Difference Methods for Stochastic Ordinary Differential Equations

By Joel N. Franklin

1. Introduction. In the theory of random signals and noise [1], [2], [3], [4] we encounter linear or nonlinear systems with forcing terms which are random processes. Such systems often may be described by sets of differential equations of the form

$$(1.1) \qquad \frac{dx_i}{dt} = f_i(x_1, \cdots, x_m) + n_i(t), \qquad x_i(0) = \xi_i \qquad (i = 1, \cdots, m),$$

where the functions f_i are deterministic and the forcing functions $n_i(t)$ are random. Each function $n_i(t)$ is "white noise," with cross-correlations

$$(1.2) \qquad En_i(t)n_j(t - \tau) = \sigma_{ij}\delta(\tau).$$

The matrix (σ_{ij}) is positive semi-definite, since

$$\sum_i \sum_j \lambda_i\lambda_j\sigma_{ij} = E\int_0^1 [\sum_i \lambda_i n_i(t)]^2 dt \geqq 0.$$

In many cases of interest, however, the matrix σ_{ij} is not positive definite. For example, if

$$\ddot{x} + \mu\dot{x} + \sin x = n(t)$$

is described as a first-order system

$$\frac{dx_1}{dt} = x_2,$$

$$\frac{dx_2}{dt} = -\sin x_1 - \mu x_2 + n(t),$$

then the matrix (σ_{ij}) has the form

$$(\sigma_{ij}) = \begin{pmatrix} 0 & 0 \\ 0 & \sigma_{22} \end{pmatrix}.$$

Numerical methods for linear systems of the form (1.1) are given in [7] and [8]. These methods do not apply to nonlinear equations.

The purpose of this paper is to provide an analysis of the random error which occurs when the stochastic, nonlinear equations (1.1) are approximated by finite-difference equations. For $\Delta t > 0$ we approximate the random vector $x(t)$ defined by (1.1) by the vector $x^{(n)}$ defined by the difference equation

$$(1.3) \qquad x^{(n+1)} = x^{(n)} + \Delta t f(x^{(n)}) + g^{(n)} \qquad (n = 0, 1, \cdots), \qquad x^{(0)} = \xi,$$

where $t = n\Delta t$ and where $g^{(n)}$ is a pseudo-random vector simulating a random

Received November 10, 1964. Revised February 18, 1965.

Gaussian vector g with

$$Eg_i = 0, \qquad Eg_ig_j = \sigma_{ij}\Delta t \qquad (i, j = 1, \cdots, m). \tag{1.4}$$

The main result is the following: Suppose that the functions $f_i(x)$ have continuous, bounded derivatives up to the fourth order. Let R be any bounded region of m-dimensional space. Let $R_1 \subset R \subset R_2$, where R_1 is a region whose boundary lies in R and where R_2 is a region which includes the boundary of R. Then we shall prove that, as $\Delta t \to 0$ with $n = [t/\Delta t]$,

$$\Pr\{x^{(n)} \in R_1\} + O(\Delta t) \leqq \Pr\{x(t) \in R\} \leqq \Pr\{x^{(n)} \in R_2\} + O(\Delta t). \tag{1.5}$$

We shall also prove convergence in distribution.

The author hopes that this paper will stimulate research on more sophisticated difference methods for stochastic differential equations. The usual error-analyses made for Adams or Runge-Kutta methods do not apply to stochastic equations (1.1), since any sample of the white-noise process $n_i(t)$ with $\sigma_{ii} > 0$ is required to be everywhere unbounded, discontinuous, and nondifferentiable.

