TIME SERIES AND HARMONIC ANALYSIS*

J. L. DOOB

UNIVERSITY OF ILLINOIS

Introduction

Although many articles on the present subject have appeared in the mathematical, statistical, and physical literature, there still seems to be some justification for one more. The statisticians have applied only small parts of the theory; the physicists have gone deeper, but write like physicists; the mathematicians have gone furthest, but write like mathematicians, only for posterity. Their work is frequently not understood, and is in general either ignored or applied in simplified forms which often are formally more formidable than the original rigorous one. The present paper attempts to give a compact outline of the harmonic analysis of stochastic processes, with applications to physical problems.*

Time series can be analyzed from two points of view.

a) A time series is a sequence of numbers, to be analyzed for trend, periodicity, prediction possibilities, and the like. The source of the series, that is, the mathematics of the background of the numbers, is ignored.

b) A time series is a sequence of numbers arising from a certain function $f(t)$ (where t is time) as t takes on a sequence of values. There are ordinarily probability parameters in $f(t)$, so that the function value $f(t_0)$ is not uniquely determined by t_0; the function values obtained determine a sample function of a stochastic process. The general properties of this process are deduced from an analysis of the origin of the series, helped by the actual sample values obtained. Trends, periods, and the like are determined in terms of the properties of the stochastic process.

Although the first point of view is superficial, it is adequate in many applications. One reason for this adequacy is the parallelism between the properties of a stochastic process (that is, the average properties of its sample functions) and the properties of almost all the individual sample functions. The harmonic analysis of an individual sample function is formally almost identical with that of the process. In other terms, the formal analysis applied to specific data is largely independent of the background, and is essentially equivalent to that applied to stochastic processes. Although the main part of this paper

* *Note added in proofs, August 10, 1948:* Since the present paper was written, almost three years ago, a number of contributions have appeared in print, particularly by French and Scandinavian writers. Because of these publications a complete reorganization of the paper might be in order. However, it was thought best to leave the paper in its original form although it has lost much of its freshness since it was written. This explains the omission of references to the work of Loève, Karhunen, and others.

is given over to the analysis of stochastic processes, the harmonic analysis of individual functions will be outlined briefly, in order to exhibit the parallelism between the two studies, and the greater simplicity of the first.

A continuously varying time parameter presents somewhat greater analytic difficulties than an integral-valued time parameter, and the rigorous treatment is less well known. Actually, after the initial spade work has been done, there is little difference between the two cases. Both will be treated below, although more details will be given in the continuous parameter case.

The functions $f(t)$ to be analyzed below will be complex-valued, and the harmonic analysis will consist of the development of representations in terms of linear combinations of terms of the form $e^{2\pi i\lambda t}$, where λ is real. The λ's involved are the frequencies present; their reciprocals are the periods. Specialization will be made to real functions when desirable, but it is clearer and simpler to carry through all the work for complex functions before specializing. The useful case for many applications turns out to be *not* that for which the Fourier integral theory can be applied to express $f(t)$ in the form

$$f(t) = \int_{-\infty}^{\infty} e^{2\pi i\lambda t} f^*(\lambda)d\lambda, \tag{0.1}$$

in which case the problem is solved, since f^* can be found by inverting (0.1). In fact, the hypotheses imposed on $f(t)$ to make the form (0.1) possible [for example, the finiteness of $\int_{-\infty}^{\infty} |f(t)|^2 dt$] imply roughly that f becomes small at $\pm\infty$, and thus preclude the existence of non-trivial stationary properties of $f(t)$, that is, average properties as $|t| \to \infty$. On the other hand, the functions $f(t)$ to be considered are precisely those, like $e^{it} + e^{\pi it}$ for example, which have such average properties.

The representation (0.1) and its inverse

$$f^*(\lambda) = \int_{-\infty}^{\infty} e^{-2\pi i\lambda t} f(t)dt \tag{0.2}$$

will, however, be fundamentally important below, when $\int_{-\infty}^{\infty} |f(t)|^2 dt$ is finite. In this case, according to the Fourier-Plancherel theorem, f^* defined by (0.2) satisfies (0.1). (The integrals must be defined as the limits in the mean of the same integrals extended over finite limits,

$$\text{l.i.m.}_{T\to\infty} \int_{-T}^{T} .)$$

1. Harmonic analysis of individual functions[1]

The harmonic analysis of an individual function $x(t)$ is accomplished, for present purposes, in terms of its correlation function $R(t)$,

$$R(t) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x(s+t)\overline{x(s)}ds. \tag{1.1}$$

[1] Only the continuous parameter case will be considered in this section. The integral parameter case is analyzed similarly. Most of the material in this section is taken from Wiener's fundamental paper, *Acta Math.*, vol. 55 (1930), pp. 117-258.

It is supposed that $R(t)$ exists for all t and is continuous when $t = 0$.[2] It is then bounded and everywhere continuous. The correlation function is analyzed harmonically in a very simple manner. In fact, it is shown to be the Fourier transform of a real bounded monotone non-decreasing function $F(\lambda)$, called variously the spectrum, spectral function, and periodogram, of $x(t)$.

$$R(t) = \int_{-\infty}^{\infty} e^{2\pi i t\lambda}\, dF(\lambda). \tag{1.2}$$

This function makes precise Schuster's idea of a periodogram. It will be seen that, roughly speaking, the increment of $F(\lambda)$ at μ measures the square of the contribution of the intensity of the term $e^{2\pi i\mu t}$ in the expansion of $x(t)$ in a linear combination of such terms, that is, the increment of $F(\lambda)$ at μ measures the strength of the frequency μ in $x(t)$. The transition from $x(t)$ to $R(t)$ obliterates phase relations and squares the moduli of intensities. In particular, if $f(t)$ is a discrete sum of exponentials

$$x(t) = \sum_n a_n e^{2\pi i\lambda_n t}, \tag{1.3}$$

the correlation function $R(t)$ and the spectral function $F(\lambda)$ are easily evaluated explicitly;

$$\begin{cases} R(t) = \sum_n |a_n|^2 e^{2\pi i\lambda_n t}, \\ F(\lambda) = \sum_{\lambda_n < \lambda} |a_n|^2. \end{cases} \tag{1.4}$$

Thus in this case $F(\lambda)$ increases by $|a_n|^2$ when λ increases through λ_n. The connection with the usual Fourier transform theory is indicated by the following relations. The integrated Fourier transform exists,

$$y(\lambda) = \int_{-1}^{1} \frac{e^{-2\pi i\lambda t} - 1}{-2\pi i t}\, x(t)\, dt + \underset{T\to\infty}{\text{l.i.m.}} \left[\int_1^T + \int_{-T}^{-1}\right] \frac{e^{-2\pi i\lambda t}}{-2\pi i t}\, x(t) dt.^{3} \tag{1.5}$$

The natural inversion formula

$$x(t) = \underset{T\to\infty}{\text{l.i.m.}} \int_{-T}^{T} e^{2\pi i t\lambda} dy(\lambda) \tag{1.6}$$

is not always true, since the indicated limit on the right side may not even exist, but it is true if a Cesaro limit in T is taken. The modification

$$x(t) = \lim_{\epsilon\to 0} \underset{T\to\infty}{\text{l.i.m.}} \int_{-T}^{T} e^{2\pi i t\lambda}\, \frac{y(\lambda+\epsilon) - y(\lambda-\epsilon)}{2\epsilon} d\lambda \tag{1.7}$$

is always true. It will be seen that (1.6) is always true for stochastic processes.

[2] It is sometimes necessary to weaken these hypotheses slightly, but these refinements will not be discussed here.

[3] The integrand is modified near 0 to ensure boundedness.

Thus (1.6) and (1.7) give the desired harmonic analysis of $x(t)$. Moreover, the following equation ties this representation up with the spectral function.

$$F(\lambda_2) - F(\lambda_1) = \lim_{\epsilon \to 0} \int_{\lambda_1}^{\lambda_2} \frac{|y(\mu + \epsilon) - y(\mu - \epsilon)^2}{2\epsilon} d\mu. \tag{1.8}$$

This shows how the increment of $F(\lambda)$ measures the intensity of the appearance of corresponding frequency bands in the symbolic representation (1.6) or in (1.7). In an average sense, $y(\mu + \epsilon) - y(\mu - \epsilon)$ is of the order of $\epsilon^{1/2}$. These facts will be clearer in the harmonic analysis of stochastic processes.[4]

2. Harmonic analysis of stochastic processes

Just as the existence of certain long-time averages is presupposed in the harmonic analysis of individual functions, so certain permanences in time are presupposed in the harmonic analysis of stochastic processes. The usual hypotheses are that the first two moments exist, and that they are unaffected by translations of the time axis, that is, that the means m and $R(t)$ defined by

$$\begin{cases} E\{x(s)\} = m, \\ E\{[x(s+t) - m]\,\overline{[x(s) - m]}\} = R(t) \end{cases} \tag{2.1}$$

are independent of s. In particular, if the process is real and Gaussian, that is, if for every finite set of values of t the corresponding $x(t)$'s determine a multivariate real Gaussian distribution, the values of m and $R(t)$ determine the process uniquely. In this case the process is temporally homogeneous:[5] all probability relations are independent of translations of the t-axis. In general, however, the processes considered may not be temporally homogeneous, and are accordingly called temporally homogeneous in the wide sense.

The hypothesis that a stochastic process is temporally homogeneous, even if only in the wide sense, is evidently very restrictive, and apparently excludes all evolutory processes, but this is not quite true, as the examples below will demonstrate. The examples are all given for integral-valued t; the analogues for the continuous parameter case are less well known, and will be described later.

a) Stationary process of uncorrelated variables.—Let $\cdots, x_n, \cdots$ be uncorrelated chance variables with common expectations and variances. The process thus determined is evidently stationary in the wide sense.

b) Stationary process of moving averages.—Define X_n by

$$X_n = \sum_{j=-\infty}^{\infty} a_j x_{n-j}, \tag{2.2}$$

[4] See Wiener and Wintner, *Amer. Jour. Math.*, vol. 63 (1941), pp. 415-426, 794-824, for an analysis of the Fourier coefficients of $x(t)$ and the corresponding analysis for stochastic processes.

[5] The expressions "stationary" and "temporally homogeneous" are used interchangeably below.

where the x_m are the variables of a stationary process (in the wide sense) of uncorrelated variables. Unless all but a finite number of the a_j vanish, restrictions must be imposed to ensure convergence (in the mean). It is sufficient if $\sum_j a_j$ and $\sum_j |a_j|^2$ converge. The X process is evidently stationary in the wide sense, and is called a stationary process of moving averages. More generally, if the x process is any stationary process in the wide sense, (2.2) defines an X process (if there is convergence) which is stationary in the wide sense, and which is actually stationary if the x process is stationary.

It is now well known that the use of moving averages as in (2.2) may introduce periodicities into the X process which were not present in the basic x process, or profoundly modify those present. A detailed analysis will be made of this phenomenon below, since the simple explanation does not seem to be generally known.

c) *Process of linear regression.*—A stationary process (in the wide sense) of linear regression is a stationary process in the wide sense for which the difference

$$X_{n+N} - a_1 X_{n+N-1} - \cdots - a_N X_n = x_{n+N} \tag{2.3}$$

is uncorrelated with the chance variables . . . , X_{n+N-2} X_{n+N-1},, for some integer N and numbers $a_1, \cdots, a_N$. The x process will then necessarily be a stationary process (wide sense) of uncorrelated variables. The chance variable X_{n+N} is obtained, when the preceding variables are known, by an experiment which gives a value of x_{n+N}. If the process has a definite starting point when $n = 1$, successive values of X_m can be found, once $X_1, \cdots, X_N$ are known by evaluation of the x_m and calculations using (2.3). The distributions of the X_m will depend on the initial values $X_1, \cdots, X_N$, but asymptotically the X_m will become uncorrelated with early values, in most cases that occur in practice, and the X_m will determine asymptotically (large M) a stationary process in the wide sense, however $X_1, \cdots, X_N$ are chosen. The process will be actually stationary in the wide sense if $X_1, \cdots, X_N$ are suitably chosen.

The process of linear regression is a special case of the process of moving averages, since in most cases (2.3) implies

$$X_n = \sum_j b_j x_{n-j} \tag{2.4}$$

for suitably chosen b_j. For example, the simplest non-trivial case of linear regression,

$$X_{n+1} = aX_n + x_{n+1}, \qquad |a| < 1, \tag{2.5}$$

leads to

$$X_n = \sum_0^\infty a^j x_{n-j}. \tag{2.6}$$

This example shows how the irregularity caused by the stochastic element x_n may just counterbalance an average tendency (in this case a systematic tendency toward zero due to the fact that $|a| < 1$) so that the resulting X_n process is stationary.

If the x process is stationary in the strict sense, the qualification "in the wide sense" can be omitted throughout the preceding discussion.

It is usually convenient in studying stationary processes to replace $x(t)$ by $x(t) - m$ if $m \neq 0$, so that the first moments of the new process vanish, and this will generally be done below. It will be seen that the correlation function $R(t)$ determines the average harmonic properties of $x(t)$, and to a considerable extent shares these properties. It corresponds exactly to the correlation function defined in (1.1) for a single function $x(t)$. It is clear, however, that the individual correlation function of a sample function $x(t)$ from a stochastic process will in general be a chance variable for each t, varying from sample function to sample function.[6] As in the harmonic analysis of individual functions, the correlation function is supposed continuous at the origin. It is then everywhere continuous, and can always be expressed as the Fourier-Stieltjes transform of a bounded monotone non-decreasing function;

$$R(t) = \int_{-\infty}^{\infty} e^{2\pi i t\lambda}\, dF(\lambda), \qquad F(-\infty) = 0, \quad F(\lambda) = F(\lambda - 0).^{7} \tag{2.7}$$

The function $F(\lambda)$ is called the spectrum or spectral function of the process. The indicated normalization will be convenient, but is not essential to the representation.

