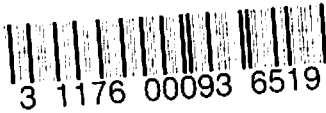


Contrails

WADD TECHNICAL REPORT 60-681

SN 73434
SUMMARY
SEPT 4, 1959-
APR. 1, 1960
C 2



Measurement and Analysis of Power Spectra and Cross-Power Spectra for Random Phenomena

JULIUS S. BENDAT

RAMO-WOOLDRIDGE

A DIVISION OF

THOMPSON RAMO-WOOLDRIDGE INC.

M275-OU4

~~NOT RELEASABLE~~
TO O.T.S.

OCTOBER 1960

WRIGHT AIR DEVELOPMENT DIVISION

Contrails

WADD TECHNICAL REPORT 60-681

SN - 73434
SUMMARY
SEPT. 4, 1959 -
APR. 1, 1960

Measurement and Analysis of Power Spectra and Cross-Power Spectra for Random Phenomena

Julius S. Bendat

Ramo-Wooldridge

A Division of
Thompson Ramo Wooldridge Inc.

October 1960

Flight Dynamics Laboratory
Contract Nr AF 33(616)-6592
Project Nr 1367
Task Nr 13579

Wright Air Development Division
Air Research and Development Command
United States Air Force
Wright-Patterson Air Force Base, Ohio

FOREWORD

This report was prepared by Julius S. Bendat of Ramo-Woolridge, Division of Thompson Ramo Woolridge Inc., Canogo Park, California, for the Northrop Corporation, Norair Division, Hawthorne, California, on Air Force Contract AF33(616)-6592, under Task Nr 13579 of Project Nr 1367, "Structural Design Criteria". The work was administered under the direction of the Flight Dynamics Laboratory, Wright Air Development Division. Mr. Leo Wack was the task engineer for the laboratory.

This report covers work from 4 September 1959 to 1 April 1960.

ABSTRACT

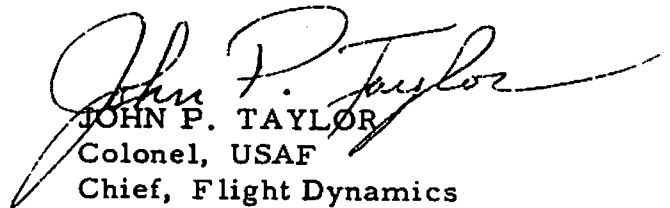
This report investigates fundamental statistical questions concerned with measuring the power spectrum (i. e., power spectral density function) and the cross-power spectrum associated with random phenomena. The analysis treats in detail a practical engineering (analog) technique for making such measurements which employs a filter and multiplier. Quantitative formulas are derived for predicting the mean square error to be expected in a set of estimates as a function of the length of the record, the bandwidth of the filter, and the true nature of the spectrum. Discrete (digital) approximations to the continuous (analog) formulas are developed based on sampling the data at equispaced intervals. Confidence limits and experimental design relations are included in the analysis.

Other material in the report explains from a broad viewpoint basic ideas of probability theory, random processes, and general matters of statistical estimation. The report examines briefly how to determine the mean value and the correlation functions for a pair of random processes. Physical applications of power spectra and cross-power spectra information are discussed relative to: (a) probability distribution of amplitude values of a random record, (b) measurement of system frequency response function in the presence of noise, (c) zero crossings and maxima probabilities of a random record.

PUBLICATION REVIEW

The publication of this report does not constitute approval by the Air Force of the findings or conclusions contained herein. It is published only for the exchange and stimulation of ideas.

FOR THE COMMANDER:


JOHN P. TAYLOR
Colonel, USAF
Chief, Flight Dynamics
Laboratory

Contrails

CONTENTS

1.	Introduction	1
2.	Probability Fundamentals	5
2.1	One Random Variable	5
2.2	Two Random Variables	8
2.3	Special Probability Distributions	12
3.	Random Processes	22
3.1	Correlation (Covariance) Structure of Weakly Stationary Random Processes	25
3.2	Spectral Decomposition of Stationary Random Processes	28
3.3	Ergodic Stationary Random Processes	36
4.	Measurement of Mean Values	39
5.	Measurement of Autocorrelation and Cross-Correlation Functions	45
6.	Measurement of Power Spectra and Cross-Power Spectra : Continuous Data	49
6.1	Power Spectra Measurements	50
6.1.1	Analysis of Bias	55
6.1.2	Analysis of Variance	56
6.1.3	Mean Square Error	57
6.1.4	Frequency Resolution	58
6.1.5	Correction for Mean and Linear Trend	59
6.2	Cross-Power Spectra Measurements	63
7.	Measurement of Power Spectra and Cross-Power Spectra : Sampled Data	67
7.1	Power Spectra Measurements	68
7.2	Summary of Formulas	75
8.	Confidence Limits and Design Relations	81
8.1	Gaussian Set of Measurements	83
8.2	Special Design Relations	86
9.	Physical Applications	89
9.1	Probability Distribution of Amplitude Values	89
9.2	Measurement of System Frequency Response Function	91
9.2.1	Noise in Measured Input	96
9.2.2	Noise in Measured Output	98
9.2.3	Confidence Limits Based on Coherence Function	100
9.3	Zero Crossings and Maxima Probabilities	101
	Bibliography	106

MEASUREMENT AND ANALYSIS OF POWER SPECTRA
AND CROSS-POWER SPECTRA FOR
RANDOM PHENOMENA

1. INTRODUCTION

The purpose of this report is to set forth clearly and concisely: (1) how to process continuous or sampled time records of random phenomena so as to obtain valid estimates of certain significant statistical quantities for single time records and joint statistical quantities for pairs of time records, and (2) how to interpret this information to yield various desirable physical properties of the statistical behavior and interrelationships of the random time records under consideration. Emphasis is placed on underlying mathematical assumptions which need to be satisfied, practical limitations in gathering and processing of the data, and careful appraisal of confidence limits to be given to the results.

In order to make this material useful to as widespread an engineering audience as possible, only simple mathematical arguments have been included here. Also, the discussion is sufficiently broad to be applicable to data covering many different kinds of physical applications, whether it be, for example, from maneuvering loads experienced at various points on an airplane during flight [Ref. 8], or from electrical noise fluctuations observed at various points in an electronic circuit during its operation [Ref. 1]. A significant difference between these two examples occurs, however, in the frequency ranges that may be involved, the maneuvering loads data generally falling below several cycles per second while electrical noise data may cover extremely high frequencies of the order of megacycles per second. This frequency difference results, in general, in different preferred methods for analyzing the data —

Manuscript released by the author March 1960 for publication as a WADD Technical Report.

Contrails

a continuous analog approach for high frequency data, a discrete digital approach for low frequency data.

For analysis of a single random time record $x(t)$, which is assumed to be a sample member of a stationary random process (i. e., an ensemble of time histories whose statistical properties averaged over the ensemble are independent of time shifts), three statistical quantities are of prime importance. These are the sample mean value (also called the average or expected value), which indicates the d-c level of the record; the sample autocorrelation function (also called the covariance function), which describes the average value of the product between values of the record a specified time interval apart as a function of this time interval; and the sample power spectral density function (also called the power spectrum), which is a way of describing how the average value of $x^2(t)$ is distributed with frequency. The power spectral density function is a real-valued function of the frequency.

The precise value of these three quantities for a single random time record is of little note. What is important is to be able to predict how repeated measurements of these same quantities will behave for future similar time histories from the same stationary random process. In this way, a single test, or a small number of such tests, may yield maximum information about the statistical probability distribution of the entire random process without the expense of making many repeated identical costly experiments. This reduces greatly the amount of data that needs to be gathered and processed, while still providing final desired useful results. There needs to be established rules for selection of the sample size required to achieve a preset confidence level in the results.

For joint analysis of two random time records, $x(t)$ and $y(t)$, which are samples from two different stationary random processes that may be related in some manner, (e. g., simultaneous records of two physical quantities which vary with time, or input vs. output in

Contrails

a physical system), one may compute a stationary sample cross-correlation function (also called the cross-covariance function), which describes the average value of the product between values of the two records a specified time interval apart as a function of this time interval; and the sample cross-power spectral density function (also called the cross-power spectrum), which describes how the average value of $x(t) y(t)$ is distributed with frequency. The cross-power spectral density function is a complex-valued function of the frequency whose real and imaginary parts give respectively the distribution with frequency of the "in phase" and "in quadrature" components of the average value of $x(t) y(t)$.

As pointed out earlier, these special results from $x(t)$ and $y(t)$ are of interest only insofar as they may be used to predict the expected behavior over the two stationary random processes being analyzed. A general problem of statistical analysis is to estimate such desired ensemble statistical quantities on the basis of appropriate calculations which are averaged over a finite period of time, and to be able to state the expected spread in results for future repeated measurements. This report defines the criteria which a satisfactory set of estimates must fulfill, and explains in some detail in Sections 6 and 7 both continuous (analog) and sampled-data (digital) methods for calculating power spectra and cross-power spectra functions. Questions of confidence limits and design relations are taken up in Section 8.

For two stationary random processes, relatively simple relations exist such that knowledge of either the cross-power spectrum or the cross-correlation function theoretically yields the other. Also, ordinary power spectra and autocorrelation functions are merely special cases of cross-power spectra and cross-correlation functions when the two random processes coincide.

A main area of application where one desires power spectra and cross-power spectra information rather than corresponding

Contrails

autocorrelation functions and cross-correlation functions occurs in connection with input-output relations for linear constant parameter engineering systems. Whereas certain input-output relations are relatively complicated convolution integrals in the time domain using correlation functions, these same input-output relations reduce to simple algebraic equations in the frequency domain in terms of power spectra and cross-power spectra. To be specific, the output power spectrum is the product of the input power spectrum with the square of the system gain factor; the cross-power spectrum between the input and the output is the product of the input power spectrum with the system frequency response function. The frequency response function is a complex-valued function of the frequency which includes both the gain factor and the phase characteristics of the system.

For certain physical applications where precise determination of a time delay may be at issue, emphasis must be placed on measurement of correlation functions rather than power spectra. In still other applications involving predicting statistical properties of the time history of a random record, such as estimating expected number of zero crossings or maximum values, or probability distribution of amplitude values, a choice between correlation functions and power spectra is a matter of taste. However, engineers generally prefer to use power spectra. Some of these physical applications are developed in Section 9.

Insofar as seems reasonable to the author, this report is self-contained and written in expository or outline form so as to cover many pertinent ideas and formulas. Deeper discussion of fundamental physical questions and more rigorous details of required mathematical proofs are available in the references listed in the Bibliography for those desiring to pursue these matters further. The author is particularly indebted to writings of Katz [Ref. 7] and Parzen [Ref. 9] for some of the statistical material in this report.

2. PROBABILITY FUNDAMENTALS

The underlying concept in probability theory is that of a set, namely, a collection of objects such that it is possible to determine of any particular object whether or not it is a member of the set. In particular, the possible outcomes of an experiment (or a measurement) represent a set of points called the sample space. These points may be grouped together in various ways, called events, and under suitable conditions probability measures may be assigned to each event. These probabilities always lie between zero and one, the probability of an impossible event being zero, of a certain event being one. For sample spaces containing a finite number of points, the probability of a particular event is simply the ratio of the number of points in the event to all possible points. For sample spaces containing an infinite number of points, a more sophisticated approach is required. In all cases, however, it is required that the probability measure be additive in the sense that if A, B, C, . . . are different events with no point in common, then

$$\text{Prob [A or B or C or . . .]} = \text{Prob [A]} + \text{Prob [B]} + \text{Prob [C]} + \dots$$

(2.1)

2.1 ONE RANDOM VARIABLE

A random variable $R(k)$ is defined as a real-valued point function of k , where k is a point from the sample space. That is, a random variable $R(k)$ represents a real number between $-\infty$ and $+\infty$ which is associated to each point k that might occur. In terms of $R(k)$ and the underlying probability measure in the sample space, one may define a (first-order) distribution function $P_R(\alpha)$ as the probability over k that $R(k) \leq \alpha$. This probability is given by the probability which is assigned to the set of points k satisfying the

Contrails

desired inequality. Observe that the set of points k satisfying $R(k) \leq \alpha$ is a subset of the totality of all points k which satisfy $R(k) \leq \infty$. In notation form

$$P_R(\alpha) = \text{Prob}_k [R(k) \leq \alpha] \quad (2.2)$$

Clearly

$$P_R(a) \leq P_R(b) \quad \text{if } a \leq b$$

$$P_R(-\infty) = 0 \quad ; \quad P_R(\infty) = 1$$

For example, let the sample space consist of tosses of a single coin where the two possible outcomes, called heads and tails, are assumed to occur with equal probability (1/2). The random variable $R(k)$ for this example takes on only two discrete values, $R(\text{heads})$ and $R(\text{tails})$, to which arbitrary real numbers may be assigned, e. g., let $R(\text{heads}) = a$ and $R(\text{tails}) = b$ where a and b are real numbers with, say, $b > a$. With these choices for $R(k)$, it follows that the distribution function

$$P_R(\alpha) = \begin{cases} 0 & ; \quad \alpha < a \\ 1/2 & ; \quad a \leq \alpha < b \\ 1 & ; \quad \alpha \geq b \end{cases}$$

As a second example, let the sample space consist of choosing a point at random in the interval $[0, 1]$, including the end points. A continuous random variable $R(k)$ for this example may be defined by the numerical value of the chosen point. The corresponding distribution function becomes

Contrails

$$P_R(\alpha) = \begin{cases} 0 & ; & \alpha < 0 \\ \alpha & ; & 0 \leq \alpha \leq 1 \\ 1 & ; & \alpha \geq 1 \end{cases}$$

If the random variable assumes a continuous range of values (which will be assumed hereafter) then a (first-order) probability density function $p_R(\alpha)$ may be defined by the differential relation

$$p_R(\alpha) d\alpha = \text{Prob} [\alpha < R(k) \leq \alpha + d\alpha] \quad (2.3)$$

Note that

$$p_R(\alpha) > 0$$

$$\int_{-\infty}^{\infty} p_R(\alpha) d\alpha = 1$$

$$P_R(\alpha) = \int_{-\infty}^{\alpha} p_R(\alpha) d\alpha \quad ; \quad \frac{dP_R(\alpha)}{d\alpha} = p_R(\alpha)$$

The probability density function $p_R(\alpha)$ should not be confused with the distribution function $P_R(\alpha)$.

Suppose $R = R(k)$ takes on values in the range $-\infty$ to $+\infty$. Then the mean value (also called expected value, average value) of R is given by the limit of the sum of assumed values when each value is multiplied by its appropriate probability of occurrence. That is,

$$E(R) = \lim_{N \rightarrow \infty} \sum_{i=1}^N \alpha_i \text{Prob} [R(k) = \alpha_i] = \int_{-\infty}^{\infty} \alpha p_R(\alpha) d\alpha = \bar{R} \quad (2.4)$$

Similarly, the expected value of any real single-valued continuous function $g(R)$ of the random variable R is given by

$$E [g(R)] = \int_{-\infty}^{\infty} g(\alpha) P_R(\alpha) d\alpha = \overline{g(R)} \quad (2.5)$$

In particular, for $g(R) = R^2$, the mean square value of R is given by

$$E [R^2] = \int_{-\infty}^{\infty} \alpha^2 P_R(\alpha) d\alpha = \overline{R^2} \quad (2.6)$$

and the variance of R is defined by the mean square value of R about its mean value, namely by,

$$\sigma^2(R) = E [(R - \bar{R})^2] = \overline{R^2} - (\bar{R})^2 \quad (2.7)$$

By definition, the standard deviation of R , denoted by σ , is the positive square root of the variance. The standard deviation is measured in the same units as the mean value.

2.2 TWO RANDOM VARIABLES

Consider next two random variables $R = R(u)$ and $S = S(v)$ where u and v are points in a suitable sample space. Let $P_R(\alpha)$ and $P_S(\beta)$ be the distribution functions associated with R and S respectively. The joint (second-order) distribution function $P_{R,S}(\alpha, \beta)$ is defined to be the probability which is associated with the subset of combined points (u, v) in the sample space satisfying the inequalities $R(u) \leq \alpha$ and $R(v) \leq \beta$. The total set of combined points (u, v) satisfies the inequalities $R(u) \leq \infty$ and $R(v) \leq \infty$. In notation form,

$$P_{R,S}(\alpha, \beta) = \text{Prob} [R(u) \leq \alpha \text{ and } S(v) \leq \beta] \quad (2.8)$$

Contrails

Clearly,

$$P_{R, S}(-\infty, \beta) = 0 = P_{R, S}(\alpha, -\infty)$$

$$P_{R, S}(\infty, \infty) = 1$$

As before, assuming the random variables to be continuous, the joint distribution function $P_{R, S}(\alpha, \beta)$ should not be confused with the joint (second-order) probability density function $p_{R, S}(\alpha, \beta)$ which is defined by the differential relation

$$p_{R, S}(\alpha, \beta) d\alpha d\beta = \text{Prob} [\alpha < R(u) \leq \alpha + d\alpha \text{ and } \beta < S(v) \leq \beta + d\beta] \quad (2.9)$$

Note that

$$p_{R, S}(\alpha, \beta) \geq 0$$
$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p_{R, S}(\alpha, \beta) d\alpha d\beta = 1$$
$$P_{R, S}(\alpha, \beta) = \int_{-\infty}^{\alpha} \int_{-\infty}^{\beta} p_{R, S}(\alpha, \beta) d\alpha d\beta$$

Also

$$P_R(\alpha) = \int_{-\infty}^{\infty} p_{R, S}(\alpha, \beta) d\beta$$
$$P_S(\beta) = \int_{-\infty}^{\infty} p_{R, S}(\alpha, \beta) d\alpha$$

Two random variables R and S are said to be independent if

$$P_{R, S}(\alpha, \beta) = P_R(\alpha) P_S(\beta) \quad (2.10)$$

It follows that

$$P_{R, S}(\alpha, \beta) = P_R(\alpha) P_S(\beta)$$

The expected value of any real single-valued continuous function $g(R, S)$ of the two random variables $R = R(u)$ and $S = S(v)$ is given by

$$E [g(R, S)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(\alpha, \beta) P_{R, S}(\alpha, \beta) d\alpha d\beta = \overline{g(R, S)} \quad (2.11)$$

For example, if $g(R, S) = (R - \bar{R})(S - \bar{S})$ where \bar{R} and \bar{S} are the mean values of R and S, respectively, this defines the covariance $\rho(R, S)$ between R and S, namely,

$$\begin{aligned} \rho(R, S) &= E [(R - \bar{R})(S - \bar{S})] = E [RS] - E [R] E [S] \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\alpha - \bar{R})(\beta - \bar{S}) P_{R, S}(\alpha, \beta) d\alpha d\beta \end{aligned} \quad (2.12)$$

A simple relation exists between the covariance of R and S and the standard deviations of R and S as expressed by the inequality

$$|\rho(R, S)| \leq \sigma(R) \sigma(S) \quad (2.13)$$

In words, the square of the covariance between R and S is less than or equal to the product of the standard deviation of R multiplied by the standard deviation of S.

Contrails

It follows from the above result that the normalized quantity

$$\Gamma (R, S) = \frac{\rho(R, S)}{\sigma(R) \sigma(S)} \quad (2.14)$$

known as the correlation coefficient, will lie between -1 and +1. Random variables R and S whose correlation coefficient is zero are said to be uncorrelated. This concept should not be confused with the previous definition of independent random variables. Note that if R and S are independent random variables, then

$$\begin{aligned} E [RS] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \alpha\beta p_{R, S}(\alpha, \beta) d\alpha d\beta \\ &= \int_{-\infty}^{\infty} \alpha p_R(\alpha) d\alpha \int_{-\infty}^{\infty} \beta p_S(\beta) d\beta = E [R] E [S] \quad (2.15) \end{aligned}$$

Hence $\rho(R, S)$ and, in turn, $\Gamma (R, S)$ equal zero so that independent random variables are also uncorrelated. The converse statement is not true in general; that is to say, uncorrelated random variables are not necessarily independent. However, for physically important situations involving two or more normally (Gaussian) distributed random variables (to be defined later), being mutually uncorrelated does imply independence.

The conditional probability density function of R, given that $S = \beta$, (i. e., given that S is between β and $\beta + d\beta$ for small $d\beta$), is defined by

$$\begin{aligned} p_R(\alpha | S = \beta) d\alpha &= \text{Prob} [\alpha < R(u) \leq \alpha + d\alpha | \text{given that } \beta < S(v) \leq \beta + d\beta] \\ &= \frac{\text{Prob} [\alpha < R(u) \leq \alpha + d\alpha \text{ and } \beta < S(v) \leq \beta + d\beta]}{\text{Prob} [\beta < S(v) \leq \beta + d\beta]} = \frac{p_{R, S}(\alpha, \beta) d\alpha d\beta}{p_S(\beta) d\beta} \end{aligned}$$

or

$$P_R(\alpha | S = \beta) = \frac{P_{R,S}(\alpha, \beta)}{P_S(\beta)} \quad \text{assuming} \quad P_S(\beta) \neq 0 \quad (2.16)$$

For independent random variables, this simplifies to

$$P_R(\alpha | S = \beta) = \frac{P_R(\alpha) P_S(\beta)}{P_S(\beta)} = P_R(\alpha) \quad (2.17)$$

In words, the conditional probability density function for R, given S, is now the same as the original probability density function for R alone.