2. The Equivalent Markov Process. To many mathematicians the definition of white noise presented in current books and papers lacks rigor. Therefore, we shall *define* the differential equations (1.1) to be equivalent to the equations (2.1) for a continuous Markov process $x_i(t)$ with initial state $x_i(0) = \xi_i$ and with limit-moments

$$\begin{aligned} &\lim_{\Delta t \to 0} E\,\frac{\Delta x_i}{\Delta t} = f_i(x_1, \cdots, x_m), \\ &\lim_{\Delta t \to 0} E\,\frac{\Delta x_i \Delta x_j}{\Delta t} = \sigma_{ij}, \\ &\lim_{\Delta t \to 0} E\,\frac{|\Delta x_i \Delta x_j \Delta x_k|}{\Delta t} = 0 \qquad (i, j, k = 1, \cdots, m). \end{aligned} \tag{2.1}$$

For all $t > 0$ these limits are supposed to be attained uniformly in x. Under these conditions, as it is proved in [2], the random vector x at time $t > 0$ has a probability-density $P(\xi \to x, t)$ defined by the Fokker-Planck partial differential equation

$$\frac{\partial P}{\partial t} = -\sum_{i=1}^{m} \frac{\partial}{\partial x_i}[f_i(x)P] + \frac{1}{2}\sum_{i,j=1}^{m}\sum \sigma_{ij}\frac{\partial^2 P}{\partial x_i \partial x_j} \tag{2.2}$$

with

$$P(\xi \to x, 0) = \delta(x - \xi). \tag{2.3}$$

The density P satisfies the relations

$$P \geqq 0, \qquad \int P\,dx = 1 \tag{2.4}$$

and the Chapman-Kolmogorov identity for $\Delta t > 0$:

$$P(\xi \to x, t + \Delta t) = \int P(\xi \to \eta, \Delta t) P(\eta \to x, t)\,d\eta. \tag{2.5}$$

If $m \leqq 2$, it is conceivable to apply numerical methods to the Fokker-Planck equation. If $m > 2$, existing numerical methods are usually impractical.

3. The Approximate Probability Density. Corresponding to a fixed time-increment $\Delta t > 0$, we shall define an approximate probability-density $Q(\xi \to x, t) = Q(\xi \to x, t; \Delta t)$. For $0 < t \leqq \Delta t$, we approximate the Fokker-Planck equation (2.2) by replacing $f_i(x)$ by the initial value $f_i(\xi)$. We then write

$$\frac{\partial Q}{\partial t} = -\sum_{i=1}^{m} f_i(\xi) \frac{\partial Q}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{m}\sum \sigma_{ij} \frac{\partial^2 Q}{\partial x_i \cdot \partial x_j} \qquad (0 < t \leqq \Delta t), \tag{3.1}$$

$$Q(\xi \to x, 0)' = \delta(x - \xi). \tag{3.2}$$

We now use the analogue of the Chapman-Kolmogorov identity to define Q at later times $t = n\Delta t$. For $n = 1, 2, \cdots$ we recursively define

$$Q(\xi \to x, (n+1)\Delta t) = \int Q(\xi \to \eta, \Delta t) Q(\eta \to x, n\Delta t)\, d\eta \tag{3.3}$$

or, equivalently, for $n = 1, 2, \cdots$

$$Q(\xi \to x, (n+1)\Delta t) = \int Q(\xi \to \eta, n\Delta t) Q(\eta \to x, \Delta t)\, d\eta. \tag{3.4}$$

By (3.1) and (3.2), the function $Q(\eta \to x, \Delta t)$ is the probability density of an m-dimensional random Gaussian variable x with mean values and second moments

$$\mu_i \equiv Ex_i = \eta_i + (\Delta t) f(\eta_i), \qquad E(x_i - \mu_i)(x_j - \mu_j) = (\Delta t)\sigma_{ij}. \tag{3.5}$$

Therefore, if we set $\eta = x^{(n)}$ and $x = x^{(n+1)}$ in (3.4), we see that

$$x^{(n+1)} = x^{(n)} + \Delta t f(x^{(n)}) + g^{(n)}, \tag{3.6}$$

where $g^{(n)}$ is an m-dimensional random Gaussian variable which is independent of $x^{(n)}$ and which has the first and second moments

$$Eg_i^{(n)} = 0, \qquad Eg_i^{(n)} g_j^{(n)} = (\Delta t)\sigma_{ij} \qquad (i, j = 1, \cdots, m). \tag{3.7}$$