Formula (2.7) for $R(t)$ is easily inverted, and the details of this inversion will be useful below. Any linear combination of $R(t_1), \cdots, R(t_\nu)$ is equal, according to (2.7), to the integral with respect to $F(\lambda)$ of the same linear combination of $e^{2\pi i t_1\lambda}, \cdots, e^{2\pi i t_\nu\lambda}$ Hence the formula can be inverted to give $F(\mu_2) - F(\mu_1)$, with proper allowance for possible discontinuities of $F(\lambda)$ at μ_1, or μ_2, by choosing the linear combination to give, under the integral sign, an expansion of the function which is 1 between μ_1 and μ_2, and otherwise vanishes. In order to make this idea precise, we define $\alpha_\mu(\lambda)$ by

$$\begin{cases} \alpha_\mu(\lambda) = 0 & \lambda < 0, \quad \lambda > \mu, \\ \mu > 0: \quad = 1/2, & \lambda = 0, \quad \lambda = \mu, \qquad \mu < 0: \ \alpha_\mu(\lambda) = -\alpha_{-\mu}(-\lambda), \\ = 1, & 0 < \lambda < \mu. \end{cases} \tag{2.8}$$

[6] If the process is stationary in the strict sense, the limit in (1.1) will exist for almost all sample functions, at each value of t, in accordance with the strong law of large numbers. The limit will be a constant independent of the sample function if the transformation group $t \to t + s$, $(-\infty < s < \infty)$, is metrically transitive in the probability measure. For further developments in this direction see the papers of Wiener and Wintner referred to in footnote 4 above.

[7] A. Khintchine, *Math. Ann.*, vol. 109 (1934), pp. 604-615. This paper laid the foundations of the theory of continuous parameter stationary processes. Wold, in *A Study in the Analysis of Stationary Time Series* (Uppsala, 1938), made the first systematic study of stationary processes of moving averages and of linear regression.

Then if $\mu_1 < \mu_2$,

$$\begin{aligned}\alpha_{\mu_2}(\lambda) - \alpha_{\mu_1}(\lambda) &= 0, && \lambda < \mu_1, \quad \lambda > \mu_2,\\ &= 1/2, && \lambda = \mu_1, \quad \lambda = \mu_2,\\ &= 1, && \mu_1 < \lambda < \mu_2.\end{aligned}$$

We can write this function, as a linear combination of terms like $e^{2\pi it\lambda}$, more precisely as a Fourier integral,

$$\alpha_{\mu_2}(\lambda) - \alpha_{\mu_1}(\lambda) = \int_{-\infty}^{\infty} \frac{e^{-2\pi is\mu_2} - e^{-2\pi is\mu_1}}{-2\pi is} e^{2\pi is\lambda} ds. \tag{2.9}$$

Inserting this in (2.7), we obtain the desired inversion,

$$\frac{F(\mu_2 + 0) + F(\mu_2)}{2} - \frac{F(\mu_1 + 0) + F(\mu_1)}{2} = \int_{-\infty}^{\infty} \frac{e^{-2\pi is\mu_2} - e^{-2\pi is\mu_1}}{-2\pi is} R(s) ds.^{8} \tag{2.10}$$

The desired expression for $x(t)$ with $E\{x(t)\} = 0$, is the analogue of (2.7),

$$x(t) = \int_{-\infty}^{\infty} e^{2\pi it\lambda} dy(\lambda) \tag{2.11}$$

for suitably defined $y(\lambda)$. The natural way to find $y(\lambda)$ is to operate on $x(t)$ just as on $R(t)$ in (2.10). Suppose first that $F(\lambda)$ has no discontinuites and define $y(\mu)$ by

$$y(\mu) = \operatorname*{l.i.m.}_{S \to \infty} \int_{-S}^{S} \frac{e^{-2\pi is\mu} - 1}{-2\pi is} x(s) ds. \tag{2.12}$$

The existence of the indicated limit in the mean follows from

$$E\left\{ \left| \left[\int_{-S'}^{S'} - \int_{-S}^{S} \right] \frac{e^{-2\pi is\mu} - 1}{-2\pi is} x(s) ds \right|^2 \right\} \tag{2.13}$$

$$= E\left\{ \left[\int_{-S'}^{S'} - \int_{-S}^{S} \right] \frac{e^{-2\pi is\mu} - 1}{-2\pi is} \cdot \frac{e^{2\pi it\mu} - 1}{2\pi it} x(s)\, \overline{x(t)} ds dt \right\}$$

$$= \int_{-\infty}^{\infty} \left[\int_{-S'}^{S'} - \int_{-S}^{S} \right] \frac{e^{-2\pi is\mu} - 1}{-2\pi is} \cdot \frac{e^{2\pi it\mu} - 1}{2\pi it} e^{2\pi i(s-t)\lambda} ds dt dF(\lambda)$$

$$= \int_{-\infty}^{\infty} \left| \left[\int_{-S'}^{S'} - \int_{-S}^{S} \right] \frac{e^{-2\pi is\mu} - 1}{-2\pi is} e^{2\pi is\lambda} ds \right|^2 dF(\lambda).$$

---

[8] The improper integral is defined by $\lim_{S \to \infty} \int_{-S}^{S}$. This proof of the Lévy inversion formula, giving a distribution function in terms of its characteristic function, is outlined because it illustrates and makes natural the later developments. The proof only requires the simple fact that the integral in (2.9) between the limits $-T$, T converges boundedly to $\alpha_{\mu_2}(\lambda) - \alpha_{\mu_1}(\lambda)$ when $T \to \infty$. The corresponding proof for integral parameter processes uses Fourier series instead of Fourier integrals.

In fact, the last integral converges to zero as S, $S' \to \infty$ because the Fourier integral for $\alpha_\mu(\lambda)$ converges boundedly to $\alpha_\mu(\lambda)$. Thus (2.12) defines a chance variable for each μ.[9] Obviously

$$E\{y(\mu)\} = 0. \tag{2.14}$$

Moreover, according to (2.13) with $S = 0, S' = \infty$,

$$E\{|y(\mu)|^2\} = \int_{-\infty}^{\infty} \alpha_\mu(\lambda)^2 dF(\lambda) = F(\mu) - F(0).$$

More generally, a similar evaluation shows that

$$E\{[y(\mu_2) - y(\mu_1)]\,\overline{[y(\mu_4) - y(\mu_3)]}\} \tag{2.15}$$

$$= \int_{-\infty}^{\infty} [\alpha_{\mu_2}(\lambda) - \alpha_{\mu_1}(\lambda)]\,\overline{[\alpha_{\mu_4}(\lambda) - \alpha_{\mu_3}(\lambda)]}\, dF(\lambda).$$

In particular, if $\mu_3 = \mu_1$, $\mu_4 = \mu_2$, this becomes

$$E\{|y(\mu_2) - y(\mu_1)|^2\} = F(\mu_2) - F(\mu_1). \tag{2.16}$$

If $F(\lambda)$ has discontinuities, we define $y(\mu)$ at the continuity points of $F(\lambda)$ by (2.12). Then (2.16) is true if $F(\lambda)$ is continuous at μ_1 and μ_2.
This equation implies that

$$\text{l.i.m.}_{\mu \downarrow \mu_0} y(\mu) = y(\mu_0 + 0), \qquad \text{l.i.m.}_{\mu \downarrow \mu_0} y(\mu) = y(\mu_0 - 0)$$

exist for every μ_0, and that, if $F(\lambda)$ is continuous at μ_0, these one-sided limits are equal to $y(\mu_0)$ with probability 1. Then we define

$$y(\mu) = y(\mu - 0)$$

[9] The measurability and integrability of $x(s)$ and similar properties used here are too deep for present study, but are not difficult to justify, and have been justified in the literature. The point is that sample functions can be treated just the same as any other function, the stochastic element being simply another parameter in the equations, which traditionally is not inserted explicitly. That is to say, whenever we write $x(t)$ we really mean that we have a function of two variables, $x(t, \xi)$, and the probability of any aggregate of sample functions $x(t)$ is simply the measure of the aggregate of values of ξ which, inserted in $x(t, \xi)$, yield those sample functions. If desired, measurability and integrability problems can be avoided temporarily by defining the integrals of $x(t)$ and similar functions as the limits in the mean of Riemann sums, instead of defining them as ordinary integrals for each sample function; but the measurability and integrability problem must be faced at some stage, and the earlier it is faced, the sooner circumlocutions can be avoided and the sooner the ordinary analytical methods can be applied with a clear conscience. On the other hand, the more recent theory that probability is not a mathematical subject, in particular not rigorous, unless the probability parameter ξ is always mentioned explicitly, and even that the rigor is enhanced if ξ is a real number varying between 0 and 1, probability measure being Lebesgue measure, is a delusion.

wherever $y(\mu)$ is not already defined; with this definition (2.16) will hold for all μ_1, μ_2, and (2.15) will also hold with no restrictions on the μ_j if each $\alpha_\mu(\lambda)$ is replaced by $\alpha_\mu(\lambda - 0)$. With this change (2.15) now shows that the y process is a process whose increments in non-overlapping intervals are uncorrelated, even if the intervals have an endpoint in common. Processes with uncorrelated increments will be studied in section 3, where it will be shown that integrals like that in (2.11) over finite intervals are definable in terms of the corresponding Riemann-Stieltjes sums, even though $y(\mu)$ will not in general be of bounded variation; the integral over an infinite interval is the limit in the mean of the integral over finite intervals. The truth of (2.11) will now be shown, with the use of this fact. From (2.12), if μ_1 and μ_2 are continuity points of $F(\lambda)$,

$$E\{x(t)\,\overline{[y(\mu_2) - y(\mu_1)]}\} = \lim_{S\to\infty}\int_{-S}^{S} \frac{e^{2\pi i s\mu_2} - e^{2\pi i s\mu_1}}{2\pi i s} E\{x(t)\overline{x(s)}\}ds \tag{2.17}$$

$$= \lim_{S\to\infty}\int_{-\infty}^{\infty}\int_{-S}^{S} e^{2\pi i(t-s)\lambda}\frac{e^{2\pi i s\mu_2} - e^{2\pi i s\mu_1}}{2\pi i s}\,ds\,dF(\lambda)$$

$$= \lim_{S\to\infty}\int_{-\infty}^{\infty}\left[\int_{-S}^{S}\frac{e^{2\pi i s\mu_1} - e^{2\pi i s\mu_2}}{2\pi i s}e^{-2\pi i s\lambda}ds\right]e^{2\pi i t\lambda}dF(\lambda).$$

The quantity in the brackets becomes, as $S \to \infty$, the inverse Fourier integral of the Fourier integral for $\alpha_{\mu_2}(\lambda) - \alpha_{\mu_1}(\lambda)$. Hence

$$E\left\{x(t)\,\overline{[y(\mu_2) - y(\mu_1)]}\right\} = \int_{\mu_1}^{\mu_2} e^{2\pi i t\lambda}dF(\lambda), \tag{2.18}$$

and it is now clear that (2.18) is true for any μ_1, μ_2, since $y(\mu) = y(\mu - 0)$ and $F(\lambda) = F(\lambda - 0)$. Let $\mu_0, \cdots, \mu_n$ be the points of a subdivision of a finite interval (a,b), $(a = \mu_0 < \cdots < \mu_n = b)$. Then if $\delta = \max.\ (\mu_j - \mu_{j-1})$, it follows from (2.18) that

$$E\left\{x(t)\int_a^b \overline{e^{2\pi i\mu s}dy(\mu)}\right\} = \lim_{\delta\to 0} E\left\{x(t)\sum_j e^{-2\pi i\mu_j s}\,\overline{[y(\mu_j) - y(\mu_{j-1})]}\right\} \tag{2.19}$$

$$= \lim_{\delta\to 0}\sum_j e^{-2\pi i\mu_j s}\int_{\mu_{j-1}}^{\mu_j} e^{2\pi i t\lambda}dF(\lambda) = \int_a^b e^{2\pi i(t-s)\lambda}dF(\lambda).$$

Hence, if $a \to -\infty$, $b \to +\infty$,

$$E\left\{x(t)\int_{-\infty}^{\infty}\overline{e^{2\pi i\mu s}dy(\mu)}\right\} = \int_{-\infty}^{\infty} e^{2\pi i(t-s)\lambda}\,dF(\lambda) = R(t-s). \tag{2.20}$$

In order to prove (2.11) we need only evaluate the expectation of the absolute value of the difference squared, getting zero.[10]

The case in which $x(t)$ is real is of special interest. In this case, according to (2.12),

$$y(\mu) + \overline{y(-\mu)} = 0 \tag{2.21}$$

if μ is a continuity point of $F(\lambda)$. If $\mu \downarrow 0$, this becomes

$$y(0+0) + \overline{y(0)} = 0, \tag{2.22}$$

and therefore, at the continuity points of $F(\lambda)$,

$$y(\mu) - y(0+0) + \overline{y(-\mu) - y(0)} = 0. \tag{2.23}$$

This implies that, at these values of λ, $F(\lambda) - F(0+0) = F(0) - F(-\lambda)$. Then (2.7) reduces to

$$R(t) = 2\int_0^\infty \cos 2\pi t\lambda dF(\lambda) - [F(0+0) - F(0)], \tag{2.24}$$

$$= \int_0^\infty \cos 2\pi t\lambda dG(\lambda), \quad G(\lambda) = 2[F(\lambda) - F(0)] - [F(0+0) - F(0)].$$

If y is separated into its real and imaginary parts, $y = y_1 + iy_2$, we find from (2.22) that $y_2(0+0) = y_2(0)$, so that the jump of $y(\mu)$ at 0, if any, must be confined to the real part. If (2.23) and this fact are combined with the given orthogonality relations of the y process, we obtain

$$\begin{cases} E\{|y_j(\mu_2) - y_j(\mu_1)|^2\} = \dfrac{F(\mu_2) - F(\mu_1)}{2}, & 0 < \mu_1 < \mu_2, \quad j = 1, 2, \\ E\{[y_j(\mu_2) - y_j(\mu_1)][y_k(\mu_4) - y_k(\mu_3)]\} = 0, & j \neq k, \\ \qquad = 0, & j = k \text{ if } 0 < \mu_1 < \mu_2 \leqq \mu_3 < \mu_4, \end{cases} \tag{2.25}$$

and, as we have already noted,

$$\begin{cases} y_2(0+0) - y_2(0) = 0, \\ E\{|y_1(0+0) - y_1(0)|^2\} = F(0+0) - F(0). \end{cases} \tag{2.26}$$

[10] The expectation of the product of two integrals of the form in (2.11) is evaluated in (3.11). This evaluation, incidentally, shows how simply the normal form (2.7) for the correlation function follows from (2.11). The normal form (2.11) was derived by Cramér [*Ark. Mat. Ast. och Fys.*, vol. 28B, no. 12 (1942), pp. 1-17]. Kolmogoroff [*C. R. Acad. Sci. URSS*, vol. 26 (1940), pp. 115-118] had already indicated that (2.11) is a special case of the well-known representation of a unitary transformation in Hilbert space in terms of a canonical resolution of the identity.