These ideas may be extended in a straight-forward manner to handle situations of three or more random variables where higher-order probability distributions would be involved.

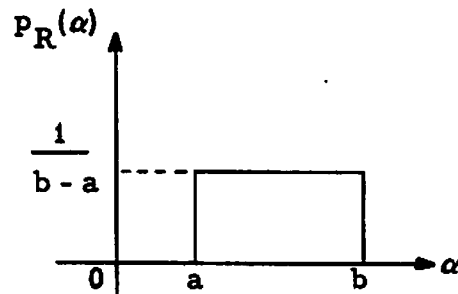
2.3 SPECIAL PROBABILITY DISTRIBUTIONS

By way of illustration, as well as because of their importance to physical problems, some special probability distributions will now be described.

(a) Uniform (Rectangular) Distribution

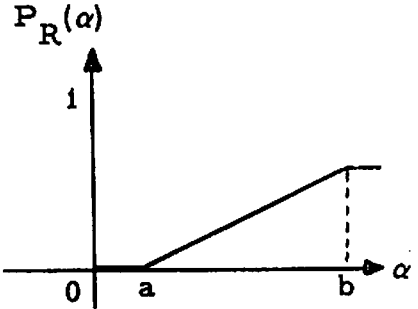
A random variable R is said to follow a uniform (or rectangular) distribution over the interval ($a \leq \alpha \leq b$) if its probability density function is given by

$$\begin{aligned} P_R(\alpha) &= \frac{1}{b-a} & ; & & a \leq \alpha \leq b \\ &= 0 & ; & & \text{otherwise} \end{aligned} \quad (2.18)$$



The corresponding uniform distribution function is

$$P_R(\alpha) = \int_{-\infty}^{\alpha} p_R(\alpha) d\alpha = \frac{\alpha - a}{b - a} \quad ; \quad a \leq \alpha \leq b$$

$$= 0 \quad ; \quad \text{otherwise} \quad (2.19)$$


The mean value m and the variance σ^2 (standard deviation = σ) of the random variable R become

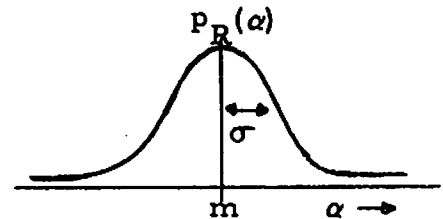
$$m = E[R] = \int_{-\infty}^{\infty} \alpha p_R(\alpha) d\alpha = \frac{1}{b - a} \int_a^b \alpha d\alpha = \frac{a + b}{2}$$

$$\sigma^2 = E[(R - m)^2] = \int_{-\infty}^{\infty} (\alpha - m)^2 p_R(\alpha) d\alpha = \frac{(b - a)^2}{12}$$

(b) Normal (Gaussian) Distribution

A random variable R is said to follow a normal (or Gaussian) distribution if its probability density function is given by

$$p_R(\alpha) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left[-(\alpha - m)^2 / 2\sigma^2 \right] \quad (2.20)$$



where m is any real constant and σ is any positive constant. It is verified easily that m and σ constitute the mean value and standard deviation of the random variable R since

Contrails

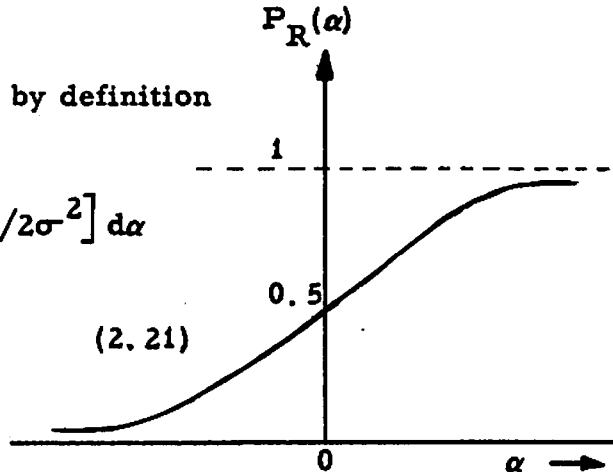
$$E [R] = \int_{-\infty}^{\infty} \alpha P_R(\alpha) d\alpha = m$$

$$E [(R - m)^2] = \int_{-\infty}^{\infty} (\alpha - m)^2 P_R(\alpha) d\alpha = \sigma^2$$

The normal distribution function is by definition

$$P_R(\alpha) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\alpha} \exp \left[-(\alpha - m)^2 / 2\sigma^2 \right] d\alpha$$

$$= \text{Prob} [R \leq \alpha]$$



and is readily available in statistical tables.

Using simple numerical methods, or from tables, it is now convenient to determine the probability that the random variable R will assume values in any desired range. In particular

$$P_R(m + k\sigma) - P_R(m - k\sigma) = \text{Prob} [m - k\sigma < R < m + k\sigma] \quad (2.22)$$

represents the probability (i. e. confidence level) that R will be within plus and minus k standard deviations of the mean value. For $k = 1, 2,$ and $3,$ the confidence results are $68.3\%, 95.4\%,$ and $99.7\%,$ respectively. Working the other way, for 80% confidence, $k = 1.3.$

The importance of the normal distribution in physical problems may be attributed in part to the Central Limit Theorem which asserts that this distribution is approximated closely by the distribution of the sum random variable of a large number of independent small random variables acting together.

Contrails

To be a bit more specific, let R_1, R_2, \dots, R_N be N mutually independent random variables whose individual distributions are not specified. Denote by m_i and σ_i^2 the mean value and variance of each R_i , $i = 1, 2, \dots, N$. Consider the sum random variable

$$R = \sum_{i=1}^N a_i R_i$$

where a_i are arbitrary fixed constants and assume that none of the $a_i R_i$ contributes significantly to the sum. Now, the mean value m and variance σ^2 of R become

$$m = E[R] = E\left[\sum_{i=1}^N a_i R_i\right] = \sum_{i=1}^N a_i E[R_i] = \sum_{i=1}^N a_i m_i$$

$$\sigma^2 = E[(R - m)^2] = E\left[\sum_{i=1}^N a_i (R_i - m_i)\right]^2 = \sum_{i=1}^N a_i^2 \sigma_i^2$$

the last result due to the mutual independence of R_i with R_j for $i \neq j$. The Central Limit Theorem states that under fairly common conditions, the sum random variable R will be normally distributed as $N \rightarrow \infty$ with the above mean value m and variance σ^2 .

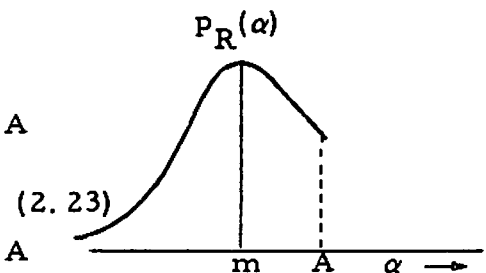
(c) Truncated Normal Distribution

A random variable R is said to follow a truncated normal distribution in the range $\alpha < A$ if its probability density function is given by

$$P_R(\alpha) = \frac{1}{\sigma B \sqrt{2\pi}} \exp\left[-(\alpha - m)^2 / 2\sigma^2\right] \quad ; \quad \alpha < A$$

$$= 0$$

$$; \quad \alpha \geq A$$



where

$$B = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^A \exp \left[-(\alpha - m)^2 / 2\sigma^2 \right] d\alpha$$

The reason for introducing the constant B is to satisfy the requirement

$$\int_{-\infty}^{\infty} P_R(\alpha) d\alpha = 1$$

Observe that to the left of the value $\alpha = A$, except for the scale factor B, the truncated normal distribution has the same shape as the original untruncated normal distribution. The parameters m and σ here do not represent the mean value and standard deviation of the truncated distribution, but refer back to the underlying untruncated distribution.

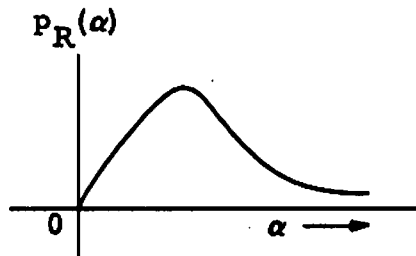
(d) Rayleigh Distribution

A random variable R which is restricted to non-negative values is said to follow a Rayleigh distribution if its probability density function is given by

$$P_R(\alpha) = \left(\frac{\alpha}{\lambda^2} \right) \exp \left(-\frac{\alpha^2}{2\lambda^2} \right) \quad ; \quad \alpha \geq 0$$

(2.24)

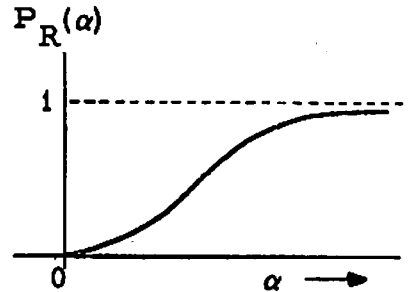
$$= 0 \quad ; \quad \alpha < 0$$



The Rayleigh distribution should not be confused with a Gaussian distribution where the random variable may take on both positive and negative values.

The corresponding Rayleigh distribution function is given by

$$P_R(\alpha) = \text{Prob} [R \leq \alpha] = 1 - \exp(-\alpha^2/2\lambda^2) \quad (2.25)$$



For a Rayleigh distribution, the mean value and mean square value are

$$E[R] = \int_0^{\infty} \alpha P_R(\alpha) d\alpha = (\pi/2)^{1/2} \lambda \approx 1.25\lambda$$

$$E[R^2] = \int_0^{\infty} \alpha^2 P_R(\alpha) d\alpha = 2\lambda^2$$

Hence the variance is now expressed by

$$\sigma^2 = E[R^2] - (E[R])^2 = \left(\frac{4-\pi}{2}\right) \lambda^2 \approx 0.43\lambda^2$$

(e) N-Dimensional Normal Distribution

Let R_1, R_2, \dots, R_n be N random variables defined over the same sample space. Denote their mean values, variances, and covariances by

$$m_i = E[R_i]$$

$$\sigma_i^2 = E\left[\left(R_i - m_i\right)^2\right]$$

$$\rho_{ij} = E\left[\left(R_i - m_i\right)\left(R_j - m_j\right)\right] \quad ; \quad \rho_{ii} = \sigma_i^2$$

Contrails

Their combined distribution is said to be on N-dimensional normal distribution if the associated N-fold probability density function is given by

$$p(\alpha_1, \alpha_2, \dots, \alpha_N) = \frac{\exp\left(\frac{-1}{2|M|} \sum_{i,j=1}^N M_{ij}(\alpha_i - m_i)(\alpha_j - m_j)\right)}{(2\pi)^{N/2} |M|^{1/2}} \quad (2.26)$$

where $\|M\|$ is the covariance matrix of the ρ_{ij} , $|M|$ is the determinant of $\|M\|$, and M_{ij} is the cofactor of ρ_{ij} in determinant $|M|$. To be more explicit,

$$\|M\| = \begin{vmatrix} \rho_{11} & \rho_{12} & \dots & \rho_{1n} \\ \rho_{21} & \rho_{22} & \dots & \rho_{2n} \\ \dots & \dots & \dots & \dots \\ \rho_{n1} & \rho_{n2} & \dots & \rho_{nn} \end{vmatrix} \quad \rho_{ii} = \sigma_i^2 \quad (2.27)$$

and the cofactor M_{ij} of any element ρ_{ij} is defined to be the determinant of order N-1 formed by omitting the ith row and jth column of $\|M\|$, multiplied by $(-1)^{i+j}$.

The outstanding feature of the N-dimensional normal distribution is that all of its properties are determined solely from knowledge of the various mean values m_i and covariances ρ_{ij} . For N = 1, the above reduces to

$$p(\alpha_1) = \frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left[-(\alpha_1 - m_1)^2 / 2\sigma_1^2\right] \quad (2.28)$$

which is the usual one-dimensional normal distribution defined previously.

Contrails

For $N = 2$, there results

$$p(\alpha_1, \alpha_2) = \frac{\exp \left\{ \frac{-1}{2(1 - \Gamma_{12}^2)} \left[\left(\frac{\alpha_1 - m_1}{\sigma_1} \right)^2 - 2\Gamma_{12} \left(\frac{\alpha_1 - m_1}{\sigma_1} \right) \left(\frac{\alpha_2 - m_2}{\sigma_2} \right) + \left(\frac{\alpha_2 - m_2}{\sigma_2} \right)^2 \right] \right\}}{2\pi \sigma_1 \sigma_2 \sqrt{1 - \Gamma_{12}^2}} \quad (2.29)$$

where $\Gamma_{12} = \frac{\rho_{12}}{\sigma_1 \sigma_2}$ is the correlation coefficient between R_1 and R_2 .

Observe that when R_1 and R_2 are uncorrelated so that $\Gamma_{12} = 0$, one obtains

$$p(\alpha_1, \alpha_2) = p(\alpha_1) p(\alpha_2) \quad (2.30)$$

which shows that R_1 and R_2 are also independent.

Similar formulas may be written down for higher order cases where $N = 3, 4, 5, \dots$ etc. For arbitrary N , it follows quite easily that if all different pairs of normally distributed random variables are mutually uncorrelated, (i. e., $\Gamma_{ij} = 0$ whenever $i \neq j$), then these random variables are mutually independent in the probability sense, that is

$$p(\alpha_1, \alpha_2, \dots, \alpha_N) = p(\alpha_1) p(\alpha_2) \dots p(\alpha_N) \quad (2.31)$$

The importance of the N -dimensional normal distribution in physical problems, analogous to the common one-dimensional normal distribution, is due in part to the Multidimensional Central Limit Theorem. This theorem yields the result that the vector sum of a large number of mutually independent N -dimensional random variables approaches an N -dimensional normal distribution under fairly general

conditions. Particular applications of this theorem, relative to zero crossing properties of random records and expected number of maxima values, for example, may be used to justify an assumption that a random record $x(t)$ and its succeeding time derivative $\dot{x}(t)$ will follow a two-dimensional normal distribution, and that $x(t)$, $\dot{x}(t)$ and $\ddot{x}(t)$ will follow a three-dimensional normal distribution.

(f) Distribution of Sums of Independent Random Variables

Suppose R_1 and R_2 are independent random variables with probability density functions $p_{R_1}(\alpha_1)$ and $p_{R_2}(\alpha_2)$ respectively. Let

$$R = a_1 R_1 + a_2 R_2 \quad (2.32)$$

be a typical composite sum record of R_1 with R_2 where $a_1 \neq 0$ and $a_2 \neq 0$ are arbitrary fixed constants (usually known in engineering problems). Then the probability density function $p_R(\alpha)$ associated with R is given by

$$p_R(\alpha) = \int_{-\infty}^{\infty} p_{R_1}(\alpha_1) p_{R_2}\left[\frac{\alpha - a_1 \alpha_1}{a_2}\right] d\alpha_1 \quad (2.33)$$

For the special case where $R = R_1 + R_2$, that is, $a_1 = a_2 = 1$, one obtains

$$p_R(\alpha) = \int_{-\infty}^{\infty} p_{R_1}(\alpha_1) p_{R_2}(\alpha - \alpha_1) d\alpha_1 \quad (2.34)$$

From the above relation, one may verify that the sum of two independent uniform distributions is no longer a uniform distribution. However, the sum of two independent normal distributions remains a normal distribution, with mean and variance equal to the sum of

Contrails

the individual means and variances. The latter result may be extended to the sum of N independent normally distributed random variables.

3. RANDOM PROCESSES

A random process $\left[{}^k x(t) \right]$, $-\infty < t < \infty$, $k = 1, 2, 3, \dots$, is an ensemble of functions of time which can be characterized through its statistical properties. See Figure 3.1. In the physical world, each particular ${}^k x(t)$, t variable, k fixed, represents the result of a single observation or experiment, and constitutes a sample function of the random process.

For example, each ${}^k x(t)$ might represent vertical wing loads on an airplane as a function of time t , the superscript index k denoting different airplanes; or each ${}^k x(t)$ might represent runway roughness at different locations k as a function of distance, the distance variable taking the place of time.

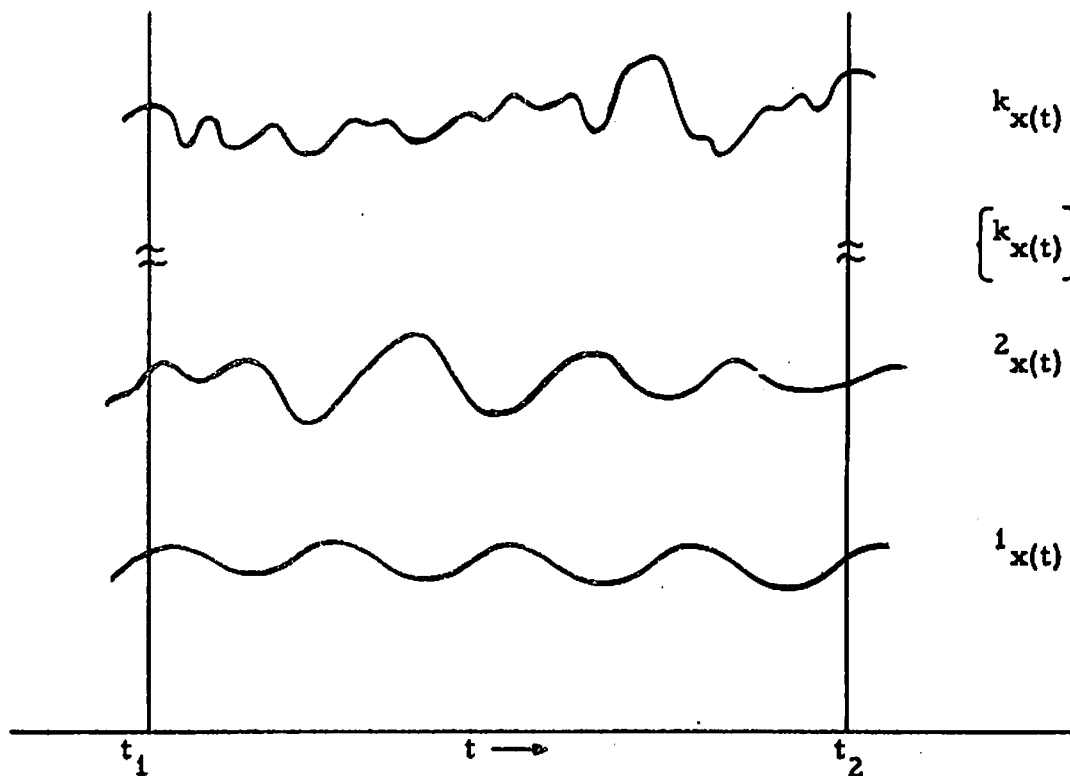


Figure 3.1. Random Process

A particular sample function, in general, would not be suitable for representing the entire random process to which it belongs. Under certain conditions to be described later, however, it turns out that for the class of ergodic random processes, it is possible to derive desired statistical information about the entire random process from appropriate analysis of a single arbitrary sample function. For the situation of a pair of random processes $\{^k x(t)\}$ and $\{^k y(t)\}$, the corresponding problem is to estimate joint statistical properties of the two random processes from proper analysis of an arbitrary pair of sample functions $^k x(t)$ and $^k y(t)$.

Consider two random processes $\{^k x(t)\}$ and $\{^k y(t)\}$. The first statistical quantities of interest are the mean values at arbitrary fixed values of t , denoted by

$$\begin{aligned} m_x(t) &= \left\langle ^k x(t) \right\rangle \text{ Av over } k & ; & \quad t \text{ fixed} \\ m_y(t) &= \left\langle ^k y(t) \right\rangle \text{ Av over } k & ; & \quad t \text{ fixed} \end{aligned} \tag{3.1}$$

Note that k is averaged out in computing these ensemble averages which are indicated by angular brackets. In general, these mean values are different at different times, that is,

$$m_x(t_1) \neq m_x(t_2) \text{ if } t_1 \neq t_2$$

$$m_y(t_1) \neq m_y(t_2) \text{ if } t_1 \neq t_2$$

The next statistical quantities of interest are the covariance functions at arbitrary fixed values of τ and t ,

$$\begin{aligned} \rho_x(\tau, t) &= \left\langle \left[x(t) - m_x(t) \right] \left[x(t + \tau) - m_x(t + \tau) \right] \right\rangle_{Av \text{ over } k} \\ \rho_y(\tau, t) &= \left\langle \left[y(t) - m_y(t) \right] \left[y(t + \tau) - m_y(t + \tau) \right] \right\rangle_{Av \text{ over } k} \\ \rho_{xy}(\tau, t) &= \left\langle \left[x(t) - m_x(t) \right] \left[y(t + \tau) - m_y(t + \tau) \right] \right\rangle_{Av \text{ over } k} \end{aligned} \quad (3.2)$$

In general, these quantities are different for different combinations of τ and t . Observe that at $\tau = 0$, (omitting the index k for simplicity in notation, but still retaining angular brackets to imply ensemble averages)

$$\begin{aligned} \rho_x(0, t) &= \left\langle \left[x(t) - m_x(t) \right]^2 \right\rangle_{Av} = \sigma_x^2(t) \\ \rho_y(0, t) &= \left\langle \left[y(t) - m_y(t) \right]^2 \right\rangle_{Av} = \sigma_y^2(t) \\ \rho_{xy}(0, t) &= \left\langle \left[x(t) - m_x(t) \right] \left[y(t) - m_y(t) \right] \right\rangle_{Av} = \rho_{xy}(t) \end{aligned}$$

Thus the covariance functions $\rho_x(0, t)$ and $\rho_y(0, t)$ represent the ordinary variances of $\{x(t)\}$ and $\{y(t)\}$ at a fixed value of t , while $\rho_{xy}(0, t)$ represents the covariance between $\{x(t)\}$ and $\{y(t)\}$. As before, different results would generally be obtained for different values of t .