For later analysis we shall require an explicit form for $Q(\xi \to x, \Delta t)$. Let U be a real, unitary matrix which diagonalizes (σ_{ij}):

$$UU^* = I, \qquad U(\sigma_{ij})U^* = \operatorname{diag}(\lambda_1, \cdots, \lambda_r, 0, \cdots, 0), \tag{3.8}$$

where $\lambda_1 \geqq \cdots \geqq \lambda_r > 0$ are the positive eigenvalues of the semi-definite matrix (σ_{ij}). We assume $r \geqq 1$. For $0 < t \leqq \Delta t$ set

$$z = U(x - \xi - f(\xi)t). \tag{3.9}$$

The equations (3.1), (3.2) now give

$$\begin{gathered} Q(\xi \to x, t) = H(z, t), \qquad H(z, 0) = \delta(z) = \prod_{i=j}^{m} \delta(z_i), \\ \frac{\partial H}{\partial t} = \frac{1}{2} \sum_{i=1}^{r} \lambda_i \frac{\partial^2 H}{\partial z_i^2}. \end{gathered} \tag{3.10}$$

Therefore,

$$Q(\xi \to x, \Delta t) = h(z_1, \cdots, z_r)\delta(z_{r+1}) \cdots \delta(z_m), \tag{3.11}$$

where

$$h(z_1, \cdots, z_r) = \frac{(2\pi\Delta t)^{-r/2}}{\sqrt{\lambda_1 \cdots \lambda_r}} \exp\left(-\frac{1}{2\Delta t}\sum_{i=1}^{r}\frac{z_i^2}{\lambda_i^2}\right). \tag{3.12}$$

An integral of Q times a function ψ takes the form

$$\begin{aligned}\int Q(\xi \to x, \Delta t)\psi(x)\, dx &= \int \cdots \int Q\psi\, dx_1 \cdots dx_m \\ &= \int \cdots \int h(z_1, \cdots, z_r)\psi(\xi + f(\xi)\Delta t + U^* z^1)\, dz_1 \cdots dz_r,\end{aligned} \tag{3.13}$$

where $z^1 = \text{col}(z_1, \cdots, z_r, 0, \cdots, 0)$.

4. The Related Difference Method. To use the difference equation (3.6) in practice, we must have some method of simulating samples from the m-dimensional Gaussian distribution with first and second moments given by (3.7). This numerical problem is discussed in [5], [6], and [8]. There is a slight complication if (σ_{ij}) is not positive definite.

We desire to write (σ_{ij}) in the form

$$(\sigma_{ij}) = TT^*. \tag{4.1}$$

If the eigenvalues λ_i and the unitary matrix U used in (3.8) are easy to compute, we may simply set

$$T = U^* \operatorname{diag}(\sqrt{\lambda_1}, \cdots, \sqrt{\lambda_r}, 0, \cdots, 0). \tag{4.2}$$

If the λ_i are unknown, we find T as follows: By successively completing squares, we may write

$$\sum_{i,j=1}^{m}\sum \sigma_{ij}x_i x_j = y_1^2 + y_2^2 + \cdots + y_r^2, \tag{4.3}$$

where, for some r indices $\alpha < \beta < \cdots < \rho$,

$$\begin{aligned} y_1 &= c_\alpha x_\alpha + (\quad)x_{\alpha+1} + \cdots + (\quad)x_m, \qquad c_\alpha \neq 0, \\ y_2 &= c_\beta x_\beta + (\quad)x_{\beta+1} + \cdots + (\quad)x_m, \qquad c_\beta \neq 0, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ y_r &= c_\rho x_\rho + (\quad)x_{\rho+1} + \cdots + (\quad)x_m, \qquad c_\rho \neq 0. \end{aligned} \tag{4.4}$$

If (4.4) is written in the form $y = Sx$, where S is a matrix with r rows and m columns, the identity (4.3) shows that $(\sigma_{ij}) = S^*S$. Therefore, we may let $T = S^*$ to satisfy (4.1).

We now set

$$g^{(n)} = \sqrt{\Delta t}\, T w^{(n)}. \tag{4.5}$$

The components of the vector $w^{(n)}$ are required to simulate independent samples from the Gaussian distribution with zero mean and unit variance. This problem is solved by the technique of Box and Muller [5]. The factorization (4.1) produces the required moments (3.7).

