SAND80-2080
Unlimited Release

UC-32

MASTER

Digital Simulation and Modeling of Nonlinear Stocastic Systems

J. Mark Richardson, James R. Rowland

Prepared by Sandia National Laboratories, Albuquerque, New Mexico 87185
and Livermore, California 94550 for the United States Department
of Energy under Contract DE-AC04-76DP00789

Printed April 1981

Sandia National Laboratories

SF 2900-Q(3-80)

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

DISCLAIMER

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Printed in the United States of America

Available from
National Technical Information Service
U. S. Department of Commerce
5285 Port Royal Road
Springfield, VA 22161

NTIS price codes
Printed copy: $6.00
Microfiche copy: A01

PAGES __1__ to __2__

WERE INTENTIONALLY

LEFT BLANK

SAND80-2080
Unlimited Release
Printed December 1980
Distribution Category UC-32

DIGITAL SIMULATION AND MODELING OF NONLINEAR STOCHASTIC SYSTEMS*

J. Mark Richardson
Safeguards Methodology Development Division 4416
Sandia National Laboratories
Albuquerque, NM 87185

James R. Rowland
Oklahoma State University
Stillwater, OK 74074

*Presented at the 1980 IEEE Conference on Cybernetics and Society, Cambridge, MA.

ABSTRACT

Digitally generated solutions of nonlinear stochastic systems are not unique but depend critically on the numerical integration algorithm used. Some theoretical and practical implications of this dependence are examined. The Ito-Stratonovich controversy concerning the solution of nonlinear stochastic systems is shown to be more than a theoretical debate on maintaining Markov properties as opposed to utilizing the computational rules of ordinary calculus. The theoretical arguments give rise to practical considerations in the formation and solution of discrete models from continuous stochastic systems. Well-known numerical integration algorithms are shown not only to provide different solutions for the same stochastic system but also to correspond to different stochastic integral definitions.

These correspondences are proved by considering first and second moments of solutions that result from different integration algorithms and then comparing the moments to those arising from various stochastic integral definitions. This algorithm-dependence of solutions is in sharp contrast to the deterministic and linear stochastic cases in which unique solutions are determined by any convergent numerical algorithm. Consequences of the relationship between stochastic system solutions and simulation procedures are presented for a nonlinear filtering example. Monte Carlo simulations and statistical tests are applied to the example to illustrate the determining role which computational procedures play in generating solutions.

CONTENTS

	Page
Introduction	9
Model Development	10
Stochastic Integral Definitions	12
Stochastic Integrals and Numerical Algorithms	16
Example	27
Conclusions	29
References	30

ILLUSTRATIONS

Figure		
1	Ensemble-Averaged Mean and Variance (RK4)	28
2	Ensemble-Averaged Mean and Variance (Euler)	29

DIGITAL SIMULATION AND MODELING OF
NONLINEAR STOCHASTIC SYSTEMS

Introduction

Computer simulations of dynamic behavior play a vital role in the analysis and design of physical systems. An important area of study in this context is the accuracy and meaning which are to be associated with the output of the computer model. This is a particularly relevant question in the field of nonlinear stochastic systems. For linear stochastic systems, the theoretical basis for analysis has a firm footing in mathematical rigor, with many of the main problems tending to be computational. However, when the system considered is nonlinear, deeper problems arise which are properly studied through the theory of stochastic differential equations.

Wiener[1] was interested in the general problem of analyzing nonlinear equations with random elements. He was also the first to define an integral with respect to a stochastic process, but his integral is defined only for nonrandom integrands.[2] Ito[3,4] showed how to extend the integral definition to include random integrands. Ito's work was motivated by the study of diffusive Markov processes and their transition probabilities. Stratonovich[5,6] also developed a stochastic integral as a means for studying diffusive Markov processes.

The Ito and Stratonovich integrals agree for linear stochastic differential equations. One is faced with the problem of interpretation of solutions, however, when nonlinear equations are studied. Both the Ito and Stratonovich theories are self-consistent, although in general they result in different solutions to the same nonlinear equation. Mortensen[7] explored this Ito-Stratonovich controversy and concluded that the choice between the Ito calculus and the Stratonovich calculus is one of personal preference, with mathematicians preferring the Ito theory because of its elegance and generality and engineers preferring Stratonovich's theory because of their familiarity with its rules. Mortensen believes that the safest answer to the stochastic modeling problem is to use a Monte Carlo computer simulation, thereby dodging the Ito-Stratonovich controversy.

McShane[8] made contributions toward unifying the theory of a stochastic calculus. He defined a stochastic integral by a modification of the procedure which Riemann used in defining the classical integral. The McShane integral exists under conditions which, in comparison with the Ito integral, are weaker regarding stochastic properties but stronger regarding continuity properties, and the Ito and McShane integrals agree when the hypotheses for the existence of both are satisfied. McShane laid the foundation for a unified theory of stochastic integration which includes both Lipschitzian and Brownian-motion processes.