Other statistical quantities can be defined over the ensemble which involve fixing three or more times, and in this way, the random processes can be described in finer and finer detail. However, if $\{x(t)\}$, $\{y(t)\}$ form a two-dimensional Gaussian distribution at a fixed value of t , then $\{x(t)\}$ and $\{y(t)\}$ are separately Gaussian.

The mean values and covariance functions listed above then provide a complete description of the underlying probability structure. For this reason, the main emphasis in this report is concerned only with these two statistical quantities and their relationships to power spectral density functions.

If the mean values $m_x(t)$ and $m_y(t)$, together with the covariance functions $\rho_x(\tau, t)$, $\rho_y(\tau, t)$, $\rho_{xy}(\tau, t)$, yield the same results for all fixed values of t [that is, are independent of time translations], then the random processes $\{x(t)\}$ and $\{y(t)\}$ are said to be weakly stationary. If all possible probability distributions involving $\{x(t)\}$ and $\{y(t)\}$ are independent of time translations, then the processes are said to be strongly stationary. Since the mean values and covariance functions are consequences only of the first-order and second-order probability distributions, it follows that the class of strongly stationary random processes is a subclass of the class of weakly stationary random processes. For Gaussian random processes, however, weak stationarity implies strong stationarity since all possible probability distributions may be derived from the mean values and covariance functions. Thus, for Gaussian random processes, these two stationary concepts coincide. Random processes which are not stationary are said to be nonstationary.

3.1 CORRELATION (COVARIANCE) STRUCTURE OF WEAKLY STATIONARY RANDOM PROCESSES

For weakly stationary random processes, $\{x(t)\}$, $\{y(t)\}$, which will be considered from henceforth, the mean values become constants independent of t , namely,

$$\begin{aligned} m_x &= \langle x(t) \rangle \\ m_y &= \langle y(t) \rangle \end{aligned} \tag{3.3}$$

Contrails

For simplicity, and without loss of generality, it will be assumed from henceforth (unless stated otherwise) that these mean values are zero.

The covariance functions for weakly stationary random processes are also independent of t , and with zero mean values, may be designated by

$$\begin{aligned}R_x(\tau) &= \langle x(t) x(t + \tau) \rangle & ; & \quad R_x(0) = \sigma_x^2 \\R_y(\tau) &= \langle y(t) y(t + \tau) \rangle & ; & \quad R_y(0) = \sigma_y^2 \\R_{xy}(\tau) &= \langle x(t) y(t + \tau) \rangle & ; & \quad R_{xy}(0) = \rho_{xy}\end{aligned} \tag{3.4}$$

where R is introduced instead of ρ to agree with engineering usage. For non-zero mean values, R is different from ρ . The quantities $R_x(\tau)$ and $R_y(\tau)$ are called the autocorrelation functions of $\{x(t)\}$ and $\{y(t)\}$, respectively, while $R_{xy}(\tau)$ is called the cross-correlation function between $\{x(t)\}$ and $\{y(t)\}$.

For arbitrary values of m_x and m_y , the covariance functions are related to the correlation functions by the equations

$$\begin{aligned}\rho_x(\tau) &= R_x(\tau) - m_x^2 \\ \rho_y(\tau) &= R_y(\tau) - m_y^2 \\ \rho_{xy}(\tau) &= R_{xy}(\tau) - m_x m_y\end{aligned} \tag{3.5}$$

Thus, correlation functions are identical with covariance functions when the mean values are zero. The reader should be careful not to confuse these un-normalized correlation functions with the correlation coefficient defined in Equation (2.14).

Contrails

From the stationary hypothesis, it follows that the autocorrelation functions $R_x(\tau)$ and $R_y(\tau)$ are even functions of τ , that is

$$\begin{aligned}R_x(-\tau) &= R_x(\tau) \\R_y(-\tau) &= R_y(\tau)\end{aligned}\tag{3.6}$$

while the cross-correlation function is neither odd nor even, but satisfies the relation

$$R_{xy}(-\tau) = R_{yx}(\tau)\tag{3.7}$$

An upper bound for the cross-correlation function is given by the inequality

$$\left| R_{xy}(\tau) \right|^2 \leq R_x(0) R_y(0)\tag{3.8}$$

the equal sign occurring only if the two processes are identically equal to each other and to a constant, a trivial case. Hence, a normalized correlation coefficient $\Gamma_{xy}(\tau)$ may be defined by

$$\Gamma_{xy}(\tau) = \frac{R_{xy}(\tau)}{\sqrt{R_x(0) R_y(0)}}\tag{3.9}$$

such that $\Gamma_{xy}(\tau)$ lies between -1 and +1. The coefficient $\Gamma_{xy}(\tau)$ measures the degree of linear dependence between $\{x(t)\}$ and $\{y(t)\}$ for a displacement τ in $\{y(t)\}$ relative to $\{x(t)\}$. The reader is cautioned not to confuse this concept with the previous definition in Equation (2.14).

In summary, the covariance structure of weakly stationary random processes $\{x(t)\}$, $\{y(t)\}$, assuming zero mean values,

may be described by the four correlation functions $R_x(\tau)$, $R_y(\tau)$, $R_{xy}(\tau)$ and $R_{yx}(\tau)$. These need be calculated only for values of $\tau \geq 0$ since symmetry properties listed above, Equations (3.6) and (3.7), yield results for $\tau < 0$.

3.2 SPECTRAL DECOMPOSITION OF STATIONARY RANDOM PROCESSES

The spectral decomposition of arbitrary random processes $\{x(t)\}$, a collection of time functions, (the superscript index k omitted for simplicity in notation), depends upon the requirement that each particular member of the random process $x(t)$, a sample time function, have a complex Fourier transform $X(f)$, where f denotes the frequency, (usually cps), such that

$$X(f) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi ft} dt \quad ; \quad j = \sqrt{-1} \quad (3.10a)$$

and conversely,

$$x(t) = \int_{-\infty}^{\infty} X(f) e^{j2\pi ft} df \quad (3.10b)$$

A sufficient set of conditions for this to occur is that $x(t)$ and its derivative $\dot{x}(t)$ be piecewise continuous in every finite interval (a, b) and that $|x(t)|$ be integrable over $(-\infty, \infty)$.

Similarly, every $y(t)$ from an arbitrary random process $\{y(t)\}$ must have a complex Fourier transform $Y(f)$ where

$$Y(f) = \int_{-\infty}^{\infty} y(t) e^{-j2\pi ft} dt \quad (3.11a)$$

Contrails

$$y(t) = \int_{-\infty}^{\infty} Y(f) e^{j2\pi ft} df \quad (3.11b)$$

Thus, the original pair of real random processes $\{x(t)\}$, $\{y(t)\}$ may be described in terms of two new complex random processes $\{X(f)\}$, $\{Y(f)\}$.

If $\{x(t)\}$ and $\{y(t)\}$ have zero mean values, which is assumed here, it follows that the ensemble averages

$$\langle X(f) \rangle = 0 = \langle Y(f) \rangle$$

Since $x(t)$ is real, it may be expressed in terms of the complex conjugate $\overline{X(f)}$ by

$$x(t) = \int_{-\infty}^{\infty} \overline{X(f)} e^{-j2\pi ft} df$$

Now, the square value

$$\begin{aligned} x^2(t) &= \int_{-\infty}^{\infty} \overline{X(f)} e^{-j2\pi ft} df \int_{-\infty}^{\infty} X(g) e^{j2\pi gt} dg \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \overline{X(f)} X(g) e^{j2\pi t(g-f)} dg df \end{aligned}$$

using a variable of integration g instead of f in the second integral to avoid confusion. The ensemble averaged value of $x^2(t)$ is thus given by

$$\langle x^2(t) \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \langle \overline{X(f)} X(g) \rangle e^{j2\pi t(g-f)} dg df \quad (3.12)$$

From separate considerations, the power spectral density function $S_x(f)$ associated with the random process $\{x(t)\}$, where f

ranges over $(-\infty, \infty)$, may be defined by the relation

$$\langle x^2(t) \rangle = \int_{-\infty}^{\infty} S_x(f) df \quad (3.13)$$

which indicates how $\langle x^2(t) \rangle$ is distributed over the doubly-infinite frequency range $(-\infty, \infty)$. In particular, $S_x(f) df$ represents the amount of "power" in $\langle x^2(t) \rangle$ lying in the frequency range $(f, f + df)$ so that $S_x(f)$ is real and non-negative for all f .

Since these last two equations for $\langle x^2(t) \rangle$ must be equivalent, one obtains

$$S_x(f) = \int_{-\infty}^{\infty} \langle \overline{X(f)} X(g) \rangle e^{j2\pi(g-f)} dg \quad (3.14)$$

which is satisfied by the requirement that

$$\langle \overline{X(f)} X(g) \rangle = S_x(f) \delta(f - g) \quad (3.15)$$

where $\delta(f - g)$ is a delta function defined by

$$\delta(f - g) = 0 \text{ when } f \neq g$$

$$\int_{-\infty}^{\infty} \delta(f - g) dg = 1 \quad (3.16)$$

The above discussion helps to justify the fact that one may prove from a deeper direct analysis (not developed here) that pairs of complex random variables $X(f)$, $X(g)$, $Y(f)$, $Y(g)$ satisfy the relations

$$\begin{aligned}
 \langle \overline{X(f)} X(g) \rangle &= S_x(f) \delta(f - g) \\
 \langle \overline{Y(f)} Y(g) \rangle &= S_y(f) \delta(f - g) \\
 \langle \overline{X(f)} Y(g) \rangle &= S_{xy}(f) \delta(f - g)
 \end{aligned}
 \tag{3.17}$$

where $S_x(f)$ and $S_y(f)$ are called the power spectral density functions of the random processes $\{x(t)\}$ and $\{y(t)\}$, respectively, while $S_{xy}(f)$ is called the cross-power spectral density function of $\{x(t)\}$ to $\{y(t)\}$. The frequency variable f ranges over $(-\infty, \infty)$.

It is now quite simple to derive the correspondance between these spectral density functions and the stationary correlation functions $R_x(\tau)$, $R_y(\tau)$, $R_{xy}(\tau)$. The results are

$$\begin{aligned}
 R_x(\tau) &= \int_{-\infty}^{\infty} S_x(f) e^{j2\pi f \tau} df \\
 R_y(\tau) &= \int_{-\infty}^{\infty} S_y(f) e^{j2\pi f \tau} df \\
 R_{xy}(\tau) &= \int_{-\infty}^{\infty} S_{xy}(f) e^{j2\pi f \tau} df
 \end{aligned}
 \tag{3.18}$$

proving that the concepts are Fourier transforms pairs. Consequently, the inverse relations yield

$$\begin{aligned}
 S_x(f) &= \int_{-\infty}^{\infty} R_x(\tau) e^{-j2\pi f \tau} d\tau \\
 S_y(f) &= \int_{-\infty}^{\infty} R_y(\tau) e^{-j2\pi f \tau} d\tau
 \end{aligned}
 \tag{3.19a}$$

$$S_{xy}(f) = \int_{-\infty}^{\infty} R_{xy}(\tau) e^{-j2\pi f\tau} d\tau \quad (3.19b)$$

No such simple relationships exist for nonstationary correlation functions.

From the symmetry properties of stationary correlation functions, it follows that

$$\begin{aligned} S_x(-f) &= S_x(f) \\ S_y(-f) &= S_y(f) \\ S_{xy}(-f) &= \overline{S_{xy}(f)} = S_{yx}(f) \end{aligned} \quad (3.20)$$

Since $S_x(f)$ and $S_y(f)$ are also real and non-negative for all f , this proves that power spectral density functions are real, non-negative, even functions of f , while cross-power spectral density functions are complex-valued functions of f .

The above relations for the real-valued power spectral density functions $S_x(f)$ and $S_y(f)$ may be simplified to

$$S_x(f) = \int_{-\infty}^{\infty} R_x(\tau) \cos 2\pi f\tau d\tau = 2 \int_0^{\infty} R_x(\tau) \cos 2\pi f\tau d\tau \quad (3.21)$$

$$S_y(f) = \int_{-\infty}^{\infty} R_y(\tau) \cos 2\pi f\tau d\tau = 2 \int_0^{\infty} R_y(\tau) \cos 2\pi f\tau d\tau$$

while

$$\begin{aligned} R_x(\tau) &= 2 \int_0^{\infty} S_x(f) \cos 2\pi f\tau df \\ R_y(\tau) &= 2 \int_0^{\infty} S_y(f) \cos 2\pi f\tau df \end{aligned} \quad (3.22)$$

Contrails

This last result shows that for the physically realizable positive frequency range where f varies only over $(0, \infty)$, the quantities $G_x(f)$ and $G_y(f)$ defined by

$$\begin{aligned} G_x(f) &= 2 S_x(f) \quad ; \quad 0 \leq f < \infty \\ G_y(f) &= 2 S_y(f) \quad ; \quad 0 \leq f < \infty \end{aligned} \tag{3.23}$$

represent the physically realizable power spectral density functions associated with $\{x(t)\}$ and $\{y(t)\}$ respectively. For mathematical calculations, the use of two-sided power spectral density functions $S_x(f)$, $S_y(f)$, defined over $(-\infty, \infty)$, and exponentials with imaginary exponents, frequently simplifies the analysis. In actual practice, one measures $G_x(f)$ and $G_y(f)$, defined over $(0, \infty)$. The reader is cautioned not to confuse these quantities. In terms of the physically realizable power spectral density functions $G_x(f)$ and $G_y(f)$, the correspondance with the stationary correlation functions $R_x(\tau)$ and $R_y(\tau)$ becomes

$$\begin{aligned} G_x(f) &= 4 \int_0^{\infty} R_x(\tau) \cos 2\pi f \tau \, d\tau \\ G_y(f) &= 4 \int_0^{\infty} R_y(\tau) \cos 2\pi f \tau \, d\tau \end{aligned} \tag{3.24}$$

while

$$\begin{aligned} R_x(\tau) &= \int_0^{\infty} G_x(f) \cos 2\pi f \tau \, df \\ R_y(\tau) &= \int_0^{\infty} G_y(f) \sin 2\pi f \tau \, df \end{aligned} \tag{3.25}$$

For the complex-valued cross-power spectral density function $S_{xy}(f)$, and the cross-correlation function $R_{xy}(\tau)$, one finds

$$S_{xy}(f) = \int_{-\infty}^{\infty} R_{xy}(\tau) \cos 2\pi f\tau \, d\tau - j \int_{-\infty}^{\infty} R_{xy}(\tau) \sin 2\pi f\tau \, d\tau \quad (3.26)$$

$$R_{xy}(\tau) = \int_{-\infty}^{\infty} S_{xy}(f) \cos 2\pi f\tau \, df + j \int_{-\infty}^{\infty} S_{xy}(f) \sin 2\pi f\tau \, df$$

Now, define the real and imaginary parts of $S_{xy}(f)$ by

$$S_{xy}(f) = C_{xy}(f) - j Q_{xy}(f) \quad (3.27)$$

where $C_{xy}(f)$ is called the co-spectrum of x to y and $Q_{xy}(f)$ is called the quad-spectrum of x to y. Observe that

$$C_{xy}(f) = \int_{-\infty}^{\infty} R_{xy}(\tau) \cos 2\pi f\tau \, d\tau = C_{xy}(-f) \quad (3.28)$$

$$Q_{xy}(f) = \int_{-\infty}^{\infty} R_{xy}(\tau) \sin 2\pi f\tau \, d\tau = -Q_{xy}(-f)$$

so that $C_{xy}(f)$ is a real-valued even function of f, while $Q_{xy}(f)$ is a real-valued odd function of f. From the symmetry relation $R_{xy}(\tau) = R_{yx}(-\tau)$, it follows that

$$C_{xy}(f) = \int_0^{\infty} [R_{xy}(\tau) + R_{yx}(\tau)] \cos 2\pi f\tau \, d\tau = \frac{1}{2} [S_{xy}(f) + S_{xy}(-f)] \quad (3.29)$$

$$Q_{xy}(f) = \int_0^{\infty} [R_{xy}(\tau) - R_{yx}(\tau)] \sin 2\pi f\tau \, d\tau = \frac{j}{2} [S_{xy}(f) - S_{xy}(-f)]$$

These relations will be needed later in Section 6.2.

From the above discussion, one sees that the spectral decomposition of the stationary random processes may be described by the three functions $S_x(f)$, $S_y(f)$, $S_{xy}(f)$, or by the four functions $S_x(f)$, $S_y(f)$, $C_{xy}(f)$ and $Q_{xy}(f)$, which need be calculated only for $f \geq 0$, since the symmetry properties, Equations (3.20) and 3.28), yield results for $f < 0$.

Analogous to the definition of the normalized correlation coefficient, Equation (3.9), the coherence function $\gamma_{xy}^2(f)$ is defined by

$$\gamma_{xy}^2(f) = \frac{|S_{xy}(f)|^2}{S_x(f) S_y(f)} \quad (3.30)$$

Since the cross-power spectral density function $S_{xy}(f)$ may be shown to satisfy the inequality

$$|S_{xy}(f)|^2 \leq S_x(f) S_y(f) \quad (3.31)$$

it follows that the coherence function lies between zero and one, and is a measure of the linear relationship between $\{x(t)\}$ and $\{y(t)\}$ at frequency f .

Certain authors prefer to use angular frequencies $\omega = 2\pi f$ in place of cyclical frequencies as is being followed in this report. This can lead to considerable confusion in factors of (2π) . A desire to preserve

$$R_x(0) = \int_{-\infty}^{\infty} S_x(f) df = \int_{-\infty}^{\infty} S_x(\omega) d\omega$$

shows that for consistency, one must have

$$S_x(f) = 2\pi S_x(\omega)$$

Similarly,

$$S_y(f) = 2\pi S_y(\omega)$$

$$S_{xy}(f) = 2\pi S_{xy}(\omega)$$
(3.32)

3.3 ERGODIC STATIONARY RANDOM PROCESSES

Consider two weakly stationary random processes $\{k_x(t)\}$ and $\{k_y(t)\}$, be variable, and two arbitrary sample functions $k_x(t)$ and $k_y(t)$, k fixed. These stationary random processes are said to be ergodic if the mean values and covariance (correlation) functions, (or spectral density functions), which are defined by certain ensemble averages, see Section 3.1, may be calculated by performing corresponding time averages on the arbitrary sample pair of functions. In this way, the underlying statistical structure of the weakly stationary random processes may be determined quite simply from an available sample pair without the need for collecting a considerable amount of data.

To be more specific, it is necessary to introduce some mathematical notation. The mean values of $k_x(t)$ and $k_y(t)$, k fixed, when computed by a time average are given by

$$k_{m_x} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} k_x(t) dt$$

$$k_{m_y} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} k_y(t) dt$$
(3.33)

Observe that the answer is no longer a function of t since t has been averaged out. In general, however, the answer is a function

of the particular record chosen — hence, the use of the parameter k .

The cross-covariance function between $k_x(t)$ and $k_y(t + \tau)$ when computed by a time average is defined by the expression

$$k_{\rho_{xy}}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \left[k_x(t) - k_{m_x} \right] \left[k_y(t + \tau) - k_{m_y} \right] dt \quad (3.34)$$

while the autocovariance functions are defined by

$$k_{\rho_x}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \left[k_x(t) - k_{m_x} \right] \left[k_x(t + \tau) - k_{m_x} \right] dt \quad (3.35)$$

$$k_{\rho_y}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \left[k_y(t) - k_{m_y} \right] \left[k_y(t + \tau) - k_{m_y} \right] dt$$

These quantities should now be compared with the previously defined ensemble mean values m_x , m_y , and ensemble covariance functions $\rho_x(\tau)$, $\rho_y(\tau)$, $\rho_{xy}(\tau)$ for stationary random processes, Equations (3.3) and 3.5). If it turns out that independent of k , (with the possible exception of a set of sample functions of zero probability)

$$\begin{aligned} k_{m_x} &= m_x \\ k_{m_y} &= m_y \end{aligned} \quad (3.36a)$$

$${}^k \rho_x(\tau) = \rho_x(\tau)$$

$${}^k \rho_y(\tau) = \rho_y(\tau) \quad (3.36b)$$

$${}^k \rho_{xy}(\tau) = \rho_{xy}(\tau)$$

then the random processes $\{x(t)\}$ and $\{y(t)\}$ are said to be weakly ergodic. If all ensemble averaged statistical properties of $\{x(t)\}$, $\{y(t)\}$, not just the means and covariances, are deducible from corresponding time averages, then the random processes are said to be strongly ergodic. Thus, strong ergodicity implies weak ergodicity, but not conversely. No distinction between these concepts exists for Gaussian random processes.