5. Analysis of the Error. The true solution $x(t)$ is a random vector with probability-density $P(\xi \to x, t)$. The approximate solution $x^{(n)}$ is a random vector with probability-density $Q(\xi \to x, n\Delta t; \Delta t)$. It is too much to hope that

$$\lim_{\Delta t \to 0; n\Delta t \to t} Q(\xi \to x, n\Delta t; \Delta t) = P(\xi \to x, t). \tag{5.1}$$

In fact, even if there is no randomness, i.e., if all $\sigma_{ij} = 0$, the relation (5.1) does not hold. If there is no randomness, both P and Q are delta functions. The spike of P occurs at the solution of the differential equation $dx/dt = f(x)$. The spike of Q occurs at the solution of the difference equation $\Delta x = f(x)\Delta t$. Although the spikes of P and Q occur at nearby points, the arithmetic difference between the two delta functions does not tend to zero.

Let R be a region, for example, the open unit sphere. We may conjecture that

$$\Pr\{x^{(n)} \in R\} \to \Pr\{x(t) \in R\} \tag{5.2}$$

as $\Delta t \to 0$ and $n\Delta t \to t$. This conjecture is also false. Again suppose that all $\sigma_{ij} = 0$. Let $x(t)$ lie on the boundary of R. If the nonrandom difference-approximation $x^{(n)}$ approaches $x(t)$ from the interior of R, then for all $\Delta t > 0$

$$\Pr\{x^{(n)} \in R\} = 1 \quad \text{but} \quad \Pr\{x(t) \in R\} = 0.$$

Therefore, we must prove a weaker result. We will show that, for all sufficiently smooth statistics or testing functions $\phi(x)$, the expected value of the random number $\phi(x^{(n)})$ approaches the expected value of the random number $\phi(x(t))$ as $\Delta t \to 0$, where $n = [t/\Delta t]$, with an error which is $O(\Delta t)$.

For any function $\phi(x)$ with continuous derivatives up to order p, we define the number $\| \phi \|_p$ to equal the least upper bound, for all x, of the absolute values of the function ϕ and of all of its partial derivatives of order $\leqq p$. We shall assume that

$$\| f_i(x) \|_4 < \infty \qquad (i = 1, \cdots, m). \tag{5.3}$$

In differential equations (1.1) of physical interest this assumption is not too unreasonable if the given functions $f_i(x)$ are redefined as smooth at any discontinuities and as bounded with bounded derivatives for very large, physically unrealizable, moduli of the state-vector x.

THEOREM. *Suppose that* $\| f_i(x) \|_4 < \infty$ $(i = 1, \cdots, m)$. *For* $t > 0$ *let* $x(t)$ *be the random vector defined by the stochastic differential equations* (1.1). *For* $\Delta t > 0$ *and* $n = [t/\Delta t]$ *let* $x^{(n)}$ *be the random vector defined by the difference equations* (1.3). *Let* $\| \phi(x) \|_4 < \infty$. *Then*

$$E\phi(x^{(n)}) = E\phi(x(t)) + O(\Delta t) \quad \text{as } \Delta t \to 0. \tag{5.4}$$

Before proving the theorem, let us show how the inequalities (1.5) follow from the theorem. Let $R_1 \subset R \subset R_2$, and let the boundaries of these regions have positive distances from each other. Let $\phi_1(x)$ be defined as

$$\phi_1(x) = 1 \quad \text{in } R_1, \qquad \phi_1(x) = 0 \quad \text{outside } R. \tag{5.5}$$

Between the boundary of R_1 and the boundary of R define $\phi_1(x)$ as some function with continuous derivatives of order $\leqq 4$, with $1 \geqq \phi_1(x) \geqq 0$. We now apply (5.4) to obtain

$$\begin{aligned} \Pr\{x^{(n)} \in R_1\} &\leqq E\phi_1(x^{(n)}) \\ &= E\phi_1(x(t)) + O(\Delta t) \\ &\leqq \Pr\{x(t) \in R\} + O(\Delta t). \end{aligned} \tag{5.6}$$