Hence the y_1 and y_2 processes are mutually uncorrelated processes, and are individually real processes with uncorrelated increments. The canonical form (2.11) reduces to

$$x(t) = 2\int_0^\infty \cos 2\pi t\lambda dy_1(\lambda) - 2\int_0^\infty \sin 2\pi t\lambda dy_2(\lambda) - [y_1(0+0) - y_1(0)]. \quad (2.27)$$

The function $G(\lambda)$, defined in (2.24) for $\lambda > 0$ only, is the appropriate spectral function in the real case. When using this spectral function, note that spectral intensities are doubled away from the origin; $dG(\lambda) = 2dF(\lambda)$ when $\lambda > 0$.

The parallelism between the harmonic analysis of individual functions and stochastic processes (and the greater formal simplicity of the latter) is now clear. The greater simplicity of the harmonic analysis of processes is indicated by the fact that (2.11) is always true, whereas the analogue for individual functions, (1.6), is not true, in general. From now on only processes will be considered.

The canonical form (2.11) shows that the $x(t)$ of a stationary process (wide sense) is a limit of linear combinations of exponentials:

$$x(t) = \sum_j \delta_j e^{2\pi i t\lambda_j}, \qquad E\{|\delta_j|^2\} = F(\lambda_{j+1}) - F(\lambda_j), \qquad \lambda_1 < \lambda_2 < \cdots, \quad (2.28)$$

where the coefficients $\{\delta_j\}$ are mutually uncorrelated, with second moments determined by the increments of the spectral function, as indicated. It has been customary in the applications to write $x(t)$ as a finite or, infinite series of the type exhibited in (2.28), that is, series of exponentials with mutually uncorrelated coefficients.[11] The present discussion shows two things:

1. There is no loss of generality in writing $x(t)$ as the sum of such a series (as a first approximation).
2. There is, however, no gain in simplicity in (2.28) over the exact form (2.11).

If t is an integral-valued parameter, (2.11) reduces to an integral over a finite range,

$$x(n) = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{2\pi i n\lambda} dy(\lambda), \quad (2.11')$$

where $y(\lambda)$ has the same properties as in the continuous parameter case, and the correlation function $R(n)$ is given by

$$R(n) = E\{x(m+n)\overline{x(m)}\} = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{2\pi i n\lambda} dF(\lambda), \quad (2.7')$$

where

$$E\{|y(\mu_2) - y(\mu_1)|^2\} = F(\mu_2) - F(\mu_1), \qquad \mu_1 < \mu_2. \quad (2.16')$$

As in the continuous parameter case, the spectral intensities are doubled, $\lambda > 0$, if the processes are real.

[11] See, for example, the valuable papers of S. O. Rice, *Bell System Technical Journal*, vol. 23 (1944), pp. 282-332; vol. 24 (1945), pp. 46-156.

The periodic properties of the correlation function $R(t)$ have frequently been used to derive information on those of $x(t)$. This is justified by the following theoretical considerations (the continuous parameter case will be considered). Corresponding to the approximation (2.28) for $x(t)$ is the approximation

$$R(t) = \sum_j e^{2\pi i t\lambda_j} [F(\lambda_{j+1}) - F(\lambda_j)] \tag{2.29}$$

for $R(t)$. This shows that $R(t)$ involves the same frequencies as $x(t)$, with intensities the mean square intensities of those of $x(t)$. [Compare with (1.3) and (1.4) and note that the sum in (2.29) is the exact correlation function of the stationary process defined by the sum in (2.28).]

3. Processes with uncorrelated increments

The fundamental role of processes with uncorrelated increments, for both integral parameter and continuous parameter processes, has been indicated, in the preceding section, in the canonical forms obtained for stationary processes, (2.11) and (2.11′). Processes with uncorrelated increments play the role for continuous parameter processes played by sequences of uncorrelated chance variables for integral parameter processes.

Processes with uncorrelated increments are those whose chance variables $y(t)$ satisfy the following hypotheses:

$$E\{|y(t) - y(s)|^2\} < \infty, \tag{3.1}$$

and

$$E\{[y(t_2) - y(t_1) - m(t_2) + m(t_1)] \overline{[y(t_4) - y(t_3) - m(t_4) + m(t_3)]}\} = 0, \quad \text{if } t_1 < t_2 \leqq t_3 < t_4, \tag{3.2}$$

where

$$E\{y(t) - y(0)\} = m(t). \tag{3.3}$$

The hypothesis (3.2) of lack of correlation between increments in non-overlapping intervals implies that, if $F(t)$ is defined by

$$\begin{aligned} F(t) &= E\{|y(t) - y(0) - m(t)|^2\}, \qquad t \geqq 0 \\ &= -E\{|y(t) - y(0) - m(t)|^2\}, \qquad t \leqq 0, \end{aligned} \tag{3.4}$$

then $F(t)$ satisfies the equation

$$F(t) - F(s) = E\{|y(t) - y(s) - m(t) + m(s)|^2\}, \qquad s < t. \tag{3.5}$$

Hence $F(t)$ is monotone non-decreasing. It is usually more convenient to deal with processes with uncorrelated increments for which the means $\{m(t)\}$ vanish; this can be done by replacing $y(t)$ with $y(t) - m(t)$.

The average change of $y_1(t) = y(t) - m(t)$ with a change h in t is of the order of $|h|^{\frac{1}{2}}$, whenever $F'(t)$ exists and does not vanish:

$$E\{|y_1(t+h) - y_1(t)|^2\} = |F(t+h) - F(t)| \sim |h|F'(t), \qquad h \to 0. \quad (3.6)$$

The class of processes with uncorrelated increments is closely related to the class of processes with independent increments (for which the only hypothesis is that the increments over any set of non-overlapping intervals are mutually independent). In fact, if the second moments of the variables in a process of the latter type are finite, (3.2) is certainly true. Two real processes which are of this type will be useful below.

Poisson process with independent increments.—It is supposed that the increments of the chance variables $\{y(t)\}$ only assume integral values, and that

$$P\{y(t) - y(s) = \nu\} = e^{-c(t-s)} \frac{(t-s)^\nu}{\nu!} c^\nu, \qquad \nu = 0, 1, 2, \cdots, \quad s < t, \quad (3.7)$$

where c is a positive constant. In this case

$$\begin{cases} m(t) = E\{y(s+t) - y(s)\} = ct, \\ F(t) = E\{|y(s+t) - y(s) - ct|^2\} = ct, \qquad t > 0. \end{cases} \quad (3.8)$$

The sample functions are (with probability 1) monotone non-decreasing in t, increasing only in isolated jumps of magnitude 1.

Gaussian process with independent increments.—It is supposed that, for any parameter values $t_1 < \cdots < t_n$, the real chance variables $y(t_2) - y(t_1), \cdots, y(t_n) - y(t_{n-1})$ have an n-variate Gaussian distribution and are mutually independent. (Note that with real Gaussian chance variables independence means the same thing as lack of correlation.) In the best known example of this type of process, $m(t) \equiv 0$ and $F(t) = ct$, where c is a positive constant. If $F(t)$ is continuous, the general case can be reduced to this special case by a change of variable (in t).

In the first of the examples above, $y(t)$ is (with probability 1) of bounded variation in every finite interval. In the second example, although the sample functions are known to be continuous, the probability that they will be of bounded variation in finite intervals is zero, if $F(t)$ is continuous,[12] and actually increases.

In both of the examples, $m(t)$ and $F(t)$ were proportional to t. Processes of subject type with these additional properties will be said to have stationary increments in the wide sense, or in the strict sense if the distribution of increments (and not merely the first two moments) is unaffected by translations of the t-axis.

[12] See, for example, Paley and Wiener, "Fourier transforms in the complex domain," *Amer. Math. Soc. Colloq. Publ.*, vol. 19, chap. 9.

As just noted, the sample functions of a process with uncorrelated increments may not be of bounded variation even in finite intervals. In spite of this fact, "integrals" of the form

$$\int_a^b f(t)dy(t) \tag{3.9}$$

are commonly used [see (2.11)]. They are easily defined in such a way that the usual formal manipulations on integrals will be admissible. It is clear that if the "integral" (3.9) is defined in terms of the usual sums then, formally at least, it becomes a chance variable satisfying

$$E\left\{\int_a^b f(t)dy(t)\right\} = 0, \quad \text{if } m(t) \equiv 0, \tag{3.10}$$

$$E\left\{\int_a^b f(t)dy(t)\cdot\int_a^b g(s)dy(s)\right\} = \int_a^b f(t)\overline{g(t)}dF(t), \quad \text{if } m(t) \equiv 0. \tag{3.11}$$

In fact, it will now be shown that, if $m(t) \equiv 0$, (3.9) can be defined to satisfy (3.10) and (3.11), to obey the standard rules for manipulation of integrals (that is, integration by parts and addition of integrals over intervals of integration with common end points), and, if $f(t)$ is Riemann-Stieltjes integrable with respect to $F(t)$, to be the limit in the mean of the usual Riemann-Stieltjes sums. The only hypothesis on $f(t)$ except at the last point is that it be measurable with respect to $F(\lambda)$ and that the integral $\int_a^b |f|^2dF$ be finite. (We assume for simplicity that F is continuous; otherwise a slight modification would be required below, in (3.12) and (3.16).)

Suppose first that $f(t) = 0$ in (a,b) except in the finite subinterval $a_1 \leqq t \leqq b_1$, in which $f(t) \equiv$ const. $= \alpha$. Then (3.9) is defined by

$$\int_a^b f(t)dy(t) = \alpha[y(b_1) - y(a_1)]. \tag{3.12}$$

If $f(t)$ is any linear combination of such functions, (3.9) is defined as the same linear combination of the y increments. Let H be the set of functions $f(t)$ measurable with respect to $F(t)$, and with $\int_a^b |f|^2dF < \infty$. The "integral" (3.9) has now been defined on a linear manifold H_0 of functions $f(t)$, and H_0 is everywhere dense in H. Moreover, it is clear that for functions in H_0, at least, (3.10) and (3.11) are true. Hence

$$E\{|\int_a^b f(t)dy(t) - \int_a^b g(t)dy(t)|^2\} \tag{3.13}$$

$$= \int_a^b |f(t) - g(t)|^2dF(t), \qquad f,g \in H_0, \quad m(t) \equiv 0.$$

Thus the transformation defined by (3.12) from the subset H_0 of Hilbert space H to a subset H'_0 of the Hilbert space H' of chance variables whose means are zero and whose dispersions are finite is distance preserving. Therefore the domain of definition of the transformation can be extended in a unique fashion to be defined on H and to take it isometrically into a subset of H'. The "integral" (3.9) is now defined for all the desired functions $f(t)$, if $m(t)$ vanishes identically. If $m(t)$ does not vanish identically, (3.9) is defined in the obvious way,

$$\int_a^b f(t)dy(t) = \int_a^b f(t)d[y(t) - m(t)] + \int_a^b f(t)dm(t), \tag{3.14}$$

where it is supposed that $m(t)$ is of bounded variation in (a,b) [or in every finite subinterval if (a,b) is an infinite interval] and that the last integral in (3.14) exists. Then (3.10) and (3.11) become

$$E\{\int_a^b f(t)dy(t)\} = \int_a^b f(t)dm(t), \tag{3.10'}$$

$$E\{\int_a^b f(t)dy(t) \cdot \int_a^b \overline{g(s)dy(s)}\} = \int_a^b f(t)\overline{g(t)}dF(t) + \int_a^b f(t)dm(t)\int_a^b \overline{g(s)}\ \overline{dm(s)}. \tag{3.11'}$$

Suppose now that $f(t)$ is real and bounded, and let $t_0, \cdots, t_n$ be the points of a subdivision of (a,b) (a,b finite) into subintervals: $a = t_0 < \cdots < t_n = b$. Let $f_1, \cdots, f_n$ be any numbers satisfying the inequalities

$$\underset{t_j \leqq t \leqq t_{j+1}}{\text{g. l. b.}} f(t) \leqq f_j \leqq \underset{t_j \leqq t \leqq t_{j+1}}{\text{l. u. b.}} f(t). \tag{3.15}$$

$$E\{|\sum_j f_j[y(t_{j+1}) - y(t_j)] - \int_a^b f(t)dy(t)|^2\} \tag{3.16}$$

$$= E\{|\sum_j \int_{t_j}^{t_{j+1}} [f_j - f(t)]\,dy(t)|^2\} = \sum_j \int_{t_j}^{t_{j+1}} |f_j - f(t)|^2 dF(t) \qquad m(t) \equiv 0.$$

Hence if $m(t) \equiv 0$ the usual Riemann-Stieltjes sums converge to the integral (3.9) (the quotation marks will hereafter be omitted) whenever $f(t)$ is Riemann-Stieltjes integrable with respect to $F(t)$. Evidently the same is true even if $m(t)$ does not vanish, provided that $m(t)$ is of bounded variation in (a,b) and that $f(t)$ is Riemann-Stieltjes integrable with respect to $m(t)$. Finally, the standard summation procedure shows that integration by parts is permissible in (3.9) if $f(t)$ is sufficiently regular.[13]

In section 2, various classes of integral parameter stationary processes were defined. Continuous parameter analogues will now be given for two of them. The analogue of the third will be given later.

[13] Integrals of type (3.9) were apparently first discussed by Wiener. For a discussion from a slightly different point of view, see the reference given in footnote 12.

Stationary process of uncorrelated variables.—A process determined by variables $\{x(t)\}$ which are uncorrelated is not useful in the continuous parameter case, since $x(t)$ is not then measurable in t except when $x(t)$ is identically constant. The type process in this connection that is useful is the process with stationary uncorrelated increments, for which

$$\begin{cases} m(t) = E\{y(t) - y(0)\} = mt, \\ F(t) = E\{|y(t) - y(0)|^2\} = ct, \qquad c \geqq 0, \end{cases} \tag{3.17}$$

where m, c are constants. Equations (3.17) lead to

$$\begin{cases} E\{y(t) - y(s)\} = m(t - s), \\ E\{|y(t) - y(s)|^2\} = c|t - s|. \end{cases} \tag{3.18}$$

Stationary process of moving averages.—Define $x(t)$ by

$$x(t) = \int_{-\infty}^{\infty} f(t - s)dy(s), \tag{3.19}$$

where the y process is one with stationary uncorrelated increments (wide sense), with m, c defined as in (3.18). It is supposed that $f(t)$ is Lebesgue measurable and that

$$\begin{cases} \int_{-\infty}^{\infty} |f(t)|\,dt < \infty, & \text{if } m \neq 0, \\ \int_{-\infty}^{\infty} |f(t)|^2 dt < \infty, & \text{if } c \neq 0. \end{cases} \tag{3.20}$$

The x process thus defined is the analogue of the X process defined by (2.2) and may therefore properly be called a stationary process of moving averages (wide sense) in the continuous parameter case. According to (3.10′) and (3.11′)

$$\begin{cases} \mu = E\{x(t)\} = m\int_{-\infty}^{\infty} f(t - s)ds = m\int_{-\infty}^{\infty} f(t)dt, \\ E\{[x(t + h) - \mu][x(t) - \mu]\} = c\int_{-\infty}^{\infty} f(t - s + h)\overline{f(t - s)}ds \\ \qquad = c\int_{-\infty}^{\infty} f(s + h)\overline{f(s)}ds = c\int_{-\infty}^{\infty} |f^*(\lambda)|^2 e^{2\pi i\lambda h}d\lambda, \end{cases} \tag{3.21}$$

where f^* is the Fourier transform of f [satisfying (0.1) and (0.2)].

