Annals of Mathematics
Vol. 43, No. 2, April, 1942

THE BROWNIAN MOVEMENT AND STOCHASTIC EQUATIONS

By J. L. Doob

(Received January 14, 1942)

The irregular movements of small particles immersed in a liquid, caused by the impacts of the molecules of the liquid, were described by Brown in 1828.[1] Since 1905 the Brownian movement has been treated statistically, on the basis of the fundamental work of Einstein and Smoluchowski. Let $x(t)$ be the x-coordinate of a particle at time t. Einstein and Smoluchowski treated $x(t)$ as a chance variable. They found the distribution of $x(t) - x(0)$ to be Gaussian, with mean 0 and variance $\alpha \mid t \mid$, where α is a positive constant which can be calculated from the physical characteristics of the moving particles and the given liquid. More exactly, such a family of chance variables $\{x(t)\}$ is now described as the family of chance variables determining a temporally homogeneous differential stochastic process: the distribution of $x(s + t) - x(t)$ is Gaussian, with mean 0, variance $\alpha \mid t \mid$, and if $t_1 < \cdots < t_n$,

$$x(t_2) - x(t_1), \cdots, x(t_n) - x(t_{n-1})$$

are mutually independent chance variables. Wiener, who was the first to discuss this stochastic process rigorously, proved in 1923 that the functions $x(t)$ of this stochastic process are continuous, with probability 1.[2] This is of course a desirable result, which makes the stochastic process somewhat more acceptable as the mathematical idealization of the Brownian movement. It was not expected, however, that the above distribution of $x(s + t) - x(s)$ would prove correct for small t. Even if the derivation did not break down for small t, the mathematical fact that $x(s + t) - x(s)$ has standard deviation $\alpha \mid t \mid$ so that $x(s + t) - x(s)$ is of the order of magnitude of $\mid t \mid^{\frac{1}{2}}$, implying that $dx(s)/ds$ cannot be finite, would suggest the desirability of modifications of the Einstein-Smoluchowski distributions. In fact it is easily seen that (with probability 1) $x(t)$ is not even of bounded variation, so that the path curves of the Einstein-Smoluchowski process have infinite length!

A different stochastic process describing the $x(t)$ was in fact derived in 1930 by Ornstein and Uhlenbeck (15),[3] and later by S. Bernstein (1), (2) and Krutkow (11), all using different methods. This new distribution of $x(s + t) - x(s)$ is

---

[1] For a historical account of the subject up to 1913, see Haas-Lorentz (6). (The numbered references will refer to the bibliography at the end of the paper.)

[2] Wiener (18, pp. 148–151) has since given a more simple proof. For a discussion of the exact meaning of such a statement concerning the continuity of paths, cf. Doob (3) and (5), §2. The result means that $x(t)$ can be treated as representing one of a multiplicity of continuous functions of t, and integrated, etc. Probability here is formally the study of measures on certain spaces of functions.

[3] Cf. also Ornstein and Wijk (16) and Wijk (17). References to work since 1913 are given in Ornstein and Uhlenbeck (15).

Gaussian, with mean 0 and variance $(\alpha/\beta)(e^{-\beta|t|} - 1 + \beta \mid t \mid)$, approximately $\alpha \mid t \mid$ for large t, but $\alpha\beta t^2/2$ for small t. (Here β is a second physically determined constant.)

The purpose of the present paper is to apply the methods and results of modern probability theory to the analysis of the Ornstein-Uhlenbeck distribution, its properties and its derivation. It will be seen that the use of rigorous methods actually simplifies some of the formal work, besides clarifying the hypotheses. A stochastic differential equation will be introduced in a rigorous way to give a precise meaning to the Langevin differential equation for the velocity function $dx(s)/ds$. This will avoid the usual embarrassing situation in which the Langevin equation, involving the second derivative of $x(s)$ is used to find a solution $x(s)$ not having a second derivative.

1. The velocity distribution

The displacement function $x(t)$, as discussed by Ornstein and Uhlenbeck, has a derivative $u(t)$, and all the probability relations needed can be derived from those of $u(t)$, as will be seen below. The distribution of $u(t)$ can be described as follows: the conditional distribution of $u(s + t)$ $(t > 0)$ for given $u(s) = u_0$, is Gaussian, with mean $u_0 e^{-\beta t}$ and variance $\sigma_0^2(1 - e^{-2\beta t})$. Here σ_0^2, β are physically determined constants. When $t \to \infty$, this distribution becomes the Maxwell distribution of velocities, furnishing stationary absolute (unconditioned) probabilities for the process, if these are desired. Using these absolute probabilities, which make the distribution easier to describe, the full description of the $u(t)$ distribution can then be stated as follows: for each t, $u(t)$ is a chance variable with a Gaussian distribution, having mean 0, variance σ_0^2; the transition probabilities are as just described; the process is a Markoff process.[4] This last fact means that the Maxwell distribution of $u(t_0)$ for each fixed t_0, and the transition probabilities just described determine the full set of probability relations of the process. Under these conditions, if $t_1 < t_2$, the pair $u(t_1)$, $u(t_2)$ has a bivariate Gaussian distribution, with zero means, equal variances σ_0^2, and correlation coefficient $e^{-\beta(t_2 - t_1)}$. This stochastic process goes back at least to Smoluchowski, although it was first derived by Ornstein and Uhlenbeck as the process describing the velocity of a particle in Brownian motion. Ornstein and Uhlenbeck were only interested in the transition probabilities. The formal manipulations made below will show that there are technical advantages in defining (unconditioned) probabilities for the $u(t)$ also. The above described process will be called the O. U. process below.

The following theorem shows that such a process is essentially determined by three fundamental properties, of which at least the first two have simple physical

[4] A process is called a Markoff process if whenever $t_1 < \cdots < t_n$, the conditional distribution of $u(t_n)$ for given values of $u(t_1), \cdots, u(t_{n-1})$ actually depends only on $u(t_{n-1})$. It is in this case, and only in this case, that the Smoluchowski equation between the transition probabilities, and the Fokker-Planck differential equations for the transitional probabilities are valid.

significance. (We can exclude Case A of the theorem, since it obviously does not fit the physical picture.)

THEOREM 1.1. *Let $u(t)$ $(-\infty < t < +\infty)$ be a one-parameter family of chance variables, determining a stochastic process with the following properties.*

1. *The process is temporally homogeneous.*[5]
2. *The process is a Markoff process.*
3. *If s, t are arbitrary distinct numbers, $u(s)$, $u(t)$ have a (non-singular) bivariate Gaussian distribution.*

Define m, σ_0^2 by

$$(1.1.1) \qquad m = E\{u(t)\}, \qquad \sigma_0^2 = E\{[u(t) - m]^2\}.^6$$

Then the given process is one of the following two types.

