CONTINUOUS STOCHASTIC PROCESSES

-oOo-

Thesis

by

Donald Allan Darling

Submitted in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California
1 9 4 7

Acknowledgement

My indebtedness to the mathematics staff, and in particular Professor Morgan Ward, is very great. The spirit of patience, understanding, and indulgence which they have exercized has been a constant source of encouragement to me. Above all, I am grateful for the way in which they have introduced to me the ineffable pleasures of mathematical research--the concise reasoning, the attainment of general conceptual maturity, the impelling curiosity of a hard problem--things only a mathematician can know.

Continuous Stochastic Processes

Table of Contents

CONTINUOUS STOCHASTIC PROCESSES

Summary

A general review of stochastic processes is given in the introduction; definitions, properties and a rough classification are presented together with the position and scope of the author's work as it fits into the general scheme.

The first section presents a brief summary of the pertinent analytical properties of continuous stochastic processes and their probability-theoretic foundations which are used in the sequel.

The remaining two sections (II and III), comprising the body of the work, are the author's contribution to the theory. It turns out that a very inclusive class of continuous stochastic processes are characterized by a fundamental partial differential equation and its adjoint (the Fokker-Planck equations). The coefficients appearing in these equations assimilate, in a most concise way, all the salient properties of the process, freed from boundary value considerations. The writer's work consists in characterizing the processes through these coefficients without recourse to solving the partial differential equations.

First, a class of coefficients leading to a unique, continuous process is presented, and several facts are proven to show why this class is restricted. Then, in terms of the coefficients, the unconditional statistics are deduced, these being the mean, variance and covariance. The most general

class of coefficients leading to the Gaussian distribution is deduced, and a complete characterization of these processes is presented. By specializing the coefficients, all the known stochastic processes may be readily studied, and some examples of these are presented; viz. the Einstein process, Bachelier process, Ornstein-Uhlenbeck process, etc. The calculations are effectively reduced down to ordinary first order differential equations, and in addition to giving a comprehensive characterization, the derivations are materially simplified over the solution to the original partial differential equations.

In the last section the properties of the integral process are presented. After an expository section on the definition, meaning, and importance of the integral process, a particular example is carried through starting from basic definition. This illustrates the fundamental properties, and an inherent paradox. Next the basic coefficients of the integral process are studied in terms of the original coefficients, and the integral process is uniquely characterized. It is shown that the integral process, with a slight modification, is a continuous Markoff process.

The elementary statistics of the integral process are deduced: means, variances, and covariances, in terms of the original coefficients. It is shown that an integral process is never temporally homogeneous in a non-degenerate process.

Finally, in terms of the original class of admissible coefficients, the statistics of the integral process are explicitly presented, and the integral process of all known continuous processes are specified.

CONTINUOUS STOCHASTIC PROCESSES

Introduction

The study of stochastic processes is concerned with the statistical description of a one parameter family of chance variables $\mathcal{U}(t)$. The parameter t is usually called the time, and is always a real parameter belonging to a certain set T; $t \in T$. This set will generally be the full real line $T = (t \,/\, -\infty \leq t \leq +\infty)$. For each $t \in T$ the chance variable $\mathcal{U}(t)$ will, in this paper, be a one dimensional random variable whose probability measure is defined over any Borel measurable set E of the one dimensional Euclidean space, and which is described by a distribution function $H(t, u)$

$$Pr\left\{\mathcal{U}(t) \in E\right\} = \int_E d_u H(t, u), \quad t \in T.$$

The case when $\mathcal{U}(t)$ may be an n-dimensional vector random variable is a more or less straightforward generalization of the one dimensional case, and all of the intrinsic features of stochastic processes are displayed in this latter case. For the more general case cf. Doob (2)* and Cramer (2).

For a given finite set of $t_j \in T$, $t_1 < t_2 \cdots < t_K$, the suite of random variables $\left\{\mathcal{U}(t_j),\ j = 1, 2, \cdots, K\right\}$ is said to form a process. If, corresponding to any process, the associated multivariate distribution

$$F(t_1, u_1; t_2, u_2; \cdots; t_K, u_K) = Pr\left\{\mathcal{U}(t_j) \leq u_j,\ j = 1, 2, \cdots K\right\}$$

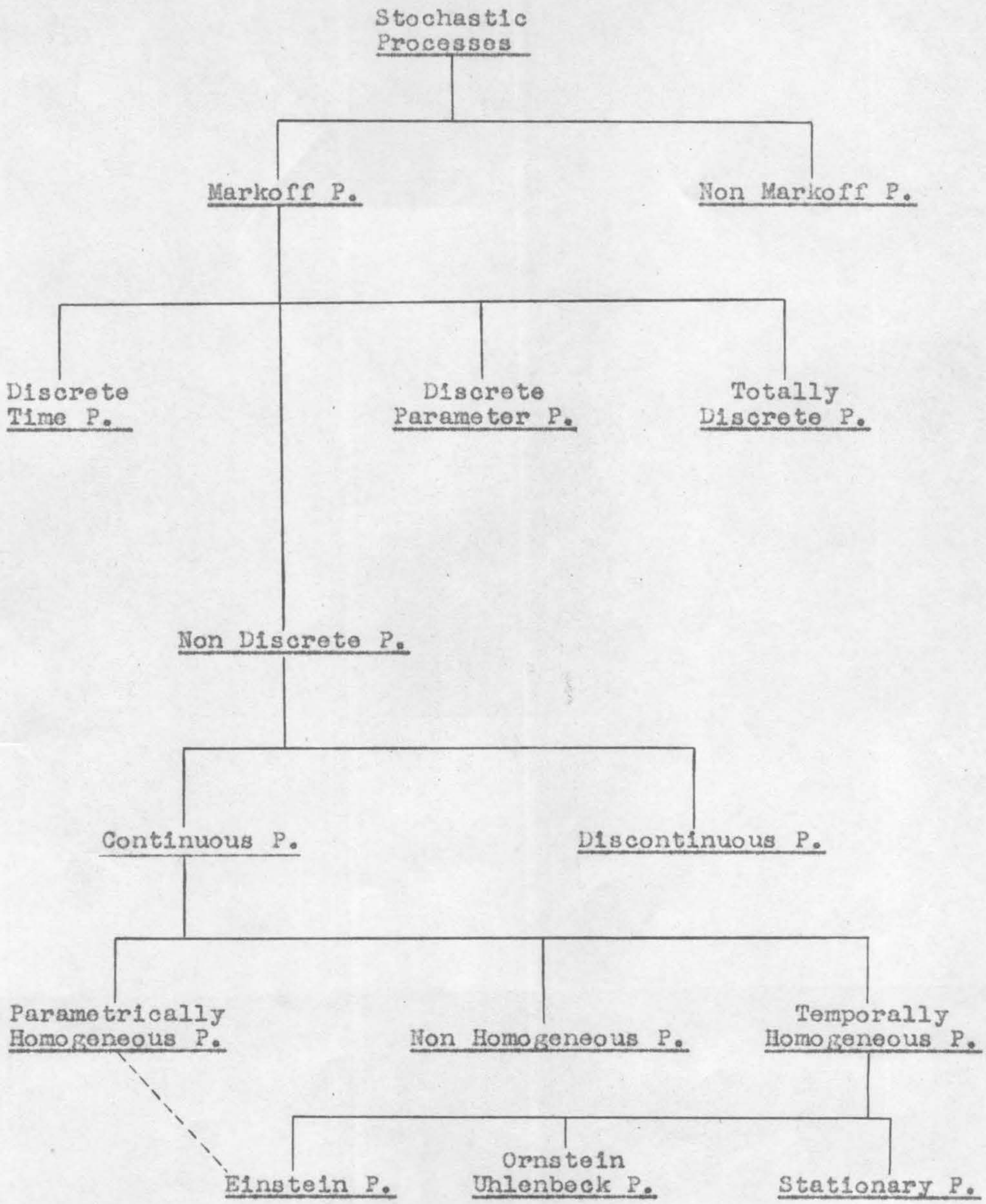
- -

*Underlined numbers in brackets refer to the bibliography at the end of the paper.

is known, then the family of chance variables $\mathcal{U}(t)$ is said to determine a stochastic process. This follows the definition of Khintchine (2).

The modern viewpoint treats the succession of chance variables $\mathcal{U}(t)$ as points in a function space. For the process composed of a finite number κ of components as above the process may be embedded in a κ dimensional Euclidean space, and this condition is generalized to a countable number of components and then to a non-denumerable number. In these generalized stochastic processes there are a number of topological difficulties that have not yet been surmounted; for a general exposition, see Doob (3). It turns out that probability considerations can be bodily carried over to measure theoretic considerations in the function space and results of great generality can be carried forward. No use will be made of these notions in this paper, which contains results of primarily analytical interest. Certain statements like "almost all (continuous) functions $\mathcal{U}(t)$, $a \leq t \leq b$, have the property P" will appear however. It is fairly simple to attach a meaning to this statement, for as shown by Paley and Wiener (1), the set of continuous functions $\mathcal{U}(t)$ can be mapped onto the real axis by associating the Fourier coefficients of $\mathcal{U}(t)$ with members in the dyadic expansion of a real number χ , and the usual Lesbegue measure can be employed.

We give now a brief outline of the various subdivisions of stochastic processes and show where the present work fits into this scheme. On the following page is a diagram of the


Stochastic Processes
Markoff P.
Non Markoff P.
Discrete Time P.
Discrete Parameter P.
Totally Discrete P.
Non Discrete P.
Continuous P.
Discontinuous P.
Parametrically Homogeneous P.
Non Homogeneous P.
Temporally Homogeneous P.
Einstein P.
Ornstein Uhlenbeck P.
Stationary P.

various processes and their interdependence, and we give a brief description of the main types.

The class of stochastic processes which most readily yield to analysis are the so-called Markoff processes, and indeed there has been virtually no work at all done on non-Markoff processes owing to the extreme complexity of the analysis. The Markoff processes are described as follows: From the previous joint distribution F it is clear that we can obtain the conditional distribution for the variable $\mathcal{U}(t_k)$ given the values of the preceding variables $\mathcal{U}(t_j)$ $j = 1, 2, 3 \cdots k-1$. We have, in fact, when $t_1 < t_2 < \cdots < t_k$

$$F(t_1, u_1;\ t_2, u_2;\ \cdots;\ t_{k-1}, u_{k-1} : t_k, u_k) = \frac{F(t_1, u_1;\ t_2, u_2;\ \cdots;\ t_k, u_k)}{F(t_1, u_1;\ t_2, u_2;\ \cdots;\ t_k, \infty)}$$

If it happens that $F(t_1 u_1; \cdots; t_{k-1}, u_{k-1} : t_k, u_k) = F(t_{k-1}, u_{k-1} : t_k, u_k)$ i.e. the conditional distribution for $\mathcal{U}(t_k)$ given the rest of the variables actually only depends upon $\mathcal{U}(t_{k-1})$, the penultimate variable, then the process is called a Markoff process. Roughly speaking, the present distribution of the variable $\mathcal{U}(t_k)$ is influenced only by the last known value it assumed, and is not modified by any earlier knowledge of it we may have.

Thus it appears that a Markoff process is completely determined if we know for any s and t, $(s<t)$ the probability that given $\mathcal{U}(s) = u$ we should have $\mathcal{U}(t) \leq v$ for any s, t, u, and v. This conditional probability, which is sometimes called a 'transition probability' is denoted by $F(s, u; t, v)$. Kolmogoroff, in his fundamental work (1),

was the first to give a comprehensive survey of the general Markoff processes and to deduce results of a broad degree of generality which covered most of the extant special treatments. Kolmogoroff's principal contribution lay in giving a rigorous postulational foundation for the temporally continuous process, which before his paper was mainly heuristic. The earlier processes, often named after their discoverers, were devised to describe certain physical phenomena such as diffusion, Brownian motion, turbulence, etc. A rather complete bibliography of the earlier work of the physicists is to be found in Ornstein-Uhlenbeck (1) and Doob (1).