Wright[9] performed a limited experiment in numerically solving a specific nonlinear stochastic differential equation. He noted heuristic correspondences among the various integral definitions and certain numerical integration algorithms. A single sample solution was generated by several numerical methods, and the behavior of this solution at a single point was studied as the integration step size was changed. This behavior provided preliminary indications of the relationships among the stochastic integrals and the numerical integration algorithms.

This paper examines the relationships of stochastic integrals and the behavior of solutions generated on the digital computer, using several well-known numerical integration methods. The next section discusses the general model to be used. The stochastic integral definitions are presented next, followed by the numerical algorithms and the correspondences based on equivalence of first and second moments. A nonlinear filtering problem is then considered in this context, and conclusions are presented in the last section.

Model Development

In modeling physical systems or analyzing equations which arise from scientific theory, differential equations of the form

$$\dot{\underline{x}}(t) = \underline{f}(\underline{x},t)$$

are often encountered. Here $\dot{\underline{x}}(t)$ is the vector of state derivatives and $\underline{f}(\underline{x},t)$ is a vector of functions which quantitatively explains the evolution of the system states with time. If the system has random inputs, then the state equation may be written

$$\underline{\dot{x}}(t) = \underline{f}(\underline{x},t) + \underline{g}(t)\underline{u}(t) \qquad (1)$$

where $\underline{g}(t)$ is a function denoting the sensitivity of the system to the random input $\underline{u}(t)$, usually modeled as white noise.

The mathematical representation for a physical signal modeled as white noise is that of a Gaussian process with zero mean and a covariance given by the Dirac δ-function, i.e.,

$$E\{\underline{u}(t)\} = 0 \; ,$$

$$E\{\underline{u}(t)\underline{u}^T(t + \tau)\} = Q \cdot \delta(\tau) \qquad (2)$$

where E denotes expectation and Q is a constant matrix expressing how $\underline{u}(t)$ is correlated with itself. Thus the white noise process has infinite variance and independent process values at any two distinct times.

In the scalar case of Eq. (1) with g(t) = 1, one is confronted with the integral

$$w(t) = \int_0^t u(s)ds \; . \qquad (3)$$

Because of the pathological nature of white noise, it is difficult to interpret Eq. (3) rigorously. For systems which are linear in noise terms, that is, the noise is additive rather than multiplicative, the difficulty is avoided by simply assuming the absolute convergence of the integral in Eq. (3) and very useful results obtain, such as covariance analysis (see Rowland and Holmes[10]). Nonlinear equations require a more critical evaluation of Eq. (3), however. One method proposed for dealing with this problem is to define w(t) directly.

Let w(t) be a Gaussian process with the following properties:

$$w(0) = 0 \; ,$$

$$E\{w(t)\} = 0 \; ,$$

$$E\{w(t)w(s)\} = q \cdot \min(t,s) \quad t,s \geq 0$$

where q is a constant.

This process was studied by Wiener and is often called the Wiener process or Brownian-motion process. Doob[11] and Parzen[12] proved several useful properties of the Wiener process. These include the facts that the sample functions are almost surely continuous, not differentiable and not of bounded variation, the process has independent increments, and the Levy oscillation property holds; i.e., if $\{a = t_0, t_1, \ldots, t_n = b\}$ is a partition of the interval $[a,b]$ and $\Delta = \max_i |t_i - t_{i-1}|$, then

$$\underset{\Delta \to 0}{\mathrm{l.i.m.}} \sum_{i=1}^{n} (w(t_i) - w(t_{i-1}))^2 = q \cdot (b - a) \; .$$

For the scalar nonlinear analog of Eq. (1), the Wiener process turns out to be much more amenable to analysis than does white noise. The nonlinear equation, written in terms of differentials rather than derivatives, is then given by

$$dx(t) = f(x,t)dt + g(x,t)dw(t) \; . \tag{4}$$

To find a solution of Eq. (4), it suffices to display a stochastic process x(t) which satisfies

$$x(t) = x(a) + \int_a^t f(x(s),s)ds + \int_a^t g(x(s),s)dw(s) \; . \tag{5}$$

If w(t) were of bounded variation, there would be no problem in interpreting x(t) in Eq. (5). However, the last integral in Eq. (5) cannot be a Lebesgue-Stieltjes integral since the Wiener process is not of bounded variation. We must therefore investigate how the second integral is defined for stochastic processes g(x,t) and w(t).

Stochastic Integral Definitions

As was mentioned earlier, Wiener was the first to define an integral with respect to a stochastic process, but his integral is defined only for nonrandom integrands. Ito showed how to extend the integral definition to include random integrands, but the integrator is less general than in the Wiener integral in that it must be a martingale. Since the Wiener process is a martingale and the function g(x,t) of Eq. (5) is random, the Ito integral is more useful than the Wiener integral.