For an arbitrary random process to be ergodic, it must first of all be stationary. Secondly, each sample function must be representative of all the others in the sense described above so that it doesn't matter which particular sample function is used in the time averaging calculations. This restriction serves to eliminate many stationary random processes from being ergodic. To repeat, a stationary random process may or may not be ergodic.

There are two important classes of random processes which one can state in advance will be ergodic. The first ergodic class is the class of stationary Gaussian random processes whose power spectral density functions are continuous. That is, no sharp lines (delta functions) appear in the power spectra corresponding to finite amounts of power at discrete frequencies. The second ergodic class (a special case of the first class) is the class of stationary Gaussian Markoff processes, a Markoff process being defined as one whose relationship to the past does not extend beyond the immediately preceding observation. The autocorrelation function of a stationary Gaussian Markoff process may be shown to be of a simple exponential form.

4. MEASUREMENT OF MEAN VALUES

Consider, first of all, the properties that should be possessed by any set of estimates. Let $\{^k x(t)\}$, $-\infty < t < \infty$, be a real stationary random process where $k = 1, 2, 3, \dots$ (perhaps even uncountable) denotes the different numbers of the random process. Suppose that z is the true value of an unknown parameter of the random process $\{^k x(t)\}$, e. g., its mean value or its power spectrum. Suppose that $^k z(T)$ is an estimate of z obtained from a measurement made on a particular finite sample $^k x(t)$ of the random process extending only over a finite time period from $0 \leq t \leq T$. How should these different possible $^k z(t)$ be related to z . For ease of notation, the superscript index k will henceforth be omitted, and expected values (averages) are tacitly understood to be taken over this index.

It seems fairly obvious to start with that, on the average, $z(T)$ should yield the true value z . In other words, for fixed T , if one takes an ensemble average over all of the possible $z(T)$ that might occur, then this procedure should give z without any error. A set of estimates having this property is said to be unbiased. To be precise,

Definition 1. A set of estimates $\{z(T)\}$ is said to be unbiased if, independent of T , the expected value is the true value, that is,

$$E z(T) = z \quad (4.1)$$

When this occurs, $z(T)$ is also said to be an unbiased estimate of z .

For a fixed T , the mere fact that the expected value of a set of estimates $\{z(T)\}$ is unbiased does not imply that any particular $z(T)$ will lie close to the true value z . There may in fact be widespread deviations from the true value. Furthermore, it may

happen that these deviations do not decrease as T is increased. To analyze these cases, for a fixed value of T , the mean square error is defined as the expected value over the set of estimates of the square of the deviations from the true value, namely,

$$E [z(T) - z]^2 \quad (4.2)$$

The expected value above represents an ensemble average over the various possible $[z(T) - z]^2$ occurring from different finite samples $z(T)$ of the random process. It appears highly desirable, from a physical point of view, to require that this mean square error should approach zero as T becomes large. Then, for large T , any particular estimate of $z(T)$ would necessarily tend to closely approximate the true value z . Estimates having this desired property are said to be consistent. In more precise terms, one writes

Definition 2. A set of estimates $\{z(T)\}$ is said to be consistent if

$$\lim_{T \rightarrow \infty} E [z(T) - z]^2 = 0 \quad (4.3)$$

When this occurs, $z(T)$ itself is also said to be a consistent estimate of z .

Observe that the mean square error

$$\begin{aligned} E [z(T) - z]^2 &= E [z(T) - E z(T) + E z(T) - z]^2 \\ &= E [z(T) - E z(T)]^2 + [E z(T) - z]^2 \end{aligned} \quad (4.4)$$

since the usual middle term has a factor equal to zero, namely,

$$E [z(T) - E z(T)] = 0$$

Contrails

Thus, the mean square error is the sum of two parts: the variance (or square of the random error) of the estimate as given by

$$\sigma^2[z(T)] = E[z(T) - E z(T)]^2 = E[z^2(T)] - [E z(T)]^2 \quad (4.5)$$

and the square of the bias (or systematic error) of the estimate as given by

$$[b z(T)]^2 = [E z(T) - z]^2 \quad (4.6)$$

It will be demonstrated in this report that great care is required to insure that both the variance and the bias will approach zero as T becomes large when estimating (i. e., measuring) the power spectrum and cross-power spectrum.

The following discussion is now concerned with estimating the mean values m_x , m_y of a pair of (weakly) stationary random processes $\{x(t)\}$, $\{y(t)\}$ by performing a finite time average on arbitrary continuous sample functions $x(t)$ and $y(t)$, which are known only for a finite time interval $0 \leq t \leq T$. By assuming certain commonly satisfied conditions for the autocorrelation functions of the random processes, it is shown that the estimates in question are unbiased and consistent. Effects due to sampling the functions at equispaced intervals of time and using digital rather than continuous data will be taken up later in Section 7.

The same analysis covers measurements of either m_x or m_y . Consider $z(t)$ as representing $x(t)$ or $y(t)$. Suppose that a single sample record $z(t)$ from a stationary random process $\{z(t)\}$ is averaged only over a finite time T . Define the sample mean value estimate by

$$m(T) = \frac{1}{T} \int_0^T z(t) dt \quad (4.7)$$

Then

$$\begin{aligned} E m(T) &= E \left[\frac{1}{T} \int_0^T z(t) dt \right] = \frac{1}{T} \int_0^T E z(t) dt \\ &= \frac{1}{T} \int_0^T m dt = m \end{aligned} \quad (4.8)$$

where the true mean value $m = E z(t)$ is independent of t . Hence, $m(T)$ is an unbiased estimate of m .

For simplicity and without loss of generality, it will be assumed unless stated otherwise that the mean value $m = 0$. Now, the mean square error over the set of estimates $\{m(T)\}$ becomes

$$E [m(T) - m]^2 = E [m(T)]^2 - m^2 = E [m(T)]^2 \quad (4.9)$$

where

$$E [m(T)]^2 = \frac{1}{T^2} \int_0^T \int_0^T E [z(u) z(v)] du dv \quad (4.10)$$

By definition, the autocorrelation function $R(\tau)$ for a stationary random process $\{z(t)\}$ is defined by

$$R(\tau) = E [z(t) z(t + \tau)] \quad (4.11)$$

From the stationary hypothesis, it follows that $R(\tau)$ is independent of t , and an even function of τ with a maximum at $\tau = 0$. It will be assumed that $R(\tau)$ is continuous and finite for all values of τ , and that all periodic components in $R(\tau)$ have been removed at the outset.

Contrails

Additional integrability properties satisfied by $R(\tau)$, which will be needed later, are assumed to be

$$\int_{-\infty}^{\infty} |R(\tau)| d\tau < \infty \quad ; \quad \int_{-\infty}^{\infty} R^2(\tau) d\tau < \infty$$
$$\int_{-\infty}^{\infty} |\tau| |R(\tau)| d\tau < \infty \quad ; \quad \int_{-\infty}^{\infty} |\tau| R^2(\tau) d\tau < \infty \quad (4.12)$$

These conditions are generally satisfied in nature, e. g., consider the exponential function $R(\tau) = \exp(-\mu|\tau|)$ where $\mu > 0$.

In terms of an arbitrary autocorrelation function $R(\tau)$, the mean square error, Equations (4.9) and (4.10), becomes

$$E[m(T)]^2 = \frac{1}{T^2} \int_0^T \int_0^T R(u-v) du dv$$
$$= \frac{1}{T} \int_{-T}^T \left(1 - \frac{|\tau|}{T}\right) R(\tau) d\tau \quad (4.13)$$

The second expression occurs from the first by substituting $\tau = u - v$, $d\tau = du$, and reversing the orders of integration between τ and v . Now,

$$\lim_{T \rightarrow \infty} T E[m(T)]^2 = \int_{-\infty}^{\infty} R(\tau) d\tau < \infty \quad (4.14)$$

provided that $R(\tau)$ and $\tau R(\tau)$ are absolutely integrable over $(-\infty, \infty)$ to justify passage to the limit inside the integral sign. In particular, Equation (4.13) shows that, for large T , the mean square error

Contrails

$$E[m(T)]^2 \approx \frac{1}{T} \int_{-\infty}^{\infty} R(\tau) d\tau \quad (4.15)$$

Hence $E[m(T)]^2$ approaches zero as T approaches infinity, proving that $m(T)$ is a consistent estimate of m .

5. MEASUREMENT OF AUTOCORRELATION AND CROSS-CORRELATION FUNCTIONS

The next statistical quantities of interest are the autocorrelation functions $R_x(\tau)$, $R_y(\tau)$, and the cross-correlation function $R_{xy}(\tau)$. As in Section 4, the mean values m_x and m_y are assumed to be zero. For continuous data $x(t)$, $y(t)$ which exists only for $0 \leq t \leq T$, the sample cross-correlation function estimate $R_{xy}(\tau, T)$ will be defined by

$$R_{xy}(\tau, T) = \begin{cases} \frac{1}{T-\tau} \int_0^{T-\tau} x(t) y(t + \tau) dt & ; \quad 0 \leq \tau < T \\ \frac{1}{T-|\tau|} \int_{|\tau|}^T x(t) y(t + \tau) dt & ; \quad -T < \tau \leq 0 \end{cases} \quad (5.1)$$

To avoid use of absolute value signs, τ will be considered positive from henceforth and the reader should supply required similar separate proofs for negative τ . The sample autocorrelation function estimates $R_x(\tau, T)$ and $R_y(\tau, T)$ are merely special cases when the two records coincide, namely,

$$R_x(\tau, T) = \frac{1}{T-\tau} \int_0^{T-\tau} x(t) x(t + \tau) dt \quad ; \quad 0 \leq \tau < T$$

$$R_y(\tau, T) = \frac{1}{T-\tau} \int_0^{T-\tau} y(t) y(t + \tau) dt \quad ; \quad 0 \leq \tau < T \quad (5.2)$$

Thus, by analyzing the cross-correlation function estimate, one derives results also for the autocorrelation function estimates.

An ensemble average over the set of possible estimates $\{R_{xy}(\tau, T)\}$ yields

$$\begin{aligned} E[R_{xy}(\tau, T)] &= \frac{1}{T-\tau} \int_0^{T-\tau} E[x(t) y(t + \tau)] dt \\ &= \frac{1}{T-\tau} \int_0^{T-\tau} R_{xy}(\tau) dt = R_{xy}(\tau) \end{aligned} \quad (5.3)$$

Hence $R_{xy}(\tau, T)$ is an unbiased estimate of $R_{xy}(\tau)$.

The mean square error here is given by

$$\begin{aligned} E[R_{xy}(\tau, T) - R_{xy}(\tau)]^2 &= E[R_{xy}^2(\tau, T)] - R_{xy}^2(\tau) \\ &= \frac{1}{(T-\tau)^2} \int_0^{T-\tau} \int_0^{T-\tau} \left\{ E[x(u) y(u + \tau) x(v) y(v + \tau)] - R_{xy}^2(\tau) \right\} dv du \end{aligned} \quad (5.4)$$

At this point, in order both to simplify the later mathematical analysis and also to agree with many physical cases of interest, it will be assumed that the random processes $\{x(t)\}$, $\{y(t)\}$ are jointly Gaussian for any set of fixed times. This restriction may be removed by substituting certain integrability conditions on the non-Gaussian parts of the random processes without altering in any essential way the results to be derived. When $\{x(t)\}$, $\{y(t)\}$ are jointly Gaussian, it follows that $\{x(t)\}$ and $\{y(t)\}$ are separately Gaussian.

For Gaussian stationary random processes with zero mean values, the statistical expression

$$\begin{aligned} E \left[x(u) y(u + \tau) x(v) y(v + \tau) \right] &= R_{xy}^2(\tau) + R_x(v - u) R_y(v - u) \\ &\quad + R_{xy}(v - u + \tau) R_{yx}(v - u - \tau) \end{aligned} \quad (5.5)$$

Hence, the mean square error

$$\begin{aligned} E \left[R_{xy}(\tau, T) - R_{xy}(\tau) \right]^2 &= \frac{1}{(T-\tau)^2} \int_0^{T-\tau} \int_0^{T-\tau} \left[R_x(v-u) R_y(v-u) \right. \\ &\quad \left. + R_{xy}(v-u+\tau) R_{yx}(v-u-\tau) \right] dv du \\ &= \frac{1}{T-\tau} \int_{-(T-\tau)}^{T-\tau} \left(1 - \frac{|y|}{T-\tau} \right) \left[R_x(y) R_y(y) + R_{xy}(y+\tau) R_{yx}(y-\tau) \right] dy \end{aligned} \quad (5.6)$$

The second expression occurs from the first by letting $\gamma = v - u$, $dy = dv$, and then reversing the order of integration between γ and u . Now,

$$\begin{aligned} \lim_{T \rightarrow \infty} T E \left[R_{xy}(\tau, T) - R_{xy}(\tau) \right]^2 \\ = \int_{-\infty}^{\infty} \left[R_x(\gamma) R_y(\gamma) + R_{xy}(\gamma + \tau) R_{yx}(\gamma - \tau) \right] d\gamma < \infty \end{aligned} \quad (5.7)$$

assuming $R_x(\gamma) R_y(\gamma)$ and $\gamma R_x(\gamma) R_y(\gamma)$ are absolutely integrable $(-\infty, \infty)$. This proves that $R_{xy}(\tau, T)$ is a consistent estimate of $R_{xy}(\tau)$ which for large T has a mean square error given by

$$E \left[R_{xy}(\tau, T) - R_{xy}^2(\tau) \right]^2 \approx \frac{1}{T} \int_{-\infty}^{\infty} \left[R_x(\gamma) R_y(\gamma) + R_{xy}(\gamma + \tau) R_{yx}(\gamma - \tau) \right] d\gamma$$

(5.8)

To summarize the work thus far, the analysis in Sections 4 and 5 has stated commonly satisfied mathematical conditions such that:

- (a) A set of estimates $\{m(T)\}$, see Equation (4.7), for measuring the mean value from continuous data is both unbiased and consistent, with mean square error for large T given by Equation (4.15).
- (b) A set of estimates $\{R_{xy}(\tau, T)\}$, see Equation (5.1), for measuring the cross-correlation function from continuous data is both unbiased and consistent, with mean square error for large T given by Equation (5.8). Autocorrelation function estimates are obtained by merely letting $x(t) = y(t)$.

The main efforts of this report devoted to power spectra and cross-power spectra measurements will now be developed in Sections 6, 7 and 8. The final Section 9 treats some important physical applications.

6. MEASUREMENT OF POWER SPECTRA AND CROSS-POWER SPECTRA : CONTINUOUS DATA

For stationary random processes with zero mean values, the real-valued cross-correlation function $R_{xy}(\tau)$ and the two-sided complex-valued cross-power spectral density function $S_{xy}(f)$, which is defined for $-\infty \leq f \leq \infty$, are related by Equations (3.18), (3.19) and (3.27), namely,

$$S_{xy}(f) = \int_{-\infty}^{\infty} R_{xy}(\tau) e^{-j2\pi f\tau} d\tau = C_{xy}(f) - jQ_{xy}(f)$$

$$R_{xy}(\tau) = \int_{-\infty}^{\infty} S_{xy}(f) e^{j2\pi f\tau} df \tag{6.1}$$

As special cases of the above, the real-valued autocorrelation functions $R_x(\tau)$, $R_y(\tau)$ yield the real-valued two-sided power spectral density functions $S_x(f)$, $S_y(f)$ through the relations

$$S_x(f) = \int_{-\infty}^{\infty} R_x(\tau) e^{-j2\pi f\tau} d\tau$$

$$S_y(f) = \int_{-\infty}^{\infty} R_y(\tau) e^{-j2\pi f\tau} d\tau \tag{6.2}$$

The problem at hand is to estimate $S_{xy}(f)$, $S_x(f)$ and $S_y(f)$ from data which is known only for a finite time interval. In order to estimate in a physical device the complex-valued function $S_{xy}(f)$, it is necessary to estimate its real-valued components, namely, the co-spectrum $C_{xy}(f)$ and the quad-spectrum $Q_{xy}(f)$. Since $S_x(f)$ and $S_y(f)$ are real-valued functions, their estimation is easier to

accomplish and to explain than $S_{xy}(f)$. Consequently, the discussion to follow begins with power spectrum measurements, after which cross-power spectrum measurements will be taken up in Section 6.2.

6.1 POWER SPECTRA MEASUREMENTS

A schematic picture of a general filter device for estimating the power spectral density function associated with a single random record, say $x(t)$, is displayed in Figure 6.1 below.

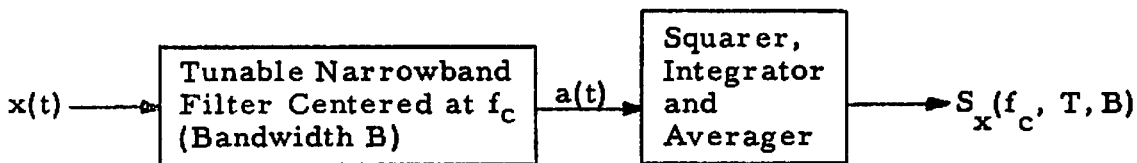


Figure 6.1 Filter Device for Measuring Power Spectrum

The input random record $x(t)$ is assumed to be of finite time-length T , and to be drawn from a stationary random process with zero mean value. The tunable narrowband discriminating filter is assumed to have a finite nonzero constant bandwidth B centered at a frequency f_c which may be varied over the frequency range of interest. It turns out that in order to obtain a consistent estimate of $S_x(f)$, one must introduce a filtering procedure which averages over a band of frequencies. The final filter output quantity $S_x(f, T, B)$ describes the time average of $x^2(t)$ in terms of its frequency components lying inside the frequency band $f_c - (B/2)$ to $f_c + (B/2)$, divided by the bandwidth B . Analog equipment of this type appears in Reference [12].

The output quantity $S_x(f_c, T, B)$ is a smoothing-over-frequency estimate of the true power spectral density function $S_x(f)$ at $f = f_c$ which would be associated with input records of infinite length and

bandwidths of zero width. The quantity $S_x(f) df$ is the infinitely long time average of the product $x^2(t)$ from frequencies lying between f and $(f + df)$. The total time average of $x^2(t)$ over all frequencies is obtained by integrating $S_x(f) df$ from $-\infty$ to $+\infty$. Because of its relationship to power dissipated in a unit resistance by a current $x(t)$, the time average of $x^2(t)$ may be considered as the "average power" in $x(t)$, and is the main justification for calling $S_x(f)$ a power spectral density function.

In an actual physical device, the bandwidth B is not zero and the record lengths T are not infinite. It is important to be able to predict within established levels of confidence how closely an actual measurement $S_x(f, T, B)$ will approximate the desired true measurement $S_x(f)$. This problem will now be discussed and some of the main analytical results will be stated.

Let the frequency response function of the narrowband filter centered at f_c be of an idealized form (see Figure 6.2)

$$\begin{aligned}
 H(f) &= \sqrt{1/2B} \quad \text{for } |f \pm f_c| \leq B/2 \\
 &= 0 \quad \text{otherwise}
 \end{aligned}
 \tag{6.3}$$

where the full bandwidth $B = B(T)$ is a function of T to be specified later. For definiteness, assume that $f_c > (B/2)$.

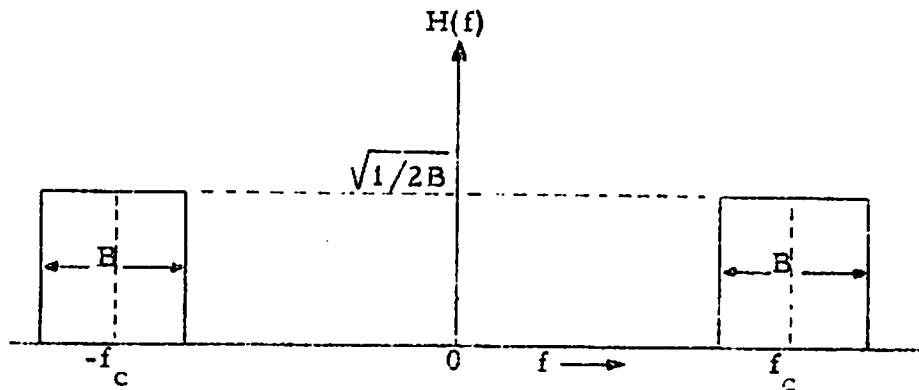


Figure 6.2 Ideal Narrowband Filter

Contrails

The inverse Fourier transform of $H(f)$ is the weighting function of the filter denoted by $h(t)$, that is,

$$h(t) = \int_{-\infty}^{\infty} H(f) e^{j2\pi ft} df$$

For a filter to be physically realizable, it is necessary that $h(t) = 0$ for $t < 0$.