Now define $\phi_2(x)$ as

$$\phi_2(x) = 1 \quad \text{in } R, \qquad \phi_2(x) = 0 \quad \text{outside } R_2, \tag{5.7}$$

with $\phi_2(x)$ varying smoothly from $\phi = 1$ to $\phi = 0$ between the boundaries of R and R_2. By (5.4)

$$\begin{aligned} \Pr\{x^{(n)} \in R_2\} &\geqq E\phi_2(x^{(n)}) \\ &= E\phi_2(x(t)) + O(\Delta t) \\ &\geqq \Pr\{x(t) \in R\} + O(\Delta t). \end{aligned} \tag{5.8}$$

This establishes the inequalities (1.5).

Proof of the Theorem. The expected values of $\phi(x^{(n)})$ and $\phi(x(t))$ are defined as

$$\begin{aligned} E\phi(x^{(n)}) &= \int Q(\xi \to x, n\Delta t; \Delta t)\phi(x)\, dx, \\ E\phi(x(t)) &= \int P(\xi \to x, t)\phi(x)\, dx. \end{aligned} \tag{5.9}$$

For $t > 0$ we have, as $\Delta t \to 0$ with $n = [t/\Delta t]$,

$$\int P(\xi \to x, n\Delta t)\phi(x)\, dx = \int P(\xi \to x, t)\phi(x)\, dx + O(\Delta t). \tag{5.10}$$

Therefore, it will suffice to prove that

$$\int [P(\xi \to x, n\Delta t) - Q(\xi \to x, n\Delta t; \Delta t)]\phi(x)\, dx = O(\Delta t). \tag{5.11}$$

We now define linear operators A and B which map functions of x into functions of ξ:

$$\begin{aligned} A\phi &= \int P(\xi \to x, \Delta t)\phi(x)\, dx, \\ B\phi &= \int Q(\xi \to x, \Delta t)\phi(x)\, dx. \end{aligned} \tag{5.12}$$

By the Chapman-Kolmogorov identity (2.5) for P and by the definition (3.3) for Q, we have

$$\begin{aligned} \int P(\xi \to x, n\Delta t)\phi(x)\, dx &= A^n\phi, \\ \int Q(\xi \to x, n\Delta t; \Delta t)\phi(x)\, dx &= B^n\phi. \end{aligned} \tag{5.13}$$

Therefore, by (5.11), we wish to prove that as $\Delta t \to 0$

$$(A^n - B^n)\phi = O(\Delta t) \quad \text{if} \quad n\Delta t \leqq t. \tag{5.14}$$

We factor the difference $A^n - B^n$ as follows:

$$A^n - B^n = \sum_{\nu=0}^{n-1} A^{n-\nu-1}(A - B)B^\nu. \tag{5.15}$$

First consider $A - B$. For any function $\psi(x)$ with $\| \psi \|_4 < \infty$ and any initial vector ξ define a function of t for $0 < t \leqq \Delta t$:

$$g(t) = \int [P(\xi \to x, t) - Q(\xi \to x, t)]\psi(x)\, dx. \tag{5.16}$$

Then

$$g(\Delta t) = (A - B)\psi. \tag{5.17}$$

First we note that

$$g(0) = \int [\delta(x - \xi) - \delta(x - \xi)]\psi(x)\, dx = 0 \tag{5.18}$$

Next we evaluate the first derivative:

$$\begin{aligned} g'(t) &= \int \left(\frac{\partial P}{\partial t} - \frac{\partial Q}{\partial t}\right)\psi(x)\, dx \\ &= \int [LP(\xi \to x, t) - MQ(\xi \to x, t)]\psi(x)\, dx, \end{aligned} \tag{5.19}$$

where, by (2.2) and (3.1),

$$\begin{aligned} LP &= -\sum \frac{\partial}{\partial x_i}[f_i(x)P] + \frac{1}{2}\sum\sum \sigma_{ij} \frac{\partial^2 P}{\partial x_i \partial x_j}, \\ MQ &= -\sum f_i(\xi)\frac{\partial P}{\partial x_i} + \frac{1}{2}\sum\sum \sigma_{ij} \frac{\partial^2 Q}{\partial x_i \partial x_j}. \end{aligned} \tag{5.20}$$