Thus (3.19) determines a stationary process in the wide sense (or strict sense if the y differences are stationary in the strict sense) and (3.21) exhibits the correlation function in the standard form (2.7).

In particular, if $f = 0$ except in the interval $(-\delta,0)$, in which $f = 1/\delta$, $x(t)$ becomes the difference quotient $[y(t+\delta) - y(t)]/\delta$, which, by (3.21), has spectral density

$$c\left|\frac{1-e^{2\pi i\lambda\delta}}{-2\pi i\lambda\delta}\right|^2 \sim c, \qquad \delta \to 0. \tag{3.22}$$

In other words, the difference quotient, as far as the frequency spectrum is concerned, behaves as if $y'(t)$ existed and had constant spectral density c. The fact that c is not integrable over $(-\infty,\infty)$ corresponds to the fact that $y'(t)$ does not actually exist. As an example of the apparent constant spectral density of y', consider the spectrum of a process determined by variables $\{Z(t)\}$, obtained by moving averages over any stationary process (wide sense) with variables $\{z(t)\}$:

$$Z(t) = \int_{-\infty}^{\infty} f(t-s)z(s)ds, \qquad E\{z(s)\} = m. \tag{3.23}$$

The correlation function is easily evaluated,

$$\begin{cases} \mu = E\{Z(t)\} = \displaystyle\int_{-\infty}^{\infty} f(t-s)E\{z(s)\}ds = m\int_{-\infty}^{\infty} f(t)dt, \\ E\{[Z(t+h)-\mu][\overline{Z(t)-\mu}]\} \\ \quad = \displaystyle\iint_{-\infty}^{\infty} f(t-s+h)\overline{f(t-s')}E\{[z(s)-m][\overline{z(s')-m}]\}dsds' \\ \quad = \displaystyle\iiint_{-\infty}^{\infty} f(t-s+h)\overline{f(t-s')}\,e^{2\pi i(s-s')\lambda}\,dsds'dF(\lambda) \\ \quad = \displaystyle\int_{-\infty}^{\infty} |f^*(\lambda)|^2\, e^{2\pi i\lambda h}dF(\lambda),^{14} \end{cases} \tag{3.24}$$

where $F(\lambda)$ is the spectral function of the $z(t)$ process. A comparison with (3.21) shows that (3.21) is a (strictly speaking improper) special case, with $y' = z$, $F' \equiv c$.

The easiest way to derive (3.21), which gives most insight into its meaning, is the way dear to the physicist's heart: if $y'(s)$ were a simple exponential, $y'(s) = e^{2\pi i\lambda s}$, (3.19) would become

$$x(t) = \int_{-\infty}^{\infty} f(t-s)e^{2\pi i\lambda s}ds = f^*(\lambda)e^{2\pi i\lambda t}. \tag{3.25}$$

[14] The definition (3.23) and the formal work in these equations is easily justified if $\int_{-\infty}^{\infty} |f(t)|\,dt < \infty$.

In other words, the exponentials are simply multiplied by $f^*(\lambda)$. Since $e^{2\pi i\lambda s}$ appears in the spectrum of $y'(s)$ with mean square intensity c, it must appear in that of $x(t)$ with mean square intensity $c|f^*(\lambda)|^2$, as indicated in (3.21), if there is any justice. There is still justice in science.

The superficially puzzling phenomenon of non-integrable spectral densities has arisen frequently in physics. For example, in a simple single-mesh LCR electrical circuit with the only driving force provided by the spontaneous thermal motion of the electrons, the current derivative dI/dt does not exist, and a formal evaluation of its spectral function will give a spectral density whose integral is infinite. In all such cases what has been involved is a function like y' above which actually does not exist, but whose spectral density can be calculated formally, and has an infinite integral. (See section 5 below for further details.)

Fourier transform of a process with stationary (wide sense) uncorrelated increments.—It will be useful below to have the Fourier transform of a process of the stated type, that is to say, to have a correspondence between pairs of these processes such that their variables $\{y(s)\}$, $\{y^*(s)\}$ satisfy the following relations:

$$\begin{cases} y'(t) = \int_{-\infty}^{\infty} e^{2\pi its} y'^*(s)ds, \qquad y'^*(t) = \int_{-\infty}^{\infty} e^{-2\pi its} y'(s)ds, \\ E\{y'(t)\} = E\{y'^*(t)\} = 0, \qquad E\{|y'(t)|^2\} = E\{|y'^*(t)|^2\}. \end{cases} \tag{3.26}$$

These symbolic equations are to be interpreted as follows. Let $f(t)$ be any Lebesgue measurable function whose square is integrable over $(-\infty, \infty)$. Then

$$\begin{cases} \int_{-\infty}^{\infty} f(t)dy(t) = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} e^{2\pi its} f(t)dt \right] dy^*(s), \quad E\{dy(s)\} = E\{dy^*(s)\} = 0, \\ \int_{-\infty}^{\infty} f(t)dy^*(t) = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} e^{-2\pi its} f(t)dt \right] dy(s), \quad E\{|dy(s)|^2\} = E\{|dy^*(s)|^2\}, \end{cases} \tag{3.27}$$

are to hold, with probability 1. Note that the integrals in the brackets are functions of s, Fourier transforms of f, whose absolute values squared are integrable over $(-\infty, \infty)$, with value $\int_{-\infty}^{\infty} |f|^2 dt$, according to the Parseval identity. Hence both sides of (3.27) are meaningful. In order to justify (3.26), therefore, $y^*(s)$ must be defined and (3.27) must be proved. We suppose that $E\{dy(s)\} = 0$ and define $y^*(t)$ by means of a special case of (3.27);

$$y^*(t) = \int_{-\infty}^{\infty} \left[\int_0^t e^{-2\pi i\tau s} d\tau \right] dy(s) = \int_{-\infty}^{\infty} \frac{1 - e^{-2\pi its}}{2\pi it} dy(s). \tag{3.28}$$

Then $E\{y^*(t)\} = 0$. In order to avoid writing down proportionality constants, we suppose that $E\{|dy(t)|^2\} = |dt|$. In order to show that $y^*(t)$ determines a

process with stationary uncorrelated increments, it is sufficient to show that for $\mu_1, \mu_2, \mu_3, \mu_4$

$$E\{[y^*(\mu_2)-y^*(\mu_1)]\overline{[y^*(\mu_4)-y^*(\mu_3)]}\}$$
$$=E\left\{\int_{-\infty}^{\infty}\left[\int_{\mu_1}^{\mu_2}e^{-2\pi i\tau s}d\tau\overline{\int_{\mu_3}^{\mu_4}e^{-2\pi i\tau s}d\tau}\right]ds\right\} \tag{3.29}$$
$$=\int_{-\infty}^{\infty}[\alpha_{\mu_2}(\lambda)-\alpha_{\mu_1}(\lambda)][\alpha_{\mu_4}(\lambda)-\alpha_{\mu_3}(\lambda)]d\lambda,$$

where $\alpha_\mu(\lambda)$ is defined by (2.8). Equation (3.29) is implied by the Parseval identity for Fourier transforms. When $\mu_1 = \mu_3$, $\mu_2 = \mu_4$, (3.29) shows that $E\{|dy^*(t)|^2\} = |dt|$. The second equation in (3.27) is true by definition of $y^*(t)$ for $f = 1$ on the interval $(0,t)$ and $f = 0$ otherwise. Since both sides of the equation are linear in f, the equation must therefore be true for functions which take on only a finite number of values, each on an interval. Finally, since any f in (3.27) can be approximated arbitrarily closely in the mean by functions of this type, and since we can go to the limit under the integral signs, the second equation in (3.27) must be true in general. Now replace $f(t)$ in this equation by its Fourier transform,

$$\int_{-\infty}^{\infty}e^{2\pi it\tau}f^*(\tau)d\tau.$$

We obtain

$$\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}e^{2\pi it\tau}f^*(\tau)d\tau\right)dy^*(t)=\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}e^{-2\pi its}\left(\int_{-\infty}^{\infty}e^{2\pi it\tau}f^*(\tau)d\tau\right)dt\right]dy(s) \tag{3.30}$$
$$=\int_{-\infty}^{\infty}f(s)dy(s),$$

which is the first equation in (3.27); the proof is now finished.

As an application of the Fourier transform of a process with uncorrelated stationary increments, we shall show how to put a stationary process of moving averages defined by (3.19) into the normal form (2.11). (The y process will be different in the equations.) Define y^*, f^* as usual. Then formally

$$x(t)=\int_{-\infty}^{\infty}f(t-s)dy(s)=\int_{-\infty}^{\infty}f(t-s)y'(s)ds \tag{3.31}$$
$$=\int_{-\infty}^{\infty}e^{2\pi its}f^*(s)y'^*(s)ds=\int_{-\infty}^{\infty}e^{2\pi its}f^*(s)dy^*(s),$$

using again the correspondence between convolutions and products of Fourier transforms. These formal manipulations must actually be correct, because the

Fourier transforms were defined to make them so, and as a matter of fact (3.31) is simply an application of the first equation in (3.27). Hence the moving average $x(t)$ can be put in the form

$$x(t) = \int_{-\infty}^{\infty} e^{2\pi i t s}\, d\tilde{y}(s), \tag{3.32}$$

where

$$\tilde{y}(s) = \int_{-\infty}^{s} f^*(\lambda) dy^*(\lambda), \qquad E\{|d\tilde{y}(s)|^2\} = c|f^*(s)|^2|ds|, \tag{3.33}$$

and this is the required form. The forms (3.31) and (3.32) show explicitly that $c|f^*(s)|^2$ is the spectral density, a fact which was deduced indirectly in (3.21).

4. Mathematical applications

Moving averages.—As a first simple application of spectral theory, we analyze the effect on the harmonic properties of a stationary process of the use of moving averages. The details will be carried out in the integral parameter case, where moving averages are frequently used for smoothing.[15] It is supposed that a stationary process is given, as determined by (2.2), except that now the basic variables may be correlated. For simplicity, suppose that $E\{x(n)\} = 0$. Then the correlation function of $X(n)$ is easily evaluated in terms of the spectral function $F(\lambda)$ of the $x(n)$ process,

$$\begin{aligned} E\{X(m+n)\overline{X(m)}\} &= \sum_{j,k} a_{n+m-j}\, \bar{a}_{m-k} E\{x(j)\overline{x(k)}\} \\ &= \sum_{j,k} \int_{-\infty}^{\infty} a_{n+m-j}\, \bar{a}_{m-k}\, e^{2\pi i(j-k)\lambda} dF(\lambda) \\ &= \sum_{j} \int_{-\infty}^{\infty} \Big|\sum_j a_j e^{-2\pi i j\lambda}\Big|^2 e^{2\pi i n\lambda} dF(\lambda).^{16} \end{aligned} \tag{4.1}$$

Thus the harmonic properties of the smoothed process, which has spectral function $\hat{F}(\lambda)$, with

$$d\hat{F}(\lambda) = \Big|\sum_j a_j e^{-2\pi i j\lambda}\Big|^2 dF(\lambda), \tag{4.2}$$

depend in an essential way on the properties of the function $\sum_j a_j e^{-2\pi i j\lambda}$. In particular, suppose that the $x(n)$ process is a process of uncorrelated variables. Then $F(\lambda)$ will be given by

$$F(\lambda) = E\{|x(0)|^2\}(\lambda + 1/2) \tag{4.3}$$

[15] See the preceding section, (3.23) and (3.24), for the continuous parameter case.

[16] The definition (2.2) and the formal work in these equations is easily justified if $\sum_j |a_j| < \infty$.

so that there are no favored frequencies; $dF(\lambda) = \text{const. } d\lambda$. On the other hand, the smoothed $X(n)$ process will stress the frequencies near the maxima of $\left|\sum_j a_j e^{-2\pi i j\lambda}\right|$. In the simplest smoothing, when

$$X_n = \frac{1}{2N+1}\sum_{j=-N}^{N} x(n+j),$$

we have, if $F(\lambda)$ is given by (4.3), so that the basic variables are not correlated,

$$d\hat{F}(\lambda) = \frac{E\{|x(0)|^2\}}{(2N+1)^2}\left|\frac{1-e^{2\pi i\lambda(2N+1)}}{1-e^{2\pi i\lambda}}\right|^2 d\lambda. \tag{4.4}$$

Effect of transition from continuous parameter to discrete parameter.—Suppose that a stationary continuous parameter process is given in canonical form,

$$x(t) = \int_{-\infty}^{\infty} e^{2\pi i t\lambda} dy(\lambda). \tag{4.5}$$

Then if $\tau > 0$,

$$x(n\tau) = \int_{-1/2}^{1/2} e^{2\pi i n\lambda} dy_\tau(\lambda), \tag{46}$$

where

$$y_\tau(\lambda) = \sum_{m=-\infty}^{\infty}\left\{y\left(\frac{m+\lambda}{\tau}\right) - y\left(\frac{m-1/2}{\tau}\right)\right\}, \qquad |\lambda| \leqq 1/2. \tag{4.7}$$

The form (4.6) is the canonical form for the integral parameter process with variables $x(n\tau)$. A frequency ω occurring with positive probability in $x(t)$ means a jump in $y(\lambda)$ when $\lambda = \omega$, and this means a jump in $y_\tau(\lambda)$ when $\lambda = \omega\tau - [\omega\tau + 1/2]$, where $[\xi]$ is the largest integer $\leqq \xi$. Thus the stressed frequency ω in the original process becomes a stressed frequency ω' in the contracted process, with

$$\frac{\omega'}{\tau} = \omega - \frac{[\omega\tau + 1/2]}{\tau}. \tag{4.8}$$

If $|\omega\tau| < 1/2$, the time frequency is therefore unchanged.