(A) *If $t_1 < \cdots < t_n$, $u(t_1), \cdots, u(t_n)$ are mutually independent Gaussian chance variables, with mean m and variance σ_0^2.*

(B) *(O. U. process) There is a constant $\beta > 0$ such that if $t_1 < \cdots < t_n$, $u(t_1), \cdots, u(t_n)$ have an n-variate Gaussian distribution, with common mean m and variance σ_0^2, and correlation coefficients determined by the equation $E\{[u(t) - m][u(s) - m]\} = \sigma_0^2 e^{-\beta|t-s|}$.*

Instead of considering $u(t)$, we can consider $(1/\sigma_0)[u(t) - m]$, which has mean 0 and variance 1. Then we shall assume in the following that $u(t)$ itself has these properties: $m = 0$, $\sigma_0^2 = 1$. Let $\rho(t)$ be the correlation function: $\rho(t) = E\{u(s + t)u(s)\}$, independent of s by Property 1. If $s < t$, the conditional distribution of $u(t)$ for given $u(s)$ has density

$$(1.1.2) \qquad \frac{1}{(2\pi)^{\frac{1}{2}}(1 - \rho^2)^{\frac{1}{2}}} \exp\left(-\frac{1}{2}\frac{[u(t) - \rho u(s)]^2}{1 - \rho^2}\right), \qquad \rho = \rho(t - s),$$

(Property 3). If $t_1 < \cdots < t_n$, $u(t_1), \cdots, u(t_n)$ then have an n-variate Gaussian distribution with density

$$(1.1.3) \qquad \frac{1}{(2\pi)^{\frac{n}{2}} \prod_1^{n-1} (1 - \rho_j^2)^{\frac{1}{2}}} \exp\left(-\frac{1}{2}u_1^2 - \frac{1}{2}\sum_1^{n-1} \frac{(u_{j+1} - \rho_j u_j)^2}{1 - \rho_j^2}\right),$$

$$\rho_j = \rho(t_{j+1} - t_j), \; u_j = u(t_j)$$

using Property 2. Now if $u_1, \cdots, u_n$ have an n-variate Gaussian distribution with density

$$(1.1.4) \qquad \frac{1}{\Delta} \exp\left(-\frac{1}{2}\sum_{i,j} a_{ij} u_i u_j\right),$$

$\Delta = \det(E\{u_i u_j\})$ is the determinant of the matrix of variances and covariances, and (a_{ij}) is the inverse of this matrix. Using these facts we can calculate $\rho(t_3 - t_1) = E\{u_1 u_3\}$ in (1.1.3) with $n = 3$, and find that $\rho(t_3 - t_1) = \rho_1\rho_3$, that is

[5] That is, the probability distributions are unaffected by translations of the t-axis.

[6] The expectation of the chance variable v will be denoted by $E\{v\}$.

$$\rho(t_3 - t_1) = \rho(t_2 - t_1)\rho(t_3 - t_2). \tag{1.1.5}$$

Then $\rho(t)$ is an even function; $|\rho(t)| \leqq 1$ (Schwarz's inequality); and according to (1.1.5) $\rho(s + t) = \rho(s)\rho(t)$ for all positive s, t. Under these conditions either $\rho(t) \equiv 0$ or there is a constant $\beta \geqq 0$ such that

$$\rho(t) = e^{-\beta|t|}. \tag{1.1.6}$$

In the present case, $\beta > 0$, by Property 3 (non-singularity of the given bivariate distributions). Evidently $\rho(t) \equiv 0$ furnishes Case A of the theorem, which certainly has the three given properties. If $\rho(t)$ is given by (1.1.6) with $\beta > 0$, we show first that the matrix (a_{ij}), the inverse of $(\rho(t_j - t_i))$ actually determines a Gaussian density distribution (1.1.4). To see this we consider the density function (1.1.3) with $\rho_j = e^{-\beta(t_{j+1}-t_j)}$. The coefficients of the quadratic form in the exponent of (1.1.3) are easily evaluated and the matrix of the form is found to be the inverse of the matrix $(e^{-\beta|t_j-t_i|})$. Thus (1.1.3) actually is the required probability density. Moreover the probability densities obtained in this way (as the t_i vary) are mutually consistent, because integrating out any variable leaves a quadratic form of the same type, without the integrated variable, but with the same rule determining the coefficients. The correlation function (1.1.6) therefore determines a stochastic process. The process obviously is a Markoff process because of the form of the probability density (1.1.3): an initial factor involving u_1 only, followed by the product of functions of pairs of adjacent variables. The proof of the theorem is now complete.

According to a theorem of Khintchine ((9) p. 608), $\rho(t)$ is the correlation function of a temporally homogeneous stochastic process if and only if it can be put in the form

$$\rho(t) = \int_0^\infty \cos \lambda t \, dF(\lambda), \tag{1.1.7}$$

where $F(\lambda)$ is monotone non-decreasing and bounded. In Case B of the theorem, (1.1.7) is true when $F(\lambda)$ is given by

$$F(\lambda) = \frac{2\beta\sigma_0^2}{\pi} \int_0^\lambda \frac{d\lambda}{\beta^2 + \lambda^2}. \tag{1.1.8}$$

In the stochastic process of Case B, the variance of $u(s + t) - u(s)$ is $2\sigma_0^2\beta|t|$ for small t:

$$E\{[u(s + t) - u(s)]^2\} = 2\sigma_0^2(1 - e^{-\beta|t|}) \sim 2\sigma_0^2\beta|t|. \tag{1.1.9}$$

Thus $u(s + t) - u(s)$ is of the order of magnitude of $|t|^{\frac{1}{2}}$, and du/dt cannot exist. Physically this means that the particles in question do not have a finite acceleration (if the given stochastic process represents the Brownian movement that closely).

THEOREM 1.2. *If $u(t)$ is the representative function of the stochastic process of Theorem* 1.1 *Case B, $u(t)$ is a continuous function of t, with probability* 1.

Let $v(t)$ be determined by the equation

$$v(t) = t^{\frac{1}{2}} u\left(\frac{1}{2\beta} \log t\right), \qquad t > 0. \tag{1.2.1}$$

Then $v(t)$ has the property that if $t_1 < \cdots < t_n$, $v(t_1), \cdots, v(t_n)$ have an n-variate Gaussian distribution. We find by direct calculation (taking $m = 0$):

$$\begin{aligned} E\{v(s + t) - v(s)\} &= 0, \\ E\{[v(s + t) - v(s)]^2\} &= \sigma_0^2 t, \\ E\{[v(s_2) - v(s_1)][v(t_2) - v(t_1)]\} &= 0, \quad (s_1 < s_2 \leqq t_1 < t_2). \end{aligned} \tag{1.2.2}$$

Then $v(t)$ determines a differential process—in fact precisely the original Einstein-Smoluchowski process. Since Wiener has proved continuity of the path functions in this case, the theorem follows.

The transition from $u(t)$ to $v(t)$ just used reduces every property of the Ornstein-Uhlenbeck stochastic process to a corresponding property of the Einstein-Smoluchowski process, and vice versa. Many properties of the individual functions of the latter process, that is, properties possessed by almost all the individual functions, in other words possessed "with probability 1," have been proved in recent years, besides the continuity property we have just used. The following theorem gives the counterparts of two of these for the O. U. process.

