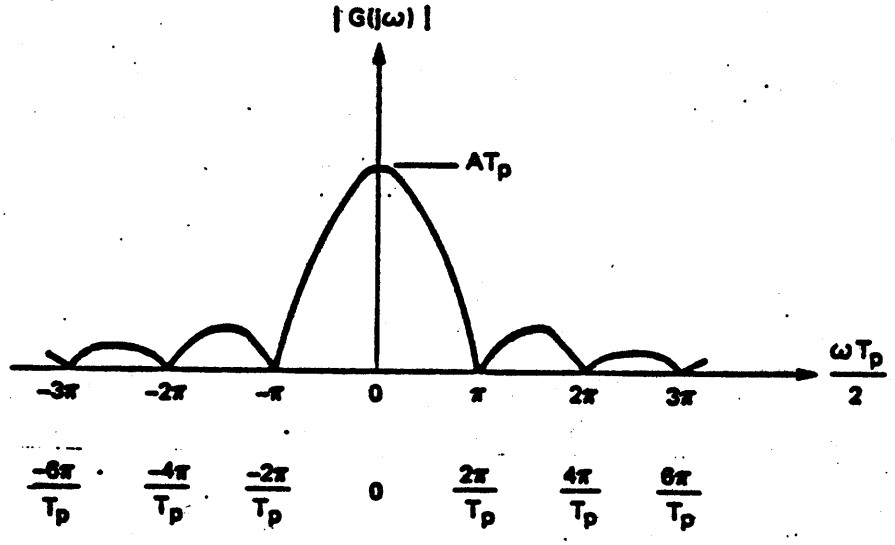
$$= \frac{2A}{\omega}\sin\left(\frac{\omega T_p}{2}\right) e^{-j\omega T_p/2} \tag{5-4}$$

The absolute amplitude of $G(j\omega)$ is

$$\left|G(j\omega)\right| = \frac{2A}{\omega}\left|\sin\left(\frac{\omega T_p}{2}\right)\right| \tag{5-5}$$

$$= AT_p\left|\frac{\sin(\omega T_p/2)}{\omega T_p/2}\right| \tag{5-6}$$

Here $\omega T_p \neq 2\pi$ because T_p is not the period of a periodic function and ω can have any value. $|G(\omega)|$ is of the form $|(\sin x)/x|$ and is plotted in Figure (5-2).



Equation 5-6 is sometimes called a "sinc" function. It can be readily calculated but also exists in tables. It has the form shown in Figure 5-2. Consequently, for <u>each</u> rectangular sampling pulse occuring in the time domain, we get <u>one</u> sinc function in the frequency domain. If these functions are combined or added in accordance

with the proper time and with an amplitude equal to that of f (t), the wave is completely synthesized. Figure 5-3 shows a sine wave constructed from three equally spaced sinc functions.

We have thus shown that a wave can be completely synthesized by two perfectly chosen data points per cycle of the highest frequency contained in the wave. The data points, of course, must be evenly spaced. Consequently, if we employ a filter having a bandwidth of 10 Hz and then sample its output for 10 seconds, we have:

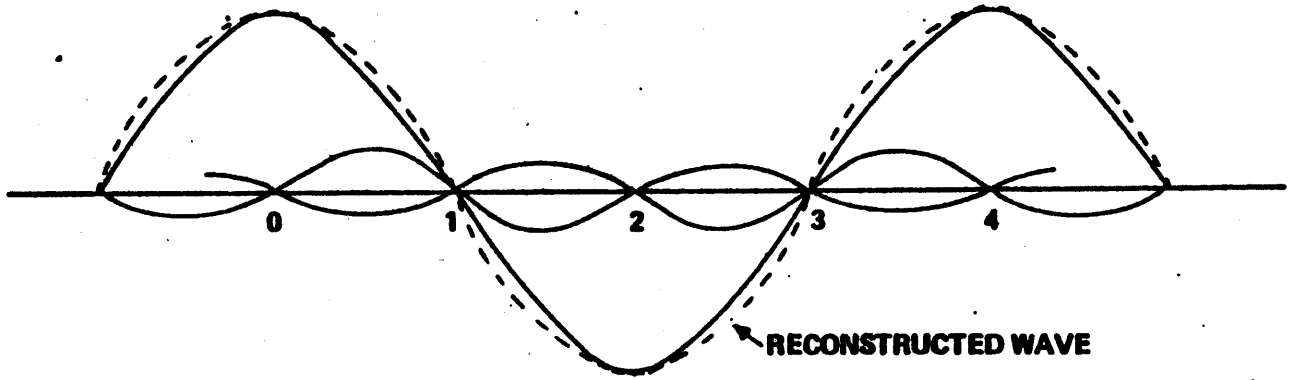
$$2BT = 2 \times 10 \text{ Hz} \times 10 \text{ sec} = 200 \text{ degrees of freedom} \qquad (5\text{-}7)$$

Before we further consider how we will use the number of degrees of freedom in a Chi-Square estimator, we will consider implications of the sampling theorem a little further.