A totally discrete Markoff process is generally called a Markoff chain after its discoverer A. A. Markoff (1). Here the set of times $\{t_j\}$ is an enumerable set, as is the set of values of the parameter $\{u_j\}$. Then the transition probabilities $F(t_i, u_i; t_j, u_j)$, $t_i < t_j$, are seen to form a finite or infinite matrix--the so-called 'stochastic matrices' whose properties were thoroughly investigated by Romanovsky (1) inter alia. The larger part of the investigations have been devoted to this topic. In general, the discrete processes can be reduced down to an urn scheme, and some simple non-Markoff processes have been studied by this device (cf. Onicescu and Mihoc (1) and Hostinsky (1)).

In the discrete process if we pass to the limit in either the time or the parameter, so that these variables become respectively continuous, we arrive at processes envisaged by Kolmogoroff, but the most interesting case is when both the parameter and the time can assume a continuous system of values (the non-discrete process). In this case powerful tools of analysis can be brought to bear upon the problem, and many of the earlier results which were obtained as asymtotic limits from the discrete processes may be attained directly.

Under this class fall two general subdivisions--the continuous and the discontinuous process. By 'continuous' is meant a process $\mathcal{U}(t)$ where the continuum of states assumed by the random variable forms, in a sense to be specified later, a continuous function. The discontinuous processes have been treated only slightly--Feller (2) has proven some general theorems on existence and uniqueness.

The continuous processes treated by Kolmogoroff (1) are characterized by certain partial differential equations. The coefficients appearing in these equations give the most succinct formulation of the process imaginable.

The present work presents a more or less complete characterization of the continuous processes in terms of these coefficients. It turns out that an extensive treatment of the unconditional (absolute) distribution of the process can be made, and results of a quite inclusive nature deduced. This is done solely in terms of the coefficients without solving the differential equations for the transition probabilities as is usually done.

The last section is devoted to a study of the integral process, which is, as its name implies, the process obtained by integrating a given process $\mathcal{U}(t)$. This process has importance in certain physical applications as well as in the law of large numbers for dependent (continuous) events. Again, a general characterization is available by means of the coefficients appearing in the differential equations characterizing the process $\mathcal{U}(t)$.

I

The Continuous Stochastic Processes

1.1 The transition probabilities

In order to satisfy the general differential relations, to be given later, the transition probabilities are presumed to satisfy certain conditions of analaticity and regularity. These are:

1. $F(s, u; t, v)$ has partial derivatives to the fourth order in s, t, u, and v, $s < t$. For fixed s and u these derivatives are uniformly bounded in s and u when $t-s \geqslant \kappa > 0$.
2. For the absolute moments

$$m^{(i)}(t, u, \Delta) = \int_{-\infty}^{\infty} |v-u|^i \, d_v F(s, u; s+\Delta, v)$$

we have

a) $m^{(i)}(t, u, \Delta)$ exists, $i = 1, 2, 3$

b) $\lim_{\Delta \to 0} m^{(i)}(t, u, \Delta) = 0$, $i = 1, 2, 3$

3. $$\lim_{\Delta \to 0} \frac{m^{(3)}(t, u, \Delta)}{m^{(2)}(t, u, \Delta)} = 0$$

These conditions as enunciated by Kolmogoroff will be collectively called conditions K. They imply that large deviations in the parameter V occur very seldom in small intervals, speaking loosely. More precisely

$$\lim_{\Delta \to 0} \int_{|v-u| > \varepsilon} d_v F(s, u; s+\Delta, v)\, dv = 0$$

for any $\varepsilon > 0$. Also, on account of Liapounoff's general inequality for the absolute moments, $\beta_1\beta_3 \geq \beta_2^2$, condition 3 implies

$$\lim_{\Delta \to 0} \frac{m^{(2)}(t,u,\Delta)}{m^{(1)}(t,u,\Delta)} = 0.$$

Since the derivative exists, we shall work hereafter with the density function $f(s,u;t,v)$ given by

$$\frac{\partial}{\partial v} F(s,u;t,v) = f(s,u;t,v)$$

It is clear that for $s=t$ the function $f(s,u;t,v)$ has a discontinuity, being a Dirac function. We have, in fact,

$$\lim_{\Delta \to 0} \int_{-\infty}^{\infty} e^{ixv} f(s,u;s+\Delta,v)\,dv = e^{ixu}$$

from the conditions K .

1.2 <u>The notion of continuity</u>

The succession of sample values $\mathcal{U}(t)$ for t varying continuously generates a statistical system of functions with an associated measure (probability) assigned in a function space. Consequently, to arrive at a notion of continuity we must utilize some mean value idea which assimilates the salient neighborhood properties of $\mathcal{U}(t)$. There are several of these available:

1. Continuity-in-probability

   A process is called continuous in probability (c. i. p.) at $t=s$ if to any two numbers $\delta > 0$ and $\varepsilon > 0$ there exists a K such that for any

$0 \leq \Delta < \kappa$ we have

$$Pr\left\{ |\mathcal{U}(s+\Delta) - \mathcal{U}(s)| \geq \varepsilon \right\} < \delta$$

If this property holds for all s in the range $T = (s | a \leq s \leq b)$ the process is c. i. p. over T. Slutsky (1) presented these ideas and proved moreover that a function c. i. p. over T is uniformly c. i. p. over T, and κ can be chosen independently of s.

2. Continuity in the mean

A process is called continuous in the mean at $t = s$ if to any $\delta > 0$ there corresponds a κ such that for any $0 \leq \Delta < \kappa$ we have

$$E\left(|\mathcal{U}(t+\Delta) - \mathcal{U}(t)| \right) < \delta$$

Again continuity in the mean is extended to the full interval by requiring this property to hold at each point of it.

3. p-continuity

A process is called p-continuous at $t = s$ if to any $\delta > 0$ there corresponds a κ such that for $0 \leq \Delta < \kappa$ we have

$$E\left(|\mathcal{U}(t+\Delta) - \mathcal{U}(t)|^p \right)^{\frac{1}{p}} < \delta$$

We see that whenever $q < p$ then p-continuity implies q-continuity. This follows from the well-known absolute moment inequality (or Hölder's inequality)

$$\beta_p^{\frac{1}{p}} \geq \beta_q^{\frac{1}{q}} \quad , \quad q < p$$

Moreover, from the Bienayme-Tchebycheff inequality c. i. p. follows from continuity in the mean. For

$$Pr.\left\{|\mathcal{U}(t+\Delta)-\mathcal{U}(t)| \geq \varepsilon\right\} \leq \frac{E\left\{|\mathcal{U}(t+\Delta)-\mathcal{U}(t)|\right\}}{\varepsilon}$$

In this paper we work with continuity in mean square (c. m. s.) where $p=2$ in 2 above. In particular our processes are c. i. p. and continuous in the mean. Khintchine (2) adopts the definition that the correlation coefficient between $\mathcal{U}(t+\Delta)$ and $\mathcal{U}(t)$ approaches 1 as $\Delta \to 0^+$; $\lim_{\Delta \to 0^+} R(t+\Delta, t) = 1$. This can be readily shown to imply c. m. s. Most other workable definitions reduce down to c. m. s.

In a particular process (the Einstein process discussed in section 2.4) Paley and Wiener (1) carried through a delicate investigation of continuity. Their main result is that for almost all functions $\mathcal{U}(t)$

$$\lim_{\Delta \to 0} \frac{|\mathcal{U}(t+\Delta)-\mathcal{U}(t)|}{\Delta^{\lambda}}$$

is zero if $\lambda < 1/2$ and is not finite (with $\overline{\lim}$) if $\lambda > 1/2$, and this uniformly in t. The case $\lambda = \frac{1}{2}$ leads to deeper questions involving the law of the iterated logarithm, etc.

The important question as to whether conditions K lead to continuity of some type will be deferred until section II.

1.3 Fundamental relationships for the transition probabilities

A necessary and sufficient condition that a non negative function $f(s,u;t,v)$ be the transition probability for a continuous stochastic process is that for $s_1 < s_2 < s_3$

$$(1) \qquad \int_{-\infty}^{\infty} f(s_1, u; s_2, v)\, dv = 1$$

$$(2) \qquad f(s_1, u; s_3, v) = \int_{-\infty}^{\infty} f(s_1, u; s_2, w)\, f(s_2, w; s_3\, v)\, dw$$

equation (2) being the Kolmogoroff-Chapman equation. The direct solution of this integral equation is apparently not feasible, and while its solution would yield all Markoff processes, little progress has been made in this direction. See, e.g., Levy (2).

However it is true that if $f(s,u;t,v)$ satisfies conditions K, then a sufficient condition that f satisfies (2) is that it satisfy the two partial differential equations

$$(3) \qquad \frac{\partial f}{\partial s} = -A(s,u)\frac{\partial f}{\partial u} - B^2(s,u)\frac{\partial^2 f}{\partial u^2}$$

$$(4) \qquad \frac{\partial f}{\partial s} = -\frac{\partial}{\partial v}\left(A(t,v) f\right) + \frac{\partial^2}{\partial v^2}\left(B^2(t,v) f\right)$$

where

$$A(s,u) = \lim_{\Delta \to 0} \frac{1}{\Delta}\int_{-\infty}^{\infty} (v-u)\, f(s,u; s+\Delta, v)\, dv$$

$$(5) \qquad 2B^2(s,u) = \lim_{\Delta \to 0} \frac{1}{\Delta}\int_{-\infty}^{\infty} (v-u)^2 f(s,u; s+\Delta, v)\, dv\, ;$$

$A(s,u)$ and $B^2(s,u)$ existing everywhere, except possibly over a non-dense set in the s,u plane, by virtue of conditions K.

The transference of the problem, so to speak, to considerations of partial differential equations leads to important questions of existence and unicity. Under much weaker restrictions than conditions K, Feller (1) has shown that (3) and (4) (generally called the Fokker-Planck equations) are a consequence of the Kolmogoroff-Chapman equation (2). Under certain restrictions on A and B^2 Feller shows that they determine exactly one continuous stochastic process. The conditions act, in a superfluous way, as a boundary condition on the differential equations.

Thus the Fokker-Planck equation (3) together with the boundary condition

$$\lim_{\Delta \to 0} \int_{|v-u|>\delta} f(s-\Delta, u; s, v)\, dv = 0 \quad , \quad \delta > 0$$

and the fact that f is a density distribution on the variable v are sufficient to ensure the existence and unicity of a process having prescribed coefficients $A(s,u)$ and $B^2(s,u)$ subject to the above-mentioned restrictions. It is not even necessary to suppose that f has derivatives or moments of any order, as demanded by conditions K.

Since the results of this paper are of a more formal nature, it will do little harm to suppose conditions K are satisfied. The results will generally be true for a wider class of processes.

1.4 Interpretation of the differential coefficients

The two functions $A(s,u)$ and $B^2(s,u)$ are seen to play a central role in the theory of continuous stochastic processes. Their intrinsic meaning is as follows: $A(s,u)$ is the mean velocity of variation of the variable $\mathcal{U}(s)$ at the displacement u, and during the time interval s to $s+ds$. $B^2(s,u)$ is the mean square velocity under the same conditions, or the 'differential dispersion' of the process, to use Kolmogoroff's terminology.

The functions A and B^2 are conditional expectations, and we may write

(6) $$A(s,u) = \lim_{\Delta\to 0} E\left(\frac{\mathcal{U}(s+\Delta) - \mathcal{U}(s)}{\Delta}\,\middle|\, \mathcal{U}(s) = u\right)$$

where $E(y|\xi)$ means the conditional expectation of the random variable y knowing the contingency ξ. Similarly for $B^2(s,u)$

(7) $$2B^2(s,u) = \lim_{\Delta\to 0} E\left(\frac{(\mathcal{U}(s+\Delta) - \mathcal{U}(s))^2}{\Delta}\,\middle|\, \mathcal{U}(s) = u\right)$$

From equation (7) we shall later deduce a condition that $\mathcal{U}(t)$ shall be c. m. s. From (7) it is apparent that for $\mathcal{U}(s)=u$ given

$$E\left((\mathcal{U}(s+\Delta) - \mathcal{U}(s))^2 \,\middle|\, \mathcal{U}(s) = u\right) = O(\Delta)\; B^2(s,u)$$

this Lipschitz condition immediately implies that for $B^2>0$ the process $\mathcal{U}(t)$ does not have a derivative. Thus, if we conceive of a particular succession of values of $\mathcal{U}(s)$ out of the multiplicity available these will form a continuous function which has a derivative almost nowhere.