THEOREM 1.3. *If $u(t)$ is the representative function of the O. U. process of Theorem 1.1 Case B,*

$$\limsup_{t\to 0} \frac{u(t) - u(0)}{(4\sigma_0^2\beta t \log\log(1/t))^{\frac{1}{2}}} = 1, \qquad \limsup_{t\to 0} \frac{u(t)}{(2\sigma_0^2 \log t)^{\frac{1}{2}}} = 1, \tag{1.3.1}$$

with probability 1.

Let $v(t)$ be defined by (1.2.1). Then Khintchine ((10) pp. 68–75) has proved

$$\limsup_{t\to 0} \frac{v(1 + t) - v(1)}{(2\sigma_0^2 t \log\log(1/t))^{\frac{1}{2}}} = 1, \qquad \limsup_{t\to \infty} \frac{v(t) - v(0)}{(2\sigma_0^2 t \log\log t)^{\frac{1}{2}}} = 1, \tag{1.3.2}$$

and (1.3.2) becomes (1.3.1) when $v(t)$ is expressed in terms of $u(t)$.

2. The distribution of displacements

It does not seem to have been realized by earlier writers that the distribution of displacements in the O. U. process can be obtained directly from that of the velocities. In fact, we have seen that as t varies, $u(t)$ considered as one of a multiplicity of continuous functions of t. Integration of $u(t)$ is therefore admissible, and will give the displacement function. If $x(t)$ is the x-coordinate of a particle at time t,

$$x(t) - x(0) = \int_0^t u(s)\, ds \tag{2.1}$$

with probability 1 (that is, neglecting the discontinuous $u(t)$ functions which have total probability 0). The main advantages of the rigorous approach to stochastic processes depending on a continuous parameter is precisely that the $u(t)$ of the process, as t varies, can be regarded as an individual function or rather, as one of many functions with whatever regularity properties the given probability distributions imply. Theorem 1.3 limits the actual upper bounds of the velocity functions $u(t)$. The following result takes advantage of the oscillations in sign.

THEOREM 2.1. *If $u(t)$ is the representative function of the O. U. process of Theorem 1.1 Case B, with $m = 0$,*

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t u(s)\, ds = \lim_{t\to\infty} \frac{x(t) - x(0)}{t} = 0, \tag{2.1.1}$$

with probability 1.

This theorem is simply the ergodic theorem applied to the $u(t)$ process to give the strong law of large numbers, (cf. Doob (4) p. 294). From (2.2.3) below, it is quite obvious that the expectation of the square of the left side of (2.1.1) goes to 0 as $t \to \infty$, so that the left side goes to 0 in the mean. The strength of (2.1.1) is that it is a statement about the path of the individual path functions, or physically, a statement about the path of a single particle. The same was true in Theorems 1.2 and 1.3.

In order to find the distribution of $x(t) - x(0)$ we proceed as follows. Riemann integrability of $u(t)$ implies that (with probability 1)

$$x(t) - x(0) = \lim_{n\to\infty} \sum_{j=1}^{n} u(tj/n)t/n. \tag{2.2.1}$$

Now the n-variate distribution of the variables summed is Gaussian. Then the sum is Gaussian, so the distribution of $x(t) - x(0)$ is also Gaussian, if it can be shown that the variance of $x(t) - x(0)$ is positive. The distribution of $x(t) - x(0)$ is thus completely determined by its first two moments, which we proceed to calculate. We shall suppose, that $E\{u(t)\} = 0$, $E\{u(t)^2\} = \sigma_0^2$. Then we find

$$E\{x(t) - x(0)\} = \int_0^t E\{u(s)\}\, ds = 0,^{7} \tag{2.2.2}$$

and, if $t > 0$,

$$\begin{aligned} E\{[x(t) - x(0)]^2\} &= \int_0^t \int_0^t E\{u(s)u(s')\}\, ds\, ds' \\ &= \sigma_0^2 \int_0^t \int_0^t e^{-\beta|s-s'|}\, ds\, ds' = \frac{2\sigma_0^2}{\beta^2} (e^{-\beta t} - 1 + \beta t). \end{aligned} \tag{2.2.3}$$

[7] By Fubini's integration theorem, we can find the expectations under the integral sign, before integrating with respect to s.

The same sort of argument shows that if $t_1, \cdots, t_n$ are any distinct numbers, the chance variables

$$\{x(t_j) - x(0), u(t_j), \qquad j = 1, \cdots, n\}$$

have a $2n$-variate Gaussian distribution, which can then be evaluated explicitly by finding the first and second moments. For example, the following equations determine the bivariate distribution of $x(t) - x(0)$, $u(t)$, $(t > 0)$:

$$\begin{gathered} E\{[x(t) - x(0)]u(t)\} = \int_0^t E\{u(t)u(s)\}\, ds = \frac{\sigma_0^2}{\beta}(1 - e^{-\beta t}), \\ E\{x(t) - x(0)\} = 0, \\ E\{[x(t) - x(0)]^2\} = \frac{2\sigma_0^2}{\beta^2}(e^{-\beta t} - 1 + \beta t), \quad E\{u(t)\} = 0, \quad E\{u(t)^2\} = \sigma_0^2. \end{gathered} \tag{2.2.4}$$

Thus the bivariate density of $x(t) - x(0)$, $u(t)$ is Gaussian, with common mean 0, and variances $(2\sigma_0^2/\beta^2)(e^{-\beta t} - 1 + \beta t)$, σ_0^2, respectively, and correlation coefficient

$$\frac{1 - e^{-\beta t}}{2^{\frac{1}{2}}(e^{-\beta t} - 1 + \beta t)^{\frac{1}{2}}}. \tag{2.2.5}$$

It is to be expected that if $s_1 < s_2 \leqq t_1 < t_2$, $x(s_2) - x(s_1)$ and $x(t_2) - x(t_1)$ become independent as $t_1 \to \infty$. In fact, these two normally distributed variables have correlation coefficient.

$$\frac{(e^{\beta s_2} - e^{\beta s_1})(e^{-\beta t_1} - e^{-\beta t_2})}{2(e^{-\beta(s_2 - s_1)} - 1 + \beta(s_2 - s_1))^{\frac{1}{2}}(e^{-\beta(t_2 - t_1)} - 1 + \beta(t_2 - t_1))^{\frac{1}{2}}}; \tag{2.2.6}$$

which goes to 0 when t_1 and t_2 become infinite.

If in this discussion only the conditional distribution functions are wanted, for $u(0) = u_0$, for example, two procedures are possible. Setting $u(0) \equiv u_0$ instead of using the initial distribution we have used above, carrying out the same type calculations as above, now would give the desired conditional probabilities. Or the conditional distributions could be calculated from the distributions just derived, since the conditional distributions of a multivariate Gaussian distribution are easily found. Theorems 1.2, 1.3 and 2.1 hold no matter what initial distribution is assigned to $u(0)$.