Definition 1 (Ito Integral)

Let z(t) be a martingale process, and suppose there exists a monotone nondecreasing function F such that, if $s < t$, then

$$E\{|z(t) - z(s)|^2\} = F(t) - F(s)$$

with probability 1. Suppose g(x,t) is a measurable function and

$$\int_{-\infty}^{\infty} E\{|g(x,t)|^2\}dF(t) < \infty .$$

If $\{a = t_0, t_1, \ldots, t_n = b\}$ is a partition of [a,b] and $\Delta = \max_i |t_i - t_{i-1}|$, then the Ito integral is defined to be

$$(I)\int_a^b g(x,t)dz(t) = \underset{\Delta \to 0}{\text{l.i.m.}} \sum_{i=1}^{n-1} g(x(t_i),t_i)(z(t_{i+1}) - z(t_i)) \qquad (6)$$

where the series converges in the mean to a random variable denoted by the integral on the left in Eq. (6).

Doob has shown that the hypotheses of the theorem imply that g(x,t) and the increments (z(t) - z(s)) are independent. From this independence and noting that $E\{z(t) - z(s)\} = 0$ for z(t) a martingale, it follows that the expected value of the Ito integral is zero. The integral is a martingale and the following equality holds:

$$E\{\int_a^t g_1(x,s)dz(s) \int_a^t g_2(x,s)dz(s)\} = \int_a^t E\{g_1(x,s)g_2(x,s)\}ds .$$

These properties explain the usefulness of the Ito integral, especially in the study of Markov processes, since moment calculations are simplified using the above facts.

In computational operations with the Ito integral, procedures from ordinary calculus can no longer be used. For instance, change of variables and differentiation require very different treatments. In particular, suppose x(t) is an Ito process determined by Eq. (4) and $\Phi(x(t),t)$ is a function of x(t) and t, with second-order partial derivatives in x(t) and t. Then $\Phi(x(t),t)$ is also an Ito process, and the so-called Ito differential rule states that

$$d\Phi = \left[\frac{\partial\Phi}{\partial t} + \frac{\partial\Phi}{\partial x}f + \frac{1}{2}q\frac{\partial^2\Phi}{\partial x^2}g^2\right]dt + q\frac{\partial\Phi}{\partial x}g\,dw \ .$$

<u>Definition 2 (Stratonovich Integral)</u>

Let z(t) be a Markov process with

$$\lim_{h\to 0} E\{(z(t+h) - z(t))/h \mid z(t) = \xi\} = a(\xi,t) \ ,$$

$$\lim_{h\to 0} E\{(z(t+h) - z(t))^2/h \mid z(t) = \xi\} = b(\xi,t) \ , \text{ and}$$

$$\lim_{h\to 0} E\{|z(t+h) - z(t)| > \delta \mid z(t) = \xi\} = 0$$

with a(z,t) and b(z,t) continuous in both arguments and b(z,t) having a continuous partial derivative $\partial b(z,t)/\partial z$. Suppose g(z,t) is continuous in t having a continuous partial derivative $\partial g(z,t)/\partial z$,

$$\int_{-\infty}^{\infty} E\{g(z,t)a(z,t)\}dt < \infty \text{ and}$$

$$\int_{-\infty}^{\infty} E\{|g(z,t)|^2 b(z,t)\}dt < \infty \ .$$

Let $\{a = t_0, t_1, \ldots, t_n = b\}$ be a partition of [a,b] and $\Delta = \max_i |t_{i+1} - t_i|$. The Stratonovich integral is defined as

$$(S)\int_a^b g(z,t)dz(t) = \underset{\Delta\to 0}{\text{l.i.m.}} \sum_{i=1}^{n-1} g(\tfrac{1}{2}(z(t_{i+1}) + z(t_i)),t_i)\cdot (z(t_{i+1}) - z(t_i)) \ . \qquad (7)$$

Although the Stratonovich integral is only defined for integrands which are functions of the integrator process, Stratonovich[5] showed how to extend the integral to more general situations by defining a multi-dimensional integral. In particular, if dx(t) and dz(t) are related by a stochastic differential equation of the form of Eq. (4), then

$$\int_a^b g(x(t),t)dz(t)$$

can be defined. If the process z(t) in Definition 2 is a Wiener process, then the function $a(z,t) = 0$ and $b(z,t) = q$, where q is the variance parameter of the Wiener process.

McShane integrals are defined in terms of "belated" partitions. Let D denote a set of real numbers with the interval [a,b] contained in D. A belated partition of the interval [a,b] is a collection of real numbers $\{t_0,t_1,\ldots,t_n;\ \tau_1,\tau_2,\ldots,\tau_n\}$ where $a = t_0 < t_1 <\cdots< t_n = b$ and τ_i is in D for each i and $\tau_i \leq t_i$.