Aliasing

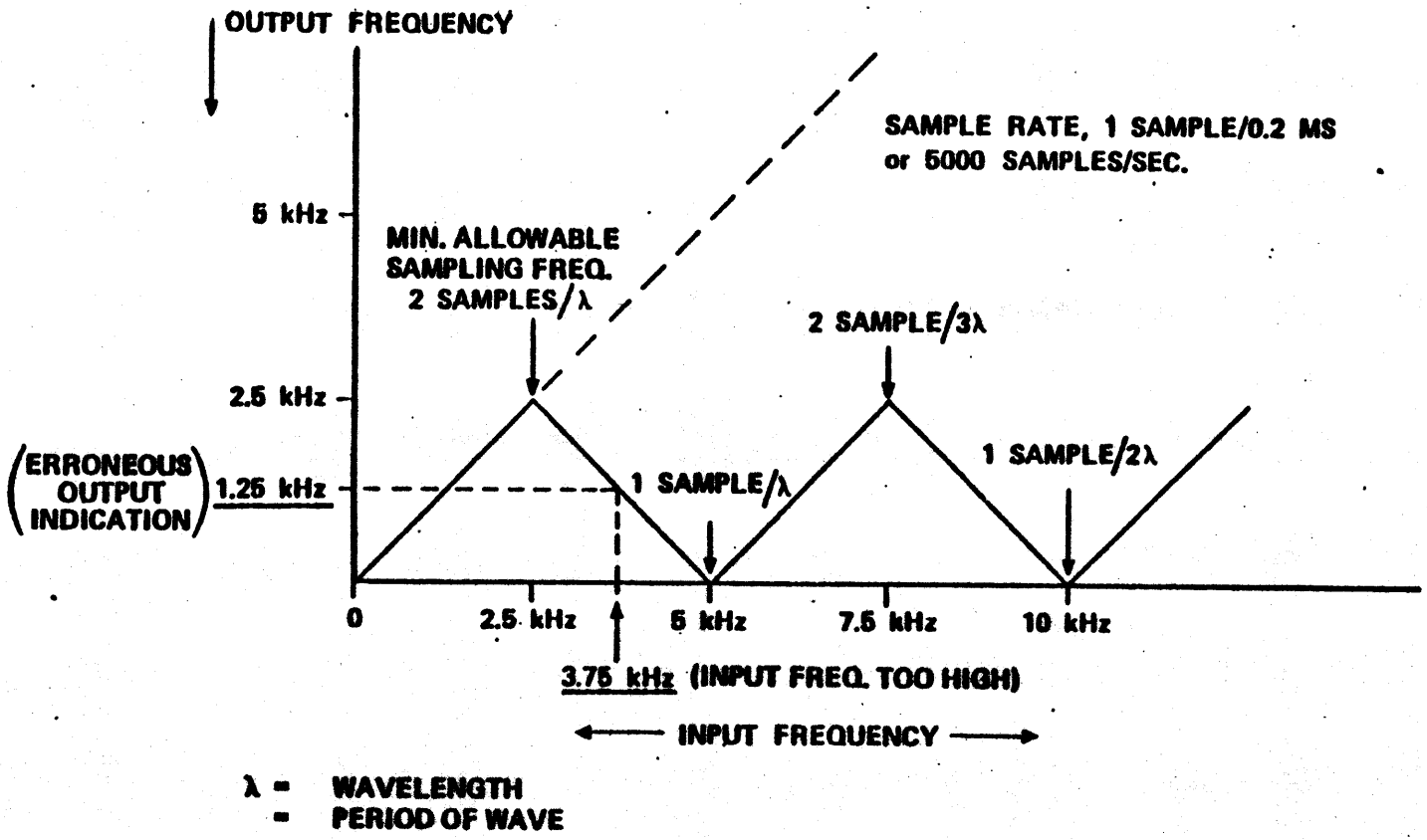
What happens when a signal is sampled with fewer than two points per wavelength? The effect is somewhat similar to that observed if a wheel rotating at a high speed is seen under stroboscopic light flashing at a different rate than the wheel rotates. The wheel will appear to be rotating at a different speed than its true rate. This phenomenon is often seen in motion pictures. If the wheel is illuminated once in each revolution (or any integral multiple thereof), it will appear as though standing still. If the wheel is illuminated at least twice every revolution, it will appear to rotate at its correct rate. There are, of course, all the conditions in between. This phenomenon, whereby the wheel appears to be rotating at an incorrect rate, is called frequency folding or aliasing. Refer to Figure 5-4. Figure 5-4 shows an input frequency of 3.75 kHz on the X axis. which is higher than the sampling rate can reconstruct. At the output, the 3.75 kHz input wave appears as if it were 1.25 kHz.

As this consequence of the sampling theory, we provide low pass filters in sampling circuitry which cut off all frequencies above 1/2 the sampling rate. These are called anti-aliasing filters.



SYNTHESIS OF A WAVE BY SINC FUNCTIONS

FIGURE 5-3



FREQUENCY ALIASING

FIGURE 5-4

Chi-Square Distribution

When estimating the mean square of a wave function, the confidence in our measurement can be determined from a knowledge of the number of statistical degrees of freedom, n.

$$n = 2\,BT \tag{5-8}$$

where B is the bandwidth of an ideal rectangular filter and T is the integrating time of a time integrator, Equation 5-8 shows that the requirement for a Chi-Square confidence estimate is identical to the sampling theorem requirement to synthesize a wave function as given in Equation 5-7. This satisfies our sensibilities if not our mathematical curiosity. The mathematical treatment is too involved for presentation here and can be found in statistical texts.

Consider a wave with a Gaussian probability density function and a true mean square amplitude of σ^2. Assume we measure a mean square voltage of s^2. Then the Chi-Square relationship exists:

$$\frac{s^2}{\sigma^2} = \frac{\chi^2}{n} \tag{5-9}$$

Now, if the mean square amplitude, s^2, is measured for a time, T, the true mean square amplitude will lie within a confidence interval determined by:

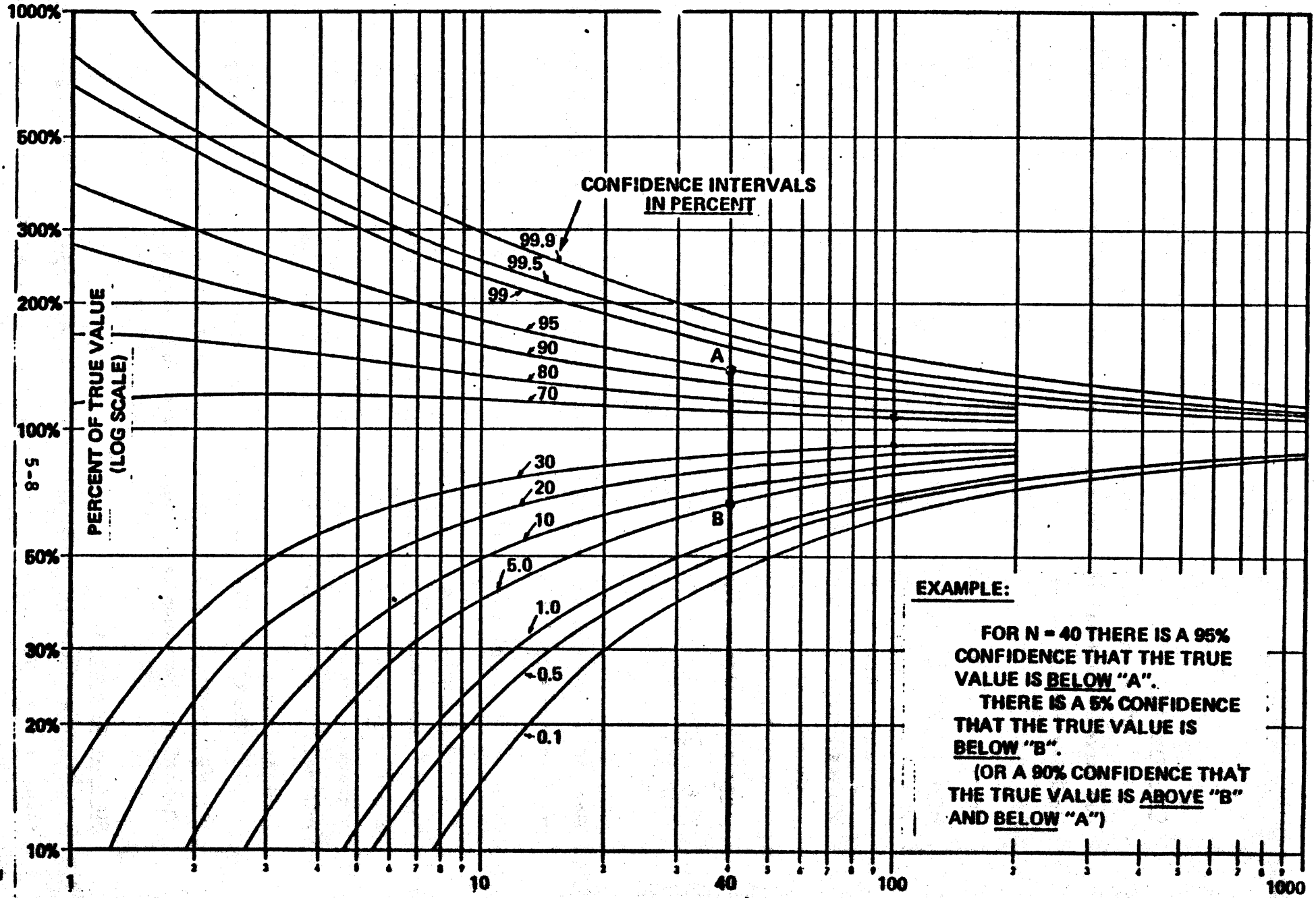
$$\sigma^2 = \left(\frac{n}{\chi^2}\right)s^2 \qquad \text{where } n = 2BT \tag{5-10}$$