Finally, there is one more fact which we shall need in the next section. Define $B(t)$ by

$$B(t) = \beta[x(t) - x(0)] + u(t) - u(0). \tag{2.2.7}$$

Then $B(t)$ has for each t a Gaussian distribution, with mean 0. Evidently the distribution of $B(s + t) - B(s)$ is independent of s. It is Gaussian, with mean 0, and the variance is easily calculated to be $2\sigma_0^2\beta \mid t \mid$. Moreover, if $s_1 < s_2 \leqq t_1 < t_2$,

$$E\{[B(t_2) - B(t_1)][B(s_2) - B(s_1)]\} = 0. \tag{2.2.8}$$

Thus the $B(t)$-process is again the Einstein-Smoluchowski process.

3. Derivation of the velocity distribution using the Langevin equation

Ornstein and Uhlenbeck base their investigation on the Langevin equation

$$\frac{du}{dt} = -\beta u(t) + A(t), \tag{3.1}$$

which is simply Newton's law of motion applied to a particle, after dividing through by the mass. The first term on the right is due to the frictional resistance or its analogue, which is supposed proportional to the velocity. The second term represents the random forces (molecular impacts). Probability hypotheses are imposed on the $A(t)$, including relations between $A(t)$ and $u(t)$, to determine the $u(t)$ distribution. Unfortunately this $u(t)$ distribution (Case B of Theorem 1.1), as we have seen, has the property that the velocity function has no time derivative. Then the solution can hardly satisfy (3.1).

Bernstein ((2) p. 361) replaces (3.1) by a finite difference equation:

$$\Delta\left(\frac{\Delta\xi_n}{\Delta t}\right) = -\beta\Delta\xi_n + \alpha_n, \qquad n = 1, 2, \cdots. \tag{3.2}$$

Here $\xi_1, \xi_2, \cdots$ is a sequence of chance variables, $\Delta\xi_n = \xi_{n+1} - \xi_n$ etc., and $\alpha_1, \alpha_2, \cdots$ is a given sequence of mutually independent chance variables. If we think of ξ_j as the analogue of $x(j\Delta t)$, the correspondence between (3.2) and (3.1) is clear. The equations of (3.2) determine definite distributions for the ξ_j in terms of those of the α_j. Bernstein shows that as $\Delta t \to 0$ the distribution of $\Delta\xi_n/\Delta t$ ($\sim \Delta x/\Delta t$) becomes the $u(t)$ distribution we have been discussing, if suitable hypotheses are made on the α_j. This approach is essentially different from that of Ornstein and Uhlenbeck in that Bernstein, as he states explicitly ((1) pp. 5, 6) is not writing a difference equation in the displacement functions $x(t)$ themselves: (3.2) determines distributions only, and these are approximated by the limiting distributions described in Theorem 1.1 Case B.

In our treatment, we shall replace the Langevin equation by a formalized differential equation for the velocity function $u(t)$. This equation is to be exact, not merely asymptotically true. The equation will be perfectly proper mathematically, so that solution by ordinary methods will provide all the information relevant to the desired distributions, and solution of more general problems, involving external forces, will require no special methods.

The problem is to find a proper stochastic analogue of the Langevin equation, remembering that we do not expect $u'(t)$ to exist. We write the equation in the following form:

$$du(t) = -\beta u(t)\, dt + dB(t), \tag{3.3}$$

and try to give these differentials a suitable interpretation. We shall suppose that the $B(t)$-process is a differential process: that is, if $t_1 < \cdots < t_n$, we suppose that

$$B(t_2) - B(t_1), \cdots, B(t_n) - B(t_{n-1})$$

are mutually independent chance variables. We also suppose temporal homogeneity, that is that the distribution of $B(s + t) - B(s)$ is independent of s. The physical meaning of these hypotheses is clear, and they will be justified further below. Equation (3.3) can be interpreted roughly in terms of small changes in momentum. An important particular case is that in which the second moments of the $B(t)$-process are finite:

$$\sigma^2(t) = E\{[B(s + t) - B(s)]^2\} < \infty. \tag{3.4}$$

The first moment $E\{B(s + t) - B(s)\}$ then exists. If this first moment vanishes, $\sigma^2(t)$ satisfies the functional equation

$$\sigma^2(s + t) = \sigma^2(s) + \sigma^2(t).$$

Then $\sigma^2(t)$ must be proportional to t: $\sigma^2(t) = t\sigma^2$. If $f(t)$ is continuous,

$$\int_a^b f(t)\, dB(t) \tag{3.5}$$

has been defined under these hypotheses (Wiener (18), pp. 151–157, Doob (3), pp. 131–134), even though the functions $B(t)$ are known not to be of bounded variation. The definition makes all the formal processes correct. For example, if $f'(t)$ exists and is continuous,

$$\int_a^b f(t)\, dB(t) = f(t)[B(t) - B(0)]\Big|_a^b - \int_a^b [B(t) - B(0)]f'(t)\, dt^{8} \tag{3.6}$$

with probability 1. The usual Riemann-Stieltjes sums converge to (3.5) in the mean. Moreover

$$\begin{aligned} &E\left\{\int_a^b f(t)\, dB(t)\right\} = 0, \\ &E\left\{\left[\int_a^b f(t)\, dB(t)\right]\left[\int_a^b g(t)\, dB(t)\right]\right\} = \sigma^2 \int^b f(t)g(t)\, dt. \end{aligned} \tag{3.7}$$

Now it can be shown even without the hypothesis of the finiteness of the second moment in (3.4) that the formal integral in (3.5) can be defined, and will satisfy (3.6). The form of the characteristic function of $B(s + t) - B(s)$ has been derived by Lévy ((14) Chapter VII) and using this it is easy to prove that the

[8] We never write $B(t)$ alone in an equation, since strictly speaking only differences like $B(t) - B(0)$ are defined. It is unnecessary to define $B(0)$ itself, although for convenience it can be taken identically 0, without affecting any of the equations to be used. Differential processes have been discussed in detail by Lévy ((12), (13), (14) Chapter VII) and Doob ((3) §3).

usual Riemann-Stieltjes sums for the integral (3.5) converge in probability. The integral is defined as the limit, and (3.6) is readily verified. On the other hand, (3.7) cannot be expected to hold, since if $f(t) = 1$ the integral becomes $B(b) - B(a)$, and we have not supposed that the expectation of this difference is finite. The special case in which the second moment is finite is the only important one for the purposes of this section, but less restrictive conditions will be needed in §5. We shall justify later the assumption that the $B(t)$ process is a differential process.