There are two types of processes of special importance: the temporally homogeneous and the parametrically homogeneous. A process is called temporally homogeneous if the functions $A(s,u)$ and $B^2(s,u)$ actually only depend on u

$$A(s,u) = A(u)$$

$$B^2(s,u) = B(u)$$

In this case the transition probability $f(s,u;t,v)$ manifestly only depends on u, v, and $t-s$. A process is called parametrically homogeneous when A and B^2 depend only on s

$$A(s,u) = \varphi_0(s)$$

$$B^2(s,u) = \psi_0(s)$$

and in this case $f(s,u;t,v)$ depends only on s, t, and $v-u$.

The temporally homogeneous case is practically the most important, for the general physical phenomena studied have a temporally constant causative mechanism.

II

Characterization of Continuous Processes

2.1 The class of admissible coefficients

The purpose of this study is to characterize continuous processes by the functions $A(s,u)$ and $B^2(s,u)$. The first step is then to find the class of functions leading to continuous processes. This is as yet an unsolved problem, but various sufficient conditions are known (cf. Feller (1) and Fortet (1)).

We first introduce the unconditional probabilities. If at a time s_0 we know the unconditional density distribution for $u(s_0)$, say $h(s_0, u)$, then the unconditional distribution at a later time s is given by

$$h(s, u) = \int_{-\infty}^{\infty} h(s_0, w) f(s_0, w; s, u)\, dw \tag{8}$$

and this distribution evidently satisfies the equation

$$\frac{\partial h(s,u)}{\partial s} = -\frac{\partial (A(s,u)h(s,u))}{\partial u} + \frac{\partial^2 (B^2(s,u)h(s,u))}{\partial u^2} \tag{9}$$

obtained from (4). It is the unconditional distribution which is of the most practical importance and whose properties are studied here.

Now from (7) we obtain by multiplying by $h(s,u)$ and integrating on u

$$2\int_{-\infty}^{\infty} B^2(s,u)h(s,u)\,du = \lim_{\Delta\to 0} \int_{-\infty}^{\infty} E\left(\frac{(u(s+\Delta) - u(s))^2}{\Delta}\,\middle|\, u(s) = u\right) h(s,u)\,du \tag{10}$$

$$= E\left(\frac{(u(s+\Delta) - u(s))^2}{\Delta}\right)$$

if the interchange of limit and integral is permissible. In this case the existence of the left-hand side of (10) will imply $\lim_{\Delta \to 0} E((u(s+\Delta)-u(s))^2)=0$ and the process will be c. m. s. In any event, the size of $B^2(s,u)$ for large u will be of critical importance, and we study this in more detail.

The function $B^2(s,u)$ is non-negative from (5), and if it is zero we obtain a degenerate process which is analyzed in detail in section 3.3. Hence let us require that $B^2(s,u) > 0$.

Now define $\varphi(s,u)$ as follows

$$\varphi(s,u) = \frac{1}{\sqrt{2}} \int_a^u \frac{dh}{B(s,h)}$$

Then $\varphi(s,u)$ is monotone increasing for each fixed s and u increasing, and we are able to prove the following proposition: For continuity of the process it is necessary that we have

$$\text{(11)} \qquad \lim_{u \to +\infty} \varphi(s,u) = +\infty \quad , \quad \lim_{u \to -\infty} \varphi(s,u) = -\infty$$

We first show that the random variable $Y(s)$ defined by

$$Y(s) = \frac{1}{\sqrt{2}} \int_a^{u(s)} \frac{dh}{B(s,h)} = \varphi(s, u(s))$$

has, for the functions $A_1(s,y)$ and $B_1^2(s,y)$ associated with it,

$$A_1(s,y) = \frac{A(s,u)}{\sqrt{2}\, B(s,u)} \quad , \quad B_1^2(s,y) = 1.$$

For
$$\frac{Y(s+\Delta)-Y(s)}{\Delta} = \frac{1}{\Delta\sqrt{2}}\int_{\mathcal{U}(s)}^{\mathcal{U}(s+\Delta)} \frac{dh}{B(s,h)}$$

$$= \frac{1}{\Delta}\left(\mathcal{U}(s+\Delta)-\mathcal{U}(s)\right)\cdot\frac{1}{\sqrt{2}\,B(s,h)}$$

where $\mathcal{U}(s+\Delta) \geq h \geq \mathcal{U}(s)$, by the mean value theorem. Taking the expectation of both sides under the assumption that $\mathcal{U}(s)=u$, and noting that $y=\varphi(s)$ has a unique inverse,

$$E\left(\frac{Y(s+\Delta)-Y(s)}{\Delta}\,\middle|\,Y(s)=y\right) = E\left(\frac{\mathcal{U}(s+\Delta)-\mathcal{U}(s)}{\Delta}\cdot\frac{1}{\sqrt{2}\,B(s,h)}\,\middle|\,\mathcal{U}(s)=u\right).$$

On letting $\Delta\to 0$ we obtain

(12) $$A_1(s,y) = \frac{A(s,u)}{\sqrt{2}\,B(s,u)}$$

Similarly for $B_1^2(s,y)$;

$$\frac{1}{\Delta}\left(Y(s+\Delta)-Y(s)\right)^2 = \frac{1}{\Delta}\left(\mathcal{U}(s+\Delta)-\mathcal{U}(s)\right)^2\cdot\frac{1}{2B^2(s,h)}$$

$$B_1^2(s,y) = 1.$$

Now we consider a process wherein $A(s,u)=0$. Then the differential equation associated with the new random variable $Y(s)$ introduced above is

$$\frac{\partial f_1(s,x;t,y)}{\partial s} = \frac{\partial^2 f_1(s,x;t,y)}{\partial x^2}$$

from the Fokker-Planck equation (3). The only non negative solution to this equation, satisfying conditions is

(13) $$f_1(s,x;t,y) = \frac{1}{2\sqrt{\pi}\,\sqrt{t-s}}\,e^{-\frac{(y-x)^2}{4(t-s)}}$$

and this moreover satisfies (1) and (2). Reverting back to the initial variable u , the transition probability for the process $u(t)$ is

$$f(s,u;t,v) = \frac{1}{2\sqrt{\pi(t-s)}\,B(s,v)}\, e^{-\frac{(\varphi(t,v)-\varphi(s,u))^2}{4(t-s)}} .$$

Suppose now that $\lim_{u\to\infty} \varphi(s,u) = \psi(s)$ which is finite for some s . Then

$$f(s,\infty;t,v) = \frac{1}{2\sqrt{\pi(t-s)}\,B(s,v)}\, e^{-\frac{(\varphi(t,v)-\psi(s))^2}{4(t-s)}} > 0$$

and for all these s the probability of an infinite saltus in the process is finite during any time interval. Thus the process is not c. m. s. or c. i. p.

Feller (1) showed, by a particular example, that conditions (11) were indispensable in order that the general boundary value problem for the parabolic differential equations should have a unique solution. However the more stringent requirement that $A(s,u)$ and $B^2(s,u)$ should be bounded themselves was used in order to carry through a complete proof of existence and unicity.

Fortet (1) was forced to require that $A(s,u)$ and $\frac{\partial A}{\partial u}$ be bounded, as well as a two dimensional Lipschitz condition on $\frac{\partial A}{\partial u}$, in order for him to carry through an extensive treatment of the continuity properties for the process. He remarks that 'if $A(t,u)$ were infinitely large with u the nature of $\mathcal{U}(t)$ could be profoundly modified.' Some remarks of Bernstein (1) seem to support this, but the whole matter is far from being precise. In point of fact, processes for which $A(s, u) = O(u)$ do exist and are well behaved.

Fortet's elegant result, with the condition on $A(s,u)$ cited above and with $B^2(s, u) = 1$, is that $\mathcal{U}(t)$ is almost surely continuous. More precisely, for any fixed $c > 1$, and for no fixed $c < 1$ we almost always have

$$\frac{\mathcal{U}(t+\Delta) - \mathcal{U}(t)}{\sqrt{\Delta}} < 2c\sqrt{-\log \Delta}$$

provided $0 < \Delta < \kappa < 1$ where κ is independent of t, and $\mathcal{U}(t)$ is supposed fixed. In particular, this specifies the critical case of the Einstein process analyzed by Paley and Wiener for which the transition probability satisfies (13): $A = 0$, $B^2 = 1$.

It does, indeed, seem difficult to exactly specify the interdependence between $A(s,u)$ and the existence, uniqueness and continuity of the process. From the preceding argument it is clear that $B^2(s, u) = O(u^2)$ from conditions (11). If we consider the $Y(s)$ process above for which $B_1^2 = 1$, and use Fortet's result that $A_1(s, y)$ in this process must be bounded, formula (12) shows that $A(s, u) = O(u)$ for large u. In this case formula (10) shows that the process is c. m. s. whenever $h(s,u)$ has a finite variance. Since by conditions κ, $f(s,u; t,v)$ has moments of order 3 and hence likewise for $h(s,u)$, the processes discussed here are always c. m. s. Again considering the $Y(s)$ process for which $B_1^2 = 1$ and A_1 is bounded, Feller's analysis shows the process exists and is unique, and hence the same is true for the $U(s)$ process in which we have $A(s, u) = O(u)$ and $B^2(s, u) = O(u^2)$.

We confine our attention to this class of coefficients for which there is definite assurance of existence, uniqueness and continuity of the process. We shall, as a matter of fact, be able to conclude these properties for the first few moments of the distribution of the process so that in particular for a process whose unconditional distribution is specified by them (e.g. the normal distribution, Poisson distribution, etc.) we shall be able to conclude unicity anew.

As a final additional condition we sometimes restrict the class of coefficients to polynomials in u, so that

$$A(s, u) = \varphi_0(s) + \varphi_1(s)\, u \tag{14}$$

$$B^2(s, u) = \psi_0(s) + \psi_1(s)\, u + \psi_2(s)\, u^2 \tag{15}$$

where the functions ψ appearing in B^2 are supposed to be such that it is positive definite. This class of coefficients will include as special cases all the known processes.

2.2 The unconditional statistics

In this section we deduce expressions for the unconditional mean, variance and covariance for a process with any coefficients A, B^2 .

The random variable $\mathcal{U}(s)$ is specified by an unconditional distribution $H(s,u) = Pr\{\mathcal{U}(s) \leq u\}$. This distribution will have, like the transition probability $F(s,u;t,v)$ derivatives up to the fourth order, and for convenience we work with the density function

$$h(s,u) = \frac{\partial}{\partial u} H(s,u).$$

This function then satisfies the two fundamental relationships (8) and (9).

The unconditional mean and variance are given respectively by

$$(16) \qquad \begin{aligned} \mu(s) &= \int_{-\infty}^{\infty} u\, h(s,u)\, du \\ \sigma^2(s) &= \int_{-\infty}^{\infty} u^2 h(s,u)\, du - \mu^2(s) \end{aligned}$$

and since f has moments up to the third by conditions K , these two quantities will exist. In addition we study the covariance $C(s,t) = E(\mathcal{U}(s)\,\mathcal{U}(t))$ in order to specify the correlation properties of the process. We have

$$C(s,t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u\,v\, h(s,u) f(s,u;t,v)\, du\, dv.$$

In equation (16) we differentiate both sides with respect to s . The equation

$$\frac{d\mu(s)}{ds} = \int_{-\infty}^{\infty} u \frac{\partial h(s,u)}{\partial s}\, du$$

will be true, under conditions K, for those values of s for which the integral on the right converges uniformly.