Figure 5-5 shows some confidence intervals for the Chi-Square distribution and more detailed information is available in statistical texts. Note from Figure 5-5 that a confidence interval is bounded by an upper and lower error for each n degrees of freedom. Thus, for any error, chosen or assumed, there exists an interval of confidence for any n. The Chi-Square distribution is therefore two dimensional.

The Chi-Square distribution is a distribution of squared values and is therefore particularly suited to the estimation of the confidence interval of squared functions such as the mean square and power spectral density.

Standard Error Estimator

The "standard error" estimate can be developed from the expected mean square error. The estimated mean square error for a time function, x (t), can


1000%
500%
300%
200%
100%
50%
30%
20%
10%
PERCENT OF TRUE VALUE
(LOG SCALE)
CONFIDENCE INTERVALS
IN PERCENT
99.9
99.5
99
95
90
80
70
30
20
10
5.0
1.0
0.5
0.1
A
B
EXAMPLE:
FOR N = 40 THERE IS A 95% CONFIDENCE THAT THE TRUE VALUE IS BELOW "A".
THERE IS A 5% CONFIDENCE THAT THE TRUE VALUE IS BELOW "B".
(OR A 90% CONFIDENCE THAT THE TRUE VALUE IS ABOVE "B" AND BELOW "A")
1
10
40
100
1000
DEGREES OF FREEDOM N
CHI - SQUARE DISTRIBUTION

Rev 5/72

be computed by time averaging over a finite time, T. This is permissible since we assume x (t) is stationary. The estimated error is:

$$\hat{\psi}_x^2 = \frac{1}{T}\int_0^T X^2(t)\,dt \qquad (5\text{-}11)$$

This is probably not the true mean square error since we are not averaging over an infinite time. It is an <u>estimate</u> as indicated by the little hat over $\hat{\psi}^2$. We can write for the true mean square error:

$$\psi_x^2 = E\left[X^2(t)\right] \qquad (5\text{-}12)$$

$\hat{\psi}_x^2$ in Equation 5-12 has a variance

$$\mathrm{Var}\left[\hat{\psi}_x^2\right] = E\left[(\hat{\psi}_x^2 - \psi_x^2)^2\right] = E\left[\hat{\psi}_x^4 - \psi_x^4\right] \qquad (5\text{-}13)$$

It can be shown that:

$$\epsilon^2 = \frac{\mathrm{Var}\left[\hat{\psi}_x^2\right]}{\psi_x^4} \approx \frac{1}{BT} \qquad (5\text{-}14)$$

which holds when the mean is <u>zero</u>.

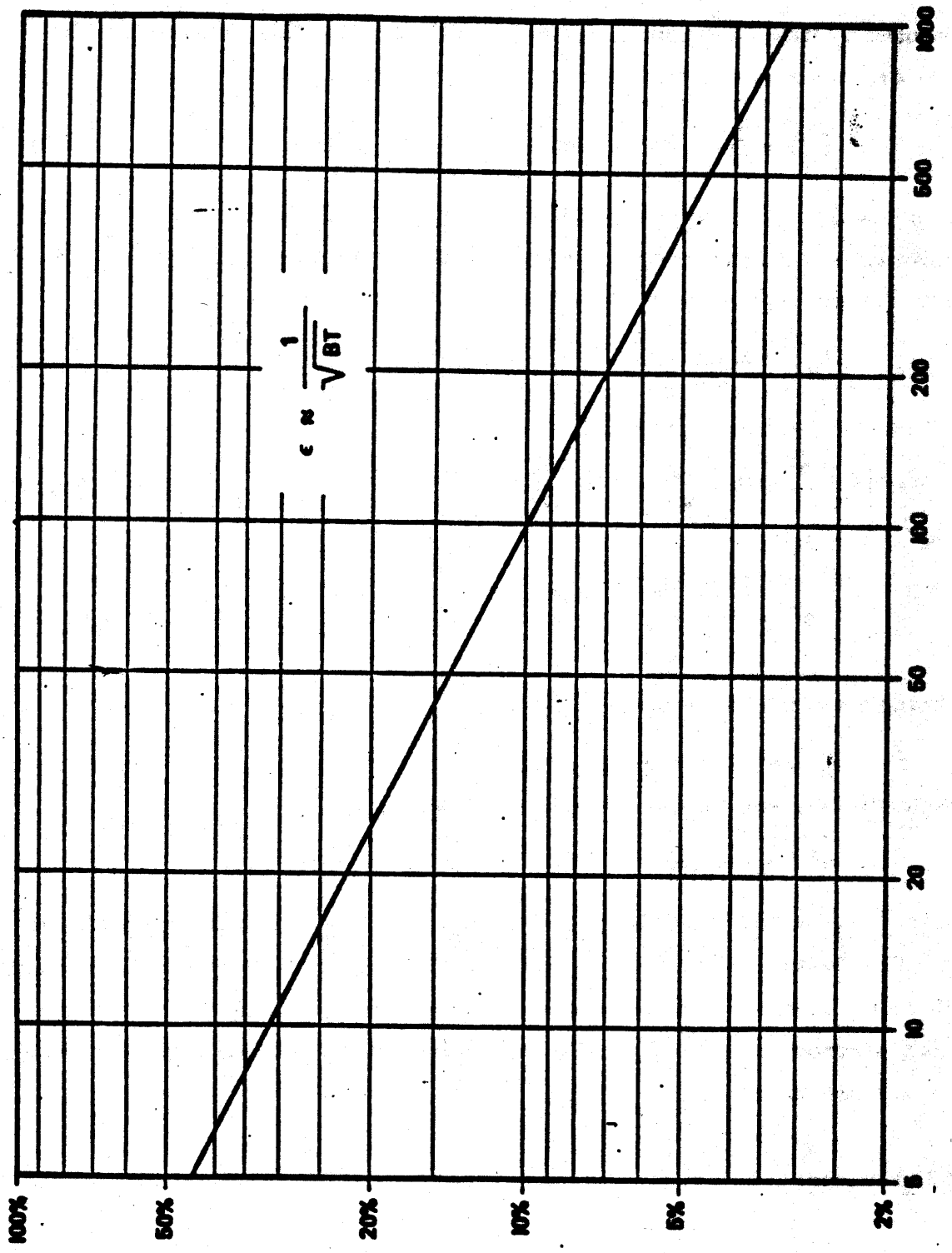
ϵ^2 is defined as the <u>normalised</u> mean square error. The square root of Equation 5-14 gives the normalised standard error.