We shall interpret an equation in differentials like (3.3) to mean the truth (with probability 1, that is for almost all functions $u(t)$) of

$$\int_a^b f(t)\,du(t) = -\beta \int_a^b f(t)u(t)\,dt + \int_a^b f(t)\,dB(t) \tag{3.8}$$

for all a, b, whenever f is a continuous function. Here the first two integrals are to be defined as the limits (in probability) of the usual Riemann or Riemann-Stieltjes sums. Equation (2.2.7) implies

$$\begin{aligned}\int_a^b f(t)\,du(t) &= -\beta \int_a^b f(t)\,dx(t) + \int_a^b f(t)\,dB(t)\\ &= -\beta \int_a^b f(t)u(t)\,dt + \int_a^b f(t)\,dB(t).\end{aligned} \tag{3.9}$$

Thus (3.3) holds for the $u(t)$ of the O. U. distribution if the $B(t)$ is defined by (2.2.7). Moreover (2.2.7) with $B(t)$ replaced by $B(t) - B(0)$ is an immediate consequence of (3.3). In this case, $B(t)$ has the property that the differences $B(s + t) - B(s)$ have finite second moments and even Gaussian distributions, but we are not making either assumption in solving (3.3).

If (3.3) is true, then (with probability 1)

$$\int_0^t e^{\beta\tau}\,du(\tau) = -\beta \int_0^t e^{\beta\tau}\,u(\tau)\,d\tau + \int_0^t e^{\beta\tau}\,dB(\tau), \tag{3.10}$$

which implies, since integration by parts is applicable,

$$u(t) = u(0)e^{-\beta t} + e^{-\beta t} \int_0^t e^{\beta\tau}\,dB(\tau) \tag{3.11}$$

for all t, with probability 1. Conversely suppose that $u(t)$ is defined by (3.11). Since $B(t)$ is known to be continuous in t except for non-oscillatory discontinuities (jumps) (Lévy (12) pp. 359–364, (13); Doob (3), pp. 134–138), the same must be true of the right side of (3.11), and therefore of $u(t)$. Then $u(t)$ is Riemann integrable with probability 1. Moreover

$$\int_a^b f(t)\,e^{-\beta t}\,d_t \int_0^t e^{\beta\tau}\,dB(\tau) = \int_a^b f(t)\,dB(t), \tag{3.12}$$

so that from (3.11)

$$\int_a^b f(t)e^{-\beta t}\, d[e^{\beta t}u(t) - u(0)] = \int_a^b f(t)\, dB(t), \tag{3.13}$$

proving incidentally that the left side exists. The left side can be simplified to

$$\beta \int_a^b f(t)u(t)\, dt + \int_a^b f(t)\, du(t), \tag{3.14}$$

and putting this into (3.13) we find that (3.8) is satisfied. Then (3.11) furnishes the complete solution of (3.3) under the stated conditions. We stress again that although (3.11) implies strong connections between the $u(t)$ and $B(t)$ processes, we have made no such hypothesis in the derivation not implicit in (3.3). Lévy ((14) pp. 166–167) has shown that the only differential processes whose path functions $B(t) - B(0)$ do not have jumps have the property that the distribution of $B(t) - B(0)$ is Gaussian. Then it is only in this case, which will lead to the O. U. process, that $u(t)$ will not have jumps.

The term $\beta u(t)$ in the Langevin equation is supposed to account for the total frictional effect, including the Doppler friction, caused by the fact that more impacts decelerate than accelerate the motion of a moving particle. The term $A(t)$ in (3.1) or $dB(t)$ in (3.3) represents the "purely random" impulses, that is, the residual effect after the frictional effect has been subtracted out. One idea running through any treatment of the Langevin equation is that this term or, sometimes, $x(t)$ itself, is independent of the given velocity at any time. This hypothesis goes back to Langevin, and has caused much controversy. We shall make the hypothesis only to the following extent. The chance variable $u(0)$ will be given various initial distributions, but will always be made independent of the $B(t)$-process for $t \geqq 0$. This means that if $0 \leqq t_1 < \cdots < t_n$ the chance variable $u(0)$ is supposed independent of the set of chance variables

$$\{B(t_{j+1}) - B(t_j), \qquad j = 1, \cdots, n-1\}.$$

We shall describe the above hypothesis in the following physical terms: *the initial velocity $u(0)$ is independent of later residual random impacts.* It would be a serious drawback to the whole treatment if when $u(0)$ is so chosen $u(t_0)$ for each $t_0 > 0$ were not independent of the $B(t)$-process for $t \geqq t_0$, that is if $u(t_0)$ were not independent of later residual random impacts for all t_0. We can prove, however, the following statement, which incidentally justifies our hypothesis that the $B(t)$-process is a differential process. *Let the $B(t)$ process be a differential process, and define $u(t)$ by* (3.11). *If the chance variable $u(0)$ is independent of the $B(t)$-process for $t \geqq 0$, then $u(t_0)$ will be independent of the $B(t)$-process for $t \geqq t_0$, for all $t_0 > 0$. Conversely suppose only that the $B(t)$-process is regular enough that the integral* (3.5) *can be defined as the limit in probability of the usual sums, and that* (3.6) *is true. Then if $u(t)$ is defined by* (3.11), *and if choosing $u(0)$ independent of the $B(t)$ process for $t \geqq 0$ implies that $u(t_0)$ will be independent of the $B(t)$-process for $t \geqq t_0$, for all $t_0 > 0$, then the $B(t)$-process is a differential process.*

PROOF. Let the $B(t)$-process be a differential process, define $u(t)$ by (3.11) and let $u(0)$ be independent of the $B(t)$-process for $t \geqq 0$. Then from (3.11) with $t = t_0$, $u(t_0)$ involves only $u(0)$ and the $B(t)$-process for $t \leqq t_0$. Then $u(t_0)$ is independent of the $B(t)$-process for $t \geqq t_0$ because the $B(t)$-process is a differential one, with differences involving t-values beyond t_0 independent of those involving t-values before t_0. Conversely suppose that choosing $u(0)$ independent of the $B(t)$-process for $t \geqq 0$ implies that $u(t_0)$ will be independent of the $B(t)$-process for $t \geqq t_0$, for all $t_0 > 0$. Then if $u(0)$ is so chosen,

$$u(0) + \int_0^{t_0} e^{\beta\tau}\, dB(\tau)$$

and therefore

$$\int_0^{t_0} e^{\beta\tau}\, dB(\tau)$$

are independent of the $B(t)$-process for $t \geqq t_0$. This fact implies that the preceding integral determines a differential process, that is, if $t_1 < \cdots < t_n$, the integrals

$$\int_{t_j}^{t_{j+1}} e^{\beta\tau}\, dB(\tau)$$

are mutually independent. Then (applying this fact to subintervals of the intervals (t_j, t_{j+1})

$$\int_{t_j}^{t_{j+1}} e^{-\beta\tau}\, d_t \int_{t_j}^{t} e^{\beta\tau}\, dB(\tau), \qquad j = 1, \cdots, n$$

are mutually independent, and these repeated integrals are simply

$$B(t_{j+1}) - B(t_j) \qquad j = 1, \cdots, n-1.$$

The latter differences are therefore mutually independent, as was to be proved.

We shall need the following lemma.