We substitute for $\frac{\partial h}{\partial s}$ the right hand side of (9)

$$\frac{d\mu(s)}{ds} = \int_{-\infty}^{\infty} u\left\{-\frac{\partial}{\partial u}\left(h(s,u)A(s,u)\right)+\frac{\partial^2}{\partial u^2}\left(h(s,u)B^2(s,u)\right)\right\}du$$

and integrate by parts. If the integrated parts vanish at the limits of integration, the second term is zero, and finally

$$\frac{d\mu(s)}{ds} = \int_{-\infty}^{\infty} A(s,u)\, h(s,u)\, du .$$

The conditions that the integrated parts vanish are

(17)
$$\begin{aligned} |h(s,u)A(s,u)| &= o\left(\tfrac{1}{u}\right) \\ \left|\tfrac{\partial}{\partial u}\left(h(s,u)B^2(s,u)\right)\right| &= o\left(\tfrac{1}{u}\right) \end{aligned} \qquad u \longrightarrow \pm\infty$$

The first of these is surely met if we restrict $A(s,u)$ to the class of functions (14), since $h(s,u)$ is supposed to have moments of order three.

In an exactly similar manner we deduce that

$$\frac{d\sigma^2(s)}{ds} = 2\int_{-\infty}^{\infty} B^2(s,u)\,h(s,u)\,du + 2\int_{-\infty}^{\infty} u\,A(s,u)\,h(s,u)\,du - 2\mu(s)\int_{-\infty}^{\infty} A(s,u)\,h(s,u)\,du$$

$$\frac{\partial}{\partial t}\,c(s,t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u\,h(s,u)\,A(t,v)\,f(s,u;\,t,v)\,du\,dv$$

provided that

(18)
$$\begin{aligned} |h(s,u)A(s,u)| &= o\left(\tfrac{1}{u^2}\right) \\ \left|\tfrac{\partial}{\partial u}\left(h(s,u)B^2(s,u)\right)\right| &= o\left(\tfrac{1}{u^2}\right) \end{aligned} \qquad u \longrightarrow \pm\infty$$

The validity of these conditions imply the former ones (17).

It is important to note that these equations can be deduced without appealing to the rather strong conditions (18). From equation (6) we have

$$\int_{-\infty}^{\infty} A(s,u)\,h(s,u)\,du = \int_{-\infty}^{\infty} \lim_{\Delta\to 0} E\left(\frac{u(s+\Delta)-u(s)}{\Delta}\,\Big|\,u(s)=u\right) h(s,u)\,du \tag{19}$$

and the quantity under the expectation is supposed continuous at $\Delta = 0$, being $A(s,u)$ for that value; hence the whole integrand on the right is essentially dominated by $|A(s,u)|\,h(s,u)$, and if this is integrable we may interchange the integration and the limiting operations in (19), giving

$$\int_{-\infty}^{\infty} A(s,u)\,h(s,u)\,du = \lim_{\Delta\to 0} E\left(\frac{u(s+\Delta)-u(s)}{\Delta}\right)$$

$$= \lim_{\Delta\to 0} \frac{\mu(s+\Delta)-\mu(s)}{\Delta}$$

$$= \frac{d\mu(s)}{ds}$$

under the sole restriction that the integral on the left converges absolutely. The same argument shows that $\mu(s)$ is continuous and differentiable for values of s where the convergence is absolute. Thus with the class of coefficients $A(s,u)$ which we are studying, and under the conditions K it appears that the unconditional mean always exists and has a derivative.

In a manner analogous to that above $\frac{d\sigma^2(s)}{ds}$ and

$\frac{\partial}{\partial t} c(s,t)$ are shown to exist under weaker conditions, and collecting the results we conclude

(20) $$\frac{d\mu}{ds} = \int_{-\infty}^{\infty} A(s,u)\, h(s,u)\, du$$

(21) $$\frac{d\sigma^2}{ds} = \int_{-\infty}^{\infty} B^2(s,u)\, h(s,u)\, du + 2\int_{-\infty}^{\infty} uAh\, du - 2\mu \int_{-\infty}^{\infty} Ah\, du$$

(22) $$\frac{\partial}{\partial t} c(s,t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u\, h(s,u)\, A(t,v)\, f(s,u;t,v)\, du\, dv$$

these formulas being valid whenever the integrals in question converge absolutely, or conditions (18) are satisfied.

If we substitute for A and B^2 the class of coefficients given by (14) we obtain the following set of simultaneous differential equations to determine μ, σ^2, and $c(s,t)$:

(23) $$\frac{d\mu(s)}{ds} = \varphi_0(s) + \varphi_1(s)\mu(s)$$

(24) $$\frac{1}{2}\frac{d\sigma^2(s)}{ds} = \psi_0(s) + \psi_1(s)\mu(s) + \psi_2(s)\, c(s,s) + \varphi_1(s)\,\sigma^2(s)$$

(25) $$\frac{\partial}{\partial t} c(s,t) = \varphi_0(t)\mu(s) + \varphi_1(t)\, c(s,t).$$

A boundary condition on the last equation is plainly $c(s,s) = \sigma^2(s) + \mu^2(s)$, and as soon as we know the initial distribution $h(s_0,u)$ everything is determined. It is, as a matter of fact, possible to solve the above system completely in quadratures.

2.3 The Gaussian distribution

In this section we deduce the class of functions A

and B^2 leading to the Gaussian distribution and show it is included in the class we have under discussion. Specifically we show that a necessary and sufficient condition that $h(s,u)$ be normally distributed is

$$A(s,u) = \varphi_0(s) + \varphi_1(s)\,u$$
$$B^2(s,u) = \psi_0(s)$$

The proof is by characteristic functions.

Suppose that $h(s,u)$ is normally distributed. A necessary and sufficient condition for this is that its characteristic function $\xi(s,p)$ is

$$\xi(s,p) = \int_{-\infty}^{\infty} e^{iup} h(s,u)\,du = e^{i\mu p - \frac{\sigma^2}{2}p^2}$$
$$\mu = \mu(s)\ , \quad \sigma^2 = \sigma^2(s)$$

On account of the exponential structure of $h(s,u)$ we can surely differentiate under the integral sign,

$$\frac{\partial \xi(s,p)}{\partial s} = \int_{-\infty}^{\infty} e^{iup} \frac{\partial}{\partial s} h(s,u)\,du = \xi\left(i\mu' p - \frac{\sigma^{2\prime}}{2} p^2\right).$$

Now replace $\frac{\partial h}{\partial s}$ by the right hand side of (9) and integrate by parts, following the same device used previously. If A and B^2 do not increase faster than any power of u, the integrated parts will vanish. This gives

(26) $$\frac{\partial \xi}{\partial s} = \int_{-\infty}^{\infty} e^{iup}\left\{-\frac{\partial}{\partial u}(Ah) + \frac{\partial^2}{\partial u^2}(B^2 h)\right\} du$$

(27) $$= ip\int_{-\infty}^{\infty} e^{iup} A h\,du - p^2 \int_{-\infty}^{\infty} e^{iup} B^2 h\,du .$$

Under conditions K, A and B^2 can be expanded in a few terms of a Taylor's series

(28)
$$A(s,u) = \varphi_0(s) + \varphi_1(s)\,u + \varphi_2(s)\,u^2 + \cdots$$
$$B^2(s,u) = \psi_0(s) + \psi_1(s)\,u^2 + \psi_2(s)\,u^3 + \cdots$$

and we note that for the derivatives of ξ with respect to p we obtain

$$\xi = \int_{-\infty}^{\infty} e^{iup} h\,du = e^{i\mu p - \frac{\sigma^2}{2}p^2}$$

$$\frac{\partial \xi}{\partial p} = i\int_{-\infty}^{\infty} u\,e^{iup} h\,du = \xi(i\mu - \sigma^2 p)$$

$$\frac{\partial^2 \xi}{\partial p^2} = i^2 \int_{-\infty}^{\infty} u^2 e^{iup} h\,du = \xi(-\mu^2 - \sigma^2 - 2i\mu\sigma^2 p + \sigma^4 p^2)$$

- - - - - - - - - - - - - - - - - - -

On substituting the expressions (28) into (27) we get for $\frac{\partial \xi}{\partial s}$, by virtue of the preceding derivatives,

$$\frac{\partial \xi}{\partial s} = ip\varphi_0\xi + p\varphi_1\xi(i\mu - \sigma^2 p) - ip\varphi_2\xi(-\mu^2 - \sigma^2 - 2i\mu\sigma^2 p + \sigma^4 p^2) + \cdots$$
$$- p^2\psi_0\xi + ip^2\psi_1\xi(i\mu - \sigma^2 p) + p^2\psi_2\xi(-\mu^2 - \sigma^2 - 2i\mu p + \sigma^4 p^2) + \cdots.$$

Now equate this to the preceding expression for

$$\frac{\partial \xi}{\partial s} = \xi\left(i\mu' p - \frac{\sigma^{2\prime}}{2}p^2\right)$$

eliminating the factor ξ ;

$$ip(\varphi_0 + \varphi_1\mu + \varphi_2(\mu^2 + \sigma^2) + \cdots) - p^2(\varphi_1\sigma^2 + \psi_0 + 2\varphi_2\mu\sigma^2 - \psi_1\mu + \psi_2\mu^2 + \cdots)$$
$$+ ip^3(\varphi_2\sigma^4 + \psi_1\sigma^2 - 2\psi_2\mu\sigma^2 + \cdots) + \cdots = ip\mu' - \frac{\sigma^{2\prime}}{2}p^2$$

Equating powers of p it is plain that we must have

$$\begin{cases} \varphi_2 = \varphi_3 = \cdots = 0 \\ \psi_1 = \psi_2 = \cdots = 0 \end{cases}$$

(29)
$$\begin{cases} \mu' = \varphi_0 + \varphi_1 \mu \\ \frac{\sigma^{2\prime}}{2} = \psi_0 + \varphi_1 \sigma^2 \end{cases}$$

and that

(30)
$$\begin{cases} A(s, u) = \varphi_0(s) + \varphi_1(s)\, u \\ B^2(s, u) = \psi_0(s) \end{cases}$$

Moreover the differential equations (29) agree with those previously found, (23) and (24), when A and B^2 are as above.

For the special case of a stable distribution ($h(s, u)$ depends only on u) this theorem was effectively proved by Doob (1) who considered only the case $\varphi_0 = 0$, $\varphi_1 = -\beta$, $\psi_0 = -\beta/2\, \sigma_0^2$, $\beta > 0$. For the case of the transition probabilities an analogous result was proved by Feldheim (1).

2.4 Some Special Processes

By specializing the coefficients in (14) and (15) we can study the unconditional properties of all the known processes.

(a) The Bachelier process. Bachelier (1) (2) in his studies on insurance and investments was led to consider the general parametrically homogeneous process.

Here we have $A = \varphi_0(s)$, $B^2 = \psi_0(s)$ and $h(s,u)$ satisfies

$$\frac{\partial h}{\partial s} = -\varphi_0(s)\frac{\partial h}{\partial u} + \psi_0(s)\frac{\partial^2 h}{\partial u^2}$$

Bachelier actually discovered this equation by intuitive reasoning from heat conduction analogies.

By the results of the preceding section the Bachelier process is always Gaussian and will thus be specified by its first two moments which we now calculate.

Equations (23), (24), and (25) become

$$\frac{d\mu(s)}{ds} = \varphi_0(s)$$

$$\frac{1}{2}\frac{d\sigma^2(s)}{ds} = \psi_0(s)$$

$$\frac{\partial}{\partial t} c(s,t) = \mu(s)\varphi_0(s)$$

$$c(s,s) = \sigma^2(s) + \mu^2(s)$$

and their solution is

$$\mu(s) = \int_{s_0}^{s} \varphi_0(y)\,dy + \mu(s_0)$$

$$\sigma^2(s) = 2\int_{s_0}^{s} \psi_0(y)\,dy + \sigma^2(s_0)$$

$$c(s,t) = \mu(s)\mu(t) + \sigma^2(s) \quad , \quad t \geq s$$

which give the complete properties of the process

$$h(s,u) = \frac{1}{2\sqrt{\pi}\,\sigma(s)}\, e^{-\frac{(u-\mu(s))^2}{4\sigma^2(s)}} .$$

It is supposed, of course, that the distribution is known at the initial time s_0, but as a matter of fact it is sufficient to merely know the first two moments of $h(s_0, u)$, i.e. $\mu(s_0)$ and $\sigma^2(s_0)$, in order to specify the process.