Linear prediction.—The problem of least-squares linear prediction in the integral parameter case can be formulated as follows. (The variables $\{x(n)\}$ are the variables of a stationary process in the wide sense.)

Coefficients $b_j = b_j^\nu$, $(j \leqq \nu)$, are to be found which minimize

$$E\{|x(n) - \sum_{j=1}^{n} b_j x(n-j)|^2\} = \int_{-1/2}^{1/2} |1 - \sum_{j=1}^{\nu} b_j e^{-2\pi i j\lambda}|^2 dF(\lambda), \tag{4.9}$$

where $F(\lambda)$ is the spectral function of the process. This can also be formulated as follows. A chance variable $\varphi_{n,\nu}$ is to be found, in the linear manifold of chance variables $\{\varphi\}$ determined by $x(n-1), \cdots, x(n-\nu)$, minimizing

$$E\{|x(n)-\varphi|^2\}. \tag{4.10}$$

Alternatively, by using the right side of (4.9) instead of the left, a function $\Phi_\nu(\lambda)$ is to be found, in the linear manifold of functions $\{\Phi(\lambda)\}$ determined by $e^{-2\pi i\lambda}, \cdots, e^{-2\pi i\nu\lambda}$, minimizing

$$\int_{-1/2}^{1/2} |1-\Phi(\lambda)|^2 dF(\lambda). \tag{4.11}$$

These last two formulations are important because they allow a direct extension to the full prediction problem, $\nu = +\infty$. In that case a chance variable $\varphi_{n,\infty}$ [a function $\Phi_\infty(\lambda)$] is to be found, in the closed linear manifold determined by the chance variables $x(n-1), x(n-2), \cdots$ (in the closed linear manifold determined by the functions $e^{-2\pi i\lambda}, e^{-4\pi i\lambda}, \cdots$) minimizing (4.10) or (4.11) for functions in the stated manifolds. Closure means closure with respect to mean square convergence for the chance variables, mean square convergence weighted by $dF(\lambda)$ for the functions of λ. These formulations are equivalent because there is an isometry between the chance variables and the functions of λ under discussion, as indicated by (4.9).

The usual argument shows that $\varphi_{n,1}, \varphi_{n,2}, \cdots$ actually exist, and are determined uniquely in their respective manifolds by orthogonality relations: $x(n)-\varphi_{n,\nu}$ is orthogonal to $x(n-1), \cdots, x(n-\nu)$; $x(n)-\varphi_{n,\infty}$ is orthogonal to $x(n-1), x(n-2), \cdots$.

If prediction is to be practical, $\varphi_{n,\nu}$ must converge in some sense to $\varphi_{n,\infty}$, as $\nu \to \infty$, and the following argument shows that there is actually convergence in the mean. In the first place, because of the stated orthogonality relations, $x(n)-\varphi_{n,\nu}$ is orthogonal to $\varphi_{n,1}, \cdots, \varphi_{n,\nu}$. Hence

$$E\{|x(n)-\varphi_{n,\nu}|^2\} = E\{|x(n)|^2\} - E\{|\varphi_{n,\nu}|^2\} \geqq 0. \tag{4.12}$$

The orthogonality relations just mentioned imply that the chance variables $\{\delta_{n,\nu}\} = \{\varphi_{n,\nu}-\varphi_{n,\nu-1}\}$ form an orthogonal set, and with the use of this fact and (4.12),

$$\sum_\nu E\{|\delta_{n,\nu}|^2\} = E\{|\sum_\nu \delta_{n,\nu}|^2\} = \lim_{\nu\to\infty} E\{|\varphi_{n,\nu}|^2\} \leqq E\{|x(n)|^2\}. \tag{4.13}$$

Hence the series $\sum_\nu \delta_{n,\nu}$ (and also the sequence $\{\varphi_{n,\nu}\}$) converge in the mean. Since $x(n)-\varphi_{n,\nu}$ is orthogonal to $x(n-N)$ when $\nu \geqq N$, $x(n) - \underset{\nu\to\infty}{\text{l.i.m.}}\varphi_{n,\nu}$ is orthogonal to $x(n-N)$ for all N. Thus this limit must be $\varphi_{n,\infty}$, as was to be shown.

The orthogonality of $x(n) - \varphi_{n,\nu}$ and $x(n-m)$, $(1 \leqq m \leqq \nu)$, can be stated in the form

$$R(m) - \sum_{j=1}^{\nu} b_j^{(\nu)} R(m-j) = \int_{-1/2}^{1/2} \left[1 - \sum_{j=1}^{\nu} \frac{b_j^{(\nu)}}{z^j}\right] z^m dF(\lambda) = 0, 1 \leqq m \leqq \nu, \quad z = e^{2\pi i\lambda}. \tag{4.14}$$

Formally, if $F(\lambda)$ is the integral of its derivative $F'(\lambda)$, the problem then becomes, according to (4.14), that of finding a function $1 - \Phi_\infty(\lambda)$,

$$1 - \Phi_\infty(\lambda) \sim 1 - \frac{b_1^{(\infty)}}{z} - \cdots \tag{4.15}$$

(where the series converges for $|z| > 1$), such that $(1 - \Phi_\infty)F'$ is of power series type,

$$(1 - \Phi_\infty)F' \sim c_0 + c_1 z + \cdots \tag{4.16}$$

(where the series converges for $|z| < 1$). This formulation is the key to the practical solution of the prediction problem.[17]

If exact prediction is possible for a finite value of ν, that is, if (4.9) can be made to vanish, $F(\lambda)$ must be constant except perhaps for jumps at the zeros of the polynomial in (4.9). Conversely, if $F(\lambda)$ is constant except for a finite number of jumps, the polynomial can be chosen to vanish at those jumps. Thus exact prediction is possible for ν finite if and only if $F(\lambda)$ is constant except for a finite number of jumps. In this case $x(n)$ is a finite sum of exponentials with mutually uncorrelated coefficients [see (4.19) below].

Kolmogoroff has shown[18] that for $\nu = \infty$ the prediction error (in the general case, for arbitrary F) is given by

$$\int_{-1/2}^{1/2} |1 - \Phi_\infty(\lambda)|^2 dF(\lambda) = e^{\int_{-\frac{1}{2}}^{\frac{1}{2}} \log F'(\lambda) d\lambda} \tag{4.17}$$

According to this equality, the mean square error vanishes, so that exact prediction is possible, if and only if the logarithmic integral diverges to $-\infty$.

Note that the practical problem of prediction consists in the evaluation of $\Phi_\nu(\lambda)$ [or $\Phi_\infty(\lambda)$]. Some of the progress made in this direction, by Wiener and perhaps also by others, has been classified. Wold and Kolmogoroff have shown, in less cautious regions, that $x(n)$ can always be written as a sum of two terms, $x(n) = x_1(n) + x_2(n)$, where, without going into the rather tricky details, the $x_1(n)$ and $x_2(n)$ processes are not correlated with each other; the $\Phi_\infty(\lambda)$ of the $x(n)$ process solves the prediction problem for the $x_1(n)$ process (giving exact

[17] Note that the conduct of the functions defined by the series in (4.15) and (4.16) as $|z| \to 1$ has been left unspecified. For the full details, see the original papers by Krein and Kolmogoroff referred to in footnote 18. In most applications the expansions are valid for $|z| > 1 - \epsilon$ and $|z| < 1 + \epsilon$ respectively, where ϵ is some positive number. This is usually true, for example when the spectral densities are rational.

[18] *Bull. Acad. Sci. URSS, Ser. Math.*, vol. 5 (1941), pp. 3-14. See also Krein, C. R. (*Dok lady*) *Acad. Sci. URSS*, vol. 46 (1945), pp. 91-94.

prediction, with no error) and for the $x_2(n)$ process; and the $x_2(n)$ process, the non-singular component, is a process of moving averages defined by a series of type (2.2) in which the coefficients having negative subscripts all vanish.[19]

The details of the continuous parameter case run parallel to those of the integral parameter case. If $F(\lambda)$ is the spectral function, exact prediction is possible if and only if

$$\int_{-\infty}^{\infty} \frac{\log F'(\lambda)}{1+\lambda^2}\, d\lambda = -\infty. \tag{4.18}$$

Types of spectra; point spectrum.—If $F(\lambda)$ increases only in jumps, say at $\{\lambda_j\}$, the canonical form (2.11) becomes a series

$$x(n) = \sum_m y_m e^{2\pi i n \lambda_m}, \tag{4.19}$$

where y_m is the jump of $y(\lambda)$ at λ_m,

$$y_m = \underset{\epsilon \downarrow 0}{\text{l.i.m.}} \, [y(\lambda_m + \epsilon) - y(\lambda_m - \epsilon)]. \tag{4.20}$$

The y_m are uncorrelated and the series converges in the mean. Exact prediction is possible in this case, according to the criterion given above. If the process parameter is continuous, (4.19) is replaced by

$$x(t) = \sum_m y_m e^{2\pi i t \lambda_m}. \tag{4.19}$$

Slutsky has proved that (with probability 1) the sample functions $x(t)$ are B_2 almost periodic in this case.[20]

As has already been remarked, any stationary process can be approximated by one with a point spectrum, by replacing the integral (2.11) or (2.11′) by an approximating sum.

Types of spectra; absolutely continuous spectral function.—If $F(\lambda)$ is the integral of its derivative $F'(\lambda)$, the canonical forms (2.11), (2.11′) can be rewritten in the following forms:

$$x(t) = \int_{-\infty}^{\infty} e^{2\pi i t \lambda}\, \varphi(\lambda) dY(\lambda). \tag{4.21}$$

$$x(n) = \int_{-1/2}^{1/2} e^{2\pi i n \lambda}\, \varphi(\lambda) dY(\lambda), \qquad |\varphi|^2 = F' \tag{4.21′}$$

[19] Wold, *A Study in the Analysis of Stationary Time Series* (Uppsala, 1938), p. 89; Kolmogoroff, *Bull. Acad. Sci. URSS, Ser. Math.*, vol. 5 (1941), pp. 3-14, and *Bull. Moscow State Univ., Math.*, vol. 2, no. 6 (1941), pp. 24-27.

[20] *Act. Sci. Ind.*, vol. 738 (1938), pp. 47-50. See also Wiener and Wintner, *Amer. Jour. Math.*, vol. 63 (1941), pp. 809-810.

Here $\varphi(\lambda)$ is any Lebesgue measurable function such that $|\varphi|^2 = F'$; $Y(\lambda)$ is the variable of a process with uncorrelated stationary increments (wide sense), determined by

$$Y(\lambda) = \int_0^\lambda \frac{dy(\lambda)}{\varphi(\lambda)} ; \tag{4.22}$$

and $y(\lambda)$ is the variable in (2.11) or (2.11'); $E\{|dY(\lambda)|^2\} = |d\lambda|$.[21] Note that there is considerable choice for $\varphi(\lambda)$; this will be useful below.

It has already been shown above that the spectrum of a stationary process of moving averages is absolutely continuous. In fact, in the discrete parameter case the spectral density of the process defined by (2.2) with uncorrelated summands is given by (4.2) with $dF(\lambda) = \text{const.}\, d\lambda$; in the continuous parameter case, the spectral density of the process defined by (3.19) is given in (3.21). *Conversely, if the spectral function $F(\lambda)$ is absolutely continuous, the process is a process of moving averages.* A Fourier integral proof will be given for the continuous parameter case; a corresponding Fourier series proof is easily carried through for the discrete parameter case.[22] In order to prove the theorem, it will be convenient to write (4.21) in the form

$$x(t) = \int_{-\infty}^{\infty} e^{2\pi i t\lambda} f^*(\lambda) dw^*(\lambda), \qquad w^* = Y, \quad f^* = \varphi. \tag{4.21''}$$

Now this is precisely the form on the right side of (3.31), which equation then shows that $x(t)$ is the moving average

$$x(t) = \int_{-\infty}^{\infty} f(t - s) dw(s), \tag{3.31'}$$

where f, w are the Fourier transforms of f^*, w^*.

It follows from this theorem that if $x_1(t), \cdots, x_N(t)$ are the variables of processes of moving averages, and if each $x_j(t)$ is uncorrelated with every $x_k(s)$, $(\neq j$; all $s, t)$, then the sum $x_1(t) + \cdots + x_N(t)$ is still a process of moving averages; in fact, the spectral density of the sum is the sum of the spectral densities, in this case.

Types of spectra; integral parameter, rational spectral density in $e^{2\pi i\lambda}$.—If $F(\lambda)$ is the integral of a rational function of $z = e^{2\pi i\lambda}$, the process is of a particularly simple type. Since $F'(\lambda)$ is the value on $|z| = 1$ of a rational function of z which is real on $|z| = 1$, it follows from the Schwarz reflection principle that to each zero (pole) at α in $|z| < 1$ there corresponds one of the same

[21] The definition of $Y(\lambda)$ must be slightly modified if $\varphi(\lambda)$ may vanish.

[22] This theorem in the discrete parameter case is due to Kolmogoroff, *Bull. Moscow State Univ., Math.*, vol. 2, no. 6 (1941), pp. 1-40, theorem 16. We assume in this discussion that $E\{x_n\} = 0$ in (2.2); the only restriction on the a_j's is then that $\sum_j |a_j|^2 < \alpha$. Correspondingly we assume that $E\{x(t)\} = 0$.

order at the inverse point $1/\bar{\alpha}$, and, since $F'(\lambda) \geqq 0$, that the zeros on $|z| = 1$ are of even order. There can be no poles on $|z| = 1$, since $F'(\lambda)$ is integrable. It follows that F' can be written in the form

$$F'(\lambda) = |F'(\lambda)| = \text{const.} \frac{\prod_j \pi(z - \alpha_j)(\bar{\alpha}_j z - 1)}{\prod_k \pi(z - \beta_k)(\bar{\beta}_k z - 1)} = \text{const.} \frac{\prod_j \pi|z - \alpha_j|^2}{\prod_k \pi|z - \beta_k|^2} \tag{4.23}$$

$$= \frac{|A(z)|^2}{|B(z)|^2}, 0 < |\alpha_j| \leqq 1, \quad 0 < |\beta_k| < |,$$

where the polynomials A, B, with no common factors, have their zeros in $0 < |z| \leqq 1$ and $0 < |z| < 1$ respectively, as indicated, and are otherwise arbitrary. If the process is a real process, $F'(\lambda)$ is an even function of λ, from which fact it follows that A and B can be assumed to have real coefficients. The variables $\{x(n)\}$ of the most general process of this type can therefore be written, if $E\{x(n)\} = 0$, in the form

$$x(n) = \int_{-1/2}^{1/2} e^{2\pi i n\lambda} \frac{A(e^{2\pi i\lambda})}{B(e^{2\pi i\lambda})} dy(\lambda), \qquad E\{|dy(\lambda)|^2\} = |d\lambda|, \tag{4.24}$$

where $y(\lambda)$ is a process with stationary uncorrelated increments. The correlation $R(n)$, given by

$$R(n) = \int_{-1/2}^{1/2} e^{2\pi i n\lambda} F'(\lambda) d\lambda = \frac{1}{2\pi i} \int_{|z|=1} z^{n-1} F'(\lambda) dz, \qquad z = e^{2\pi i\lambda}, \tag{4.25}$$

is the coefficient of $1/z^n$ in the Laurent expansion of a rational function of z, convergent in an annulus enclosing $|z| = 1$. The values of $R(n)$ therefore decrease exponentially when $|n| \to \infty$; $x(m)$ and $x(m + n)$ thus become uncorrelated rapidly as n increases.