Introduce the adjoint operators L^* and M^*:

$$\begin{aligned} L^*\psi(x) &= \sum f_i(x)\frac{\partial \psi}{\partial x_i} + \frac{1}{2}\sum\sum \sigma_{ij} \frac{\partial^2 \psi}{\partial x_i \partial x_j}, \\ M^*\psi(x) &= \sum f_i(\xi)\frac{\partial \psi}{\partial x_i} + \frac{1}{2}\sum\sum \sigma_{ij} \frac{\partial^2 \psi}{\partial x_i \partial x_j}. \end{aligned} \tag{5.21}$$

Then, for $0 < t \leqq \Delta t$, (5.19) yields

$$g'(t) = \int [P(\xi \to x, t)L^*\psi - Q(\xi \to x, t)M^*\psi]\, dx. \tag{5.22}$$

In passing from (5.19) to (5.22) we have used the fact that our solutions P and Q of the Fokker-Planck equations (2.2) and (3.1) tend to zero as $\| x \| \to \infty$. Therefore,

$$g'(0) = L^*\psi - M^*\psi = 0 \quad \text{for} \quad x = \xi. \tag{5.23}$$

Next,

$$(5.24)\qquad g''(t) = \int [P(\xi \to x, t)(L^*)^2 \psi - Q(\xi \to x, t)(M^*)^2\psi]\, dx.$$

But

$$(5.25)\qquad |(L^*)^2\psi| \leqq \mu\|\psi\|_4 \quad \text{and} \quad |(M^*)^2\psi| \leqq \mu\|\psi\|_4$$

for some constant μ independent of the function ψ and of the point ξ. Therefore,

$$(5.26)\qquad |g''(t)| \leqq 2\mu\|\psi\|_4 \qquad (0 < t \leqq \Delta t).$$

Incidentally, $g''(0) \neq 0$ if the $f_i(x)$ are not constant. But for some t in the range $0 < t < \Delta t$,

$$(5.27)\qquad g(\Delta t) = g(0) + \Delta t g'(0) + \tfrac{1}{2}(\Delta t)^2 g''(t).$$

Therefore, by (5.17), (5.18), (5.23), and (5.26),

$$(5.28)\qquad |(A - B)\psi| \leqq \mu(\Delta t)^2 \|\psi\|_4 .$$

From the estimate (5.28) and from the factorization (5.15) it is clear that we must examine the effect of the iterated transformation B^ν on the norm $\|\ \ \|_4$. Again let $\psi(x)$ be any function with $\|\psi\|_4 < \infty$. We will show that there is a constant λ which is independent of ψ, of Δt, and of ξ such that, for $0 < \Delta t \leqq \epsilon$ independent of ξ and of ψ,

$$(5.29)\qquad \|B\psi\|_4 \leqq (1 + \lambda \Delta t)\|\psi\|_4 .$$

From the definition (5.12) and the identity (3.13) we have

$$(5.30)\qquad B\psi = \int \cdots \int h(z_1, \cdots, z_r)\, \psi(\xi + f(\xi)\Delta t + U^* z^1)\, dz_1 \cdots dz_r .$$

The Gaussian function h is positive and has integral $= 1$. Define a new variable $y = \xi + \Delta t f(\xi)$. Since $\|f_i\|_4 < \infty$, the function $B\psi$ has continuous partial derivatives of order $\leqq 4$ with respect to $\xi_1, \cdots, \xi_m$ and also with respect to $y_1, \cdots, y_m$. From (5.30) we have

$$|B\psi| \leqq \int \cdots \int h\, |\psi|\, dz_1 \cdots dz_r \leqq \|\psi\|_0 .$$

Differentiating (5.30) $p \leqq 4$ times with respect to $y_i, y_j, \cdots$, we find

$$(5.31)\qquad \frac{\partial^p B\psi}{\partial y_i \cdots} = \int \cdots \int h \frac{\partial^p}{\partial y_i \cdots} \psi(y + U^* z^1)\, dz_1 \cdots dz_r .$$