LEMMA 3. *Suppose that $a < 1$, and let $x_0, x_1, \cdots$ be mutually independent chance variables with a common distribution function. If there is a chance variable y with a Gaussian distribution such that the distribution function of $\sum_{j=0}^{n-1} a^{n-j}x_j$ approaches that of y as $n \to \infty$, then the x_j have Gaussian distributions.*

Many of the hypotheses of the lemma are unnecessary, but its statement is general enough for our purposes, and the proof will apply to a situation to be discussed in §5, where the distribution of y will not be Gaussian. The hypotheses imply that the distribution of $\sum_m^n a^j x_j$ approaches that of $a^{m-1}y$ as $n \to \infty$. If $\varphi(t)$ is the characteristic function of x_1 and $\psi(t)$ that of y, writing $\sum_1^n a^j x_j$ in the form $ax_1 + \sum_2^n a^j x_j$ shows that

$$\psi(t) = \varphi(at) \cdot \psi(at).$$

Solving for φ we find that it is the characteristic function of a Gaussian distribution, as was to be proved.

In the physical picture under discussion, further conditions on the solution of (3.3) are known. In fact the Brownian movement is simply a visible example of molecular or near molecular movement. The general principles of such movements are therefore applicable, and the principle of equipartition of energy leads to the Maxwell distribution of velocities. Let k be the Boltzmann constant, and T the absolute temperature. We can formulate the significance of the Maxwell distribution (as much as we shall need it) as follows.

M_1. *Tendency towards the Maxwell distribution.* Whatever the initial distribution of $u(0)$, the transition probabilities have the property that when $t \to \infty$ the distribution function of $u(t)$ converges to the Gaussian distribution function with mean 0 and variance kT/m. (Here m is the mass of the moving particle.)

M_2. *Stability of the Maxwell distribution.* If $u(0)$ is independent of later residual random impacts, and if it has the Gaussian distribution described in M_1, $u(t)$ will have this same distribution for every positive t.

These two statements are closely related, but neither apparently can be deduced from the other without further assumptions. Since these principles act the part of a *deus ex machina* in a discussion of the Langevin equation, we shall use them as little as possible. It will usually be sufficient to use a weakened form of M_1:

M_1'. There is an initial distribution of $u(0)$, such that the transition probabilities have the property that when $t \to \infty$ the distribution function of $u(t)$ converges to the Gaussian distribution function with mean 0 and variance kT/m. It is understood here as before that $u(0)$ is to be independent of later residual random impacts.

Conditions M_1 and M_2 restrict the possibilities for the $B(t)$-process. In fact suppose that condition M_1' is satisfied. Then (3.11) shows that

$$e^{-\beta t} \int_0^t e^{\beta \tau} \, dB(\tau)$$

is nearly Gaussian for large t, with mean 0 and variance kT/m. We write this integral as a sum, replacing t by nt:

$$e^{-\beta n t} \int_0^{nt} e^{\beta \tau} \, dB(\tau) = \sum_0^{n-1} e^{-\beta t(n-j)} x_j, \tag{3.15}$$

where

$$x_j = \int_{jt}^{(j+1)t} e^{\beta(\tau - jt)} \, dB(\tau). \tag{3.16}$$

Since the $B(t)$-process is a differential process, and is temporally homogeneous, the x_j are mutually independent, with identical distributions. According to the lemma, the right side of (3.15) cannot become Gaussian for large t unless the distribution of x_1 is Gaussian. Thus, since t is arbitrary in the above discussion,

$$\int_s^t e^{\beta \tau} \, dB(\tau)$$

has a Gaussian distribution for all s, t. Since the chance variables

$$\int_{jt/n}^{(j+1)t/n} e^{\beta\tau}\, dB(\tau), \qquad j = 1, \cdots, n \tag{3.17}$$

are mutually independent and Gaussian, the chance variable

$$\sum_{0}^{n-1} e^{-\beta jt/n} \int_{jt/n}^{(j+1)t/n} e^{\beta\tau}\, dB(\tau) \tag{3.18}$$

also has a Gaussian distribution. When n becomes infinite, (3.18) becomes $B(t) - B(0)$, with probability 1. The latter difference thus has a Gaussian distribution, with mean 0. The $B(t)$-process therefore has finite second moments $\sigma^2(t) = t\sigma^2$ as defined in (3.4). According to (3.7) the last term in (3.11), which we now know has a Gaussian distribution, has mean 0 and variance $\sigma^2(1 - e^{-2\beta t})/2\beta$. Then $u(t) - e^{-\beta t}u(0)$ has this same distribution. The variance becomes $\sigma^2/2\beta$ when $t \to \infty$, and therefore, according to M_1', $\sigma^2 = 2\beta kT/m$. Thus condition M_1' completely determines the $B(t)$-process. We show next that condition M_2 determines this same $B(t)$-process. In fact suppose condition M_2 is true, and assign to $u(0)$ the distribution of that condition. Then $u(0)$ is independent of the integral in (3.11), and in (3.11), $u(t)$ (which has a Gaussian distribution, according to condition M_2) is expressed as the sum of two independent chance variables, of which the first is Gaussian. The characteristic function of the second is the quotient of the characteristic functions of two Gaussian distributions, and is therefore the characteristic function of a Gaussian distribution. Thus the expression

$$e^{-\beta t} \int_0^t e^{\beta\tau}\, dB(\tau) \tag{3.19}$$

has a Gaussian distribution for all t, and this implies, as above, that $B(t) - B(0)$ has a Gaussian distribution, with variance $\sigma^2 t$. The variances on the right side of (3.11) add up to that on the left, giving an equation for σ^2:

$$\frac{kT}{m} = e^{-2\beta t}\frac{kT}{m} + \frac{1 - e^{-2\beta t}}{2\beta}\sigma^2. \tag{3.20}$$

Then $\sigma^2 = 2\beta kT/m$ as above.

We can now finally derive the O. U. velocity process as the solution of the Langevin equation. Suppose the $B(t)$-process is the one derived in the preceding paragraphs, and choose the chance variable $u(0)$ to be independent of the $B(t)$-process for $t \geqq 0$. Then $u(0)$ is independent of the integral in (3.11), and this means that the conditional distribution of $u(t)$ for $u(0) = u_0$ is Gaussian, with mean 0 and variance $kT(1 - e^{-2\beta t})/m$. Moreover, (3.11) implies

$$u(s + t) = u(s)e^{-\beta t} + e^{-\beta(s+t)} \int_s^{s+t} e^{\beta\tau}\, dB(\tau). \tag{3.21}$$

As we have seen, $u(s)$ is independent of the $B(t)$-process as far as it appears in (3.21) and therefore is independent of the integral. Thus the transition probabilities from s to $s + t$ are the same as those from 0 to t, which are precisely those of the O. U. process. Incidentally it follows that the full condition M_1 is satisfied. Finally, if $u(0)$ is not only supposed independent of the $B(t)$-process, for $t \geqq 0$, but also is supposed to have a Gaussian distribution with mean 0 and variance kT/m, the same will be true of $u(t)$ (as can be calculated from (3.11)) and condition M_2 is thus satisfied. We can summarize all our results as follows.

THEOREM 3. *Let the $B(t)$-process be a temporally homogeneous differential process. Then* (3.11) *furnishes the solution of* (3.3). *The following conditions on the solution are equivalent.*

(i) *The solution satisfies condition* M_1'.