The linear prediction problem is easily solved explicitly for spectra of this type. Let $A(z)$, $B(z)$ be given by

$$\begin{cases} A(z) = a_0 z^a + \cdots + a_a, & a_0 a_a \neq 0, \\ B(z) = b_0 z^b + \cdots + b_b, & b_0 b_b \neq 0. \end{cases} \tag{4.26}$$

The $F'(\lambda)$ can be written in the form

$$F'(\lambda) = \frac{|A(z)|^2}{|B(z)|^2} = \frac{\sum_j a_j z^{a-j} \sum_j \bar{a}_j z^j}{\sum_j b_j z^{b-j} \sum_j \bar{b}_j z^j} z^{b-a}. \tag{4.27}$$

The function $1 - \Phi_\infty$ which solves the prediction problem can be chosen to satisfy (4.15) and (4.16) as follows:

$$1 - \Phi_\infty = \frac{B(z)a_0}{A(z)b_0} z^{a-b}. \tag{4.28}$$

An expansion of type (4.15) is valid for $|z| \geqq 1$, if $A(z)$ does not vanish on $|z| = 1$, since then $1 - \Phi_\infty$ will be analytic outside and on the circle, with value 1 at ∞. An expansion of type (4.16) is valid for $(1 - \Phi_\infty)F'$, since

$$(1 - \Phi_\infty)F' = \frac{\sum_j \bar{a}_j z^j a_0,}{\sum_j \bar{b}_j z^j b_0} \tag{4.29}$$

which defines a function of z analytic for $|z| \leqq 1$. If $A(z)$ has zeros of modulus 1, (4.15) cannot be checked without going more deeply into the significance of (4.15) than was done above. The definition (4.28) is still admissible, however, since the integrability conditions on $1 - \Phi_\infty$ are certainly satisfied and $\Phi_\infty(\lambda)$ is in the proper manifold; the orthogonality relations which $1 - \Phi_\infty(\lambda)$ should satisfy are then easily checked directly.

The prediction error can now be evaluated explicitly, as follows:

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} |1 - \Phi_\infty|^2 F'(\lambda)d\lambda = \int_{-\frac{1}{2}}^{\frac{1}{2}} \left|\frac{a_0}{b_0}\right|^2 d\lambda = \left|\frac{a_0}{b_0}\right|^2. \tag{4.30}$$

With the use of (4.27) and (4.28),

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} \log [|1-\Phi_\infty| F'(\lambda)]^2 d\lambda = \int_{-\frac{1}{2}}^{\frac{1}{2}} \log \left[\left|\frac{a_0}{b_0}\right|^2 F'(\lambda)\right] d\lambda = \log \left|\frac{a_0}{b_0}\right|^2 + \int_{-\frac{1}{2}}^{\frac{1}{2}} \log F'(\lambda)d\lambda. \tag{4.31}$$

On the other hand, $[(1 - \Phi_\infty)F']^2$ is, according to (4.29), the boundary function on $|z| = 1$ of a function which is analytic and non-vanishing in $|z| < 1$. Hence

$$\frac{1}{2\pi i}\int_{|z|=1} \frac{\log [(1 - \Phi_\infty)F']^2}{z} dz = 2 \log [(1 - \Phi_\infty)F']\big|_{z=0} = 2 \log \left|\frac{a_0}{b_0}\right|^2 \tag{4.32}$$

and since the real part of the left side of (4.32) is the same as the left side of (4.31) it follows that

$$\log \left|\frac{a_0}{b_0}\right|^2 = \int_{-\frac{1}{2}}^{\frac{1}{2}} \log F'(\lambda)d\lambda, \tag{4.33}$$

which verifies (4.17) in this particular case.

In the case under discussion, it is possible to find $N - 1$ auxiliary variables $x_2(n), \cdots, x_N(n)$, where $N = \max(a + 1, b)$, such that the N-dimensional process with determining variables $\{x_1(n) = x(n), \cdots, x_N(n)\}$ is stationary (wide sense, or strict sense if the $x_1(n)$ process is stationary in the strict sense) and has the following very simple prediction properties. The linear least-squares prediction of each $x_j(n)$ in terms of the complete past of all variables: $x_k(m)$ for $k \leqq N$, $m < n$, is simply a linear combination of $x_1(n-1), \cdots, x_N(n-1)$[23]. Conversely, if $N - 1$ such auxiliary variables can be found, the spectral function of the process must either be the integral of a rational function of $z = e^{2\pi i\lambda}$ or differ from such a function by a function increasing only in a finite number of jumps. If there are J jumps, $N = \max(a + J + 1, b + J)$, where a, b have the same meaning as above; if $A(z) \equiv 0$, set $a = -1$, $b = 0$.

In particular, suppose that the numerator polynomial is of zero degree, $A \equiv$ const. Then, according to the evaluation (4.28) of Φ_∞, the linear least-squares prediction of $x(n)$ in terms of its complete past is a linear combination $\sum_j b_j x(n-j) = \varphi_{n,\infty}$ of $x(n-1), \cdots, x(n-N)$, $N = b$. In this case the chance variables $\{\varphi_{n,\nu}\}$ defined above in the discussion of prediction are all the same for $\nu \geqq N$. This special case is worth examining in more detail at the risk of some repetition. The regression equation works out very elegantly, with the use of the normal form (4.24):

$$b_0x(n) + b_1x(n-1) + \cdots + b_bx(n-b) = \int_{-1/2}^{1/2} \frac{b_0z^n + \cdots + b_bz^{n-b}}{B(z)}\, dy(\lambda)$$

$$= \int_{-1/2}^{1/2} z^{n-b}dy(\lambda) = \eta_n. \qquad (4.34)$$

$$E\{x(n-m)\overline{\eta(n)}\} = E\left\{\int_{-1/2}^{1/2} \frac{z^{n-m}}{\overline{B(z)}}\, dy(\lambda) \overline{\int_{-1/2}^{1/2} z^{n-b}dy(\lambda)}\right\} = \int_{-1/2}^{1/2} \frac{z^{b-m}}{\overline{B(z)}}\, d\lambda \qquad (4.35)$$

$$= \frac{1}{2\pi i}\int_{|z|=1} z^{b-m-1}\left[\frac{1}{\overline{b_0}z^b} + \text{higher powers of } \frac{1}{z}\right] dz = 0, \qquad m \geqq 1,$$

and

$$E\{\eta_n\overline{\eta_m}\} = E\{\int_{-1/2}^{1/2} z^{n-b}dy(\lambda) \overline{\int_{-1/2}^{1/2} z^{m-b}dy(\lambda)}\} = \int_{-1/2}^{1/2} z^{n-m}d\lambda = \delta_{n,m}, \qquad (4.36)$$

[23] Doob, *Annals of Math. Stat.*, vol. 15 (1944), pp. 229-282, theorem 3.9. A stationary N-dimensional process (wide sense) is one whose first and second moments are stationary; the expectations

$$E\{x_j(s)\} = m_j$$
$$E\{[x_j(s+t) - m_j]\,\overline{[x_k(s) - m_k]}\} = r_{jk}(t)$$

are independent of s. The referenced paper treats only real Gaussian processes. However, the transition to complex processes is inconsequential, and, as stated in the introduction in that paper, the hypothesis that the processes are Gaussian is irrelevant to the present considerations. In the real Gaussian case, an N-dimensional process with the simple prediction properties described is a Markoff process. See Cramér, *Ann. of Math.*, vol. 41 (1940), pp. 215-230, for the basic theory of N-dimensional stationary processes.

where $\delta_{n,m}$ is the Kronecker δ. Thus the η_n determine a stationary unself-correlated process and η_n is uncorrelated with $x(n-1), x(n-2), \cdots$. The $x(n)$ process is thus a process of linear autoregression and (4.34) (divided through by b_0) is the regression equation; the mean square error is $1/|b_0|^2$.

Types of spectra, continuous parameter, rational spectral density.—If $F(\lambda)$ is the integral of a rational function of λ, an argument similar to that used in the integral parameter case shows that $F'(\lambda)$ can be put into the form

$$F'(\lambda) = \frac{|A(\lambda)|^2}{|B(\lambda)|^2} = \frac{|\sum_j a_j \lambda^{a-j}|^2}{|\sum_j b_j \lambda^{b-j}|^2}, \tag{4.23'}$$

where A, B are polynomials with no common factor, of degrees a, b respectively, whose zeroes have positive or zero imaginary parts, and positive imaginary parts respectively. Since $F'(\lambda)$ is integrable, $b > a$. The canonical form becomes, if $E\{x(t)\} = 0$,

$$x(t) = \int_{-\infty}^{\infty} e^{2\pi i t\lambda} \frac{A(\lambda)}{B(\lambda)}\, dy(\lambda), \qquad E\{|dy(\lambda)|^2\} = |d\lambda|; \tag{4.24'}$$

the y process is one with stationary uncorrelated increments. If the process is real, $A(\lambda)$ and $B(\lambda)$ are even, that is to say, $A(i\xi)$ considered as a polynomial in ξ can be assumed to have real coefficients.

The correlation function $R(t)$ is the Fourier transform of the rational function F' and is therefore a sum of exponentials. Thus $R(t)$ decreases exponentially when $|t| \to \infty$; $x(s)$ and $x(s+t)$ become uncorrelated rapidly as $|t|$ increases.

Just as in the corresponding integral parameter case, the linear least-squares prediction problem can be simplified by the use of auxiliary variables. In fact, b $-$ 1 auxiliary families of chance variables $\{x_2(t), \cdots, x_b(t)\}$, $(-\infty < t < \infty)$, can be found with the following properties: The b-dimensional process with determining variables $\{x_1(t) = x(t), \cdots, x_b(t)\}$ is stationary (wide sense, or strict sense if the $x(t)$ process is stationary in the strict sense). The linear least-squares prediction of each $x_j(t)$ in terms of all the $x_k(\tau)$ for $k \leqq b$, $\tau \leqq s < t$ is simply a linear combination of $x_1(s), \cdots, x_b(s)$.[24] Conversely, if $N - 1$ such auxiliary variables can be chosen, the spectral function of the process must be either the integral of a rational function of λ or differ from such a function by a function increasing only in a finite number of jumps. If there are J jumps, $N = b + J$, where b has the meaning given above.

In particular, suppose that the numerator polynomial $A(\lambda)$ is of degree zero, $A \equiv$ const. Then it has been shown[25] that $x(t)$ will have $b-1$ derivatives, all determining stationary processes, and that they can be chosen as the auxiliary variables, $x_j(t) = x^{(j-1)}(t)$. The linear least-squares prediction of $x(t)$ in terms of $x(\tau)$ for $\tau \leqq s < t$ is therefore a linear combination of $x(s)$, $x'(s)$,

[24] Doob, *Annals of Math. Stat.*, vol. 15 (1944), pp. 229-282, theorem 4.9.
[25] Doob, *op. cit.*, theorem 4.9 and corollary.

$\cdots, x^{(b-1)}(s)$. Moreover (when $s \to t$) this prediction formula becomes a linear differential equation [corresponding to the difference equation (4.34)],

$$b_0 \frac{x^{(b)}(t)}{(2\pi i)^b} + b_1 \frac{x^{(b-1)}(t)}{(2\pi i)^{b-1}} + \cdots + b_b x(t) = \tilde{y}'(t), \qquad E\{|d\tilde{y}(t)|^2\} = |dt|. \quad (4.34')$$

Here the chance variable $\tilde{y}(t)$ is a chance variable determining a process with uncorrelated stationary increments, the exact analogue of the process involved in (4.34). Unfortunately, we have seen that $\tilde{y}'(t)$ may not exist, and (4.34') must therefore be interpreted symbolically; if (α,β) is any finite interval, and if $f(t)$ is any function defined and continuous in that interval, then

$$\begin{aligned} \frac{b_0}{(2\pi i)^b}\int_\alpha^\beta f(t)dx^{(b-1)}(t) &= \frac{b_1}{(2\pi i)^{b-1}}\int_\alpha^\beta f(t)x^{(b-1)}(t)dt \\ &+ \cdots + b_b\int_\alpha^\beta f(t)dx(t) = \int_\alpha^\beta f(t)d\tilde{y}(t) \end{aligned} \quad (4.37)$$

with probability 1.[26] The $y'(t)$ in (4.34') is uncorrelated with $x(s)$ for $s < t$, that is, $y(t + h) - y(t)$ is uncorrelated with these variables if $h > 0$, and it is, therefore clear that this $x(t)$ process is the proper analogue of the process of linear autoregression in the discrete parameter case.

Equation (4.37) can easily be obtained directly, with the use of the normal form (4.24'):

$$\begin{aligned} &\frac{b_0}{(2\pi i)^b}\int_\alpha^\beta f(t)dx^{(b-1)}(t) + \cdots + b_b\int_\alpha^\beta f(t)x(t)dt \\ &= \int_{-\infty}^{\infty}\left[\int_\alpha^\beta f(t)e^{2\pi i t\lambda}dt\right]dy(\lambda) = \int_\alpha^\beta f(t)d\tilde{y}(t), \end{aligned} \quad (4.38)$$

where the $\tilde{y}$ process is a Fourier transform of the y process [see (3.27)]. This simple application shows the formal simplifications gained by the use of these normal forms.

Spectra with rational density functions are those most frequently encountered in the applications, because of their intimate connection with finite systems of linear difference and differential equations. In the next section this connection will be illustrated in detail.