In applications, particularly in the study of time series, harmonic analysis, etc., the correlation between $\mathcal{U}(t)$ and $\mathcal{U}(s)$ is important. This is easily computed from the above quantities, and we have, by definition

$$R(s,t) = \frac{c(s,t) - \mu(s)\mu(t)}{\sigma(s)\,\sigma(t)} .$$

Substituting the calculated values into this expression we obtain

$$R(s,t) = \frac{\sigma(s)}{\sigma(t)} , \quad t \geqslant s$$

$$= \frac{\left(\int_{s_0}^{s} \psi_0(y)\, dy + \sigma^2(s_0) \right)^{1/2}}{\left(\int_{s_0}^{t} \psi_0(y)\, dy + \sigma^2(s_0) \right)^{1/2}}$$

and the correlation is always positive. It is remarkable that this is independent of $\varphi_0(t)$.

It is seen that the only case in which, for $s \to \infty$, we get a limiting distribution is when $\int_{s_0}^{\infty} \psi_0(y)\, dy$ converges, for otherwise we should have a normal distribution with infinite variance. But in this case the limiting distribution is not independent of the initial distribution, for if $\int_{s_0}^{\infty} \psi_0(y)\, dy = K$, we have $R(s_0, \infty) = \frac{1}{\sqrt{K}} > 0$.

This case stands in distinction to those processes where $h(s, u) \longrightarrow h(u)$ for $s \to \infty$ and $h(u)$ is a distribution function independent of the variable $\mathcal{U}(s_0)$. Such a limiting distribution is called a stable (or stationary) process.

A special case of the Bachelier process is the Einstein (1) process, which is both temporally and parametrically homogeneous. In this case A and B^2 are constants which we take to be 0 and 1 respectively. We also put $\mu(s_0) = 0$ and $\sigma^2(s_0) = 0$ so that $h(s_0, u)$ is the singular distribution; i.e. $\mathcal{U}(s_0) = 0$ with probability 1. Then from the above

$$h(s, u) = \frac{1}{2\sqrt{\pi (s - s_0)}} e^{-\frac{u^2}{4(s-s_0)}}, \quad s > s_0$$

$$R(s, t) = \sqrt{\frac{s - s_0}{t - s_0}}, \quad t \geq s > s_0 .$$

This process, which is the earliest known example, is supposed to describe the Brownian motion of a particle in the absence of friction and inertia, and we return to it in section III.

(b) The Ornstein-Uhlenbeck process. Ornstein and Uhlenbeck (1) discussed the process which bears their name, after Doob (1). The process appears, however, to go back to Smoluchowski (1) who considered it in connection with problems in diffusion. The process is the most general Gaussian process, and the coefficients are given by (30). The above writers, however, treated only the stable distribution. For the general process the equation of the statistics are

(31) $$\frac{d\mu(s)}{ds} = \varphi_0(s) + \varphi_1(s)\mu$$

(32) $$\frac{1}{2}\frac{d\sigma^2(s)}{ds} = \psi_0(s) + \varphi_1(s)\sigma^2$$

(33) $$\frac{\partial c(s,t)}{\partial t} = \varphi_0(t)\mu(s) + \varphi_1(t)\, c(s,t)$$

using (23), (24), and (25). These equations are easily solved

$$\mu(s) = e^{\int_{s_0}^{s}\varphi_1(y)dy}\left(\mu(s_0) + \int_{s_0}^{s}\varphi_0(x)e^{-\int_{s_0}^{x}\varphi_1(y)dy}dx\right)$$

$$\frac{1}{2}\sigma^2(s) = e^{2\int_{s_0}^{s}\varphi_1(y)dy}\left(\sigma^2(s_0) + \int_{s_0}^{s}\psi_0(x)e^{-2\int_{s_0}^{x}\varphi_1(y)dy}dx\right)$$

$$c(s,t) = e^{\int_{s}^{t}\varphi_1(y)dy}\left(c(s,s) + \mu(s)\int_{s}^{t}\varphi_0(x)e^{-\int_{s}^{x}\varphi_1(y)dy}dx\right)$$

where $c(s,s) = \sigma^2(s) + \mu^2(s)$ as before. Since this is a Gaussian distribution, $h(s,u)$ is completely determined by these quantities. It also might be mentioned, in passing, that the transition probability $f(s,u;t,v)$ is also uniquely determined since it is a bivariate conditional normal distribution, and we know the correlation from $c(s,t)$.

When φ_0, $\varphi_1 \neq 0$, and $\psi_0 > 0$ are constants, we arrive at the temporally homogeneous Gaussian process discussed at length by Doob (1) and (2). This is supposed to represent the velocity distribution of a particle under Brownian motion and under the influence of dissipative Doppler forces. Its (transitional) properties were deduced by Ornstein and Uhlenbeck (1) starting from a stochastic differential equation governing the motion, known as the Langevin equation. All of the salient properties may be deduced from the above equations very easily.

It is customary to require that the variance $\sigma^2(s)$ be a constant $= \sigma_0^2$, and we put $\varphi_0 = 0$, $\varphi_1 = -\beta$. Then (31), (32), and (33) give

$$
\begin{aligned}
\frac{d\mu}{ds} &= -\beta\mu \\
0 &= \psi_0 - \beta\sigma^2 \qquad (34) \\
\frac{\partial}{\partial t} c(s,t) &= -\beta\, c(s,t)
\end{aligned}
$$

and since $\beta^2 = \psi_0 > 0$ we must have $\beta > 0$ from (34). Solving these equations we get

$$
\begin{aligned}
\mu(s) &= \mu(s_0)\, e^{-\beta(s-s_0)} \\
\sigma^2(s) &= \sigma_0^2 \\
c(s,t) &= e^{-\beta(t-s)}\left(\sigma_0^2 + \mu^2(s)\right), \quad t > s
\end{aligned}
$$

and for the correlation

$$
\begin{aligned}
R(s,t) &= \frac{e^{-\beta(t-s)}\left(\sigma_0^2 + \mu^2(s) - \mu(s)\mu(t)\right)}{\sigma_0^2} \\
&= e^{-\beta(t-s)}, \qquad t \geq s.
\end{aligned}
$$

As $t \to \infty$, s and s_0 fixed, $R(s,t) \to 0$, $\mu(s) \to 0$ and $\sigma^2(s) = \sigma_0^2$, and hence the limiting distribution does exist,

$$\lim_{s \to \infty} h(s,u) = \frac{1}{\sqrt{2\pi}\,\sigma_0} e^{-\frac{u^2}{2\sigma_0^2}}$$

corresponding to the Maxwellian equipartition law for molecular motion. This is an example of a stationary process.

(c) <u>A third example.</u> As a final example we consider the temporally homogeneous process in which

$$\varphi_0 = 0 \quad , \quad \varphi_1 = \beta$$

$$\psi_0 = \psi_1 = 0 \quad , \quad \psi_2 = \alpha$$

so that

$$A(s,u) = \beta u$$

$$B^2(s,u) = \alpha u^2$$

Then the unconditional mean, variance, and covariance are given by

$$\frac{d\mu}{ds} = \beta\mu$$

$$\frac{1}{2}\frac{d\sigma^2}{ds} = \alpha\left(\sigma^2(s) + \mu^2(s)\right)$$

$$\frac{\partial c(s,t)}{\partial t} = \beta\, c(s,t)$$

of which the solution is

$$\mu(s) = \mu(s_0)\, e^{\beta(s-s_0)}$$

$$\frac{1}{2}\sigma^2(s) = e^{2\alpha(s-s_0)}\sigma^2(s_0) + \frac{\mu^2(s_0)}{2}\cdot\frac{\alpha}{\alpha-\beta}\left(e^{2\alpha(s-s_0)} - e^{2\beta(s-s_0)}\right), \quad \beta \neq \alpha$$

$$= e^{2\alpha(s-s_0)}\left(\sigma^2(s_0) + \alpha\mu^2(s_0)(s-s_0)\right) \quad , \quad \beta = \alpha$$

$$C(s,t) = C(s,s)\, e^{\beta(t-s)}$$

$$R(s,t) = \frac{\sigma(s)}{\sigma(t)} \cdot e^{\beta(t-s)}$$

and if, in particular $-\beta = \alpha > 0$, $\sigma^2(s)$ becomes

$$\tfrac{1}{2}\sigma^2(s) = e^{2\alpha(s-s_0)}\sigma^2(s_0) + \frac{\mu^2(s_0)}{2} \sinh 2\alpha(s-s_0).$$

Now suppose initially that $\mu^2(s_0) + \sigma^2(s_0) \neq 0$. Then for $t \to \infty$, $\sigma^2(t) \to \infty$ and $R(s,t) \to 0$ with s fixed. Hence a stable distribution (having a finite variance and thus satisfying conditions K) does not exist. In view of the two preceding examples it appears that may or may not have a limiting value of zero when the limiting distribution fails to exist, and its behavior gives no information as to the existence of limiting distributions. However, in certain processes it may do so, as shown by Itô (1).

A remarkable feature of this process is that if $\mu(s_0)$ and $\sigma^2(s_0)$ are both zero; i.e. $\mathcal{U}(s_0) = 0$ with probability 1, then $\mu(s)$ and $\sigma^2(s)$ remain zero for all s --that is, $\mathcal{U}(s) = 0$ with probability 1 everywhere. This state of affairs thus shows that if the process once in its history attained the value zero, it would remain at this value thereafter. Now if the process does not start off at $u = 0$ its limiting variance is infinite, and we conclude that the probability of the variable $\mathcal{U}(t)$ having a zero in any finite t interval is zero.

We can deduce the distribution of this process quite easily, for making the transformation given previously, it is seen that the variable $y(s) = \frac{1}{\sqrt{2\alpha}} \log \mathcal{U}(s)$ is a Bachelier process whose distribution we have already deduced. We get then

$$h(s, u) = \frac{1}{2\sqrt{\pi}\,\sigma(s)\sqrt{2\alpha}\; u}\, e^{-\frac{\left(\frac{1}{\sqrt{2\alpha}}\log u - \mu(s)\right)^2}{4\sigma^2(s)}}$$

wherein the anomalies of the process at $u=0$ are manifest. The origin of this singular behavior goes back, of course, to the fact that $\beta^2(s,u) = 0$ when $u = 0$.

The usual way of discussing the properties of a process is to actually solve the original Fokker-Planck equations for the transitional probabilities, use these to find the means, correlations, etc. The ideas outlined in this section are seen to give the characteristics directly, however.

III

The Integral Process

3.1 The integral of a random process

In the theory of discrete random variables an important question is whether the law of large numbers is satisfied; i.e. whether the mean value of the sum of a large number of variables approaches the sum of the mean values in a probability sense. In the case of independent variables this question has been completely solved (see Cramer (1)); and in the theory of Markoff chains certain results are known.

Since we have seen that the stochastic processes considered here are (almost always) continuous, we should be able to arrive at an analogue to the ordinary Riemannian integral, and the corresponding questions about the law of large numbers should have a counterpart for the integrals. Also in certain physical considerations the idea is important--thus the temporally homogeneous Gaussian process is supposed to represent the velocity distribution of Brownian particles, and the integral process will then give the distribution of displacements of the particles.