Therefore, if $p \leqq 4$,

$$(5.32)\qquad \left| \frac{\partial^p B\psi}{\partial y_i \cdots} \right| \leqq \|\psi\|_p .$$

But

$$(5.33)\qquad \frac{\partial}{\partial \xi_i} = \sum_{i=1}^{m} \frac{\partial y_j}{\partial \xi_i} \frac{\partial}{\partial y_j} = \frac{\partial}{\partial y_i} + \Delta t \sum_{i=1}^{m} \frac{\partial f_j(\xi)}{\partial \xi_i} \frac{\partial}{\partial y_j}.$$

The required inequality (5.29) now follows from repeated application of (5.32) and (5.33).

From (5.29) we find, for $0 < \Delta t \leqq \epsilon$,

$$\| B^\nu \phi \|_4 \leqq (1 + \lambda \Delta t)^\nu \| \phi \|_4 . \tag{5.34}$$

By (5.28) and (5.34)

$$|(A - B) B^\nu \phi | \leqq \mu (\Delta t)^2 \| B^\nu \phi \|_4 \leqq \mu (\Delta t)^2 (1 + \lambda \Delta t)^\nu \| \phi \|_4 . \tag{5.35}$$

By the definition (5.12),

$$| A \psi | \leqq \int P(\xi \rightarrow x, \Delta t) \, dx \, \| \psi \|_0 = \| \psi \|_0 . \tag{5.36}$$

From (5.36) and (5.35) we have

$$| A^{n-1-\nu} (A - B) B^\nu \phi | \leqq \mu (\Delta t)^2 (1 + \lambda \Delta t)^\nu \| \phi \|_4 . \tag{5.37}$$

From (5.15) we conclude, for $n \Delta t \leqq t$,

$$\begin{aligned} |(A^n - B^n) \phi | &\leqq \mu (\Delta t)^2 \| \phi \|_4 (\lambda \Delta t)^{-1} \{ (1 + \lambda \Delta t)^n - 1 \} \\ &\leqq \mu \| \phi \|_4 \lambda^{-1} (e^{\lambda t} - 1) \Delta t = O(\Delta t). \end{aligned} \tag{5.38}$$

This completes the proof of the theorem.

Let us define the distribution functions

$$\begin{aligned} F(a_1, \cdots, a_m) &= \Pr \{ x_1 < a_1, \cdots, x_m < a_m \}, \\ F_{\Delta t}(a_1, \cdots, a_m) &= \Pr \{ x_1^{(n)} < a_1, \cdots, x_m^{(n)} < a_m \}. \end{aligned} \tag{5.39}$$

At points of continuity of the distribution function F we have the following result:

THEOREM. *If $(a_1, \cdots, a_m)$ is a point of continuity of the distribution function F, then*

$$\lim_{\Delta t \rightarrow 0} F_{\Delta t}(a_1, \cdots, a_m) = F(a_1, \cdots, a_m). \tag{5.40}$$

Proof. For any two vectors u and v we define the inequality $u < v$ to mean $u_i < v_i \; (i = 1, \cdots, m)$.

At almost all pairs of points $u < v$ the increment function $\Pr \{ u < x < v \}$ is continuous [9]. We will first show that, at any pair of points of continuity of the increment function, we have

$$\Pr \{ u < x^{(n)} < v \} \rightarrow \Pr \{ u < x < v \} \quad \text{as } \Delta t \rightarrow 0. \tag{5.41}$$

Let $\epsilon > 0$, and let $e = (\epsilon, \epsilon, \cdots, \epsilon)$. By (1.5),

$$\Pr \{ u < x^{(n)} < v \} \leqq \Pr \{ u - e < x < v + e \} + O(\Delta t),$$

$$\Pr \{ u < x^{(n)} < v \} \geqq \Pr \{ u + e < x < v - e \} + O(\Delta t).$$

Keeping e fixed and letting $\Delta t \rightarrow 0$ we find

$$\limsup_{\Delta t \rightarrow 0} \Pr \{ u < x^{(n)} < v \} \leqq \Pr \{ u - e < x < v + e \},$$

$$\liminf_{\Delta t \rightarrow 0} \Pr \{ u < x^{(n)} < v \} \geqq \Pr \{ u + e < x < v - e \}.$$