[26] The integrals in (4.37) are ordinary integrals except for the one on the right, and the first on the left. The former has already been defined, and the latter is defined as the limit in the mean of the usual Riemann-Stieltjes sums, a limit which exists because of the existence of the corresponding limits for the other terms in the equation; see also the direct derivation of (4.37) given below.

5. Physical application

The following discussion shows how the ideas of the preceding sections are applied to physical problems involving the Brownian movement, that is, to problems involving spontaneous thermal motion. In these problems, no matter what the background, there is one common basis if the mechanisms involved are linear; non-linear problems are considerably more complicated.

It is supposed that events occur in a random manner at a uniform rate, in such a way that the numbers of events occurring in any succession of non-overlapping time intervals are mutually independent. These hypotheses are equivalent to the hypothesis that the times of the events go in accordance with the Poisson process with stationary independent increments, discussed in section 3. The probability that n events occur in a time interval of length t is given by

$$e^{-ct}\frac{c^n t^n}{n!}, \qquad n = 0, 1, \cdots, \tag{5.1}$$

where c is a positive constant, the rate at which the events occur. In the applications considered, each event has a certain intensity u, and has an effect $uf(t)$, t time units after its occurrence (linearity of mechanism). The number $uf(t)$ may be the current in an electric circuit at time $t_0 + t$ ascribable to the passage of an electron at time t_0 from filament to plate of a vacuum tube in the circuit; the momentum of a pendulum suspended in air ascribable to the impact on it of an air molecule t seconds before; etc. It is supposed that u is a chance variable with finite first and second moments. The simplest case is that in the shot effect, the first example above, when each event is like any other event (unless it is assumed that electrons can go over in groups), so that u is a chance variable taking on some value u_0 with probability 1. In the pendulum example, u is evidently more complicated.[27] The total effect is given by the sum of the effects of each previous event (linearity of mechanism):

$$\vartheta(t) = \sum_j f(t - t_j)u_j, \tag{5.2}$$

where $t_1, t_2, \cdots$ are the times of the events occurring before time t and $u_1, u_2, \cdots$ are the intensities at these times. The u_j are independent chance variables, each having the distribution of u. These events are the jumps of the sample function of a $y(t)$ process with stationary (strict sense) increments, and (5.2) can therefore be written in the form

$$\vartheta(t) = \int_{-\infty}^{0} f(t - s)dy(s). \tag{5.3}$$

[27] In the pendulum problem the effect depends on the state of the mechanism, that is, on the velocity of the molecule relative to the pendulum bob. Thus there are two complications, the variation in molecular velocity and the variation in pendulum velocity. The effect of impacts is equivalent to an average slowing down proportional to the pendulum bob velocity plus impacts distributed symmetrically on the two sides of the pendulum. It is actually the latter impacts which are considered here.

Thus $\vartheta(t)$ determines a stationary process (strict sense) of moving averages (see section 3); we set $f(t) = 0$ for $t < 0$. It will be useful to evaluate the following first and second moments:

$$\begin{array}{ll} \displaystyle\int_0^\infty f(t)dt = a, & \displaystyle\int_0^\infty |f(t)|^2 dt = b, \\ E\{u\} = \alpha & E\{u^2\} = \beta. \end{array} \tag{5.4}$$

$$\begin{aligned} &E\{y(s+t) - y(s)\} = c\alpha t, \\ &E\{|y(s+t) - y(t) - c\alpha t|^2\} = c\beta t. \end{aligned} \tag{5.5}$$

$$\begin{aligned} &E\{\vartheta(t)\} = c\alpha \int_0^\infty f(t)dt = ca\alpha, \\ &\{[\vartheta(t+h) - ca\alpha]\,\overline{[\vartheta(t) - ca\alpha]}\} = c\beta \int_{-\infty}^\infty e^{ih\lambda} |f^*(\lambda)|^2 d\lambda, \end{aligned} \tag{5.6}$$

where f^* is the Fourier transform of f [see (3.21)]. The evaluation of the dispersion of $\vartheta(t)$,

$$E\{|\vartheta(t) - ca\alpha|^2\} = c\beta \int_0^\infty |f(s)|^2 ds, \tag{5.7}$$

when applied to the shot effect and similar phenomena in which $u \equiv u_0$, so that $\beta = \alpha^2$, is known to physicists as Campbell's theorem.[28] The fact that the expectation is also the time average,

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T |\vartheta(t) - ca\alpha|^2 dt = E\{|\vartheta(t) - ca\alpha|^2\} \tag{5.8}$$

(with probability 1), is simply the strong law of large numbers applied to this particular process.[29]

The characteristic functions of $y(t) - y(0) - c\alpha t$ and $\vartheta(t) - ca\alpha$ are easily evaluated in terms of the characteristic function of u, $\psi(z)$:

$$\log E\{e^{iz[y(t) - y(0) - c\alpha t]}\} = ct[\psi(z) - 1 - i\alpha z] = -ct\beta \frac{z^2}{2} + c\,o(z^2), \tag{5.9}$$

$$\begin{aligned} \log E\{e^{iz[\vartheta(t) - ca\alpha]}\} &= c \int_0^\infty \Big(\psi[zf(s)] - 1 - iz\alpha f(s)\Big) ds = \\ &- c\beta b \frac{z^2}{2} + c \int_0^\infty o[f(s)^2 z^2]\, ds. \end{aligned} \tag{5.10}$$

Now in the applications, c is large compared to the time constants of the physical system involved, and this effectively smooths the y and ϑ dis-

[28] See A. Khintchine, *Bull. Acad. Sci. URSS* (1938), pp. 313-322, for a general treatment of Campbell's theorem and extensions which go far beyond that given here.

[29] The strong law of large numbers is applicable, since $\vartheta(t)$, and therefore $|\vartheta(t) - c\,a\alpha|^2$, define stationary processes in the strict sense. The fact that the limit in (5.8) is constant can be deduced from the fact that the transformation of chance variables induced by translations of the time axis is metrically transitive. [See Doob, *Trans. Amer. Math. Soc.*, vol. 43 (1937), theorem 3.2.]

tributions. In order to see this, let φ_y, φ_ϑ be the characteristic functions of $[y(t) - y(0) - c\alpha t]/c^{1/2}$ and of $[\vartheta(t) - ca\alpha]/c^{1/2}$ respectively. Then

$$\begin{cases} \log \varphi_y(z) = \dfrac{-\beta t z^2}{2} + c\, o\left(\dfrac{z^2}{c}\right) \to \dfrac{-\beta t z^2}{2}, \\ \log \varphi_\vartheta(z) = \dfrac{-\beta b z^2}{2} + c\int_0^\infty o\left[\dfrac{f(s)z^2}{c}\right] ds \to \dfrac{-\beta b z^2}{2}, \qquad c \to \infty. \end{cases} \tag{5.11}$$

In other words, the y and ϑ variables are asymptotically Gaussian. An extension of the preceding argument to the consideration of n, t values simultaneously shows that the y and ϑ processes are asymptotically Gaussian processes. Clearly the limiting y process has stationary (strict sense) independent increments and the limiting ϑ process is stationary (strict sense).[30] In physical applications it has been customary to take the y process as Gaussian in the first place. The ϑ process is then automatically Gaussian. Thus in all these applications there will be a solution either of the form (5.3), where $y(s)$ determines a process with stationary independent Gaussian increments, or possibly of a sum of independent terms of this same form. (It can be shown that such a sum can be reduced to a single term of this form.) This standard form will be verified in an important particular case, leading to rational spectral densities, in the concluding part of this section.

The use of Gaussian distributions is sometimes justified a posteriori by the fact that the Maxwell distribution of particle velocities is Gaussian. The preceding discussion shows that *there is no mathematical necessity for this Gaussian hypothesis*, however. There are many other $y(s)$ processes with independent stationary (strict sense) increments, and with finite second moments.

Spontaneous currents in electrical circuits.—In was realized years ago that the spontaneous thermal motion of electrons in a conductor must result in spontaneous surges of current, dependent on the absolute temperature of the conductor.[31] In the simplest treatment of electrical networks, the inductances, resistances, and capacitances are lumped as separate units. In such a treatment it is natural to suppose that each resistor R is actually a generator with internal resistance R and of appropriate output characteristics. Let $y(t)$ be the electromotive impulse delivered by the generator between times 0 and t. The actual electromotive force $F(t)$, whose integral is $y(t)$ in normal usage, is considered to a first approximation to be unselfcorrelated, with zero means:

$$\begin{cases} E\{F(t)\} = 0, \\ E\{F(t)F(s)\} = 0 \quad \text{if } t \neq s. \end{cases} \tag{5.12}$$

[30] See a series of papers by A. Blanc-Lapierre, *C. R. Acad. Sci. Paris*, 1943-1946, which deduce this result in the same general way.

[31] See, for example, H. A. Lorentz, *Les théories statistiques en thermodynamique* (Leipzig-Berlin, 1916), pp. 50, 98-99. A detailed discussion of this and related phenomena will be found in E. B. Moullin's book *Spontaneous Fluctuations of Voltage* (Oxford, 1938). Recent valuable treatments have been given by M. C. Wang and G. E. Uhlenbeck, *Reviews of Modern Physics*, vol. 17 (1945), pp. 323-342, and by S. O. Rice, *Bell System Technical Journal*, vol. 23 (1944), pp. 282-332; vol. 24 (1945), pp. 46-156.

Actually, $y(t)$ is considered, rather than $F(t)$, and, as suggested by (5.12), it is supposed that the $y(t)$ process is a process with uncorrelated increments and zero means.[32] From physical considerations it is clear that the increments of $y(t)$ are stationary (strict sense). Moreover, the spontaneous voltages in separate resistors are certainly mutually independent, and have the same average properties for identical resistors. Hence $y(t)$ considered as a function of R must be a process with independent stationary (strict sense) increments also,

$$\begin{cases} E\{y(s+t) - y(s)\} = 0, \\ E\{|y(s+t) - y(s)|^2\} = \text{const. } R|t|. \end{cases} \tag{5.13}$$

Some light can be shed on the constant in (5.13) on a priori grounds. The constant must surely be constant only at a particular value of the absolute temperature T, and clearly must actually be a multiple of T. Moreover, it is a well-known principle of molecular physics that wherever there is a T, Boltzmann's constant k cannot be far behind. Thus the constant should ideally be some simple multiple of kT. A more detailed analysis, involving the principle of equipartition of energy, can be used to evaluate this simple multiple. It is 2.[33]

$$\begin{cases} E\{y(s+t) - y(s)\} = 0, \\ E\{|y(s+t) - y(s)|^2\} = 2kTR|t|. \end{cases} \tag{5.13'}$$

Finally, it is assumed that the $y(s)$ process is a Gaussian process, in accordance with the general principle developed in the beginning of this section.

The dispersion evaluation (5.13') has been seen in section 3 to have a simple frequency interpretation; the real form of the frequency spectrum of the electromotive force (remember that both this force and the frequency spectrum are only symbolic) has constant density $4kTR$. Nyquist deduced this frequency spectrum directly, in order to explain the spontaneous voltages measured by Johnson.[34]

The problem of analyzing the spontaneous thermal currents in an electrical network is the problem of solving the Lagrange equations in the mesh charges:

$$\sum_n [L_{mn}Q''_n(t) + R_{mn}Q'_n(t) + G_{mn}Q_n(t)] = E_m(t), \tag{5.14}$$

where the voltages $\{E_m(t)\}$ are chosen to give the determining voltages ascribed to the resistors. The solution thus gives the solution to any system of coupled harmonic oscillators with this type of driving force, for example, to a mechanical system subjected to external molecular impacts.

[32] Strictly speaking, an $F(t)$ process satisfying (5.12) and not vanishing identically is not measurable, and therefore $F(t)$ cannot be integrated. Corresponding to this fact is the fact already noted several times above that $y'(t)$ does not exist in general.

[33] See, for example, H. A. Lorentz, *op. cit.*, p. 49; the value of this constant will be deduced by essentially the same method at the end of this paper.

[34] J. B. Johnson, *Phys. Rev.*, vol. 32 (1928), pp. 97-109; H. Nyquist, *Phys. Rev.*, vol. 32 (1928), pp. 110-113.

Wang and Uhlenbeck[35] solved the system (5.14) with the given stochastic voltage sources in the sense that they found the transition probabilities for the $2N$-dimensional stochastic process determined by the variables $Q_1(t), \cdots, Q_N(t), Q'_1(t), \cdots, Q'_N(t)$. This process is Gaussian and asymptotically temporally homogeneous ($t \to \infty$), and in fact actually so if the appropriate particular solution of the system (5.14) is chosen. Wang and Uhlenbeck found the probability distribution by solving the Fokker-Planck equation for the transition probabilities, using the fact that this process is a Markoff process. The system (5.14) itself was used

a) implicitly, to justify by the fact of its linearity the use of the Fokker-Planck equation (that is, the inference that the Q, Q' process is a Markoff process);
b) explicitly, to derive the coefficients in the differential equation.

The following method gets at the distribution by finding the specific solution for the charges and currents. Two alternative approaches will be given, the first to exhibit an approach which can be used in all linear problems, the second one specifically adapted to the system (5.14).

A slightly different problem will be considered first, which shows the principle of the method to be used. *Given any two points of the network*, A, B, *find the spontaneous voltage generated between them.* Let the resistors in the network be $R_1, \cdots, R_N$ and let $z_n(\omega)$ be the transfer impedance from R_n to AB at frequency ω.[36] Then if $z(\omega)$ is the impedance between A and B at frequency ω, the steady-state voltage at AB due to the voltage $e^{2\pi i\omega t}$ at R_n is $\frac{z(\omega)}{z_n(\omega)} e^{2\pi i\omega t}$, and more generally the steady-state voltage across AB due to a voltage $E_n(t)$ at R_n is

$$V_n(t) = \int_{-\infty}^{\infty} A_n(t-s)E_n(s)ds, \tag{5.15}$$

where

$$A_n(t) = \int_{-\infty}^{\infty} e^{2\pi i\omega t} \frac{z(\omega)}{z_n(\omega)}\, d\omega.^{37} \tag{5.16}$$

[35] *Op. cit.*

[36] The transfer impedance is defined as follows: a voltage source with output $e^{2\pi i\omega t}$ and internal resistance R_n placed in the circuit instead of R_n gives rise to a steady-state short-circuit current $e^{2\pi i\omega t}/z_n(\omega)$ between A and B.