In terms of an arbitrary $h(t)$, the corresponding frequency response function

$$H(f) = \int_{-\infty}^{\infty} h(t) e^{-j2\pi ft} dt$$

It follows for real $h(t)$ that $H(-f) = \overline{H(f)}$, the complex conjugate of $H(f)$. This is the reason why the ideal narrowband filter of Figure 6.2 where $H(f) = \overline{H(f)}$ has a theoretical pass-band in the negative frequency region with $H(-f) = H(f)$. Note also that $H(f)$ is defined so that

$$\int_{-\infty}^{\infty} |H(f)|^2 df = 1$$

In words, $|H(f)|^2$ has unit area over the doubly infinite frequency range for any bandwidth B .

For an input $x(t)$, the filter output $a(t)$ is given by

$$a(t) = \int_0^T x(u) h(t - u) du = \int_{-\infty}^{\infty} x(u) h(t - u) du \quad ; \quad t \geq 0$$

since $x(u)$ is zero outside the range $(0, T)$. This output is now squared, then integrated over all positive time and average by T ,

Contrails

to yield as a smoothed estimate for the time average of $x^2(t)$ in the bandwidth B,

$$S_x(f_c, T, B) = \frac{1}{T} \int_0^{\infty} a^2(t) dt = \frac{1}{T} \int_{-\infty}^{\infty} a^2(t) dt$$

since $a(t)$ is zero for negative t . For large T , observe that

$$S_x(f_c, T, B) \approx \frac{1}{T} \int_0^T a^2(t) dt$$

By tuning the filter to different center frequencies f_c , one obtains $S_x(f, T, B)$ for all f .

Further analysis (not given here) shows that

$$S_x(f_c, T, B) = \int_{-\infty}^{\infty} |H(f)|^2 S_x(f, T) df \quad (6.4)$$

where

$$S_x(f, T) = \int_{-\infty}^{\infty} \tilde{R}_x(\tau, T) e^{-j2\pi f\tau} d\tau = \int_{-T}^T \tilde{R}_x(\tau, T) e^{-j2\pi f\tau} d\tau \quad (6.5)$$

Here, the inverse Fourier transform of $S_x(f, T)$ yields a special autocorrelation function estimate

$$\begin{aligned} \tilde{R}_x(\tau, T) &= \int_{-\infty}^{\infty} S_x(f, T) e^{j2\pi f\tau} df \\ &= \frac{1}{T} \int_0^{T-|\tau|} x(t) x(t+|\tau|) dt \quad \text{for } |\tau| \leq T \\ &\quad \text{(otherwise zero)} \end{aligned} \quad (6.6)$$

which is not the same as Equation (5.1). Observe that the expected value of $\tilde{R}_x(\tau, T)$, as defined above, is given by

$$E[\tilde{R}_x(\tau, T)] = \left(1 - \frac{\tau}{T}\right) R_x(\tau) \quad (6.7)$$

Hence, this expression for $\tilde{R}_x(\tau, T)$ is not an unbiased estimate of $R_x(\tau)$.

On setting $\tau = 0$, one derives the relations

$$\tilde{R}_x(0, T) = \int_{-\infty}^{\infty} S_x(f, T) df = \frac{1}{T} \int_0^T x^2(t) dt \quad (6.8)$$

which shows how $S_x(f, T)$ distributes the "power" in $x^2(t)$ over the doubly infinite frequency range from $-\infty$ to $+\infty$.

An alternative way of expressing $S_{xy}(f, T)$ is to define an unbiased autocorrelation estimate by [see Equation (5.1)],

$$R_x(\tau, T) = \frac{1}{T-|\tau|} \int_0^{T-|\tau|} x(t) x(t+|\tau|) dt = \frac{T}{T-|\tau|} \tilde{R}_x(\tau, T) \quad (6.9)$$

Now

$$S_x(f, T) = \frac{1}{T} \int_{-T}^T (T-|\tau|) R_x(\tau, T) e^{-j2\pi f\tau} d\tau \quad (6.10)$$

These equations, together with Equation (6.4), provide the basis for the discrete digital technique to be described later in Section 7.

6.1.1 Analysis of Bias

For the bias problem, one may prove that $S_x(f_c, T, B)$, (as obtained from Figure 6.1), provides an asymptotically unbiased estimate of $S_x(f_c)$ as $T \rightarrow \infty$ if the bandwidth $B \rightarrow 0$ as $T \rightarrow \infty$. This condition on B as a function of T is therefore assumed.

From Equations (6.4) and (6.5),

$$S_x(f_c, T, B) = \int_{-\infty}^{\infty} |H(f)|^2 \left[\int_{-T}^T \tilde{R}_x(\tau, T) e^{-j2\pi f \tau} d\tau \right] df \tag{6.11}$$

By definition, Equation (4.6), the bias term is given by

$$b[S_x(f_c, T, B)] = E[S_x(f_c, T, B)] - S_x(f_c) \tag{6.12}$$

At this point, some detailed mathematical analysis must be carried out which is not developed here. The final result is expressed in the important asymptotic formula below which assumes that $R_x(\tau)$, $\tau R_x(\tau)$ and $\tau^2 R_x(\tau)$ are all absolutely integrable functions over $(-\infty, \infty)$. This result is that at any frequency f ,

$$\lim_{T \rightarrow \infty} B^{-2} \left| b[S_x(f, T, B)] \right| \approx \frac{1}{24} \left| S_x''(f) \right| \tag{6.13}$$

where $S_x''(f)$ is the second derivative of $S_x(f)$ with respect to f as given by

$$S_x''(f) = -4\pi^2 \int_{-\infty}^{\infty} \tau^2 R_x(\tau) e^{-j2\pi f \tau} d\tau \tag{6.14}$$

From the above, it follows that the bias term approaches zero as T approaches infinity provided that B approaches zero.

In fact, for large T,

$$\left| b \left[S_x(f, T, B) \right] \right| \approx \frac{B^2}{24} \left| S_x''(f) \right| \quad (6.15)$$

No apparent limitation exists from the above formula on how rapidly B should approach zero as T approaches infinity. It will be shown in the next section that B may not approach zero too rapidly if the variance in the estimate is to be small.

6.1.2 Analysis of Variance

For the variance problem, one may prove that the estimate $S_x(f_c, T, B)$, (as obtained from Figure 6.1), has a variance which approaches zero as $T \rightarrow \infty$ if the product $BT \rightarrow \infty$. Since the bias approaches zero as $T \rightarrow \infty$ only if $B \rightarrow 0$, these two statements taken together imply that B should approach zero slower than T approaches infinity.

By definition, Equation (4.5), the variance is given by

$$\sigma^2 \left[S_x(f_c, T, B) \right] = E \left[S_x(f_c, T, B) - E S_x(f_c, T, B) \right]^2 \quad (6.16)$$

After a considerable amount of careful mathematical analysis, one may derive the following important result. At any frequency f,

$$\begin{aligned} \lim_{T \rightarrow \infty} BT \sigma^2 \left[S_x(f, T, B) \right] &\approx S_x^2(f) \quad ; \quad f \neq 0 \\ &\approx 2 S_x^2(0) \quad ; \quad f = 0 \end{aligned} \quad (6.17)$$

Thus, for large T,

$$\begin{aligned} \sigma^2 [S_x(f, T, B)] &\approx (1/BT) S_x^2(f) \quad ; \quad f \neq 0 \\ &\approx (2/BT) S_x^2(0) \quad ; \quad f = 0 \end{aligned} \tag{6.18}$$

These equations show that the variance approaches zero as $T \rightarrow \infty$ is provided that $BT \rightarrow \infty$. This result combined with the previous result for the bias term gives the two parts required for a mean square error analysis of power spectrum (and cross-power spectrum) measurements. Observe that at the zero frequency point, $f = 0$, the right-hand side is increased by a factor of two over the general result which is valid for $f \neq 0$. In the sequel, formulas will refer to cases where $f \neq 0$, and should be modified by this factor of two if $f = 0$.

6.1.3 Mean Square Error

The mean square error of the power spectrum estimate $S_x(f, T, B)$ at any frequency $f \neq 0$ is given by the expression, see Equation (4.4),

$$\begin{aligned} E [S_x(f, T, B) - S_x(f)]^2 &= \sigma^2 [S_x(f, T, B)] + b^2 [S_x(f, T, B)] \\ &\approx \frac{S_x^2(f)}{BT} + \left(\frac{B^2 |S_x''(f)|}{24} \right)^2 \quad \text{for large T} \end{aligned} \tag{6.19}$$

using Equations (6.15) and (6.18). It is clear that the mean square error approaches zero as $T \rightarrow \infty$ if B, considered as a function of T, is restricted so that $B \rightarrow 0$ and $BT \rightarrow \infty$.

For example, suppose

$$B = cT^{\alpha-1} \quad ; \quad c > 0, \quad 0 < \alpha < 1 \quad (6.20)$$

Here, $B \rightarrow 0$ as $T \rightarrow \infty$, and $BT \rightarrow \infty$ as $T \rightarrow \infty$.

Equation (6.19) above is one of the more important statistical results in this report since it indicates the mean square error to be expected in estimating $S_x(f)$ [or $S_y(f)$] using any given finite B and finite T . Further analysis of this result will be taken up in Section 8.

6.1.4 Frequency Resolution

Another important practical question in power spectrum measurements is to determine how closely estimates should be taken along any frequency range of interest. It is clear that if these points are spaced too closely together, the results would be highly correlated and considerable extra unnecessary work would be involved. On the other hand, for points spaced too far apart, considerable information may be lost. It is important to determine the smallest frequency interval Δf that can be resolved in power spectrum measurements in the sense that estimates taken at this frequency interval apart will be essentially uncorrelated. For the idealized narrowband filters shown in Figure 6.2, a choice of $\Delta f = B$ represents the minimum resolution attainable. Two different peaks in a power spectrum which are less than B cps apart may be blurred together and not distinguished from one another. Two peaks which are further than B cps apart, however, would be separated.

In actual practice, since realizable filters do not have sharp cut-off edges as pictured in Figure 6.2, a more reasonable figure to use for the resolution is $2B$ cps instead of B cps.

Thus, for high resolution, the bandwidth B should be made as small as possible. This is also desirable, as stated earlier, in order for the estimates to have a low bias. Low bias and high

resolution are consequently complementary properties, both being consequences of narrowband filtering. From the point of view of reducing the variance in the estimates, however, for a given record length T , the bandwidth B should be made as large as possible since the variance is inversely proportional to the BT product. Thus, the choice of B is quite critical. If T is not restricted in length, then it is possible to attain arbitrarily high resolution and small bias as well as arbitrarily low variance. To accomplish this objective, one should let B approach zero and T approach infinity, but in such a way that B approaches zero at a slower rate than T approaches infinity.

The quantity $2BT$ represents the number of degrees of freedom associated with a finite record T seconds long and restricted to a frequency bandwidth $(0, B)$ cps wide, in the sense that the record can be reconstructed from its samples taken $(1/2B)$ seconds apart on the time scale. Thus, $2BT$ numbers completely determine a record which is T seconds long.

6.1.5 Correction for Mean and Linear Trend

The previous analysis has assumed that the input random record $x(t)$ is a sample member from a stationary random process with zero mean value. If the mean value is not zero, then the power spectral density function will exhibit a large peak (theoretically infinite) at zero frequency. Considerable distortions will occur in measurements of the power spectra curve at low frequencies by feeding the record directly into the analog device of Section 6.1 without correcting for this non-zero mean value.

A second correction may be needed to subtract out a slowly varying linear trend (i. e. non-zero slope of $x(t)$ with respect to time) about which the random record may be oscillating. This may be due to the recording equipment, or to an actual change in the random record over a long observation time. Whatever the cause,

it is clear that a better estimate of the power spectra curve can be obtained by taking proper account of this linear trend in the data.

Let $\xi(t)$ represent an input random record from a random process $\{\xi(t)\}$ which may need to be corrected for a non-zero mean value and for a linear trend. In particular, suppose that

$$\xi(t) = m_x + \alpha_x \left(t - \frac{T}{2} \right) + x(t) \quad ; \quad 0 \leq t \leq T \quad (6.21)$$

where m_x denotes the mean value of $\xi(t)$ over its length $[0, T]$, the parameter α_x denotes the average slope of the record $\xi(t)$ with respect to time t , and the final term $x(t)$ represents a sample record from a stationary random process $\{x(t)\}$ with zero mean value and zero slope. Observe that if m_x and α_x equal zero, then $\xi(t)$ becomes $x(t)$.

The parameters m_x and α_x may be estimated from $\xi(t)$ by the easily derived formulas

$$m_x = \frac{1}{T} \int_0^T \xi(t) dt \quad (6.22)$$

$$\alpha_x = \frac{1}{(T/3)(2T/3)} \left[\int_{(2T/3)}^T \xi(t) dt - \int_0^{(T/3)} \xi(t) dt \right] \quad (6.23)$$

These relations lead to a simple analog device for determining $x(t)$ from $\xi(t)$ as sketched in Figure 6.3. This output $x(t)$ can now be fed into the power spectral analyzer circuit of Figure 6.1 so as to yield estimates of $S_x(f)$.

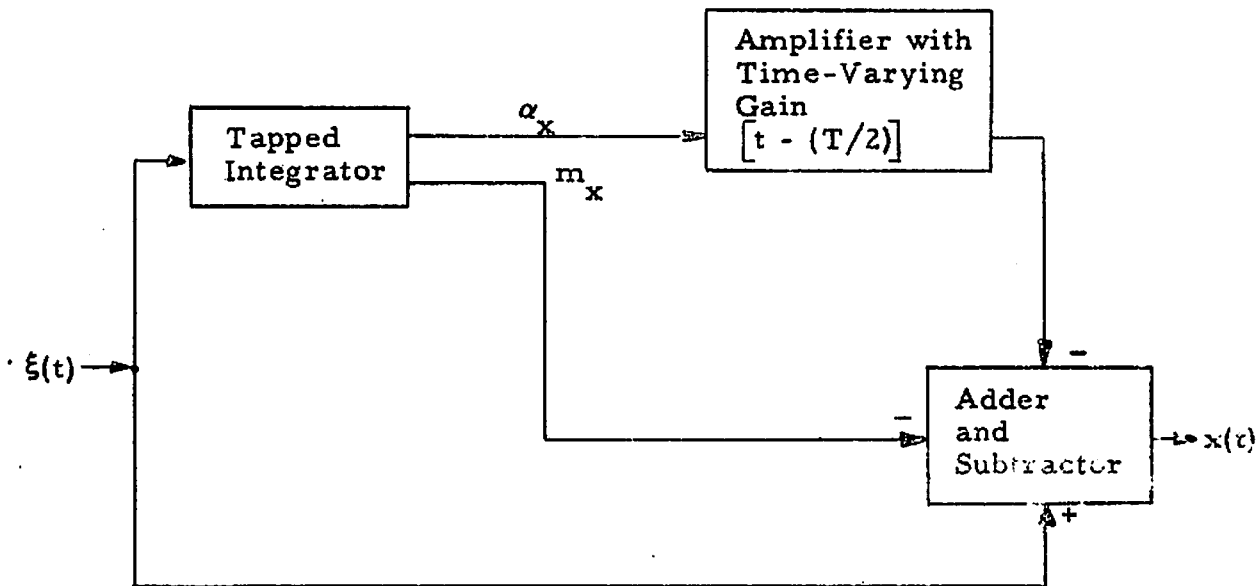


Figure 6.3 Circuit for Removing Mean Value and Linear Trend

For later use in digital calculations, as well as for its own interest, one should prove that the autocorrelation function estimate, see Equation (6.9),

$$\begin{aligned}
 R_x(\tau, T) &= \frac{1}{T-\tau} \int_0^{T-\tau} x(t) x(t+\tau) dt \quad ; \quad 0 \leq \tau \leq T \\
 &= \frac{1}{T-\tau} \int_0^{T-\tau} \xi(t) \xi(t+\tau) dt - m_x^2 - \frac{1}{12} \lambda(\tau, T) \alpha_x^2 \quad (6.24)
 \end{aligned}$$

where

$$\lambda(\tau, T) = T^2 \left[1 - 2(\tau/T) - 2(\tau/T)^2 \right] \quad (6.25)$$

For $\tau \ll T$, note that

$$\lambda(\tau, T) \approx T^2$$

Similarly, let $\eta(t)$ represent an input random record from a second random process $\{\eta(t)\}$ such that

$$\eta(t) = m_y + \alpha_y \left(t - \frac{T}{2} \right) + y(t) \quad ; \quad 0 < t < T \quad (6.26)$$

where m_y denotes the mean value of $\eta(t)$ over its length $[0, T]$, the parameter α_y denotes the average slope of the record $\eta(t)$ with respect to time t , and the final term $y(t)$ represents a sample record from a stationary random process $\{y(t)\}$ with zero mean value and zero slope. As before,

$$m_y = \frac{1}{T} \int_0^T \eta(t) dt \quad (6.27)$$

$$\alpha_y = \frac{1}{(T/3)(2T/3)} \left[\int_{(2T/3)}^T \eta(t) dt - \int_0^{(T/3)} \eta(t) dt \right] \quad (6.28)$$

$$\begin{aligned} R_y(\tau, T) &= \frac{1}{T-\tau} \int_0^{T-\tau} y(t) y(t+\tau) dt \quad ; \quad 0 \leq \tau \leq T \\ &= \frac{1}{T-\tau} \int_0^{T-\tau} \eta(t) \eta(t+\tau) dt - m_y^2 - \frac{1}{12} \lambda(\tau, T) \alpha_y^2 \end{aligned} \quad (6.29)$$

where $\lambda(\tau, T)$ is given by Equation (6.25).

Finally, a cross-correlation function estimate $R_{xy}(\tau, T)$ is given by

$$\begin{aligned}
 R_{xy}(\tau, T) &= \frac{1}{T-\tau} \int_0^{T-\tau} x(t) y(t + \tau) dt \quad ; \quad 0 \leq \tau \leq T \\
 &= \frac{1}{T-\tau} \int_0^{T-\tau} \xi(t) \eta(t + \tau) dt - m_x m_y - \frac{T}{2} (m_x \alpha_y - m_y \alpha_x) - \lambda(\tau, T) \alpha_x \alpha_y
 \end{aligned}$$

(6.30)

This formula includes the autocorrelation function estimates $R_x(\tau, T)$ and $R_y(\tau, T)$ as special cases, and reduces to the usual expression involving calculation of only the first term on the right-hand side when the quantities m_x , m_y , α_x and α_y equal zero. Other situations when some but not all of these quantities equal zero are also readily obtainable.

6.2 CROSS-POWER SPECTRA MEASUREMENTS

A schematic picture of a filter device for estimating the cross-power spectral density function associated with two random records $x(t)$ and $y(t)$ is displayed in Figure 6.4 below. Physically realizable real-valued estimates are obtained of the co-spectrum $C_{xy}(f)$ and the quad-spectrum $Q_{xy}(f)$ which can later be combined to yield the cross-power spectrum $S_{xy}(f)$ from the defining relation

$$S_{xy}(f) = C_{xy}(f) - jQ_{xy}(f) \quad (6.31)$$

The input random records $x(t)$ and $y(t)$ are assumed to be of finite time-length T and to be drawn from stationary random processes with zero mean values. The two separate identical tunable narrowband filters are assumed to have a finite nonzero constant bandwidth B centered at a frequency f_c which may be varied over the frequency range of interest. To estimate the co-spectral density

function, $C_{xy}(f)$, the in-phase frequency components in the filter outputs are multiplied together, then integrated and averaged. This is completely analogous to what was done previously in individual power spectra measurements, the multiplier circuit now performing the same role as the previous squarer circuit. To estimate the quad-spectral density function, $Q_{xy}(f)$, one of the filter outputs is passed through a 90° phase shifter before being multiplied by the output of the other filter. The product is then integrated and averaged as before. This yields the average product of the 90° out-of-phase frequency components in the two random functions, a proper physical interpretation of the quad-spectrum. The absolute value and phase angle of the cross-spectrum may be determined by vectorially combining the co-spectrum and quad-spectrum.