Suppose we are given a random variable $u(t)$ which forms a continuous Markoff process in $a \leq t \leq b$. We subdivide the interval into n parts $a = t_0 < t_1 < \cdots < t_n = b$ and choose a point τ_i in the i^{th} interval at random (e.g. with a rectangular law of distribution) and form the sum

$$S = \sum_{i=1}^{n} \mathcal{U}(\tau_i)(t_i - t_{i-1})$$

Then S is a random variable depending not only on the mode of subdivision and the points τ_i chosen, but upon the random variable $\mathcal{U}(t)$ itself. If there exists a random variable J such that for any given $\delta > 0$, $\varepsilon > 0$, an η can be found for which $\max_{(i)} |t_i - t_{i-1}| < \eta$ will imply

(35) $$P_r \{ |S - J| \geq \varepsilon \} < \delta$$

then the random variable J is called the 'integral in probability' of $\mathcal{U}(t)$, and we write

$$J = \int_a^b \mathcal{U}(t)\, dt$$

The process $\mathcal{U}(t)$ is then called 'integrable in probability,' and it is clear how we should define integrable in the mean (or in the mean square). The expression (35) would be simply $E(|S-J|) < \delta$ (or $E((S-J)^2)^{1/2} < \delta$). As before, these will imply integrability in probability.

The principal result proven by Slutsky (2) is that a process continuous in the mean is integrable in probability (over a finite range). See Fan (1) for a detailed account.

In this section we consider the integral

$$X(s) - X(s_0) = \int_{s_0}^{s} \mathcal{U}(y)\,dy$$

where $\mathcal{U}(y)$ is a Markoff process considered in the preceding section. We study the properties of $X(t)$, showing it is again a Markoff process, and characterize it by the functions $A(s,u)$ and $B^2(s,u)$ associated with the process $\mathcal{U}(s)$. In order to attain a non-trivial process $X(t)$ an important modification must be made in the integral. This is almost a priori evident, for as previously shown, the general Markoff process $\mathcal{U}(t)$ has a derivative almost nowhere, whereas $X(t)$, being an integral process, presumably has a derivative.

3.2 <u>Example of an integral process</u>

Before embarking on the general theory, the integral of a special process is analyzed by the finite Darboux sum limit. This will illustrate some of the mathematical properties and the attendant physical implications more clearly than a purely analytical approach. This will also clearly exemplify the nature of the necessary modification mentioned in the preceding paragraph.

Suppose we have given a set of random variables $\mathcal{U}(t)$ defined for all $t \geq 0$, mutually independent and identically distributed. For convenience we let their common mean and variance be 0 and $\sigma^2 < \infty$ respectively. Also, let their distribution function be $F(u)$ and characteristic function be $\varphi(s)$

$$F(u) = P_r\{\mathcal{U}(t) \le u\}$$

$$\varphi(s) = \int_{-\infty}^{\infty} e^{isu}\, dF(u)$$

Now $\mathcal{U}(t+s)$ and $\mathcal{U}(t)$ are supposed independent, and hence the process $\mathcal{U}(t)$ will present a totally discontinuous function of t. Nevertheless, in a formal way, we try to find the 'integral' of $\mathcal{U}(t)$. To this end we put

$$t_j = j\,h \quad , \quad j = 0, 1, 2, \cdots$$
$$h = \text{const.} > 0$$

and consider the random variable $X(t_n)$

(36) $$X(t_n) = \sum_{j=0}^{n}{}' \, \mathcal{U}(t_j)\,h + X(0)$$

Now the process $X(t_k)$ is a discrete time Markoff process, for $X(t_{k+1}) = X(t_k) + \mathcal{U}(t_{k+1})\,h$, and thus the properties of $X(t_{k+1})$, if we are given values of $X(t_k)$, $X(t_{k-1})$ etc. depend only on $X(t_k)$ plus an uncorrelated increment $\mathcal{U}(t_{k+1})\,h$. Thus the possibility of attaining a continuous Markoff process, by allowing $h \to 0$, presents itself.

Let $\Phi(t_n, s)$ be the characteristic function for the variable $X(t_n)$. By the addition theorem for independent random variables

(37) $$\Phi(t_n, s) = (\varphi(sh))^n \Phi(0, s)$$

and for convenience we put $\Pr\{X(0) = 0\} = 1$ so that $\Phi(0, s) = 1$. By hypothesis $\mathcal{U}(t)$ has zero mean and finite variance σ^2, hence $\varphi(t)$ can be expanded in a few terms in a power series, giving

$$\Phi(t_n, s) = \left(1 - \frac{\sigma^2 s^2 h^2}{2} + o(s^2 h^2)\right)^n$$

To arrive at a continuous process we allow $h \to 0$ and $n \to \infty$ in such a way that $nh = t_n \to t$, a continuous variable. Then we get

(38) $$\Phi(t, s) = \lim_{h \to 0} \left(1 - \frac{\sigma^2 s^2 h^2}{2} + o(s^2 h^2)\right)^{\frac{t}{h}}$$

$$= 1$$

uniformly in $|s| \leq R < \infty$. Then $\Pr\{X(t) = 0\} = 1$ for all t so that the variable $X(t)$ has no variation, being always equal to its initial value zero. This stands in complete accordance with the law of large numbers, for we have essentially found the mean value of an unlimited number of random variables each having mean zero.

However instead of defining $X(t_n)$ as in (36) let us put

$$X(t_n) = \sum_{j=1}^{n}{}' \mathcal{U}(t_j)\sqrt{h} + X(0)$$

then the above expression (38) for $\Phi(t,s)$ becomes

$$\Phi(t,s) = \lim_{h \to 0} \left(1 - \frac{\sigma^2 s^2 h}{2} + o(s^2 h)\right)^{\frac{t}{h}}$$

$$= e^{-\frac{s^2\sigma^2 t}{2}}$$

and hence the density distribution for $X(t)$ is

$$f(t,x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-isx - \frac{s^2\sigma^2 t}{2}}\, ds$$

(39) $$= \frac{1}{\sqrt{2\pi t}\,\sigma} e^{-\frac{x^2}{2\sigma^2 t}}$$

which is the fundamental Einstein process.

This procedure clearly exemplifies why the process has no derivative, and in the ordinary sense is not the integral of a random process. The equation (39) is simply defined as the continuous stochastic process connected with the uncorrelated impulse process $\mathcal{U}(t)$, without reference to its mode of origin as the above pseudo integral.

We see that the variable $X(t)$ satisfies a Lipschitz condition

(40) $$X(t+h) - X(t) = O(\sqrt{h})$$

in agreement with the result of Paley and Wiener (1) noted before. This fact on the order of $X(t+h) - X(t)$ has received much attention from various authors in the past.

Doob (2) calls it 'either an imperfection in nature or in the mathematics,' depending on one's viewpoint, while Khintchine (2) terms it a 'continuous idealization' of the natural phenomenon. In any event, no matter how the rationalization is attempted, the non existance of the derivative must be regarded as a serious defect in the statistical description of continuous random phenomena.

The derivation of the elementary stochastic process by the various writers is not done by the above method. Instead an artifice is employed which effectively masks the true physical picture. This is done as follows: It is noted that $t_{n+m} = t_n + t_m$, and from (37) we have

$$\Phi(t_{n+m}, s) = \Phi(t_n, s)\,\Phi(t_m, s)$$

$$\Phi(t_n + t_m, s) = \Phi(t_n, s)\,\Phi(t_m, s)$$

for all n and m. Now replace the discrete variable t_n by a continuous variable τ, and require the same functional equation to hold for it:

$$\Phi(\tau_1 + \tau_2, s) = \Phi(\tau_1, s)\,\Phi(\tau_2, s)$$

The solution of this equation is

$$\Phi(\tau, s) = e^{\tau g(s)}$$

and it remains only to find $g(s)$. It is shown by Kolmogoroff (2) and others that

$$g(s) = -\frac{s^2\sigma^2}{2} + \Xi(s)$$

where

$$\Xi(s) = \int_{-\infty}^{\infty} \frac{e^{isx} - 1 - isx}{x^2} \, d\Omega(x)$$

$\Omega(x)$ being bounded, monotone, and continuous at $x = 0$. If the process is to be continuous, however, $\Xi(s) = 0$ and the two derivations for $\Phi(t,s)$ agree--see Cramer (1). Also the case when the initial distribution $F(u)$ may have an infinite variance is treated slightly by Levy (1).

The general integral process is shown, in the following section, to have the same properties as the integral in the above process, i.e. it will have zero variance. In order to arrive at a non trivial result we shall effectively have to employ the pseudo integral which has an increment $\sqrt{h}$ instead of h itself. This is exemplified by actually finding the integral of the Einstein process $X(t)$ above. Now $X(t)$ is continuous, but by (40) not of bounded variation in any t interval. Paley and Wiener have shown, nevertheless, that its (Stieltjes) integral does exist.

Again proceeding formally, define $Z(t_n)$ by

(41)
$$Z(t_n) = \sum_{K=0}^{n}{}' X(t_K)\sqrt{h} + Z(0)$$

$$= \sum_{K=0}^{n}{}' \sum_{j=0}^{K}{}' U(t_j)\,h + Z(0)$$

then $Z(t_n)$ is a discrete time non-Markoff process, for $Z(t_{n+1}) = 2Z(t_n) - Z(t_{n-1}) + U(t_{n+1})h$ and the dependence goes back to the two preceding observations. Let $\Phi_1(t_n, s)$ be the characteristic function for $Z(t_n)$ and choose $\Phi_1(0, s) = 1$ as before. Then

$$\Phi_1(t_n, s) = \prod_{K=0}^{n} \prod_{J=0}^{K} \varphi(sh)$$

$$= (\varphi(sh))^{\frac{n(n+1)}{2}}$$

$$= \left(1 - \frac{s^2\sigma^2h^2}{2} + o(s^2h^2)\right)^{\frac{n^2}{2} + \frac{n}{2}}$$

Again let $h \to 0$, $n \to \infty$, and $nh = t_n \to t$, giving

$$\Phi_1(t, s) = \lim_{h \to 0} \left(1 - \frac{s^2\sigma^2h^2}{2} + o(s^2h^2)\right)^{\frac{t^2}{2h^2} + \frac{t}{2h}}$$

$$= e^{-\frac{\sigma^2 s^2}{2} \cdot \frac{t^2}{2}}$$

and this has for a distribution function

$$f(t, x) = \frac{1}{\sqrt{\pi}\, t\sigma} e^{-\frac{x^2}{t^2\sigma^2}}$$

In terms of the 'integrated' variable $X(t)$ the $Z(t)$ process is again seen to have no derivative

$$Z(t+h) - Z(t) = X(t+h)\, O(\sqrt{h})$$

In equation (41) it was necessary to have the time increment $\sqrt{h}$ again instead of h, for we should have again obtained a variable with zero variance--hence the $X(t)$ process has no true non-trivial integral. An interesting question arises as to whether the variable $Z(t)$, being a limit of a discrete non-Markoff process, is a continuous non-Markoff process. This question is answered negatively in the next section.

3.3 The integral formulas

Starting with the fundamental formula

$$X(s) - X(s_0) = \int_{s_0}^{s} \mathcal{U}(t)\,dt \tag{42}$$

which was developed in section 3.1, we study the process $X(s)$, and attempt to characterize it by the properties of $\mathcal{U}(t)$, which we presume to be a continuous stochastic process associated with the functions $A(s,u)$ and $B^2(s,u)$.

We first show the $X(s)$ process is a Markoff process. Suppose that we know $X(s_1)=x_1$, $X(s_2)=x_2$ and that $s_1 < s_2 < s_3$ and let us calculate $E(X(s_3) \mid X(s_1)=x_1, X(s_2)=x_2)$. From (42)

$$X(s_3) = \int_{s_2}^{s_3} \mathcal{U}(t)\,dt + X(s_2)$$

and we can calculate the expectation under the integration sign by Fubini's theorem

$$\begin{aligned} &E\left(X(s_3) \mid X(s_1)=x_1,\ X(s_2)=x_2\right) \\ &\qquad = x_2 + \int_{s_2}^{s_3} E\left(\mathcal{U}(t) \mid X(s_1)=x_1,\ X(s_2)=x_2\right) dt \end{aligned} \tag{43}$$

Now $\mathcal{U}(t)$ is a Markoff process and hence

$$E\left(\mathcal{U}(t) \mid X(s_1)=x_1,\, X(s_2)=x_2\right) = E\left(\mathcal{U}(t) \mid X(s_2)=x_2\right)$$

Thus the left hand side of (43) is independent of the fact that $X(s_1)=x_1$, and the $X(s)$ process is accordingly a Markoff process.