[37] This standard result is clear formally from the fact that if $E_n(s)$ is written as a Fourier integral,

$$E_n(s) = \int_{-\infty}^{\infty} e^{2\pi i\omega s} E_n^*(\omega)\, d\omega,$$

the voltage across AB is

$$V_n(t) = \int_{-\infty}^{\infty} E_n^*(\omega) \frac{z(\omega)}{z_n(\omega)} e^{2\pi i\omega t}\, d\omega,$$

since the factor $z(\omega)/z_n(\omega)$ is introduced at each frequency. This evaluation of V_n is equivalent to (5.16) because of the relation between Fourier transforms and convolutions.

In the present application, $E_n(t)$ corresponds to a real Gaussian process with stationary independent increments, whose variables $\{y_n(t)\}$ satisfy

$$E\{dy_n(t)\} = 0, \qquad E\{|dy_n(t)|^2\} = 2R_nkT_n|dt|, \tag{5.17}$$

where T_n is the absolute temperature of R_n, in the sense that

$$E_n(t)dt = dy_n(t). \tag{5.18}$$

Hence

$$V_n(t) = \int_{-\infty}^{\infty} A_n(t-s)dy_n(s).^{38} \tag{5.19}$$

It has already been seen that this integral makes sense if $|A_n(t)|^2$ is integrable. In the present application it is well-known that each z_n is rational, with zeros only in the upper half-plane. It follows from elementary residue theory that $A_n(t)$ vanishes for $t < 0$, and decreases exponentially as $t \to +\infty$. The full voltage across AB is therefore given by

$$V(t) = \sum_n \int_{-\infty}^{t} A_n(t-s)dy_n(s). \tag{5.20}$$

The fact that the upper limit of integration is effectively t could have been deduced from physical principles, since the solution at time t could hardly depend on the thermal voltages at later times.

Thus $V(t)$ has the form indicated by the general argument initiating this section. The correlation function of the stationary (strict sense) Gaussian $V(t)$ process can now be written down, with the use of the general analysis of processes of moving averages made in section 3 [see (3.21)],

$$\begin{cases} E\{V(t)\} = 0, \\ E\{V(t+h)\overline{V(t)}\} = 2k\sum_n T_nR_n \int_{-\infty}^{\infty} \dfrac{|z(\omega)|^2}{|z_n(\omega)|^2} e^{2\pi i\omega h}\, d\omega. \end{cases} \tag{5.21}$$

Since $V(t)$ is actually real, it is customary to use the real form of the frequency spectrum,

$$E\{V(t+h)\overline{V(t)}\} = 4k\sum_n T_nR_n \int_0^{\infty} \frac{|z(\omega)|^2}{|z_n(\omega)|^2} \cos 2\pi\omega h\, d\omega, \tag{5.22}$$

so that the spectral density (real form) of the $V(t)$ process is

$$4k\sum_n T_nR_n \frac{|z(\omega)|^2}{|z_n(\omega)|^2}.$$

[38] This is of course only one of the particular solutions, chosen to give a stochastically steady state.

This evaluation of the density is due to F. C. Williams, who showed that if $T_n = T$ is independent of n it reduces to Nyquist's formula,

$$4kT \text{ (real part of } z\text{)}.^{39} \tag{5.24}$$

The solution of (5.14) for the charges and currents is essentially the same problem as that just treated. For simplicity we shall suppose that all the resistors have the same temperature, T. If $1/\zeta_{mn}$ is the transfer impedance between the mth and nth meshes,[40] a particular solution to (5.14) is given by

$$\begin{cases} Q_m(t) = \sum_n \int_{-\infty}^{t} A_{mn}(t-s)E_n(s)ds, \\ Q'_m(t) = \sum_n \int_{-\infty}^{t} A'_{mn}(t-s)E_n(s)ds, \end{cases} \tag{5.25}$$

where

$$\begin{cases} A_{mn}(t) = \int_{-\infty}^{\infty} e^{2\pi i\omega t} \dfrac{\zeta_{mn}{}^{(\omega)}}{2\pi i\omega}\, d\omega,^{41} \\ A'_{mn}(t) = \int_{-\infty}^{\infty} e^{2\pi i\omega t} \zeta_{mn}{}^{(\omega)} d\omega. \end{cases} \tag{5.26}$$

These equations are the analogues of (5.15) and (5.16). In the present case,

$$E_n(t)dt = dy_n(t), \tag{5.27}$$

where $\{y_n(t)\}$ are the variables of a real Gaussian process of stationary independent increments, with

$$\begin{cases} E\{dy_n(t)\} = 0, \\ E\{|dy_n(t)|^2\} = 2kTR_{mn}|dt|, \\ E\{dy_m(t)dy_n(t)\} = 2kTR_{mn}|dt|. \end{cases} \tag{5.28}$$

Each y_n process is a sum of independent processes corresponding to the resistors in the nth mesh. This means that the y_n process and the y_m process are not mutually independent if the mth and nth meshes have common resistors. This explains the last line of (5.28). The correlation matrices

$$\begin{aligned} E\{Q(t+h)\overline{Q(t)}\} &: (E\{Q_m(t+h)Q_n(t)\}) \\ E\{Q(t+h)\overline{Q'(t)}\} &: (E\{Q_m(t+h)Q'_n(t)\}) \\ E\{Q'(t+h)\overline{Q'(t)}\} &: (E\{Q'_m(t+h)Q'_n(t)\}) \end{aligned}$$

[39] H. Nyquist, *Phys. Rev.*, vol. 32 (1928), pp. 110-113; F. C. Williams, *Jour. Inst. Elec. Eng.*, vol. 81 (1937), pp. 751-760.

[40] The matrix (ζ_{mn}) is the inverse of the matrix of complex mutual impedances between the meshes; ζ_{mn} is the transfer admittance.

[41] It is easily verified that $\zeta_{mn}{}^{(\omega)}/\omega$ remains finite when $\omega \to 0$.

are easily calculated, applying the methods already used in finding the correlation functions of processes of moving averages. The results are, if the matrices R, ζ, ζ^* are given by

$$R = (R_{mn}), \qquad \zeta = (\zeta_{mn}), \qquad \zeta^* = (\zeta^*_{mn}), \qquad \zeta^*_{mn} = \bar{\zeta}_{nm},$$

$$\begin{cases} E\{ Q(t+h)\,\overline{Q(t)}\} = 2kT\displaystyle\int_{-\infty}^{\infty} e^{2\pi i\omega h}\,\frac{\zeta R\zeta^*}{(2\pi\omega)^2}\,d\omega \\ E\{Q(t+h)\,\overline{Q'(t)}\} = 2kT\displaystyle\int_{-\infty}^{\infty} e^{2\pi i\omega h}\,\frac{\zeta R\zeta^*}{2\pi i\omega}\,d\omega, \\ E\{Q'(t+h)\overline{Q'(t)}\} = 2kT\displaystyle\int_{-\infty}^{\infty} e^{2\pi i\omega h}\,\zeta R\zeta^* d\omega. \end{cases} \tag{5.29}$$

The preceding discussion has been indirect. The quickest way to gain real insight into the solutions of the system (5.14) is to solve the system explicitly. This can be done very simply, as follows. Definite the N-dimensional matrices L, R, G by

$$L = (L_{mn}), \qquad R = (R_{mn}), G = (G_{mn}), \tag{5.30}$$

and the $2N$-dimensional matrices U, V, W by

$$U = \begin{pmatrix} I & 0 \\ O & L \end{pmatrix}, \qquad V = \begin{pmatrix} 0 & -I \\ G & R \end{pmatrix}, \qquad W = U^{-1}V = \begin{pmatrix} 0 & -I \\ LG & L^{-1}R \end{pmatrix} \tag{5.31}$$

Let Q be a $2N$-dimensional vector with components $Q_1(t), \cdots, Q_N(t), Q'_1(t), \cdots, Q'_N(t)$, and finally let $y(s)$ be the $2N$-dimensional vector with components $(0, \cdots, 0, y_1(s), \cdots, y_N(s)$, where the y processes are defined as above, satisfying (5.27) and (5.28). Then the system (5.14), with the stochastic voltages inserted in place of the E_n, can be written in the form

$$UQ'(t) + VQ(t) = y'(t). \tag{5.32}$$

The general solution of this symbolic equation is

$$Q(t) = e^{-W(t-t_0)}Q(t_0) + \int_{t_0}^{t} e^{-W(t-S)}U^{-1}dy(s), \tag{5.33}$$

where t_0 is any initial point. If $t_0 \to -\infty$, this becomes the stationary solution

$$Q(t) = \int_{-\infty}^{t} e^{-W(t-s)}\,U^{-1}dy(s). \tag{5.34}$$

Here we are using the fact that R, L are symmetric and positive definite to ensure that the transient term disappear and that the integral in (5.34) con-

verge. (The integrand decreases exponentially as the argument goes to infinity.) The particular solution (5.34) is the one obtained by a different method above. This solution is a stationary process of moving averages, in which $Q(t)$ involves only past thermal voltages, as required by the physics background. This property is shared by any solution (5.33) if $Q(t_0)$ is chosen to be independent of future thermal voltages, that is, independent of $y(\tau)$ differences for $\tau > t_0$. If this is done, $Q(t)$ will depend only on $Q(t_0)$ and on the thermal voltages at times between t_0 and t. Now suppose the initial conditions have been so chosen. [In particular, one such choice will give (5.34), at least for $t \geqq t_0$.] Then, if $t > t_1 > t_0$,

$$Q(t) = e^{-W(t-t_1)}Q(t_1) + \int_{t_1}^{t} e^{-W(t-s)}U^{-1}dy(s), \tag{5.35}$$

and the difference

$$Q(t) - e^{-W(t-t_1)}Q(t_1)$$

depends only on thermal voltages after time t_1. Hence this difference is independent of $Q(s)$ for $s \leqq t_1$, and we have proved that the probability distribution of $Q(t)$, if $Q(\tau)$ is preassigned at $t_1 > t_2 > \cdots \geqq t_0$ has mean value $e^{-W(t-t_1)}Q(t_1)$ and dispersion that of the last term in (5.35). Thus this distribution is entirely unaffected by the values assigned to $Q(t_2), \cdots$. In other words, the Q process is a Markoff process.[42] The details of the calculation of the correlation matrix of the process will be omitted. The matrix has been found above by a different method.

As an application of these methods, which will clarify the significance of the results, we solve the Lagrange equation for a single mesh consisting of a resistance connected across a condenser,

$$RQ'(t) + Q/C = y', \qquad E\{dy(t)\} = 0, \qquad E\{|dy(t)|^2\} = 2kTR|dt|. \tag{5.36}$$

(Note again that this equation is interpreted symbolically, since y' does not exist.) The stationary solution is

$$Q(t) = \frac{1}{R}\int_{-\infty}^{t} e^{-\frac{(t-s)}{RC}}\,dy(s). \tag{5.37}$$

In order to see the significance of the spectrum more clearly, we write $Q(t)$ in the normal form (2.11), using (3.30),

$$Q(t) = \int_{-\infty}^{\infty} \frac{e^{2\pi its}}{\frac{1}{C} + 2\pi iRs}\,dy^*(s) \qquad E\{dy^*(s)\} = 0 \tag{5.38}$$

$$E\{|dy^*(s)|^2\} = 2kTR|ds|,$$

[42] The form (5.34) is essentially in a normal form of an N-dimensional Gaussian-Markoff process; see Doob, *Annals of Math. Stat.*, vol. 15 (1944), theorem 4.3.

where y^* is the Fourier transform of y. Then the spectral density in the complex form is

$$F'(\lambda) = \frac{2kTR}{\frac{1}{C^2} + (2\pi R\lambda)^2} \tag{5.39}$$

and the correlation function is

$$E\{Q(t+h)Q(t)\} = 2kTR\int_{-\infty}^{\infty} \frac{e^{2\pi i\lambda h}}{\frac{1}{C^2} + (2\pi R\lambda)^2}\, d\lambda. \tag{5.40}$$

In particular,

$$E\{Q(t)^2\} = 2kTR\int_{-\infty}^{\infty} \frac{d\lambda}{\frac{1}{C^2} + (2\pi R\lambda)^2} = kTC. \tag{5.41}$$

This equation is in accordance with the principle of equipartition of energy, since.

$$E\{Q(t)^2/C\} = kT. \tag{5.42}$$

Conversely, suppose that we had stopped at (5.13), when it had been deduced that the y process was a Gaussian process of stationary independent increments with zero means, and $E\{|dy(t)|^2\} = \text{const.}\ |dt|$ but with the value of the constant left undetermined. Then (5.42) could be used to obtain the constant, $2kRT$.

Formally, it might be expected that (5.38) could be differentiated, giving

$$Q'(t) = \int_{-\infty}^{\infty} \frac{e^{2\pi ist} 2\pi is}{\frac{1}{C} + 2\pi iRs}\, dy^*(s), \tag{5.43}$$

which might be expected to imply that $Q'(t)$ determines a stationary process with non-integrable spectral density

$$2kTR\,\frac{(2\pi\lambda)^2}{\frac{1}{C^2} + (2\pi R\lambda)^2}. \tag{5.44}$$

In fact, non-integrable spectral densities have been derived in essentially this way, and have caused some confusion. These operations have not been mathematically correct, however. According to (5.36), $Q'(t)$ does not exist, since $y'(t)$ does not exist, and corresponding to this fact the integral in (5.43) does not converge. In physical terms, there is no current in this mesh; the charges

do not change smoothly enough! If there were an inductance in the mesh also, (5.36) would be replaced by

$$LQ''(t) + RQ'(t) + Q(t)/C = y'(t), \tag{5.45}$$

and in this case there would be a current $Q'(t)$, but the current derivative $Q''(t)$ would not exist. Mathematically speaking, then, there is no difficulty with these non-integrable spectral densities which have arisen to plague investigators; they pertain to the spectra of non-existent functions, pertaining to specters, not spectra. From a physical standpoint, it should be noted that the reasoning which led to the mathematical hypothesis that $y(t)$ was the variable of a process with independent increments cannot be taken too seriously; the increments of $y(t)$ either as time increases or along the length of a resistor are only independent, or nearly so, if the intervals of measurement are not too close together. A more reasonable hypothesis at the next stage of approximation, even from the point of view of classical physics, might well make the y process regular enough to afford derivatives, in which case the non-integrable spectral densities would not arise, or at least would arise at a different level. Such a modification would lose some of the formal elegance of the theory as outlined here, but enough is known about the most general process with stationary increments to give a complete spectral theory with no essentially new difficulties.