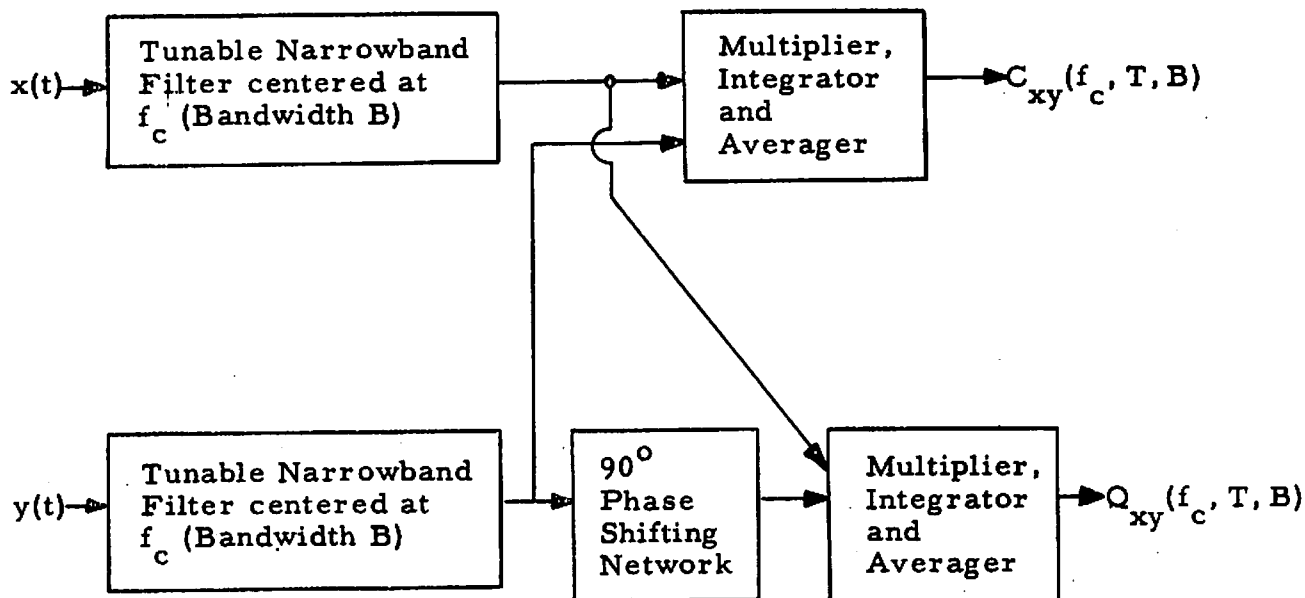


Figure 6.4 Cross-Power Spectral Density Analyzer

A complete analysis can now be carried out which will indicate the bias and variance to be associated with the estimates $C_{xy}(f, T, B)$ and $Q_{xy}(f, T, B)$ that would be obtained using Figure 6.4. Many parts of the analysis are quite similar to what was sketched previously for individual power spectra estimates. Since a detailed mathematical analysis of this type is not deemed to be appropriate for this report, only the main conclusions will be summarized below.

The bias terms for any frequency f are bounded above by

$$b [C_{xy}(f, T, B)] \leq \frac{B^2}{24} \left| S''_{xy}(f) \right| \tag{6.32}$$

$$b [Q_{xy}(f, T, B)] \leq \frac{B^2}{24} \left| S''_{xy}(f) \right|$$

where $S''_{xy}(f)$ is the second derivative of $S_{xy}(f)$ with respect to f , and is related to $R_{xy}(\tau)$ by the expression

$$S''_{xy}(f) = -4\pi^2 \int_{-\infty}^{\infty} \tau^2 R_{xy}(\tau) e^{-j2\pi f\tau} d\tau \tag{6.33}$$

The variance terms for any frequency $f \neq 0$ are bounded above by

$$\sigma^2 [C_{xy}(f, T, B)] \leq \frac{S_x(f) S_y(f)}{BT} \tag{6.34}$$

$$\sigma^2 [Q_{xy}(f, T, B)] \leq \frac{S_x(f) S_y(f)}{BT}$$

At $f = 0$, the right-hand sides above should be multiplied by a factor of two.

Thus, one finds that a mean square error analysis for co-spectrum and quad-spectrum estimates is closely analogous to a mean square error analysis of individual power spectrum. In particular, the conflicting demands on B to be small for low bias (and high resolution), and to be large for low variance are the same as previously.

Furthermore, if the actual available records are not $x(t)$ and $y(t)$, but $\xi(t)$ and $\eta(t)$, respectively, where $\xi(t)$ and $\eta(t)$ have non-zero mean values and non-zero slopes during the time of observation, then prior corrections must be made as indicated in Section 6.1.5.

This completes the main discussion on how to estimate power spectra and cross-power spectra from continuous data, and how to evaluate the expected mean square error of the measurements. Some further statistical error analysis will be developed in Section 8. Before taking up these matters, however, the next section will be devoted to writing down the appropriate sampled data (digital) formulas for calculating power spectra and cross-power spectra.

7. MEASUREMENT OF POWER SPECTRA AND CROSS-POWER SPECTRA : SAMPLED DATA

The first point that needs to be determined in connection with transforming continuous data into discrete data by sampling at equi-spaced intervals of time is how many data points should be read. It is clear that sampling at points which are too close together will yield highly correlated redundant data and greatly increase both the labor and cost of calculations. On the other hand, sampling at points which are too far apart may lose significant information at high frequencies since a sampling interval

$$\Delta t = h \quad (7.1)$$

induces a frequency cutoff

$$f_0 = \frac{1}{2h} = \frac{1}{2\Delta t} \quad (7.2)$$

known as the Nyquist frequency. Frequencies in the data which are above f_0 cps becomes "aliases" of frequencies which are below f_0 cps, and cannot be distinguished from one another.

Figure 7.1 below shows how a high frequency sine wave sampled at regular intervals appears as a lower frequency sine wave.

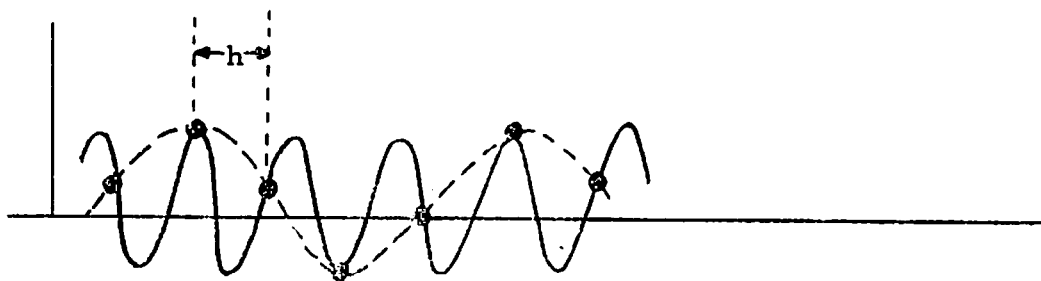


Figure 7.1 Slow Sampling of Fast Sine Wave

In order to avoid difficulties from this source, one should choose the sampling interval h such that its associated Nyquist frequency cutoff f_0 equals or exceeds the maximum frequency for which the power spectral density function is significant. Physically, this requires that the sampling interval h should contain at least two samples per cycle of the highest frequency of any importance in the record. Consideration should also be given to the removal of power at possible higher confounding frequencies by use of appropriate filters.

In actual practice, the sampling interval h [and hence the corresponding cutoff frequency $f_0 = (1/2h)$] may be determined empirically by examining a few specially chosen points to be certain that the power spectrum contains no appreciable power in the upper $(1/3)$ or $(1/2)$ of the frequency range $[0, f_0]$. Assuming h and f_0 to be properly chosen, it follows that summations of the power spectrum over $(-\infty, \infty)$ may be replaced by summations over $(-f_0, f_0)$ with negligible aliasing error.

Appropriate formulas for estimating power spectrum and cross-power spectrum using discrete data may now be obtained directly from formulas for continuous data by merely replacing integrals by finite summations. The case of an individual power spectrum measurement $S_x(f, T, B)$ for estimating $S_x(f)$ from discrete data will now be explained in some detail, after which similar formulas will be written down without proof for the other estimates $S_y(f, T, B)$, $C_{xy}(f, T, B)$ and $Q_{xy}(f, T, B)$.

7.1 POWER SPECTRA MEASUREMENTS

Suppose that a finite continuous record $\xi(t)$ is available, $0 \leq t \leq T$, having a mean value m_x , an average slope α_x , and a residual term $x(t)$ which is a sample from a stationary random process with zero mean and zero slope such that

Contrails

$$\xi(t) = m_x + \alpha_x \left(t - \frac{T}{2}\right) + x(t) \quad ; \quad 0 \leq t \leq T \quad (7.3)$$

The record $\xi(t)$ should be sampled at N equally spaced time intervals a distance h apart, where h has been selected so as to produce a sufficiently high frequency cutoff f_0 . Let

$$\xi_n = \xi(nh) \quad ; \quad n = 1, 2, 3, \dots, N \quad ; \quad T = Nh \quad (7.4)$$

Compute the sample mean value

$$m_x = \frac{1}{T} \int_0^T \xi(t) dt \approx \frac{1}{N} \sum_{n=1}^N \xi_n \quad (7.5)$$

and the sample slope term

$$\begin{aligned} \alpha_x &= \frac{1}{(T/3)(2T/3)} \left[\int_{(2T/3)}^T \xi(t) dt - \int_0^{(T/3)} \xi(t) dt \right] \\ &\approx \frac{1}{h\nu(N-\nu)} \left[\sum_{n=N-\nu}^N \xi_n - \sum_{n=1}^{\nu} \xi_n \right] \end{aligned} \quad (7.6)$$

where

$\nu = [N/3]$ = largest integer contained in $(N/3)$. For example, $\nu = 33$ if $N = 100$.

Compute next the raw autocorrelation function expression

$$R_{\xi}(\tau, T) = \frac{1}{T-\tau} \int_0^{T-\tau} \xi(t) \xi(t+\tau) dt \quad ; \quad \tau > 0$$

which for sampled data becomes

$$R_{\xi}(rh, Nh) \approx \frac{1}{N-r} \sum_{n=1}^{N-r} \xi_n \xi_{n+r} \quad ; \quad r \geq 0 \quad (7.7)$$

where

$$\tau = rh \quad ; \quad T = Nh \quad (7.8)$$

The autocorrelation function estimate for $x(t)$ is now given by [see Equation (6.24)],

$$R_x(rh, Nh) = R_{\xi}(rh, Nh) - m_x^2 - \frac{1}{12} \lambda_r \alpha_x^2 \quad (7.9)$$

where

$$\lambda_r = N^2 h^2 [1 - 2(r/N) - 2(r/N)^2] \quad (7.10)$$

To simplify succeeding notation, write $R_x(rh)$ in place of $R_x(rh, Nh)$.
Now, consider the quantity

$$\begin{aligned} S_x(\alpha, T) &= \frac{1}{T} \int_{-T}^T (T - |\tau|) R_x(\tau, T) e^{-j2\pi\alpha\tau} d\tau \\ &\approx h \sum_{r=-N}^N (1 - |r|/N) R_x(rh) e^{-j2\pi\alpha rh} \end{aligned} \quad (7.11)$$

By hypothesis,

$$S_x(\alpha, T) \approx 0 \quad \text{for} \quad |\alpha| > f_0 \quad (7.12)$$

Contrails

The power spectrum estimate $S_x(f, T, B)$ is given by

$$\begin{aligned} S_x(f, T, B) &= \int_{-\infty}^{\infty} |H(\alpha)|^2 S_x(\alpha, T) d\alpha \\ &\approx \int_{-f_0}^{f_0} |H(\alpha)|^2 S_x(\alpha, T) d\alpha \end{aligned} \quad (7.13)$$

where

$$|H(\alpha)|^2 = \frac{1}{2B} \quad ; \quad |\alpha \pm f| \leq (B/2), \quad \text{zero otherwise.} \quad (7.14)$$

To derive an alternative equivalent expression for the above, define a weighting function $\omega(\alpha)$ by

$$\omega(\alpha + f) = \frac{1}{2B} \quad ; \quad |\alpha + f| \leq (B/2), \quad \text{zero otherwise} \quad (7.15)$$

$$\omega(\alpha - f) = \frac{1}{2B} \quad ; \quad |\alpha - f| \leq (B/2), \quad \text{zero otherwise}$$

It follows that

$$S_x(f, T, B) = \int_{-f_0}^{f_0} S_x(\alpha, T) [\omega(\alpha + f) + \omega(\alpha - f)] d\alpha \quad (7.16)$$

Note that $\omega(\alpha) = \omega(-\alpha)$, and for $|f| < f_0$,

$$\int_{-f_0}^{f_0} \omega(\alpha) d\alpha = \frac{1}{2B} \left[\int_{-f-(B/2)}^{-f+(B/2)} d\alpha + \int_{f-(B/2)}^{f+(B/2)} d\alpha \right] = 1 \quad (7.17)$$

Contrails

Next, consider $\omega(\alpha)$ to be periodic of period $2f_0$. Then its complex Fourier expansion may be written

$$\omega(\alpha) = \sum_{n=-\infty}^{\infty} A_n e^{j2\pi n h \alpha} = \frac{1}{2B} ; \quad |\alpha| \leq B/2 \quad (7.18)$$

where $h = 1/2f_0$ and

$$A_n = \frac{1}{2f_0} \int_{-f_0}^{f_0} \omega(\alpha) e^{-j2\pi n h \alpha} d\alpha = \frac{h}{2} \left(\frac{\sin n\pi B h}{n\pi B h} \right) \quad (7.19)$$

Substitute Equations (7.11) and (7.18) into Equation (7.16).

This proves

$$\begin{aligned} S_x(f, T, B) &= h \int_{-f_0}^{f_0} \sum_{r=-N}^N (1 - |r|/N) R_x(rh) e^{-j2\pi r h \alpha} \\ &\quad \left\{ \sum_{n=-\infty}^{\infty} A_n \left[e^{j2\pi n h (\alpha+f)} + e^{j2\pi n h (\alpha-f)} \right] \right\} d\alpha \\ &= 2 \sum_{r=-N}^N (1 - |r|/N) R_x(rh) A_r \cos 2\pi r h f \quad (7.20) \end{aligned}$$

since

$$\int_{-f_0}^{f_0} e^{-j2\pi \alpha h (r-n)} d\alpha = \begin{cases} 2f_0 & \text{of } n = r \\ 0 & \text{of } n \neq r \end{cases}$$

Equation (7.20) with A_r given by Equation (7.19) states how to calculate the power spectrum estimate $S_x(f, T, B)$. The amount

of work required to use this formula will be quite costly if N is large (of order several hundred to several thousand) mainly because of having to calculate the different correlation functions. This labor can be reduced if the coefficients A_r beyond a certain point were to become arbitrarily small. The coefficient

$$A_r = \frac{h}{2} \left(\frac{\sin r\pi Bh}{r\pi Bh} \right) = \frac{h}{2} \left(\frac{\sin r\pi/J}{r\pi/J} \right) \quad (7.21)$$

where

$$J = \frac{1}{Bh} = \frac{2f_0}{B} \quad (7.22)$$

The number J represents the largest number of uncorrelated spectral estimates that can fill a frequency interval $(-f_0, f_0)$ with small filters of bandwidth B . The number $(J/2)$ thus represents the maximum number of uncorrelated estimates in the positive realizable frequency range $(0, f_0)$ using filters of bandwidth B , and J is called the resolution number. For a given value of f_0 , arbitrarily high resolution (and small bias) are achieved by making J large. In practice J will generally be considerably smaller than N .

Observe that

$$A_r = 0 \text{ if } r = J \quad (7.23)$$

This suggests defining a new coefficient θ_r by

$$\begin{aligned} \theta_r &= (1 - |r|/N) A_r \text{ for } |r| \leq J \\ &= 0 \text{ for } |r| > J \end{aligned} \quad (7.24)$$

The corresponding weighting function $\hat{\omega}(f)$ for an approximating filter using these new coefficients becomes

$$\hat{\omega}(f) = \sum_{r=-J}^J \theta_r e^{j2\pi r h f} \quad (7.25a)$$

and should be compared with the previous weighting function $\omega(f)$ for the ideal filter, Equation (7.18), where

$$\omega(f) = \frac{1}{2B} \quad ; \quad |f| \leq B/2 \quad (\text{zero otherwise}) \quad (7.25b)$$

Assuming $J < N$ so that the factor $1 - |r|/N \approx 1$ for $|r| \leq J$, a separate analysis reveals that $\hat{\omega}(f)$ is a reasonable approximation to $\omega(f)$, which reduces the resolution by at most a factor of order two. See Figure 7.2.

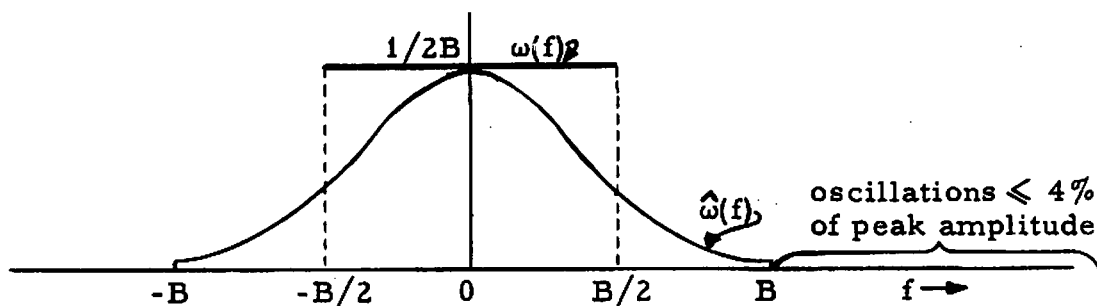


Figure 7.2 Comparison of Ideal and Approximating Filters

In place of Equation (7.20), the spectral estimate now becomes

$$S_x(f, T, B) \approx 2 \sum_{r=-J}^J \theta_r R_x(rh) \cos 2\pi r h f$$

$$= h \sum_{r=-J}^J (1 - |r|/N) \left(\frac{\sin r\pi/J}{r\pi/J} \right) R_x(rh) \cos 2\pi r h f \quad (7.26)$$

The bias in this estimate, given by the result analogous to the continuous case, is simply

$$b[S_x(f, T, B)] \approx \frac{B^2}{24} |S_x''(f)| \approx \frac{1}{6} \left(\frac{f_0}{J} \right)^2 |S_x''(f)| \quad (7.27)$$

The variance in the estimate, derived also from the continuous case, is

$$\sigma^2[S_x(f, T, B)] \approx \left(\frac{1}{BT} \right) S_x^2(f) \approx \left(\frac{J}{N} \right) S_x^2(f) \quad ; \quad f \neq 0$$

$$\approx \left(\frac{2}{BT} \right) S_x^2(0) \approx \left(\frac{2J}{N} \right) S_x^2(0) \quad ; \quad f = 0 \quad (7.28)$$

This, to achieve an arbitrarily low variance for a given J, one should make N as large as possible. Both low bias and low resolution will occur in numerical computations by taking J and N large with $J < N$.

This completes the present treatment on how to estimate the power spectral density function $S_x(f)$ from a digital processing of $S_x(f, T, B)$, and how to evaluate the errors involved for given choices of h, J and N.

7.2 SUMMARY OF FORMULAS

For future ease of computation, appropriate formulas for estimating both power spectra and cross-power spectra from sampled data will be summarized below. In view of the preceding discussion, the notation should be self-explanatory.