Since $\mathcal{U}(t)$ has been proven c. m. s., it is integrable in probability, at least, by Slutsky's theorem. We have

(43) $$\frac{X(s+\Delta)-X(s)}{\Delta} = \frac{1}{\Delta}\int_s^{s+\Delta} \mathcal{U}(t)\,dt$$

and we use this formula in finding the functions $A_1(s,x)$ and $B_1^2(s,x)$ associated with the process $X(s)$. Taking the expectation of both sides of (43) under the condition that $X(s)=x$ we find

$$E\left(\frac{X(s+\Delta)-X(s)}{\Delta} \,\Big|\, X(s)=x\right) = \frac{1}{\Delta}\int_s^{s+\Delta} E\left(\mathcal{U}(t) \mid X(s)=x\right)dt$$

and we obtain, on passing to the limit,

(44) $$A_1(s,x) = E\left(\mathcal{U}(s) \mid X(s)=x\right)$$

Likewise

$$\frac{\left(X(s+\Delta)-X(s)\right)^2}{\Delta} = \frac{1}{\Delta}\int_s^{s+\Delta}\int_s^{s+\Delta} \mathcal{U}(t)\,\mathcal{U}(y)\,dt\,dy$$

(45) $$E\left(\frac{\left(X(s+\Delta)-X(s)\right)^2}{\Delta} \,\Big|\, X(s)=x\right)$$

$$= \frac{1}{\Delta}\int_s^{s+\Delta}\int_s^{s+\Delta} E\left(\mathcal{U}(t)\,\mathcal{U}(y) \mid X(s)=x\right)dt\,dy$$

and on passing to the limit

$$B_1^2(s, x) \equiv 0.$$

Thus the differential dispersion of the integral process vanishes, and we study the consequences of this.

Assume that for a process $X(s)$ we have $B^2(s,x) \equiv 0$. Choose any bounded, single valued, non-constant function $\psi(s, x)$ satisfying the equation

$$\frac{\partial \psi(s,x)}{\partial s} + A(s,x)\,\frac{\partial \psi(s,x)}{\partial x}.$$

This is manifestly possible, since if ψ is a solution, so is $c_1 e^{c_2 \psi}$. We assume $A \neq 0$, and if, for instance $A = a_1(s)\,a_2(x)$ we could choose

$$\psi(s,x) = \int_{K_1}^{s} a_1(z)\,dz + \int_{K_2}^{x} \frac{dw}{a_2(w)}.$$

Consider the random variable $Y(s)$ defined by

$$Y(s) = \psi(s, X(s)).$$

We now show that the coefficients connected with the process $Y(s)$ vanish identically: $A_1(s,y) \equiv 0$, $B_1^2(s,y) \equiv 0$. Following a device used previously

$$\frac{Y(s+\Delta) - Y(s)}{\Delta} = \frac{\psi(s+\Delta, X(s+\Delta)) - \psi(s, X(s))}{\Delta}$$

$$= \frac{\Delta \frac{\partial \psi}{\partial s} + (X(s+\Delta) - X(s))\frac{\partial \psi}{\partial x} + \Delta^2 \frac{\partial^2 \psi}{\partial s^2}}{\Delta}$$

$$+ \frac{\Delta(X(s+\Delta) - X(s))\frac{\partial^2 \psi}{\partial s \partial x} + (X(s+\Delta) - X(s))^2 \frac{\partial^2 \psi}{\partial x^2}}{\Delta}$$

$$+ o(1)$$

$$\frac{Y(s+\Delta) - Y(s)}{\Delta} = \frac{\partial \psi}{\partial s} + \frac{X(s+\Delta) - X(s)}{\Delta}\frac{\partial \psi}{\partial x}$$

$$+ \frac{(X(s+\Delta) - X(s))^2}{\Delta}\frac{\partial^2 \psi}{\partial x^2} + o(1)$$

Now take the expectation of both sides under the assumption that $Y(s) = y = \psi(s,x) = \psi(s, X(s))$; $x = X(s)$

$$E\left(\frac{Y(s+\Delta) - Y(s)}{\Delta} \middle| Y(s) = y\right) = \frac{\partial \psi}{\partial s} + E\left(\frac{X(s+\Delta) - X(s)}{\Delta} \middle| X(s) = x\right)\frac{\partial \psi}{\partial x} + o(1)$$

the remaining terms being small since $B^2(s,x) = 0$.

Letting $\Delta \to 0$ we finally have

$$A_1(s,y) = \frac{\partial \psi}{\partial s} + A(s,x)\frac{\partial \psi}{\partial x}$$

$$= 0.$$

In an exactly similar manner

$$\frac{(Y(s+\Delta) - Y(s))^2}{\Delta} = \frac{(X(s+\Delta) - X(s))^2}{\Delta}\left(\frac{\partial \psi}{\partial x}\right)^2 + o(1)$$

the order of the remaining terms being at greatest $(X(s+\Delta) - X(s))$ or Δ, both of which go to zero. Taking the expectation of both sides under the same conditions as previously, and letting $\Delta \to 0$

$$B_1^2(s,y) = B^2(s,x)\left(\frac{\partial \psi}{\partial x}\right)^2$$

$$= 0$$

which proves the assertion.

For this process, equations (20), (21), and (22) show that

$$\mu(s) = \text{const.}$$
$$\sigma^2(s) = \text{const.}$$
$$C(s,t) = \text{const.}$$

and moreover $h(s,u) = h(u)$ by equation (9). These conditions hold no matter what initial distribution we choose. Thus, choose the initial distribution $\mu(s_0) = 0$, $\sigma^2(s_0) = 0$, i.e. $P_r\{Y(s) = 0\} = 1$, and we have $P_r\{Y(s) = 0\} = 1$ for all s. Hence the process has no variation.

Now we return to the original variable $X(s)$ for which we have assumed $B^2(s,x) = 0$. Either $A(s,x) \equiv 0$, in which case the variable $X(s)$ has no variation, or as above we can find a random variable $Y(s) = \psi(s, X(s))$ so that it has no variation. Hence the variable $X(s)$ is such that $\psi(s, X(s)) = 0$ for almost all functions $X(s)$. Thus, in the s,x plane an equation $\psi(s,x) = 0$ determines a curve $x = x(s)$ which is almost surely followed by the process $X(s)$. Thus for each s, $X(s)$ is a random variable with zero dispersion, and it takes on its mean value with probability 1. We conclude that for almost all functions

$$X(s) = \int_{s_0}^{s} \mu(t)\,dt + \mu(s_0)$$

This result is in complete harmony with the Einstein process discussed in the preceding section, for which we deduced that the variable almost surely retained its initial value. As in the discussion of that process, however, we can make a modification and establish a new process. This modification consists of the same scale transformation--namely, instead of defining $B_1^2(s,x)$ as in (45) we divide by Δ^2 to obtain

$$B_1^2(s,x) = \lim_{\Delta \to 0} \frac{1}{\Delta^2} \int_s^{s+\Delta}\int_s^{s+\Delta} E\left(\mathcal{U}(t)\mathcal{U}(w) \middle| X(s) = x\right) dt\,dw$$

$$= E\left(\mathcal{U}^2(s) \middle| X(s) = x\right)$$

This expression, in connection with formula (44), enables us to deduce a process $X(s)$ which has no derivative, and yet in the sense given previously is an integral process.

It is noted that in the formula for $A_1(s,x)$ and $B_1^2(s,x)$ we obtain expressions involving the conditional expectation of $\mathcal{U}(s)$, $\mathcal{U}^2(s)$ under the assumption $X(s) = x$, and we must have at least this knowledge of the joint distribution of $\mathcal{U}$ and X in order to specify the coefficients A and B_1^2. Consequently the process $X(s)$ is not uniquely determined by that of $\mathcal{U}(s)$. However, if $\mathcal{U}(s)$ is independent of $X(s)$ we shall have

$$A_1(s,x) = E(\mathcal{U}(s)) = \mu(s)$$

$$B_1^2(s,x) = E(\mathcal{U}^2(s)) = \sigma^2(s) + \mu^2(s)$$

and since μ and σ^2 have been deduced in terms of $A(s,u)$ and $B^2(s,u)$, the new coefficients are determined in terms of the original ones. It is seen that here $A_1(s, x)$ and $B_1^2(s, x)$ depend only on s, and hence we have a Bachelier process, and ipso facto a Gaussian process. This situation will occur, for instance, if the velocity of the variable is independent of its displacement, as in the case of Brownian motion, etc.

The difficulty in representing jointly the velocity $\mathcal{U}$ and the displacement X, where both are supposed to be Markoff processes, and the former the derivative of the latter, is in view of the discussion above an inherent one, and in fact represents a serious fault in the structure of the stochastic process. Bass (1) and (2) has studied the joint distribution of $X(s)$ and $\mathcal{U}(t)$ by considering them jointly as a two dimensional vector variable, which should satisfy an extension of the Fokker-Planck equation. He discovers effectively the paradox with the Δ increments above, and concludes a satisfactory physical scheme cannot be so represented.

3.4 The integral statistics

Despite the fact that the coefficients $A_1(s, x)$ and $B_1^2(s, x)$ are not determined from the $\mathcal{U}(s)$ process, the elementary statistics can be found. We denote, in integral process, the mean, variance, and covariance by $\mu_1(s)$, $\sigma_1^2(s)$ and $C_1(s, t)$:

$$\mu_1(s) = E(X(s))$$

$$\sigma_1^2(s) = E(X^2(s)) - \mu_1^2(s)$$

$$C_1(s, t) = E(X(s)X(t)).$$

In addition, the mixed covariance $C_{xu}(s,t) = E(X(s)\,\mathcal{U}(t))$ will be needed.

By using the formula

(46) $$X(s) = X(s_0) + \int_{s_0}^{s} \mathcal{U}(t)\,dt$$

we can deduce these quantities. Taking the expectation of both sides of (46)

$$E(X(s)) = E(X(s_0)) + \int_{s_0}^{s} E(\mathcal{U}(t))\,dt$$

(47) $$\mu_1(s) = \mu_1(s_0) + \int_{s_0}^{s} \mu(t)\,dt$$

and we get a differential expression for $\mu_1(s)$

(48) $$\frac{d\mu}{ds} = \mu(s)$$

and using (20)

$$\frac{d^2\mu_1}{ds^2} = \int_{-\infty}^{\infty} A(s,u)\,h(s,u)\,du .$$

For the variance we obtain, by squaring (46)

(49) $$(X(s))^2 = (X(s_0))^2 + 2\int_{s_0}^{s} X(s_0)\,\mathcal{U}(w)\,dw + \int_{s_0}^{s}\int_{s_0}^{s} \mathcal{U}(w)\,\mathcal{U}(y)\,dw\,dy$$

and taking the expectations

$$\sigma_1^2(s) + \mu_1^2(s) = \sigma_1^2(s_0) + \mu_1^2(s_0) + 2\int_{s_0}^{s} C_{XU}(s_0,w)\,dw$$

$$+ \int_{s_0}^{s}\int_{s_0}^{s} C(w,y)\,dw\,dy$$

$$\frac{d\sigma_1^2(s)}{ds} + \frac{d\mu_1^2(s)}{ds} = 2C_{xu}(s_0,s) + 2\int_{s_0}^{s} c(w,s)\,dw. \tag{50}$$