(ii) *The solution satisfies condition* M_1.

(iii) *The solution satisfies condition* M_2.

(iv) $B(t) - B(0)$ *has a Gaussian distribution, with mean* 0 *and variance* $\sigma^2 t = t2\beta kT/m$.

If the above conditions are satisfied, $u(t) - e^{-\beta t}u(0)$ *will have a Gaussian distribution with mean* 0 *and variance* $kT(1 - e^{-2\beta t})/m$; *if* $u(0)$ *is independent of the $B(t)$-process for* $t \geqq 0$, $u(s)$ *is independent of the $B(t)$-process for* $t \geqq s$ *for all* $s > 0$, *and the transition probabilities of the $u(t)$-process are those of the O. U. velocity process. If in addition* $u(0)$ *has the Gaussian distribution with mean* 0 *and variance* kT/m, *the $u(t)$-process becomes the O. U. process, with* $m = 0$, $\sigma_0^2 = kT/m$.

The Langevin equation gives a physical interpretation to every property of the O. U. process. It is interesting to verify that as $h \to 0$ the correlation coefficient of the pair $B(s + h) - B(s)$, $u(t)$ (any s, t) goes to 0. In this sense then, $dB(s)$, the effect of the residual random impacts at time s, is independent of the velocity at any particular time t. Since in (3.11) $u(t)$ is written in terms of the $B(t)$-process, $u(t)$ is of course not independent of this process.

We have written $u(t)$ in terms of the $B(t)$-process. It is easy to write $x(t)$ in terms of the $B(t)$ process by combining (2.1) with (3.11):

$$x(t) = x(0) + \frac{1 - e^{-\beta t}}{\beta} u(0) + \frac{1}{\beta} \int_0^t [1 - e^{-\beta(t-\tau)}]\, dB(\tau). \tag{3.22}$$

Instead of finding the distributions of the displacement and velocity processes, and their correlations, as at the beginning of the paper, we could easily derive the desired results using (3.11) and (3.22). The various expectations can be calculated using (3.7).

In physical applications, the correlation function $E\{u(s)u(s + t)\}$ is sometimes wanted as a time average. Now the transformation S_h taking $B(t) - B(0)$ into $B(t + h) - B(h)$ preserves the $B(t)$ probability relations (temporal homogeneity), and the family of transformations $\{S_h\}$ is well known to be metrically transitive.[9] Then applying the ergodic theorem to the function $u(0)u(h)$, considered as a function of the $B(t)$, we find that

[9] Cf. for example Doob, (3) p. 125.

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t u(s)u(s+h)\,ds = E\{u(0)u(h)\} = \frac{kT}{m} e^{-\beta|h|}, \tag{3.23}$$

with probability 1, that is for almost all functions $u(t)$. The ergodic theorem was applied to the $B(t)$-process in essentially this way by Wiener ((18) p. 169) who has been interested in the harmonic analysis of functions like the $u(t)$ discussed here. The work of this paper verifies in this particular case the importance Wiener gave to the functions of the $B(t)$-process of the type (3.5).

There is no difficulty in extending the above results to bound particles. For example, the Langevin equation of the harmonically bound particle is

$$\frac{du}{dt} = -\beta u - \omega^2 x + A(t), \tag{3.24}$$

which in our treatment becomes

$$du = -\beta u\,dt - \omega^2 x\,dt + dB. \tag{3.25}$$

The usual methods of solving the differential equation (3.24) are still applicable to (3.25) and again the distribution of u turns out to be Gaussian.[10] The distribution of displacements is then obtained as above.

4. The $B(t)$-impact process

When the $B(t)$-process and the initial conditions on $u(0)$ are given, the solution $u(t)$ is determined by (3.11). Conversely if the solution $u(t)$ is known, $B(t)$ is determined by the equation

$$B(t) - B(0) = \beta \int_0^t u(s)\,ds + u(t) - u(0) \tag{4.1}$$

which is derived immediately from (3.3). The O. U. velocity distribution for the $u(t)$-process can therefore be given only by the $B(t)$-process described in §3. We shall investigate the possibility that a different choice of the $B(t)$-process might have led to a different velocity process compatible with the known physical conditions like M_1 and M_2. If we suppose that $u(0)$ can be chosen so that the velocity at each moment is independent of subsequent residual random impacts, then we have seen that the $B(t)$-process must be differential, and is then uniquely determined by conditions M_1' or M_2. Any velocity process other than the O. U. process satisfying the Langevin equation and M_1' or M_2 would therefore imply dependence between velocity and later residual impacts. This is really another way of saying that the frictional resistance cannot be considered as proportional to the velocity. Before going further we put a condition going back to Maxwell in its modern setting. We formulate a hypothesis M_3 as follows.

M_3. In two or more dimensions (using any orthogonal axes) the velocity components are mutually independent.

[10] Cf. Ornstein and Wijk (16) and Wijk (17). The $B(t, \Delta)$ used in these papers corresponds formally to our $dB(t)$. The difference is that it is possible to give a precise description of the $B(t)$-distribution.

In conjunction with the following lemma, due to Kaç ((8) p. 278), hypothesis M_3 implies that all quantities linear in the displacement or velocity functions have Gaussian distributions.

LEMMA.[11] *Let* $(x_1, y_1), \cdots, (x_n, y_n)$ *be* $2n$ *chance variables with the property that the sets of chance variables*

$$\{x_j \cos\theta + y_j \sin\theta, j = 1, \cdots, n\}\{-x_j \sin\theta + y_j \cos\theta, j = 1, \cdots, n\}$$

are mutually independent for each value of θ. *Then* $(x_1, \cdots, x_n)$ *have an n-variate Gaussian distribution or a singular Gaussian distribution.*

We can combine the Maxwell hypotheses to obtain another justification of the O. U. velocity process.

THEOREM 4. *Let the* $B(t)$*-process be any process such that the distribution of* $B(t_2) - B(t_1)$ *or of any quantity depending on such differences is unaffected by translations of the t-axis, and that the integral* (3.5) *can be defined as the limit in probability of the usual sums, with* (3.6) *valid. Then if* $u(t)$ *is defined by* (3.11), *and if conditions* M_2 *and* M_3 *are satisfied, the* $B(t)$*-process must be precisely that finally obtained in* §3, *leading to the O. U. velocity process.*

Suppose that condition M_2 is satisfied, and let $u(0)$ be fixed as in that condition. Just as in §3, (3.11) then implies that the integral

$$B^*(t) = \int_0^t e^{\beta\tau}\, dB(\tau)$$

has a Gaussian distribution with mean 0 and variance $(kT/m)(e^{2\beta t} - 1)$. If condition M_3 is also satisfied, $B^*(t_2) - B^*(t_1)$, and more generally any finite set of such differences, has a one or more dimensional Gaussian distribution. Using the fact that the distribution of $e^{-\beta s}[B^*(s + t) - B^*(s)]$ is the same as that of $B^*(t)$, in evaluating the expectations in the following equation

$$E\{B^*(s + t)^2\} = E\{[B^*(s) + [B^*(s + t) - B^*(s)]]^2\}, \tag{4.2}$$

we find that $B^*(s) = B^*(s) - B^*(0)$ and $B^*(s + t) - B^*(s)$ are uncorrelated. These two variables are therefore independent. Going further, similar calculations show that any differences $B^*(t_2) - B^*(t_1)$, $B^*(s_2) - B^*(s_1)$ with $0 \leqq s_1 < s_2 \leqq t_1 < t_2$ are independent. Using the fact (derived from condition M_3) that any finite set of differences has a multivariate Gaussian distribution, the $B^*(t)$-process is thus a differential process. This means, by a method we have used above, that the $B(t)$-process is a differential process, leading to the O. U. velocity distribution, because condition M_2 is satisfied.