$$\sqrt{\epsilon^2} = \epsilon \approx \frac{1}{\sqrt{BT}} \qquad (5\text{-}15)$$

The estimator, ϵ, and the Chi-Square estimates are approximately equivalent when the Chi-Square degrees of freedom are bounded by 70% on the upper bound and 30% on the lower. The confidence interval for the standard error, ϵ, is 68% that the error, $1/\sqrt{BT}$, is plus or minus the calculated value. Figure 5-6 is a plot of the estimator, ϵ.

<u>Use of Estimators</u>

The use of the Chi-Square distribution allows a more refined estimate of confidence within much smaller limits than the standard error, ϵ, does. However, one needs a Chi-Square table or chart. The value of ϵ can be computed without a table or chart being available, but has the rather large confidence interval of 68%.



ESTIMATION UNCERTAINTY VERSUS BT PRODUCT

FI .E 5-6

Although Figure 5-6 indicates BT products down to 5, the BT product should always be $BT >> 1$; $BT \geq 20$ should be achieved if possible. The same thing applies to a Chi-Square distribution. An attempt to keep n = degrees of freedom ≥ 10 should always be made. Both estimators show diminishing returns. However, when BT is very large, for example, the Chi-Square distribution in Figure 5-5 shows very little improvement in confidence when $n = 2BT$ goes from 500 to 1000. Yet, this can represent a <u>doubling</u> of the analysis time or a two-fold reduction in frequency resolution.

The standard error formula for estimating the confidence interval for amplitude probability density must be modified from that given in Equation 5-15. This arises from the fact that the confidence interval improves as the number of events per unit time increases. Assuming a Gaussian amplitude distribution, it is clearly recognized that the number of events when $\sigma = 0$ is much larger than when $\sigma = \pm 3$ (3 times rms). Accordingly, a factor must be inserted in the formula for the normalized standard error to compensate for the reduction of the number of events as σ increases. A suitable formula is:

$$\epsilon \approx \frac{0.3}{\sqrt{BTWp(x)}} \qquad (5\text{-}16)$$

where B is the bandwidth of the process being analyzed, T is the true averaging time, W is the width of the "amplitude window" stated in terms of σ, and $p(x)$ is the Gaussian amplitude probability at the amplitude in consideration. If the window makes 10 contiguous steps from $\sigma = 0$ to $\sigma = 1$, then, $W = 0.1$. At an amplitude of $\sigma = \pm 1$, $p(x)$ is 24.2% or .242. At an amplitude of $\sigma = \pm 2$, $p(x)$ is .054. The bandwidth, B, is assumed to be idealized and white. T assumes true integration for each step of the window. The value 0.3 is derived empirically from test data since no exact expression is known. The value 0.3 becomes considerably larger in digital computing techniques.

For mean and correlation measurements, the value of ϵ should be modified for Equation 5-15 to:

$$\epsilon = \frac{1}{\sqrt{2BT}} \qquad (5\text{-}17)$$

For power spectral density measurements, ϵ should become $\epsilon = 1/\sqrt{\pi BT}$ in those cases where the analyzer filter approximates a single tuned circuit rather than the idealized rectangular straight-sided type.

Section 5

SOME STATISTICS

What is an "estimator" as pertaining to statistics?

Sketch a $\frac{\sin nx}{nx}$ function.

What is the theoretical number of samples per wavelength required to reconstruct a wave?

Define aliasing.

When would one use a Chi Square distribution as an estimator?

What is the mathematical expression for the standard error estimator?

Section 5

SOME STATISTICS

What is an "estimator" as pertaining to statistics?

Sketch a $\frac{\sin nx}{nx}$ function.

What is the theoretical number of samples per wavelength required to reconstruct a wave?

Define aliasing.

When would one use a Chi Square distribution as an estimator?

What is the mathematical expression for the standard error estimator?