We now let $e \rightarrow 0$. Because of the continuity of the increment function at $u < v$, we conclude

$$\limsup_{\Delta t \rightarrow 0} \Pr \{ u < x^{(n)} < v \} \leqq \Pr \{ u < x < v \} \leqq \liminf_{\Delta t \rightarrow 0} \Pr \{ u < x^{(n)} < v \}.$$

The required limit (5.41) follows.

Let a be a point of continuity of the distribution function F. Let an arbitrarily small number $\alpha > 0$ be given. Choose $u < a < v$ so that

$$\Pr\{u < x < v\} \geqq 1 - \alpha \tag{5.42}$$

and so that $u < v$ *and* $u < a$ are points of continuity of the increment function. By the limit (5.41), there is a number $\delta > 0$ so small that

$$\Pr\{u < x^{(n)} < v\} \geqq 1 - 2\alpha \quad \text{if } \Delta t \leqq \delta. \tag{5.43}$$

From the last two formulas, we have

$$\Pr\{x_1 \leqq u_1 \quad or \quad x_2 \leqq u_2 \quad or \cdots or \quad x_m \leqq u_m\} < \alpha \tag{5.44}$$

and, if $\Delta t \leqq \delta$,

$$\Pr\{x_1^{(n)} \leqq u_1 \quad or \quad x_2^{(n)} \leqq u_2 \quad or \cdots or \quad x_m^{(n)} \leqq u_m\} < 2\alpha. \tag{5.45}$$

By (5.44), we have

$$0 \leqq F(a) - \Pr\{u < x < a\} < \alpha.$$

Similarly, if $\Delta t \leqq \delta$, we find from (5.45)

$$0 \leqq F_{\Delta t}(a) - \Pr\{u < x^{(n)} < a\} < 2\alpha.$$

But, by (5.41), there is a positive number $\delta_1 \leqq \delta$ so small that

$$|\Pr\{u < x^{(n)} < a\} - \Pr\{u < x < a\}| < \alpha \quad \text{if} \quad \Delta t \leqq \delta_1 .$$

The last three formulas imply

$$|F(a) - F_{\Delta t}(a)| < 4\alpha \quad \text{if} \quad \Delta t \leqq \delta_1 .$$

Since $\alpha > 0$ is arbitrarily small, we have established the limit (5.40).

California Institute of Technology
Pasadena, California

1. W. B. Davenport, Jr. & W. L. Root, *An Introduction to the Theory of Random Signals and Noise*, McGraw-Hill, New York, 1958. MR **19**, 1090.
2. David Middleton, *An Introduction to Statistical Communication Theory*, International Series in Applied Physics, McGraw-Hill, New York, 1960. MR **22** # 9334.
3. Emanuel Parzen, *Stochastic processes*, Holden-Day, San Francisco, Calif., 1962. MR **25** # 2628.
4. Nelson Wax, (Ed.), *Selected Papers on Noise and Stochastic Processes*, Dover, New York, 1954. MR **15**, 970.
5. G. E. P. Box & M. E. Muller, "A note on the generation of random normal deviates," *Ann. Math. Statist.*, v. 29, 1958, pp. 610–611.
6. J. N. Franklin, "Deterministic simulation of random processes," *Math. Comp.*, v. 17, 1963, pp. 28–59. MR **26** # 7125.
7. J. N. Franklin, "The covariance matrix of a continuous autoregressive vector time series," *Ann. Math. Statist*, v. 34, 1963, pp. 1259–1264. MR **27** # 5297.
8. J. N. Franklin, "Numerical simulation of stationary and nonstationary Gaussian random processes," *SIAM Rev.*, v. 7, January, 1965.
9. M. Loève, *Probability Theory*, (2nd ed.), The University Series in Higher Mathematics, Van Nostrand, New York, 1960. MR **23** # A670.