Contrails

Consider a finite pair of continuous records

$$\xi(t) = m_x + \alpha_x \left(t - \frac{T}{2}\right) + x(t) \quad ; \quad 0 \leq t \leq T \quad (7.29)$$

$$\eta(t) = m_y + \alpha_y \left(t - \frac{T}{2}\right) + y(t) \quad ; \quad 0 \leq t \leq T$$

which are sampled at points $\Delta t = h$ time-units apart, with a corresponding frequency cutoff $f_0 = (1/2h)$. Let the sample values be

$$\left. \begin{aligned} \xi_n &= \xi(t_0 + nh) \\ \eta_n &= \eta(t_0 + nh) \end{aligned} \right\} \quad n = 1, 2, \dots, N \quad ; \quad T = Nh \quad (7.30)$$

(t_0 is an arbitrary time origin which does not enter into the calculations). The simplest quantities to compute are

$$m_x = \frac{1}{N} \sum_{n=1}^N \xi_n$$

$$m_y = \frac{1}{N} \sum_{n=1}^N \eta_n$$

(7.31)

$$\alpha_x = \frac{1}{h\nu(N-\nu)} \left[\sum_{n=N-\nu}^N \xi_n - \sum_{n=1}^{\nu} \xi_n \right]$$

$$\nu = \left[\frac{N}{3} \right]$$

$$\alpha_y = \frac{1}{h\nu(N-\nu)} \left[\sum_{n=N-\nu}^N \eta_n - \sum_{n=1}^{\nu} \eta_n \right]$$

Assume in what follows that the resolution number $J < N$ so as to reduce the number of required correlation calculations. Divide the work in two stages. First, compute the raw correlation function expressions

$$\left. \begin{aligned}
 R_{\xi}(rh) &= \frac{1}{N-r} \sum_{n=1}^{N-r} \xi_n \xi_{n+r} \\
 R_{\eta}(rh) &= \frac{1}{N-r} \sum_{n=1}^{N-r} \eta_n \eta_{n+r} \\
 R_{\xi\eta}(rh) &= \frac{1}{N-r} \sum_{n=1}^{N-r} \xi_n \eta_{n+r} \\
 R_{\eta\xi}(rh) &= \frac{1}{N-r} \sum_{n=1}^{N-r} \eta_n \xi_{n+r}
 \end{aligned} \right\} r = 0, 1, 2, \dots, J \quad (7.32)$$

The sample correlation functions are then given by

$$\begin{aligned}
 R_x(rh) &= R_{\xi}(rh) - m_x^2 - \frac{1}{12} \lambda_r \alpha_x^2 \\
 R_y(rh) &= R_{\eta}(rh) - m_y^2 - \frac{1}{12} \lambda_r \alpha_y^2 \\
 R_{xy}(rh) &= R_{\xi\eta}(rh) - m_x m_y - \frac{r}{2} (m_x \alpha_y - m_y \alpha_x) - \frac{1}{12} \lambda_r \alpha_x \alpha_y \\
 R_{yx}(rh) &= R_{\eta\xi}(rh) - m_x m_y + \frac{r}{2} (m_x \alpha_y - m_y \alpha_x) - \frac{1}{12} \lambda_r \alpha_x \alpha_y
 \end{aligned} \tag{7.33}$$

where

$$\lambda_r = N^2 h^2 \left[1 - 2(r/N) - 2(r/N)^2 \right] \quad ; \quad r = 0, 1, 2, \dots, J \quad (7.34)$$

For later use, calculate the quantities

$$\phi_{xy}(rh) = \frac{1}{2} \left[R_{xy}(rh) + R_{yx}(rh) \right] \quad (7.35)$$

$$\psi_{xy}(rh) = \frac{1}{2} \left[R_{xy}(rh) - R_{yx}(rh) \right]$$

The various spectral quantities may now be estimated at an arbitrary frequency f in the interval $(0, f_0)$ by the formulas

$$S_x(f, T, B) = h \left[R_x(0) + 2 \sum_{r=1}^{J-1} D_r R_x(rh) \cos 2\pi r h f \right] \quad (7.36)$$

$$S_y(f, T, B) = h \left[R_y(0) + 2 \sum_{r=1}^{J-1} D_r R_y(rh) \cos 2\pi r h f \right]$$

$$S_{xy}(f, T, B) = C_{xy}(f, T, B) - j Q_{xy}(f, T, B) \quad (7.37)$$

where

$$C_{xy}(f, T, B) = h \left[\phi_{xy}(0) + 2 \sum_{r=1}^{J-1} D_r \phi_{xy}(rh) \cos 2\pi r h f \right] \quad (7.38)$$

$$Q_{xy}(f, T, B) = 2h \sum_{r=1}^{J-1} D_r \psi_{xy}(rh) \sin 2\pi r h f$$

The factor

$$D_r = (1 - |r|/N) \left(\frac{\sin r\pi/J}{r\pi/J} \right) \approx \left(\frac{\sin r\pi/J}{r\pi/J} \right) \quad \text{if } J \ll N \quad (7.39)$$

The resolution number J represents twice the actual resolution in the frequency interval $(0, f_0)$. In other words, if the frequency interval $(0, f_0)$ is divided into J equally spaced points f_0/J apart, then spectral estimates at alternate points $2f_0/J$ apart will be essentially uncorrelated. Thus, to reduce the labor above, it is sufficient to calculate the various spectral estimates at only J equally spaced points, recognizing that $J/2$ of these points should then furnish independent estimates. To be specific, choose

$$f = k(f_0/J) \quad ; \quad k = 0, 1, 2, \dots, J \quad (7.40)$$

This completes the summary of computing formulas that are appropriate for estimating power-spectra and cross-power spectra from discrete data. Other formulas developed by Tukey and Blackman [Ref. 2] give slightly different estimates because of using a different smoothing filter than the one involved here, but the bias errors and variance errors are essentially the same as derived here. To a first order of approximation, the bias and variance errors for $f \neq 0$ are given by

$$\begin{aligned} b[S_x] &= \frac{1}{6} \left(\frac{f_0}{J} \right)^2 |S''_x| \quad ; \quad \sigma^2[S_x] = \left(\frac{J}{N} \right) S_x^2 \\ b[S_y] &= \frac{1}{6} \left(\frac{f_0}{J} \right) |S''_y| \quad ; \quad \sigma^2[S_y] = \left(\frac{J}{N} \right) S_y^2 \\ b[C_{xy}] &\cong \frac{1}{6} \left(\frac{f_0}{J} \right) |S''_{xy}| \quad ; \quad \sigma^2[C_{xy}] \cong \left(\frac{J}{N} \right) S_x S_y \\ b[Q_{xy}] &\cong \frac{1}{6} \left(\frac{f_0}{J} \right) |S''_{xy}| \quad ; \quad \sigma^2[Q_{xy}] \cong \left(\frac{J}{N} \right) S_x S_y \end{aligned} \quad (7.41)$$

At $f = 0$, the variance expressions should be multiplied by two.

In the next section of this report, the continuous error formulas of Section 6 will be assumed. The reader desiring to relate the results to the discrete cases of this section has merely to substitute

$$T = Nh = N/2f_0 \quad ; \quad B = 2f_0/J = 1/Jh \quad (7.42)$$

Note that the number

$$k = 2BT = 2N/J \quad (7.43)$$

represents the number of degrees of freedom in the sense of distributing the $J/2$ independently resolved frequencies in the interval $(0, f_0)$ equally over the N observations.

8. CONFIDENCE LIMITS AND DESIGN RELATIONS

From Equation(6.19) for continuous data, the mean square error of the estimate $S(f, T, B)$, [which will be taken as representative of S_x , S_y , C_{xy} and Q_{xy} as well], is given by

$$E [S(f, T, B) - S(f)]^2 \approx \frac{|S(f)|^2}{BT} + \left(\frac{B^2}{24}\right)^2 |S''(f)| \quad (8.1)$$

The mean square percentage error of the estimate denoted by e^2 , is defined by the mean square error divided by the square of the true value. Hence

$$\begin{aligned} e^2 &= e^2 [S(f, T, B)] = \frac{E [S(f, T, B) - S(f)]^2}{S^2(f)} \\ &\approx \frac{1}{BT} + \frac{B^4}{576} \left| \frac{S''(f)}{S(f)} \right|^2 \end{aligned} \quad (8.2)$$

Let the quantity

$$\lambda(f) = \left| \frac{S(f)}{S''(f)} \right|^{1/2} \quad (8.3)$$

Then $\lambda(f)$ has units of frequency (cps), and is called the "spectral bandwidth" of the random process $\{x(t)\}$ under consideration. In terms of $\lambda(f)$,

$$e^2 \approx \frac{1}{BT} + \frac{1}{576} \left[\frac{B}{\lambda(f)} \right]^4 \quad (8.4)$$

This equation enables one to make quantitative statements about the mean square percentage error e^2 in measuring a power spectrum $S(f)$ for given values of B , T and $\lambda(f)$. The latter quantity $\lambda(f)$ demands some apriori knowledge of the spectrum which one is trying to measure.

If e^2 is large, then any particular individual measurement $S(f, T, B)$ would not be likely to fall close to the true value $S(f)$. However, if e^2 is small, then all individual measurements of $S(f, T, B)$ would tend to closely approximate $S(f)$. Thus, to guarantee in advance that an arbitrary measurement represents well the true measurement, one should try to make e^2 as small as possible through prior choice of B and T .

Returning to Equation (8.4), suppose that the "spectral bandwidth" $\lambda(f)$ is known (or can be reasonably estimated) for the random process under consideration. Suppose also that the bandwidth B of the discriminating filter, and the record length T , can be set to any desired design values. Then, in order to nearly always be able to separate peaks in the true spectrum $S(f)$ which may be a spectral bandwidth $\lambda(f)$ apart, it appears reasonable to select B so that

$$B \ll \frac{\lambda(f)}{2} \quad (8.5)$$

This choice of B (together with a proper T as found below) will then guarantee, with a low probability of error, that if $S(f)$ has two distinct peaks which are $\lambda(f)$ cps apart, then these two peaks can be resolved by taking measurements of $S(f)$ at intervals of B cps apart. Another way of looking at this statement is to say that measurements of $S(f)$ at intervals of B cps apart will practically always distinguish peaks which are $2B$ cps apart.

Assuming B to satisfy $B \leq \lambda(f)/2$, the second term in Equation (8.4) becomes negligible and Equation (8.4) reduces to the simple relation

$$e^2 \approx \frac{1}{BT} \quad (8.6)$$

In particular, for $e = 0.10$, corresponding to a root mean square percentage error of 10 %, the value of the product BT should be

$$BT \approx 100 \quad \text{or} \quad T \approx (100/B) \quad (8.7)$$

To illustrate these last formulas, suppose that $\lambda(f) \geq 40$ cps. First, from Equation (8.5), choose $B = 20$ cps. Then from Equation (8.7), choose $T = 5$ seconds. It follows that different measurements of $S(f)$ taken 20 cps apart will now resolve peaks which are 40 cps apart, and the rms percentage error in the measurements will be at most 10 percent.

If B is not small compared to $\lambda(f)$, then the original formula of Equation (8.4) must be used to calculate the rms percentage error. For example, suppose that $B = 2\lambda(f)$ at a particular value of f . Then, for $B = 20$ cps and $T = 5$ sec, the same two values considered in the previous paragraph, it now follows that $e^2 \approx 0.038$ and $e \approx 19.5$ percent. It is clear from this example how important it is to have $B \leq 0.50 \lambda(f)$ for all f , if this is possible.

8.1 GAUSSIAN SET OF MEASUREMENTS

Assume next that a set of power spectrum measurements $\{S(f, T, B)\}$ at a particular value of f follows a normal (Gaussian) distribution. To simplify the notation, let

$$\begin{aligned} S^* &= S(f, T, B) & S &= S(f) \\ m &= E S^* & b &= |m - S| \\ \sigma^2 &= E [S^* - m]^2 \end{aligned} \quad (8.8)$$

Contrails

Observe that the bias $b \neq 0$ when the true value S is expected to differ from the mean value m .

For a normal distribution, the probability p ($0 \leq p \leq 1$) that S^* will fall within any specified deviation $\pm \alpha_p \sigma$ from the mean value m is given by the normal probability integral

$$p = \text{Prob} \left[\left| S^* - m \right| \leq \alpha_p \sigma \right] = \int_{m - \alpha_p \sigma}^{m + \alpha_p \sigma} \frac{1}{\sigma \sqrt{2\pi}} \exp \left[\frac{-(s-m)^2}{2\sigma^2} \right] ds \quad (8.9)$$

For example, if $\alpha_p = 1, 2,$ or 3 , then $p = 68.3\%, 95.4\%,$ and 99.7% , respectively.

From the inequality

$$\left| S^* - S \right| \leq \left| S^* - m \right| + \left| m - S \right| = \left| S^* - m \right| + b$$

it follows that

$$\text{Prob} \left[\left| S^* - S \right| \leq \alpha_p \sigma + b \right] \geq p \quad (8.10)$$

In words, the probability that S^* will differ from the true value S by less than $(\alpha_p \sigma + b)$ is at least p . The p percent Gaussian range of percentage error will now be defined by the quantity

$$\Delta_p = \frac{\alpha_p \sigma + b}{S} \quad (8.11)$$

and satisfies the relation

$$\text{Prob} \left[\left| \frac{S^* - S}{S} \right| \leq \Delta_p \right] \geq p \quad (8.12)$$

Physically, this relation states that the probability that the true percentage error, $|(S^* - S)/S|$, will be less than Δ_p is at least p .

For the idealized filter under consideration in this report,

$$\frac{\sigma}{S} \approx \frac{1}{\sqrt{BT}} \quad ; \quad \frac{b}{S} \approx \frac{1}{24} \left(\frac{B}{\lambda} \right)^2 \quad (8.13)$$

Hence, Equation (8.11) becomes

$$\Delta_p \approx \frac{\alpha_p}{\sqrt{BT}} + \frac{1}{24} \left(\frac{B}{\lambda} \right) \quad (8.14)$$

If B is chosen so that $B \leq 0.50\lambda$, then only the first term above is significant, whereupon

$$\Delta_p \approx \frac{\alpha_p}{\sqrt{BT}} = \alpha_p e \quad \text{where} \quad e = \frac{1}{\sqrt{BT}} \quad (8.15)$$

To illustrate this last statistical result, suppose that $BT = 400$. Then the probability is 95% (corresponding to $\alpha_p = 2.0$) that the percentage error Δ_p is less than 10%. From another point of view, this formula is of considerable practical importance in designing experiments in that it states how large the BT product should be to guarantee that an arbitrary estimate S^* will be in error by at most Δ_p percent at a confidence level of p percent. For example, suppose that the error is to be no more than $\Delta_p = 30\%$ at an 80% confidence level ($\alpha_p = 1.3$). Then the BT product from Equation (8.15) should be $BT \approx 18.7$. For complete resolution of 20 equally spaced points on a frequency range of interest, using a digital technique, since $BT = N/J$, see Equation (7.43), where $(J/2)$ represents the number of resolved frequencies, it follows that $J = 40$. Hence $N = 750$ data points are needed for frequencies $f \neq 0$ while $N = 1500$ for $f = 0$.

Suppose that the tunable filter (see Figure 6.1) is tuned in a uniform continuous fashion over some wide frequency interval \tilde{B} in the time \tilde{T} . Then, the average amount of time T that the input record $x(t)$ spends within the narrow discriminating filter bandwidth B (for any center frequency) satisfies the relation

$$\dot{\beta} = \frac{\tilde{B}}{\tilde{T}} = \frac{B}{T} \quad (8.16)$$

where $\dot{\beta}$ is the sweep rate (cps/sec). Solving for T , and substituting in $e^2 = (1/BT)$, one obtains

$$e^2 = \frac{\dot{\beta}}{B^2} \quad (8.17)$$

which indicates how the mean square error e^2 varies as a function of $\dot{\beta}$ and B . Observe that for an rms error $e \leq 10\%$, $\dot{\beta}$ should satisfy

$$\dot{\beta} \leq 0.01 B^2 \text{ cps/sec} \quad (B \text{ in cps}) \quad (8.18)$$

For example, if $B = 20$ cps, then $\dot{\beta} \leq 4$ cps/sec in order to keep the rms error below 10%.

8.2 SPECIAL DESIGN RELATIONS

The mean square error equation

$$e^2 \approx \frac{1}{BT} + \frac{1}{576} \left[\frac{B}{\lambda(f)} \right]^4 \quad (8.4)$$

may be solved for T explicitly as a function of the other parameters B , e , and $\lambda = \lambda(f)$. One obtains

Contrails

$$T = \frac{576}{B \left[576 e^2 - (B/\lambda)^4 \right]} \quad (8.19)$$

Considering T as a function of B alone, for fixed values of e and λ , one may now determine the value B_{\min} of B which minimizes T , and the value T_{\min} of T at this minimum. The result is

$$\begin{aligned} B_{\min} &\approx 3.38 \lambda e^{1/2} \\ T_{\min} &\approx 0.37 / \lambda e^{5/2} \end{aligned} \quad (8.20)$$

Note that for $e = 0.10$,

$$\begin{aligned} B_{\min} &\approx 1.07 \lambda \\ T_{\min} &\approx (117/\lambda) \end{aligned} \quad (8.21)$$

Any other choice of B for a given e and λ will require a longer T . In particular, this is true if one set $B = 0.50 \lambda$, as recommended in Equation (8.5). This is verified from Equation (8.7) since the corresponding T satisfies $BT = 100$ or $T \approx (200/\lambda)$.

Alternately, one can consider e^2 in Equation (8.19) as a function of B alone, for fixed values of T and λ , and determine the value B_{\min} of B which minimizes e , and the value e_{\min} of e at this minimum. The result here is

$$\begin{aligned} B_{\min} &\approx 2.70 \lambda^{4/5} / T^{1/5} \\ e_{\min}^2 &\approx 0.60 / \lambda^{4/5} T^{4/5} \end{aligned} \quad (8.22)$$

Contrails

Note that for $T = (200/\lambda)$,

$$\begin{aligned} B_{\min} &\approx 0.94\lambda \\ e_{\min}^2 &\approx 0.0067 \quad ; \quad e_{\min} \approx 8.2 \text{ percent} \end{aligned} \tag{8.23}$$

Any other choice of B for a given T and λ will result in a larger error e . In particular, suppose $B = 0.50\lambda$, as recommended in Equation (8.5). Then, for $T = (200/\lambda)$, it follows from Equation (8.6) that $e \approx 10$ percent.

All matters being considered however, especially the desire to resolve peaks in $S(f)$ which may be $\lambda(f)$ apart, it is deemed to be of practical importance to choose $B = 0.50\lambda$ if this can be done. Then e , B and T are related quite simply by $e^2 = (1/BT)$. Otherwise, one should refer to the general expression of Equation (8.4), or the special results of Equations (8.20) and (8.22), in order to determine appropriate experimental design values for e , B and T .

9. PHYSICAL APPLICATIONS

To conclude this report, a few out of many important physical applications will be discussed briefly which involve power spectra and cross-power spectra information. These are: Probability distribution of Amplitude Values, Measurement of System Frequency Response Function, and Zero Crossings and Maxima Probabilities.

9.1 PROBABILITY DISTRIBUTION OF AMPLITUDE VALUES

Consider a random record $x(t)$ which is a representative member of a stationary ergodic Gaussian random process with zero mean value. From the ergodic property, the time-wise behavior of $x(t)$ over a long period of time will exhibit the same statistical characteristics as corresponding ensemble averages at various fixed times. As a consequence, it follows that the probability density function associated with the amplitude values of $x(t)$ that will occur over a long time interval is given here by the Gaussian formula for zero mean value, namely,

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2} \quad (9.1)$$

where the variance σ^2 equals

$$\begin{aligned} \sigma^2 &= \left\langle x^2(t) \right\rangle_{Av} \quad ; \quad \text{independent of } t \\ &\approx \frac{1}{T} \int_0^T x^2(t) dt \quad \text{for large } T \\ &= \int_{-\infty}^{\infty} S_x(f) df = 2 \int_0^{\infty} S_x(f) df \end{aligned} \quad (9.2)$$

the quantity $S_x(f)$ denoting the two-sided power spectral density function of $x(t)$ as defined over $(-\infty, \infty)$. Statistical procedures for estimating $S_x(f)$ from finite data were developed in Sections 6, 7 and 8.

Thus, the probability density function $p(x)$ is completely characterized through knowledge of $S_x(f)$ since $S_x(f)$ alone determines σ . This important result places knowledge of $S_x(f)$ at the forefront of much work in analysis of random records obeying a normal distribution.

If the mean value of $x(t)$ is not zero, then the underlying probability density function is given by the general Gaussian formula

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-m)^2/2\sigma^2} \quad (9.3)$$

where the mean value

$$m = \langle x(t) \rangle_{Av} \quad ; \quad \text{independent of } t$$

$$\approx \frac{1}{T} \int_0^T x(t) dt \quad \text{for large } T \quad (9.4)$$

and

$$\sigma^2 = \langle [x(t) - m]^2 \rangle_{Av}$$

$$= \int_{-\infty}^{\infty} \tilde{S}_x(f) df \quad (9.5)$$

the quantity $\tilde{S}_x(f)$ denoting the power spectral density associated with the zero mean value portion of $x(t)$.

Working with either an arbitrary probability density function, or with the normal probability density function $f(x)$ from Equation (9.1) or (9.3), the probability over all time t that $x(t)$ lies below or above any prescribed value x_0 may be computed from the formulas

$$\begin{aligned}\text{Prob}_t [x(t) \leq x_0] &= \int_{-\infty}^{x_0} p(x) dx \\ \text{Prob}_t [x(t) > x_0] &= \int_{x_0}^{\infty} p(x) dx\end{aligned}\tag{9.6}$$

For a large positive constant k , one may show that for a Gaussian $p(x)$,

$$\text{Prob}_t [x(t) > m + k\sigma] \approx \frac{1}{k\sqrt{2\pi}} e^{-k^2/2}\tag{9.7}$$

In particular, for $k = 3$, $\text{Prob}_t [x(t) > m + 3\sigma] \approx 0.002$. Other values may be calculated quite easily and are available in tables.

These results are not valid for random records which violate the basic hypothesis of being representative members of stationary ergodic Gaussian random processes, and the reader is cautioned against using these formulas indiscriminately.

9.2 MEASUREMENT OF SYSTEM FREQUENCY RESPONSE FUNCTION

For a constant parameter linear system, it is well-known that such a system can be characterized by a weighting function $h(\tau)$, which, by definition, yields the response of a system to a unit impulse function τ time units after the impulse occurs. See Figure 9.1.

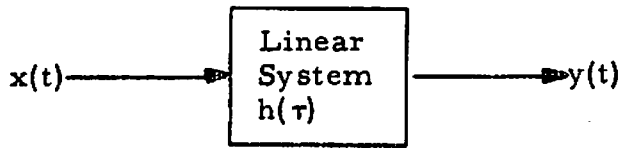


Figure 9.1 Linear System

For the system to be physically realizable it is necessary that

$$h(\tau) = 0 \quad \text{for } \tau < 0 \quad (9.8)$$

since a system can not respond to an impulse before it occurs.