PROBLEMS IN ANALYZING RANDOM WAVES

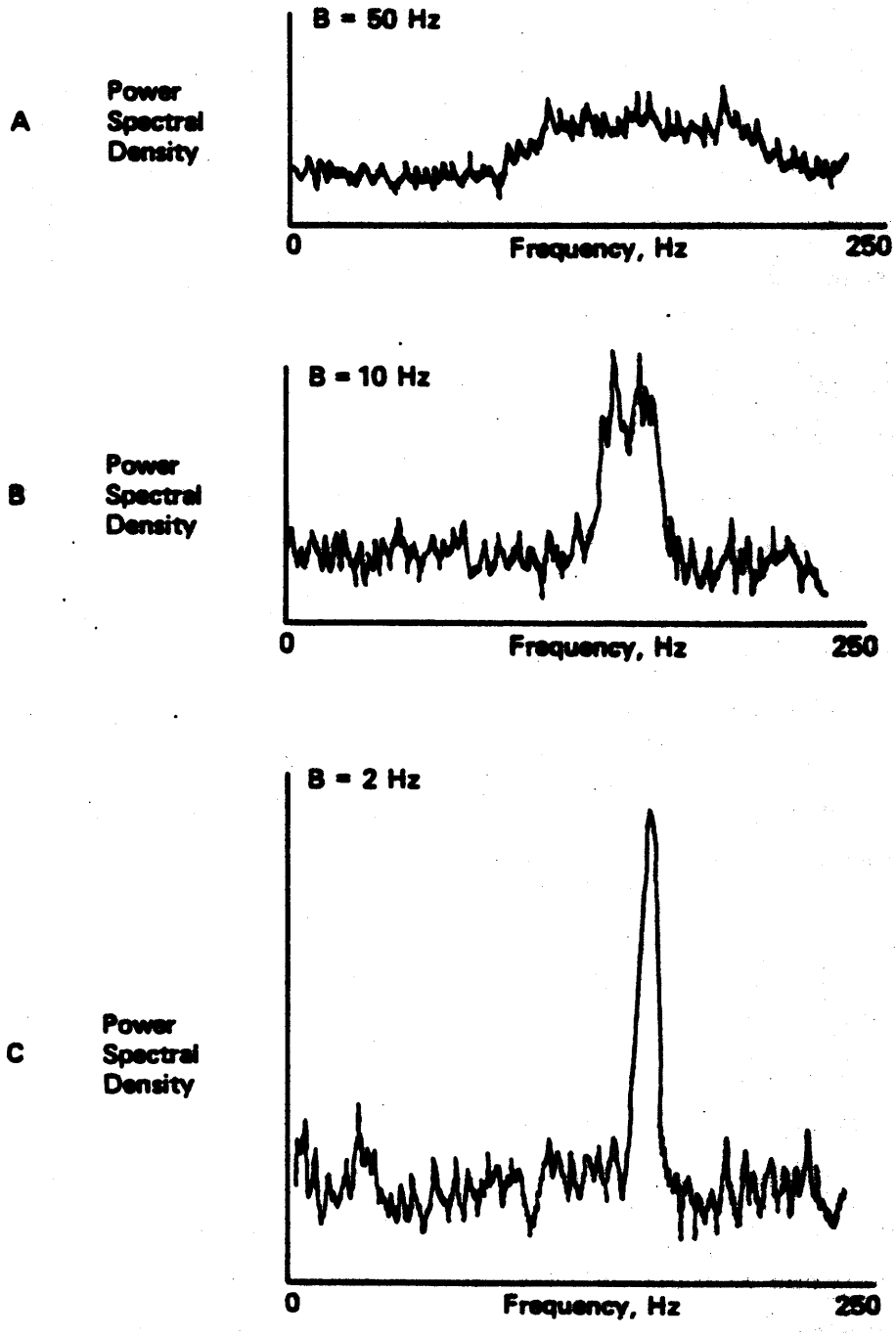
Randomness

The first problem in analyzing an arbitrary wave is to determine if it is periodic or random. A very narrow random peak, caused by a high Q resonance, is sometimes difficult to distinguish from a periodic component. If the vibration data is purely periodic, the fact is usually obvious by observing an oscilloscope picture but, if it is a mixture of sine and random, this fact is not always obvious. A simple test which is effective in many cases is to tune in the peak which is suspected to be a sinusoid with an analyzer and successively switch in two or more filters of different bandwidths. Since the amplitude of a sinusoid is independent of the filter bandwidth, whereas a random wave is not, the presence of a sinusoid in random noise is often revealed. In addition, the "bandwidth" of a sinusoidal component will always be the same as the filter bandwidth and decrease as filter bandwidth is decreased. The bandwidth of a narrow band random signal will not decrease below its effective width no matter how narrow the filter becomes. Figure 12-1 shows the results of such a test. This method will not work when the random response peak is less than the filter bandwidth. For the case of random spectral peaks with center frequencies of less than perhaps 50 Hz, the structural resonance may have a bandwidth of less than 1 Hz making it very difficult to distinguish from a sine wave.

In the latter case, a quantitative test for randomness may be useful, although somewhat tedious. As has been seen from Figure 12-1, the existence of a sinusoid in an otherwise random signal will be revealed by a sharp peak in a properly resolved power spectrum. However, a narrowband random signal representing a high Q resonance in a structural vibration response will also show up as a sharp peak in the power spectrum. The two cases can be distinguished from one another by examining repeated records of the mean square value for the spectral peak at different times. If the peak is a sinusoid, the mean square value should be the same except for observation and instrument errors. On the other hand, mean square measurements taken at different

Power Spectra for a Sine Wave Mixed with a Random Signal Obtained Using Three Different PSD Analyzer Filter Bandwidths, B.

RMS Amplitude of Sine Wave = 0.05 Volts
RMS Amplitude of Random Signal = 1.0 Volts



ACTUAL POWER SPECTRA PLOTS

FIGURE 12-1

times will show a dispersion or scatter if the signal is random. The following steps are suggested for such a test:

1. Using the narrowest filter available, tune the PSD analyzer over the peak of interest so that the peak is isolated from the rest of the signal.
2. Make a number of mean square measurements by averaging over each of a number, N, of equally long segments of the sample record. The averaging time, T_a, should be equal to $T_a = T/N$ where T is the total length of the sample record. N should be less than 0.1 BT where B is the bandwidth between the -3 dB points of the filter in the analyzer. See Figure 12-2.
3. If an RC type continuously averaging system is used, the mean square values can be recorded as in Figure 12-2. Equally spaced discrete values of the mean square are selected and the <u>expected</u> normalized variance, ϵ^2, is computed from:

$$\epsilon^2 = \frac{1}{BT_a} \tag{12-1}$$

4. Calculate the <u>actual</u> normalized variance, $\hat{\epsilon}^2$, from the formula:

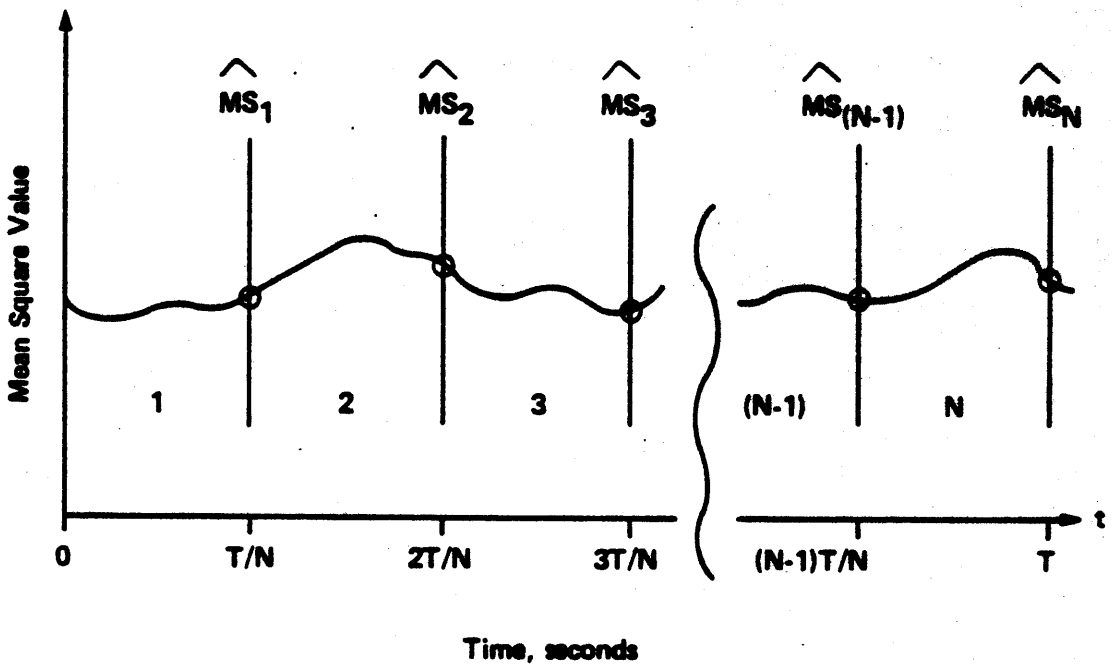
$$\hat{\epsilon}^2 = \frac{N}{N-1}\left\{\frac{N\sum_{i=1}^{N}\hat{MS}_i}{\left[\sum_{i=1}^{N}\hat{MS}_i\right]^2} - 1\right\} \tag{12-2}$$

5. Determine the ratio:

$$R_\epsilon = \hat{\epsilon}^2/\epsilon^2 \tag{12-3}$$

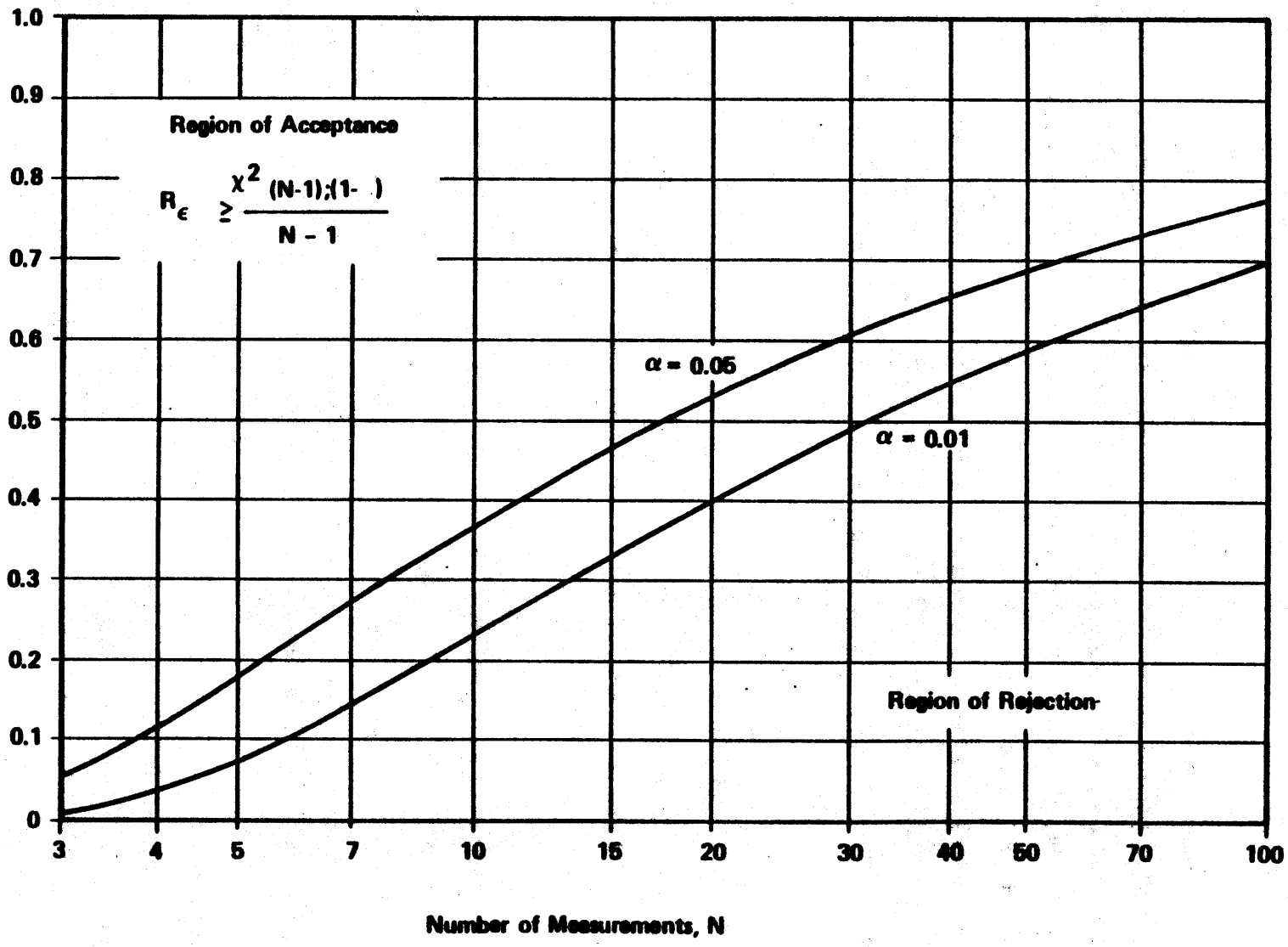
If R_ϵ is statistically equivalent to unity, the signal in question is random. If significantly less than unity, it is considered to be periodic. Figure 12-3 is the criterion for decision, where α is the level of significance for the decision. Thus, if the number of measurements is 20 and the variance ratio is 0.6, there is a 5% probability ($\alpha = 0.05$) that the decision is wrong. Therefore, the decision is in the region of acceptance.

Given a sample record of length T seconds, a continuous RC averaged mean square value measurement may be reduced to a collection of equivalent true averaged mean square value measurements as follows: Let the RC averaging time constant K be short compared to the record length T; that is, K << T. Now, divide the continuous measurement into N equal intervals such that each interval is about 3K or 4K long, as shown below. The level of the continuous measurement at the end of each interval will constitute a discrete measurement based upon an equivalent true averaging time of T_a = 2K. The individual readings should be 3K or 4K apart to assure that they are statistically independent.