It is easily seen from counterexamples that Theorem 4 is no longer correct if condition M_1 is supposed instead of condition M_2.

5. Velocity processes not subject to Maxwell's laws

In all the above work the role of the Maxwell velocity distribution has been fundamental. In certain studies, however, other distributions play a somewhat

[11] The result is stated slightly incorrectly by Kaç.

analogous role.[12] It is interesting to note that the Langevin equation can be solved to give a distribution whose transition probabilities are asymptotically any of the symmetric stable distributions classified by Lévy ((14) §30, §56, §57). Such a distribution has characteristic function

$$e^{-\sigma_0^2|z|^\gamma}$$

where σ_0^2 is a positive parameter and $0 < \gamma \leqq 2$. The Gaussian distribution is obtained when $\gamma = 2$. The parameter σ_0^2 plays the role of the variance, although the second moment is never finite when $\gamma < 2$. The velocity process we shall derive will be called the O. U. (γ) process. It is the O. U. process when $\gamma = 2$. The O. U. (γ) process can be described as follows.

1. The process is temporally homogeneous, that is translations of the t-axis do not affect the probability distributions.
2. The process is a Markoff process.
3. For each fixed t, $u(t)$ has a symmetric stable distribution with parameter value σ_0^2, exponent γ. The conditional distribution of $u(s + t)$ for $u(s) = u_0$ is the stable distribution symmetric about $u_0e^{-\beta t}$, with parameter value $\sigma_0^2(1 - e^{-\gamma\beta|t|})$ and exponent γ.

We can obtain this process as a solution of the Langevin equation by choosing the $B(t)$-process properly. In fact, let the $B(t)$-process be the temporally homogeneous differential process in which $B(s + t) - B(s)$ has a symmetric stable distribution with exponent γ and parameter value $\sigma^2 t$. Let $u(t)$ be the corresponding solution of the Langevin equation, given by (3.11). If y is the sum of two independent chance variables with stable symmetric distributions, having parameter values σ_1^2, σ_2^2, and with the same exponent γ then y also has a symmetric stable distribution, with the same exponent, γ, and with parameter value $\sigma_1^2 + \sigma_2^2$. From this fact it is simple to check that the integral (3.5) in the present case has a symmetric stable distribution with exponent γ and parameter value.

$$\int_a^b |f(t)|^\gamma \, dt.$$

If $u(0)$ is given a symmetric stable distribution independent of the $B(t)$-process for $t \geqq 0$, with parameter value $\sigma^2/\gamma\beta$, the distribution of $u(t)$ can be calculated, using characteristic functions, and is found to be symmetric and stable, with exponent γ and parameter value $\sigma^2/\gamma\beta$. The $u(t)$ thus defined determines an O. U. (γ) process, with the above three properties, setting $\sigma_0^2 = \sigma^2/\gamma\beta$.

We shall not spend any time on the details of the analysis of the O. U. (γ) process, since the work runs parallel to that for the case $\gamma = 2$, already discussed. There are, however, a few essential differences. If $v(t)$ is determined by the equation

$$(1.2.1) \qquad v(t) = t^{1/\gamma} u\left(\frac{1}{\gamma\beta} \log t\right), \qquad t > 0,$$

[12] Cf. Holtzmark (7).

the $v(t)$ process can be analyzed using (3.11). The $v(t)$-process has the same distribution as the $B(t)$-process just described. The continuity properties of the velocity process can now be derived from those of the $v(t)$-process, which are known. When $\gamma < 2$, the velocity function $u(t)$ is no longer a continuous function of t with probability 1, but is certain to have discontinuities. These discontinuities are however non-oscillatory (jumps).[13] We omit the details of the analogue of Theorem 1.3. Theorem 2.1 is still true if $\gamma \geqq 1$. The considerations of §3 have their obvious counterparts here. Lemma 3 played an essential role, but its statement and proof are correct if the variable y of the lemma is supposed to have a symmetric stable distribution and if the conclusion is that the x_j have a symmetric stable distribution with the same exponent as y.

UNIVERSITY OF ILLINOIS
AND
INSTITUTE FOR ADVANCED STUDY

BIBLIOGRAPHY

1. S. BERNSTEIN, Comptes Rendus de l'Académie des Sciences de l'URSS, N.S. (1934) pp. 1–9.
2. S. BERNSTEIN, Comptes Rendus de l'Académie des Sciences de l'URSS, N.S. (1934), pp. 361–364.
3. J. L. DOOB, Transactions of the American Mathematical Society 42 (1937), pp. 107–140.
4. J. L. DOOB, Duke Mathematical Journal 6 (1940), pp. 290–306.
5. J. L. DOOB, Transactions of the American Mathematical Society 47 (1940), pp. 455–486.
6. G. L. DE HAAS-LORENTZ, *Die Brownsche Bewegung und einige verwandte Erscheinungen*, Braunschweig (1913).
7. J. HOLTZMARK, Annalen der Physik 58 (1919), pp. 577–630.
8. M. KAÇ, American Journal of Mathematics 61 (1939), pp. 726–728.
9. A. KHINTCHINE, Mathematische Annalen 109 (1934), pp. 604–615.
10. A. KHINTCHINE, Ergebnisse der Mathematik 4 No. 3.
11. G. KRUTKOW, Physikalische Zeitschrift der Sowjet-Union 5 (1934), pp. 287–300
12. P. LÉVY, Pisa Annali Series 2 vol. 3 (1934), pp. 337–366.
13. P. LÉVY, Pisa Annali Series 2 vol. 4 (1935), pp. 217–218.
14. P. LÉVY, *Théorie de l'addition des variables aléatoires*, Paris 1937.
15. L. S. ORNSTEIN AND G. E. UHLENBECK, Physical Review 36 (1930), pp. 823–841.
16. L. S. ORNSTEIN AND W. R. VAN WIJK, Physica 1 (1934), pp. 235–254, errata p. 966.
17. W. R. VAN WIJK, Physica 3 (1936), pp. 1111–1119.
18. N. WIENER AND R. E. A. C. PALEY, *Fourier Transforms in the Complex Domain*, American Mathematical Society Colloquium Publications Vol. XIX.

[13] For further details, cf. Lévy (14) Chapter VII.