If $x(t)$ is an input to this system, and $y(t)$ the resulting output, then the output is given as a weighted linear sum over the entire (infinite) past history of the input as expressed by

$$y(t) = \int_0^{\infty} h(\tau) x(t - \tau) d\tau \quad (9.9)$$

If the system operates on $x(t)$ only for a finite fixed time T , then

$$y(t) = \int_0^T h(\tau) x(t - \tau) d\tau \quad (9.10)$$

If $x(t)$ exists only for $t \geq 0$, then

$$y(t) = \int_0^t h(\tau) x(t - \tau) d\tau \quad (9.11)$$

Instead of using $(h(\tau))$, the system may be characterized by its frequency response function $Y(jf)$ which is defined as the Fourier

Contrails

transform of $h(\tau)$, namely,

$$Y(jf) = \int_0^{\infty} h(\tau) e^{-j2\pi f\tau} d\tau \quad (9.12)$$

assuming $h(\tau) = 0$ for $\tau < 0$ as needed for physical realizability. The frequency response function should not be confused with the transfer function of the system as defined by the Laplace transform of $h(\tau)$, and denoted by

$$Y_1(p) = \int_0^{\infty} h(\tau) e^{-p\tau} d\tau \quad (9.13)$$

Note that $Y(jf)$ differs formally from $Y_1(p)$ by merely replacing p by $j2\pi f$.

The frequency response function is a complex-valued function of f such that

$$Y(jf) = |Y(jf)| e^{j\phi(f)} \quad (9.14)$$

where $|Y(jf)|$, the absolute value of $Y(jf)$ measures the amplitude response (gain) of the system to an input sinusoidal exciting frequency f , while $\phi(f)$ indicates the corresponding phase shift.

Consider a complex-valued sinusoidal input

$$x(t) = e^{j2\pi f_0 t} \quad (9.15)$$

From Equations (9.9) and (9.12), the response $y(t)$ is

$$y(t) = Y(jf_0) e^{j2\pi f_0 t} \quad (9.16)$$

For a real-valued sinusoidal input

$$x(t) = \sin(2\pi f_0 t + d) \quad (9.17)$$

the real-valued output is given by

$$y(t) = |Y(jf)| \sin[2\pi f_0 t + d + \phi(f_0)] \quad (9.18)$$

These relations show how knowledge of both the gain factor and the phase shift term are needed to describe the system's operation.

From physical realizability requirements, the frequency response function, the gain factor, and the phase shift term satisfy the symmetry properties

$$Y(-jf) = \overline{Y(jf)}$$

$$|Y(-jf)| = |Y(jf)| \quad (9.19)$$

$$\phi(-f) = -\phi(f)$$

If one linear system, described by $Y_1(jf)$, is followed by a second linear system, described by $Y_2(jf)$, then the overall system may be described by $Y(jf)$ where

$$Y(jf) = Y_1(jf) Y_2(jf) \quad (9.20)$$

Hence

$$|Y(jf)| = |Y_1(jf)| |Y_2(jf)| \quad (9.21)$$

$$\phi(f) = \phi_1(f) + \phi_2(f)$$

Thus, on cascading two linear systems, the gain factors multiply and the phase shift factors add together.

Contrails

Assume now that the input $x(t)$ is a representative member from a stationary random process with zero mean value. Then, the same property is true for $y(t)$, and the ordinary power spectral density functions $S_x(f)$ and $S_y(f)$ are related to $|Y(jf)|$ by the simple (real-valued) formula

$$S_y(f) = |Y(jf)|^2 S_x(f) \quad (9.22)$$

This is an important result and is frequently quoted. At any fixed frequency f , knowledge of two of these quantities determines the third. The phase shift term $\phi(f)$, however, is still in doubt since this formula involves only the gain factor.

By a straight-forward approach, one may verify that the entire frequency response function $Y(jf)$ is related to the input power spectral density function $S_x(f)$, and to the cross-power spectral density function $S_{xy}(f)$ between the input and the output by another simple (complex-valued) formula.

$$S_{xy}(f) = Y(jf) S_x(f) \quad (9.23)$$

Thus, if

$$S_{xy}(f) = |S_{xy}(f)| e^{j\theta(f)} \quad ; \quad Y(jf) = |Y(jf)| e^{j\phi(f)} \quad (9.24)$$

one obtains

$$\begin{aligned} |S_{xy}(f)| &= |Y(jf) S_x(f)| \\ \theta(f) &= \phi(f) \end{aligned} \quad (9.25)$$

Note that the coherency function becomes here, Equation (3.30),

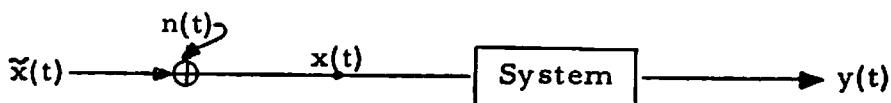
$$\gamma_{xy}^2(f) = \frac{|S_{xy}(f)|^2}{S_x(f) S_y(f)} = 1 \quad (9.26)$$

indicating complete linear dependence between x and y at every frequency.

The main limitations in applying the above formulas are due to failing to satisfy requirements that the system is of a constant parameter linear type, and that the input random process is stationary. No such simple relation exists for time varying linear systems, for non-linear systems, or for nonstationary random processes.

Consider now the additive effects of noise on frequency-response function estimates, and the crucial role played by the resulting coherency function. Two cases will be explored depending upon whether uncorrelated noise occurs in the input or in the output.

9.2.1 Case 1. Noise in Measured Input



For this case, the linear time response is given by

$$y(t) = \int_0^{\infty} h(\tau) x(t - \tau) d\tau \quad (9.27)$$

where

$$x(t) = \tilde{x}(t) + n(t) \quad (9.28)$$

represents the contaminated input, the quantity $\tilde{x}(t)$ representing the desired input signal, and $n(t)$ the noise. The spectral relationships obtained by measuring spectra and cross-spectra yield (for uncorrelated noise),

$$S_y(f) = |Y(jf)|^2 [S_x(f)] \tag{9.29}$$

$$S_{xy}(f) = S_{\tilde{x}y}(f) + S_{ny}(f) = Y(jf) S_x(f)$$

where

$$S_x(f) = S_{\tilde{x}}(f) + S_n(f) \tag{9.30}$$

The coherence function between $\tilde{x}(t)$ and $y(t)$ becomes

$$\gamma_{\tilde{x}y}^2(f) = \frac{S_{\tilde{x}y}(f)}{S_x(f)} = \frac{1}{1 + [S_n(f)/S_{\tilde{x}}(f)]} < 1 \tag{9.31}$$

showing that the coherence is now below unity, and is inversely proportional to the input noise-to-signal power ratio.

If $S_n(f) \ll S_{\tilde{x}}(f)$, then

$$\gamma_{\tilde{x}y}^2(f) \approx 1 - [S_n(f)/S_{\tilde{x}}(f)] \tag{9.32}$$

Observe that in terms of the coherence function,

$$|Y(jf)|^2 = \frac{S_y(f)}{S_x(f)} = \gamma_{\tilde{x}y}^2(f) \left[\frac{S_y(f)}{S_{\tilde{x}}(f)} \right] \tag{9.33}$$

Also

$$|Y(jf)| = \frac{|S_{xy}(f)|}{S_x(f)} = \gamma_{\tilde{x}y}^2(f) \frac{|S_{\tilde{x}y}^2(f)|}{S_{\tilde{x}}(f)} \quad (9.34)$$

if one assumes $S_{xy}(f) = S_{\tilde{x}y}(f)$. Thus the coherence function shows the extent to which the gain factor of the frequency response function has been affected by the noise. To estimate the phase shift term of the frequency response function, one may use the estimate of the phase obtained by the contaminated cross-spectrum measurement, $S_{xy}(f)$, since both the real and imaginary parts are changed to the same extent. In equation form,

$$\text{Phase of } S_{xy}(f) = \text{Phase of } Y(jf) \quad (9.35)$$

9.2.2 Case 2. Noise in Measured Output



For this case, the linear time response is given by

$$y(t) = \tilde{y}(t) + n(t) = \int_0^{\infty} h(\tau) x(t - \tau) d\tau + n(t) \quad (9.36)$$

where $n(t)$ represents the output noise, assumed to be uncorrelated with $x(t)$. The spectral relations become

Contrails

$$S_y(f) = |Y(jf)|^2 S_x(f) + S_n(f) = S_{\tilde{y}}(f) + S_n(f)$$

$$S_{xy}(f) = S_{x\tilde{y}}(f) = Y(jf) S_x(f) \quad (9.37)$$

Hence, the coherence function between $x(t)$ and $y(t)$ is

$$\gamma_{xy}^2(f) = \frac{|Y(jf)|^2 S_x(f)}{S_y(f)} = \frac{1}{1 + [S_n(f)/S_{\tilde{y}}(f)]} < 1 \quad (9.38)$$

where

$$S_{\tilde{y}}(f) = |Y(jf)|^2 S_x(f) \quad (9.39)$$

As before, the coherence function is below unity, the reduction being inversely proportional now to output noise-to-signal power ratio.

If $S_n(f) \ll S_{\tilde{y}}(f)$, then

$$\gamma_{xy}^2(f) \approx 1 - [S_n(f)/S_{\tilde{y}}(f)] \quad (9.40)$$

It appears from the above cross-spectrum measurement that, independent of the noise term, a useful estimate of the complete frequency response function (both gain and phase) is given by

$$Y(jf) = \frac{S_{xy}(f)}{S_x(f)} \quad (9.41)$$

It will be pointed out below, however, that when the coherence function is less than unity, reduced confidence limits must be placed on the results which are obtained.

9.2.3 Confidence Limits Based on Coherence Function

For both cases considered above, and for combinations thereof, an estimate of the true frequency response $Y(jf)$ may be obtained by measuring $S_{xy}(f)$ and $S_x(f)$. To distinguish between the true value of $Y(jf)$ and a particular estimate of the true value which would be measured in practice, let

$$\hat{Y}(jf) = \frac{\hat{S}_{xy}(f)}{\hat{S}_x(f)} = \hat{G}(f) e^{j\hat{\phi}(f)} \quad (9.42)$$

represent the estimate in question where $\hat{G}(f)$ denotes the estimate of the true gain factor $G(f)$, and $\hat{\phi}(f)$ denotes the estimate of the true phase shift term $\phi(f)$.

Results of Goodman [Ref. 4], quoted by Katz [Ref. 7] and partially displayed graphically by Press [Ref. 3], demonstrate that to a very close approximation,

$$\begin{aligned} \text{Prob} \left[\left| \frac{\hat{G}(f) - G(f)}{G(f)} \right| < \sin e \text{ and } \left| \hat{\phi}(f) - \phi(f) \right| < e \right] \\ \approx 1 - \left[\frac{1 - \gamma_{xy}^2(f)}{1 - \gamma_{xy}^2(f) \cos^2 e} \right]^k \end{aligned} \quad (9.43)$$

where $\gamma_{xy}^2(f)$ is the coherence function and k is the number of degrees of freedom, see Equation (7.43),

$$k = 2BT = 2N/J \quad (9.44)$$

with terms as defined previously.

This formula is of considerable practical importance in determining the confidence level at which the gain and phase can be estimated to a desired error for a given value of γ^2 and k . For example, suppose one wants to estimate the phase ϕ to within 0.1 radian ($e = 0.1$) and the gain G to within 10% error ($\sin e \approx 0.1$) at a 90% confidence level (Prob = 0.9). The table below shows the required number of degrees of freedom k corresponding to various assumed coherences γ^2 .

γ^2	0.9	0.8	0.7	0.6	0.5	($e = 0.1$ Prob = 0.9)
k	27	58.5	100	156	232	

In practice, one does not know in advance what the coherence function will be and can only roughly estimate it from the measurements. This restricts the application of the above work to some extent. However, a conservative choice is usually desirable, and the above formula and table shows clearly that accurate measurement of a frequency response function is strongly dependent upon the value of the coherence function.

9.3 ZERO CROSSINGS AND MAXIMA PROBABILITIES

Consider a random record $x(t)$ whose behavior over a long period of time exhibits many random oscillations. The expected number of zero crossings per unit time (usually seconds) of the record, denoted by N_0 , gives an indication of its "apparent frequency". For example, a 60 cps sine wave has 120 zeros per second. For a random record, the situation is, of course, more complex but still knowledge of N_0 , in addition to other quantities, helps to characterize the random record. This type of information

and certain of its extensions discussed below is particularly useful for fatigue analysis and reliability prediction of structures under random loading and vibration.

At an arbitrary level, say, $x = \alpha$, the expected number of crossings per unit time through the interval $(\alpha, \alpha + d\alpha)$, where $d\alpha$ is arbitrary small, will be denoted by N_α . It follows that the expected number of times per unit time that $x(t)$ exceeds the value α (i. e., crosses the line $x = \alpha$ with positive slope) is given by $(1/2)N_\alpha$, since $x(t)$, on the average, passes the value α half of the time with positive slope and half of the time with negative slope. When $\alpha = 0$, N_α reduces to N_0 , the expected number of zero crossings per unit time; the quantity $(1/2)N_0$ represents the expected number of zero crossings per unit time with positive or negative slopes.

General probability formulas may be written down for evaluating N_α . In practice, however, useful simple formulas have been obtained which apply only to situations where the random record $x(t)$ is assumed to be a sample member from a stationary ergodic Gaussian random process with zero mean value, governed by Equation (9.1). Analysis of this important case is due to Rice [Ref. 11], and yields the following result:

$$N_\alpha = 2 \left(\frac{\sigma_{\dot{x}}}{\sigma_x} \right) e^{-\alpha^2 / 2\sigma_x^2} \quad (9.45)$$

where

$$\sigma_x^2 = \int_{-\infty}^{\infty} S_x(f) df = 2 \int_0^{\infty} S_x(f) df$$

$$\sigma_{\dot{x}}^2 = \int_{-\infty}^{\infty} f^2 S_x(f) df = 2 \int_0^{\infty} f^2 S_x(f) df \quad (9.46)$$

Contrails

Physically, σ_x^- represents the rms value of $x(t)$, and $\sigma_{\dot{x}}^-$ represents the rms value of $\dot{x}(t)$. Thus, setting $\alpha = 0$,

$$N_0 = 2(\sigma_{\dot{x}}^-/\sigma_x^-) \quad (9.47)$$

These formulas depend upon knowledge of the power spectrum $S_x(f)$ in a surprisingly simple manner.

By an analogous but more complicated analysis, Rice derived a further property about the expected number of maxima (or minima) of $x(t)$ per second, denoted by M . Since the quantity M represents either the number of positive peaks or the number of negative peaks, which may be expected to occur equally often on the average, the expected number of both positive and negative peaks per second is given by $2M$. The expression for M turns out to be simply

$$M = (\sigma_{\ddot{x}}^-/\sigma_{\dot{x}}^-) \quad (9.48)$$

where

$$\sigma_{\dot{x}}^{-2} = \int_{-\infty}^{\infty} f^2 S_x(f) df = 2 \int_0^{\infty} f^2 S_x(f) df \quad (9.49)$$

$$\sigma_{\ddot{x}}^{-2} = \int_{-\infty}^{\infty} f^4 S_x(f) df = 2 \int_0^{\infty} f^4 S_x(f) df$$

Physically, $\sigma_{\ddot{x}}^-$ represents the rms value of $\ddot{x}(t)$.

The probability that a positive peak will fall between $(\alpha, \alpha + d\alpha)$ can also be calculated. In terms of a standard variable z with zero mean and unit variance,

$$z = (x/\sigma) \quad ; \quad \sigma^2 = \int_{-\infty}^{\infty} S_x(f) df \quad (9.50)$$

the probability density function $W(z)$ that a positive peak will fall between z and $z + dz$ is expressed by the formula

$$W(z) = \frac{k_1}{\sqrt{2\pi}} e^{-z^2/2k_1^2} + \left(\frac{N_0}{2M}\right) z e^{-z^2/2} \left[1 - P_n(z/k_2)\right] \quad (9.51)$$

where

$$k_1 = \sqrt{1 - (N_0/2M)^2} \quad \frac{N_0}{2M} = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_z^2} \quad (9.52)$$

$$k_2 = \frac{k_1}{(N_0/2M)}$$

and

$$P_n(z/k_2) = \frac{1}{\sqrt{2\pi}} \int_{z/k_2}^{\infty} e^{-y^2/2} dy \quad (9.53)$$

Note that $P_n(z/k_2)$ is the probability for a standard normal distribution with zero mean and unit variance that the value (z/k_2) will be exceeded. This integral is readily available in statistical tables.

The shape of $W(z)$ is determined by the parameter $(N_0/2M)$. It can be shown from basic considerations that $(N_0/2M)$ always falls between zero and unity. If $(N_0/2M) = 0$, then $W(z)$ reduces to a standard normal probability density function; if $(N_0/2M) = 1$, then $W(z)$ becomes a Rayleigh probability density function.

In terms of $W(z)$ the probability $P_p(z)$ that a positive peak chosen at random from among all the possible positive peaks will exceed the value z is given by the formula

$$\begin{aligned}
 P_p(z) &= \int_z^{\infty} W(z) dz \\
 &= P_n(z/k_1) + \left(\frac{N_0}{2M} \right) e^{-z^2/2} \left[1 - P_n(z/k_2) \right]
 \end{aligned}
 \tag{9.54}$$

using the P_n of Equation (9.53). Graphs of $W(z)$ and $P_p(z)$ are plotted in Reference [6].

From the above, it should be noted that the actual number of positive peaks per second which would exceed the value $\alpha = z\sigma$, denoted by M_α , may be estimated by the formula

$$M_\alpha = MP_p(\alpha/\sigma) = MP_p(z) \tag{9.55}$$

Finally, for large values of α relative to σ , one may verify

$$M_\alpha \approx (N_0/2) e^{-\alpha^2/2\sigma^2} \tag{9.56}$$

showing that for large α , the expected number of maxima per second lying above the line $x = \alpha$ is equal to the expected number of times per second that $x(t)$ crosses the line $x = \alpha$ with positive slope, an intuitively satisfying result.

This concludes the present discussion and the report.

BIBLIOGRAPHY

1. Bendat, J. S., Principles and Applications of Random Noise Theory, John Wiley and Sons, Inc., New York, 1958.
2. Blackman, R. B., and J. W. Tukey, The Measurement of Power Spectra, Dover Publications, Inc., New York, 1958.
3. Coleman, T. L., Press, H. and M. T. Meadows, "An Evaluation of Flexibility on Wing Strains in Rough Air by Means of Experimentally Determined Frequency-Response Functions with an Assessment of Random-Process Techniques Employed", Technical Note 4291, Langley Aeronautical Laboratory, Langley Field, Va., July, 1958.
4. Goodman, N. R., "On the Joint Estimation of the Spectra, Cospectrum and Quadrature Spectrum of a Two-Dimensional Stationary Gaussian Process", Scientific Paper No. 10, Engineering Statistics Laboratory, New York University, 1957.
5. Grenander, U., and M. Rosenblatt, Statistical Analysis of Stationary Time Series, John Wiley and Sons, Inc., New York, 1957.
6. Huston, W. B., and T. H. Skopinski, "Probability and Frequency Characteristics of Some Flight Buffet Loads", Technical Note 3733, Langley Aeronautical Laboratory, Langley Field, Va., August, 1956.
7. Katz, S., "Spectra and Cross-Spectra of Stationary Time Series: Measurement and Interpretation", American Cyanamid Co., New York, 1959.

BIBLIOGRAPHY (cont'd)

8. Mayer, J. P. and H. A. Hamer, "Applications of Power Spectral Analysis Methods to Maneuver Loads Obtained on Jet Fighter Airplanes During Service Operations", NACA Report RM L56J15, Langley Aeronautical Laboratory, Langley Field, Va., January, 1957.
9. Parzen, E., "On Consistent Estimates of the Spectrum of a Stationary Time Series", and "On Choosing an Estimate of the Spectral Density Function of a Stationary Time Series", *Annals Math. Statistics*, Vol. 28, pp. 329-343, June 1957; Vol. 28, pp. 921-932, December, 1957.
10. Press, H. and J. W. Tukey, Power Spectral Methods of Analysis and Their Application to Problems in Airplane Dynamics, Vol. IV of AGARD Flight Test Manual, Part IV C, Enoch J. Durbin, ed., North Atlantic Treaty Organization.
11. Rice, S. O., "Mathematical Analysis of Random Noise", *Bell System Technical Journal*, vol. 23, pp. 282-332, July, 1944; vol. 24, pp. 46-156, January, 1945. (This article appears also in Selected Papers on Noise and Stochastic Processes, edited by N. Wax, Dover Publ., New York, 1954.)
12. Smith, F. B., "Analog Equipment for Processing Randomly Fluctuating Data", *Aero. Eng. Rev.*, vol. 14, pp. 113-119, May, 1955.