ANALYSIS OF CONTINUOUS MEAN SQUARE VALUE MEASUREMENTS

FIGURE 12-2

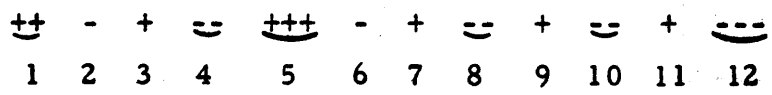


ACCEPTANCE REGIONS FOR RANDOMNESS TEST

FIGURE 12-3

Stationarity

The stationarity of a process can easily be determined by some simple measurements. An analyzer is not required. With an rms meter, plot a mean square curve as in Figure 12-2. Divide the curve into N segments with N discrete mean square measurements. Determine the median for the measurements; that is, the value for which half the measurements are higher and half lower. Identify each measurement higher than the median by a (+) and each lower by a (-). Arrange the measurements in the proper time sequence. Count the number of "runs". A "run" is defined as a sequence of identical symbols which is followed and preceded by a different symbol or no symbol at all. As an example:



the number of runs is twelve, the number of segments, N, is 20. Entering the chart of Figure 12-4 at 12 and 20, it is evident that the sequence is stationary with a 90% confidence interval.

Spectrum Resolution Considerations

In the analysis of the power spectrum from actual structures, it is seen from previous discussions that the statistical quality of measurement improves as the bandwidth of the analysis filter is broadened and the averaging time is increased. However, the analysis time suffers with increases in averaging time and the ability to resolve resonances and peaks in the power spectrum suffers with wide filter bandwidths. Therefore, the selection of the proper bandwidth for an analysis becomes an art. To optimize the analysis time and resolution, one must virtually know the results of an analysis before analyzing. Quite often this may be resolved by making quick-look scans of the data prior to performing the final analysis.

Analysis Rate of a Constant Bandwidth Analyzer

If the filter buildup time $T_{100\%} \approx 4/B$ seconds and we then follow the detector with a true averaging circuit, we would not want to move the filter

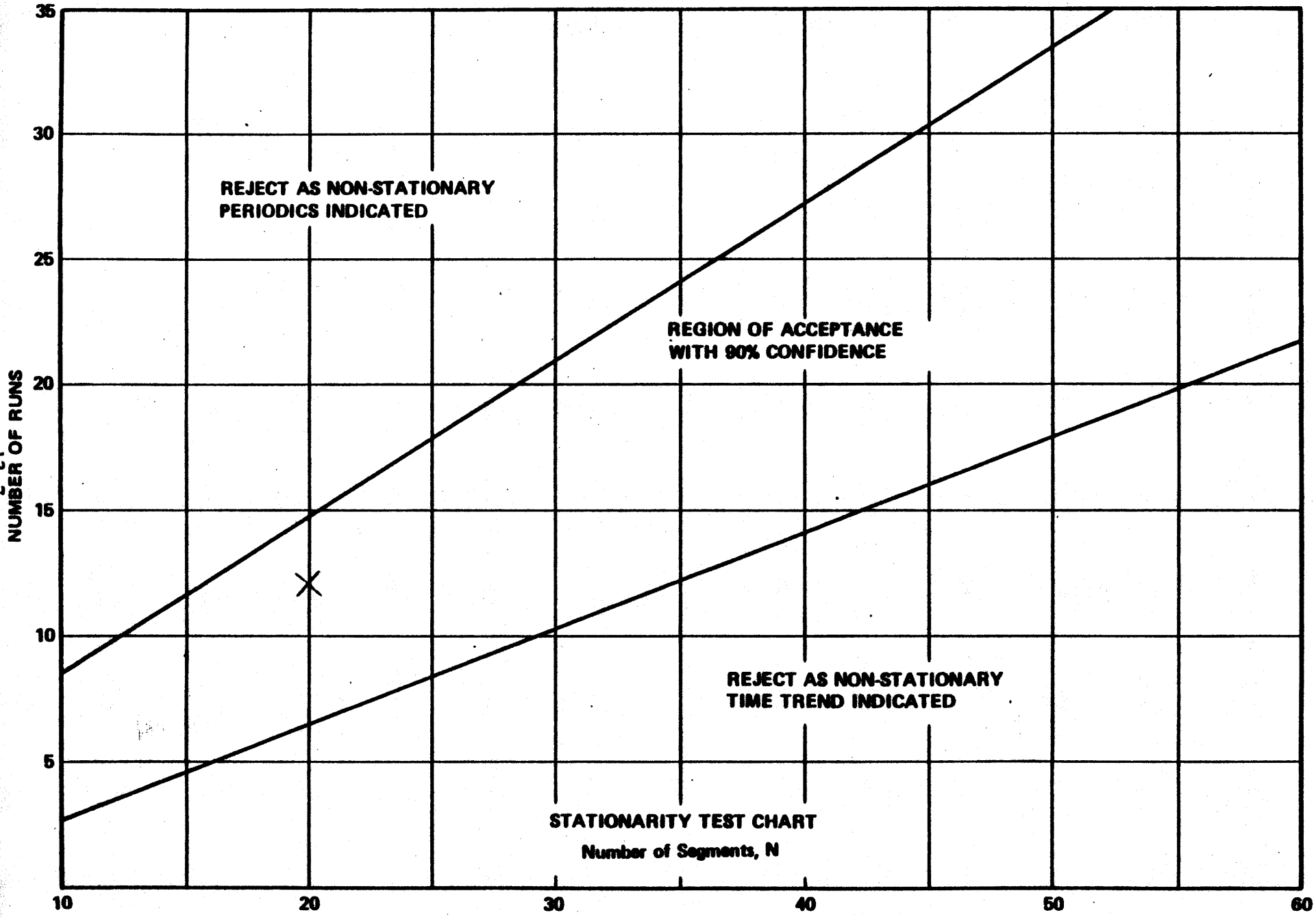

REJECT AS NON-STATIONARY
PERIODICS INDICATED
REGION OF ACCEPTANCE
WITH 90% CONFIDENCE
REJECT AS NON-STATIONARY
TIME TREND INDICATED
STATIONARITY TEST CHART
Number of Segments, N
NUMBER OF RUNS
35
30
25
20
15
10
5
10
20
30
40
50
60

FIGURE 12-4.

any faster than one bandwidth for every 4/B seconds. We would then desire an average value of the power from the filter after this time, or the data would be influenced by transient signals during the buildup time. An expression for analysis rate is:

$$AR \leq \frac{B}{\frac{4}{B} + T} \qquad (12\text{-}4)$$

Where:

AR = analysis rate in Hz/sec

B = filter noise bandwidth in Hz

T = averaging time in seconds

Upon further examination of the above equation and consideration of the practical values of T for a given B, based on desirable statistical confidence levels, we see that the 4/B term will make very little difference in the final scan rate value. In other words, as B becomes larger, T usually becomes smaller and the 4/B part becomes negligible compared to T. For a small B, T becomes large, therefore, Equation 12-4 becomes:

$$AR \leq \frac{B}{T} \qquad (12\text{-}5)$$

If RC averaging is used, we must allow at least 4RC time constants before the output of the averaging circuit is producing a correct averaged value. Therefore, we would not want to sweep the analyzer filter any faster than one bandwidth for every 4RC time constants or:

$$AR \leq \frac{B}{4RC} \qquad (12\text{-}6)$$

Let's now compare Equations 12-5 and 12-6. Remember after 4RC seconds that RC = T/2. AR for true averaging with the equivalent RC time constant is twice that of an analyzer using RC averaging from:

$$AR_T \leq \frac{B}{2RC} \text{ (using Equation 12-5)}$$

Where:

AR_T = analysis rate using true averaging

and,

$$AR_{RC} \leq \frac{B}{4RC} \text{ (using Equation 12-6)}$$

Where:

AR_{RC} = analysis rate using RC averaging

Therefore: $$\frac{AR_T}{AR_{RC}} = \frac{B/2RC}{B/4RC} = 2$$

This becomes a consideration when trying to shorten the analysis time. It is also interesting to note that the analysis rate is related to the statistical confidence levels:

Since $n = 2BT$ or $B = \frac{n}{2T}$ and using Equation 12-5,

$$AR \leq \frac{B}{T}, \text{ we find } AR(T) = B = \frac{n}{2T} \text{ or}$$

$$n = 2AR(T)^2 \qquad (12\text{-}7)$$

In another sense, since

$$n = 4BRC$$

$$4RC = n/B$$

and

$$AR \leq \frac{B}{4RC}$$

we have

$$AR \leq B/n/B$$

$$AR \leq B^2/n \qquad (12\text{-}8)$$

so we see that analysis rate is directly related to statistical confidence.

Analysis Rate and Time for Constant Bandwidth Analyzer

The above analysis rates state that the analysis filter should be scanned using a linear sweep rate, or:

$$\frac{df}{dt} = AR \qquad (12\text{-}9)$$

then $\int_{f_1}^{f_2} df = AR \int_{0}^{A_t} dt$, solving $t\Big|_0^{A_t}$ provides

$$A_t = \frac{1}{AR}\left(f_2 - f_1\right) \qquad (12\text{-}10)$$

Where: A_t = total analysis time (sec.)

f_2 = highest analysis frequency (Hz)

f_1 = lowest analysis frequency (Hz)

AR = analysis rate (Hz/sec.)

Thus the total analysis time is equal to the total frequency range divided by the analysis rate.

A sample problem of "on-line" analysis using the equations developed for a constant bandwidth analyzer follows:

1. Require a ±10% statistical accuracy; 10% = 100 ϵ (from Equation 12-1)
2. Analysis frequency range, 10 Hz to 2000 Hz
3. Analyzer bandwidth, B, required for desired resolution is 10 Hz (noise bandwidth)

Since the error desired is less than 20%, we can use:

$$\epsilon = \frac{1}{\sqrt{BT}}, \quad BT = \frac{1}{\epsilon^2} = \frac{1}{.01} = 100$$

Since B = 10, $T = \frac{100}{10}$ = 10 second averaging time.

If the analyzer uses RC averaging, then:

$$\epsilon = \frac{1}{\sqrt{2BRC}}, \quad RC = \frac{1}{2\epsilon^2 B} = 5 \text{ seconds}$$

$$AR = \frac{B}{T} = \frac{10}{10} = 1 \text{ Hz/sec (for true averaging)}$$

$$AR = \frac{B}{4RC} = \frac{10}{20} = 0.5 \text{ Hz/sec (for RC averaging)}$$

$$A_t = \frac{2000 - 10}{AR} = \frac{1990}{1} = 1990 \text{ seconds or } 33 \text{ minutes}$$

(for true averaging)

$$A_t = \frac{2000 - 10}{AR} \quad \frac{1990}{0.5} = 3980 \text{ seconds or } 66 \text{ minutes}$$

(for RC averaging)

Dynamic Range

A consideration which is very important yet often overlooked in wave analysis is the dynamic range capabilities of the analyzing system. This matter is of such importance that an illustration is warranted.

Consider the case where a band of random frequencies up to 20 kHz is being analyzed with a 2 Hz filter. To make the example easier, we shall assume that the 20 kHz spectrum is white, which it undoubtedly will not be. We will assume that 3-sigma limiting is permissible and that the voltage of the signal is 1.0 V rms. The dynamic range of the analyzer is 60 dB.

The amplitude adjustment of the input signal must then accommodate a three volt peak signal. This signal should just be on the verge of overloading the analyzer. When this adjustment is made, the voltage out of the 2 Hz filter becomes 3 millivolts peak which is 60 dB under the input signal. Consequently, the analyzer has no dynamic range left and the analysis should not be made since it is worthless. This state of affairs comes about because the output voltage is dependent on bandwidth.

$$\frac{\text{bandwidth of output}}{\text{bandwidth of input}} = \frac{2}{2000} = .001 = -60 \text{ dB} \qquad (12\text{-}11)$$

Therefore, the output of the 2 Hz filter would be at the noise floor of the analyzer. There is no way out of this dilemma except to bandwidth limit the input or to use a wider filter. Suppose we limit the input signal to a 10 kHz band. We will have gained 3 dB and the range of our analysis will be restricted to that 3 dB. Suppose we choose a 10 Hz filter instead of the 2 Hz. Then we will have gained 14 dB. If both means were employed, we would have an actual dynamic range of 17 dB, which would probably be acceptable if the input signal were white or nearly

white. Of course, we could gain almost an additional 4 dB if we let the analyzer overload at 2-sigma. We would lose 5% of the information but, even so, the solution is not to be recommended unless all else fails. The answer to this is in thoughtful compromise and, indeed, it may be necessary to make one or two trial analyses before a reasonable solution is found.

Section 12

PROBLEMS IN ANALYZING RANDOM WAVES

How can one tell when the signal to be analyzed is random?

The analysis rate should not be greater than the filter bandwidth divided by time. What happens if we exceed this rate?

How is the dynamic range of an analyzer affected by the bandwidth of the input signal? Explain the implications of the ratio "bandwidth of filter/ bandwidth of input".