For the covariance $C_1(s,t)$ we use the expression

$$(X(t)-X(s))^2 = \int_s^t\int_s^t \mathcal{U}(w)\,\mathcal{U}(y)\,dw\,dy$$

$$E\left((X(t)-X(s))^2\right) = \int_s^t\int_s^t E(\mathcal{U}(w)\,\mathcal{U}(y))\,dw\,dy$$

$$E(X(t))^2 - 2E(X(t)X(s)) + E(X(s))^2 = \int_s^t\int_s^t c(w,y)\,dw\,dy$$

$$2C_1(s,t) = \sigma_1^2(t) + \mu_1^2(t) + \sigma_1^2(s) + \mu_1^2(s) - \int_s^t\int_s^t c(w,y)\,dw\,dy \tag{51}$$

$$2\frac{\partial C_1(s,t)}{\partial t} = \frac{d\sigma_1^2(t)}{dt} + \frac{d\mu_1^2(t)}{dt} - 2\int_s^t c(w,t)\,dw \tag{52}$$

$$\frac{\partial^2 C_1(s,t)}{\partial s\,\partial t} = c(s,t).$$

In order to study the mixed covariances we proceed as follows

$$X(t) - X(s_0) = \int_{s_0}^{t} \mathcal{U}(w)\,dw$$

$$X(t)\,\mathcal{U}(r) - X(s_0)\,\mathcal{U}(r) = \int_{s_0}^{t} \mathcal{U}(w)\,\mathcal{U}(r)\,dw$$

and hence taking expectations

$$C_{xu}(t,r) = C_{xu}(s_0,r) + \int_{s_0}^{t} c(w,r)\,dw \tag{53}$$

The only undetermined quantity in the above formulas are the expressions $C_{xu}(s_0, r)$, and in order to determine this we must know the initial joint distribution of $U(s)$ and $X(s)$; i.e. $h(s_0; x, u)$. It is only in this case that we have the complete initial conditions specified. If $f(s_0, u; t, v)$ is the transition probability for the process $U(s)$ we immediately get

$$C_{xu}(s_0, r) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u x \, h(s_0; x, v) f(s_0, v; r, u) \, du \, dv \, dx$$

For the expression $C_{xu}(t, s_0)$ we get a simpler expression from (53)

$$C_{xu}(t, s_0) = C_{xu}(s_0, s_0) + \int_{s_0}^{t} C(w, s_0) \, dw$$

in which we know

$$C_{xu}(s_0, s_0) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x u \, h(s_0; x, u) \, dx \, du .$$

For the special case when $\mu(s) = \mu_1(s) = 0$ some of these expressions for the covariance were presented, without proof, by Loève (1). At least one of Loève's results is incorrect--it is said that a necessary and sufficient condition that $U(s)$ and $X(t)$ be independent (i.e. $C_{xu}(t, s) = \mu_1(t)\mu(s)$) is that $C_1(s, t)$ should be a function of st alone, say $D(st)$. Assuming they are independent, we get from (53)

$$\mu(r)\left(\mu_1(t) - \mu_1(s)\right) = \int_{s_0}^{t} C(w, r) \, dw$$

and on using (47)

$$\mu(r) \int_{s_0}^{t} \mu(w)\, dw = \int_{s_0}^{t} C(w, r)\, dw$$

Differentiating both sides with respect to t we obtain $C(t,r) = \mu(t)\,\mu(r)$, and substituting this in (51) we obtain

$$2\, C_1(s,t) = \sigma_1^2(t) + \sigma_1^2(s) + 2\,\mu_1(s)\,\mu_1(t)$$

which is not a function of ts alone for an arbitrary process $X(s)$. We have proved, however, that if $X(s)$ and $\mathcal{U}(t)$ are independent, then $\mathcal{U}(s)$ and $\mathcal{U}(t)$ are independent, which is prima facie a credible result.

We now prove that the integral process $X(s)$ of a temporally homogeneous process $\mathcal{U}(s)$ is temporally homogeneous only if the $\mathcal{U}(s)$ process is degenerate, i.e. $A(s,u) = const.$, $B^2(s,u) = 0$, $P_r\{\mathcal{U}(s) = k\} = 1$. For if the process $X(s)$ is temporally homogeneous then the expressions $E\left(\left(X(s) - X(s_0)\right)^k\right)$ for $k = 0, 1, \cdots$ must be functions of $s - s_0$ only:

$$f_k(s, s_0) = E\left(\left(X(s) - X(s_0)\right)^k\right)$$

and we must have, since f_k is a function of $s - s_0$ only,

$$\frac{\partial f_k}{\partial s} + \frac{\partial f_k}{\partial s_0} = 0$$

For $\kappa = 1$

$$f_1 = E \int_{s_0}^{s} \mathcal{U}(w)\, dw$$

$$= \int_{s_0}^{s} \mu(w)\, dw$$

$$\frac{\partial f_1}{\partial s} + \frac{\partial f_1}{\partial s_0} = \mu(s) - \mu(s_0)$$

and hence $\mu(s) = const. = \mu(s_0)$. Then $f_\kappa(s, s_0) = \mu(s_0)(s - s_0)$ irrespective of the choice of $\mu(s_0)$. In an identical manner we show that $\sigma^2(s) = \sigma^2(s_0) = const.$ for any initial distribution. Hence choose $\sigma^2(s_0) = 0$; then $\sigma^2(s) = 0$ for all s . We have already characterized this phenomenon in section 3.3, and since $\mu(s) = \mu(s_0)$ we must have $A(s, u) = Const.$, $B^2(s, u) = 0$ as asserted.

3.5 Special coefficients in the integral process

It is possible to deduce the unconditional statistics $\mu_1(s)$, $\sigma_1^2(s)$, $c_1(s,t)$ and $C_{xu}(s,t)$ specifically for the class of coefficients $A(s,u)$ and $B^2(s,u)$ given by (14) and (15). We obtain, on substituting these coefficients in (48), (49), and (50)

$$\frac{d^2\mu_1(t)}{dt^2} = \varphi_0(t) + \varphi_1(t)\frac{d\mu_1(t)}{dt}$$

$$\frac{\partial}{\partial t} C_{xu}(s,t) = C_{xu}(s_0,t) + 2\varphi_0(t)\mu_1(s)$$

$$\frac{d^2}{dt^2}\left(\sigma_1^2(t) + \mu_1^2(t)\right) = 2\frac{\partial}{\partial t} C_{xu}(s_0,t) + 2\left(\sigma^2(t) + \mu^2(t)\right)$$

$$2\frac{\partial c_1(r,t)}{\partial t} = \frac{d}{dt}\sigma_1^2(t) + 2\mu_1\frac{d\mu_1}{dt} - 2\int_r^t c(w,t)\, dw\, dt$$

and these expressions in conjunction with the differential equations (24), (25), and (26) specifying the functions $\sigma^2(t)$, $\mu(t)$, and $\rho(s,t)$, enable one to find the statistics for the integral process.

We note also that the integral of a Gaussian process is again Gaussian (cf. Doob (1)), so that in particular if the coefficients are of the form (30) the above moments will completely determine the probability distribution $h_1(s,x)$ for the integral process .

Finally, we can obtain by purely algebraical operations with the above statistics, a useful relationship between the correlation coefficients of a process and its integral process. We obtain, by elementary calculations,

$$2\left(1-R_1(t,r)\right)\sigma_1(t)\,\sigma_1(r) = \left(\sigma_1(t)-\sigma_1(r)\right)^2 - \int_r^t\!\!\int_r^t \sigma(w)\,\sigma(y)\,R(w,y)\,dw\,dy$$

an expression which is free of the mean values.

Bibliography

Bachelier, E.

1. 'Theorie de la speculation.' Ann. de l'Ecole norm. sup. 17 (1900) p 21

2. 'Calcul des probabilitiés,' Paris, (1912)

Bass, J.

1. 'Sur la structure des fonctiones aléatoires.' C. R. Acad. Sci. Paris 220 (1945) p 190-192

2. 'Quelques conséquences méchaniques de l'equation de structure d'un corpuscule aléatoire.' C. R. Acad. Sci. Paris 220 (1945) p 272-274

Bernstein, S.

1. 'Les équationes differentielles stochastique.' Actualities Scientifiques 737 p 7 Paris, Hermann edit.

Cramer, H.

1. 'Random variables and probability distributions.' Cambridge Univ. Press (1938)

2. 'On the theory of stationary random processos.' Ann. of Math. 41 (1940) p 215-230

Doob, J.

1. 'The Brownian movement and stochastic equations.' Ann. of Math. 43 (1942) p 351-369

2. 'The elementary Gaussian processes.' Ann. of Math. Statistics 15 (1944) p 229-282

3. 'Probability in function space.' Bull. Amer. Math. Soc. 53 (1947) p 15-30

Einstein, A.

1. 'Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flussigkeiten suspendierten Teilchen.' Ann. der Physik 17 (1905) p 549

Fan, K.

1. 'Sur l'approximation et l'integration des fonctiones aléatoires.' Bull. Soc. Math. France 72 (1944) p 97-117

Feldheim, E.

1. 'Sur les probabilities en chaîne.' Math. Ann. 112 (1936) p 775-780

Feller, W.

1. 'Zur Theorie der Stochastichen Prozesse; Existenz und Eindeutigkeitssatze.' Math. Ann. 113 (1936) p 113-160

2. 'On the integro-differential equations of purely discontinuous Markoff processes.' Trans. Amer. Math. Soc. 48 (1940) p 488-515

Fortet, R.

1. 'Les fonctiones aléatoires de type du Markoff associées a certaines équationes lineares aux dérivées partielles du type parabolique.' Jour. Math. Pures Appl. (9) 22 (1943) p 177-243

Hostinsky, B.

1. 'Sur les probabilities en chaîne.' Acad. Sci. Paris 202 (1936) p 2029-2031

Itô, K.

1. 'On the ergodicity of certain stationary processes.' Proc. Imp. Acad. Tokyo 20 (1944) p 54-55

Khintchine, A.

1. 'Asymtotische Gesetze der Wahrscheinlichkeitsrechnung.' Ergebnisse der Mathematik v. 2 nr. 4 (1933)

2. 'Korrelationstheorie der stationären Stochastichen Prozesse.' Math. Ann. 109 (1934) p 604-615

Kolmogoroff, A.

1. 'Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung.' Math. Ann. 104 (1931) p 415-458

2. 'Sulla forma generali di un processo stocastico omogeneo.' Rend. R. Acad. Lincei (6) 15 (1932) p 805-808, 866-869

Levy, P.

1. 'Sur les integrales dont les éléments sont des variables aléatoires indépendantes.' Annali R. Sci. Norm. Sup. Pisa (2) 3 (1934) p 337-366

2. 'Sur certaines solutiones de l'equation de Chapman.' C. R. Acad. Sci. Paris 205 (1937) p 1355-1357

Loeve, M.

1. 'Sur la covariance d'une fonctione aléatoire.' C. R. Acad. Sci. Paris 220 (1945) p 295-296

Markoff, A.

1. 'Wahrscheinlichkeitsrechnung.' Leipzig (1912)

Onicescu, O. and Mihoc, G.

1. 'Sopra li leggi limite della probabilità.' Giorn. Ist. Ital. Attuari 7 (1936) p 54-69

Ornstein, L. and Uhlenbeck, G.

1. 'On the theory of Brownian motion.' Phys. Rev. 36 (1930) p 823-841

Paley, R. and Wiener, N.

1. 'Fourier transforms in the complex domain.' Amer. Math. Soc. Colloquium Publ. Vol. 19 (1934)

Romanovsky, V.

1. 'Recherches sur les chaînes de Markoff.' Acta Math. 66 (1935) p 147-251

Slutsky, E.

1. 'Qualche proposizione relativa alla teoria della funzioni aleatoire.' Giorn. Ist. Ital. Attuari 8 (1937) p 183-199

2. 'Sur les fonctiones éventuelles continues, integrable et derivables dans le sens stochastique.' C. R. Acad. Sci. 187 (1928) p 878-880

v. Smoluchowski, M.

1. 'Drei Vorträge uber Diffusion Brownsche Molekularbewegung und Koagulation von Kolloidteilchen.' Physik. Zeitschr. 17 (1916) p 557