Definition 3 (McShane Integral)

Let D be a set of real numbers and [a,b] a closed interval contained in D. Let $\{t_0,t_1,\ldots,t_n;\ \tau_1,\tau_2,\ldots,\tau_n\}$ be a belated partition of D with $\Delta = \max |t_{i+1} - t_i|$. Let z(t) be a stochastic process on [a,b] satisfying, for some constant K,

$$|E\{z(t) - z(s)|z(\tau),\ \tau \leq s < t\}| \leq K(t - s)$$

and

$$E\{|z(t) - z(s)|^2|z(\tau),\ \tau \leq s < t\} \leq K(t - s) ,$$

both with probability 1. If g(x,t) is a measurable process on D which is L_2-bounded and L_2-continuous with probability 1, then the McShane integral is defined to be

$$(M) \int_a^b g(x,t)dz(t) = \lim_{\Delta \to 0} \sum_{i=0}^{n-1} g(x(\tau_i),\tau_i)(z(t_{i+1}) - z(t_i))$$

where the convergence is in probability.

It is seen from the definitions that the integrator process is more general for the McShane integral than for the Ito integral; in particular, the McShane integrator does not have to be a martingale. The integrand for McShane's integral is not as general, however, since it is required to be L_2-bounded and L_2-continuous and the Ito definition only requires mean-square integrability. McShane unified the theory of ordinary and stochastic integrals in the sense that his integral exists and is equal to the ordinary integral when the system inputs are well-behaved time functions. This situation is not true for the Ito integral. When the regions of definition overlap, McShane's integral is the same as whichever of Ito's or the ordinary integral exists.

Stochastic Integrals and Numerical Algorithms

Studying the stochastic integral definitions in the previous section, one notices a rather profound conceptual difference in these definitions and in the definition of the Riemann integral. This technical difference arises because of the irregularity of stochastic processes as compared with deterministic functions. The point of evaluation of the integrand of a Riemann integral, defined as the limit of Riemann sums, is determined by the values of the integrand within each subinterval arising from a partition. The bounds of the function within each subinterval determine the point of functional evaluation. This approach is not true of stochastic integrals. Instead, the evaluation point of the integrand within each subinterval is specified by the definition. The fixed point of evaluation also differs among the various definitions of stochastic integrals. This circumstance gives rise to many interesting features of these integrals.

The necessary properties for the integrand and integrator processes vary somewhat in the definitions. Also, the properties which the integrals themselves enjoy are different, in some cases profoundly so. But perhaps the most fundamental difference distinguishing the stochastic from the deterministic is that the value of the integral is affected by the evaluation point. The extremely erratic behavior of the stochastic integrator processes involved, along with the rather surprising fact that second-order terms do not vanish in the limit as they do in the deterministic case, helps to explain this phenomenon.

Because of the discrepancies within the theory of stochastic integrals and the differences between it and the deterministic theory, one is thus led to the possibility that numerical solutions of stochastic integrals may not provide consistent results. With the increasing utility of digital computers and the greater understanding of stochastic phenomena at all levels, a deeper understanding of these relationships becomes necessary.

From familiarity with the integral definitions and some numerical integration schemes, one can make intuitive correspondences among definitions and digital integration procedures. The purpose of this section is to investigate more thoroughly some of these correspondences and to determine if there is justification for the supposed correlation between these widely divergent areas.

Since this report is concerned with stochastic integrals, attention will be restricted to equations of the form $dx(t) = g(x,t)dw(t)$, with the stochastic process $w(t)$ a Wiener process and $g(x,t)$ a random function. The deterministic analog of this equation is $\dot{x}(t) = g(x,t)$, with $g(x,t)$ no longer random. Solving this equation involves computing

$$\int g(x,t)dt \ ,$$

and, in a similar manner, one can investigate the results of employing numerical integration procedures in the evaluation of the stochastic integral

$$\int g(x,t)dw(t)$$

arising from the above stochastic differential equation.

The Euler method of numerical integration approximates the differential equation with a step function and evaluates the integral of that step function; that is, the equation is assumed constant over each integration step length, with the constant value over a step length determined by the functional value at the initial point of each subinterval. The approximation is given by

$$x_{i+1} = x_i + g(x_i,t_i)(w_{i+1} - w_i) \tag{8}$$

where

$$x_i = x(t_i) \ .$$

From Eq. (8), one can calculate the statistics of the solution of a stochastic equation which has been solved by Euler's method. Specifically,

$$E\{x_{i+1}\} = E\{x_i\} + E\{g(x_i,t_i)\}E\{w_{i+1} - w_i\} = E\{x_i\} \tag{9}$$

since $g(x_i,t_i)$ is independent of the Wiener process increment and the Wiener process has a mean value of zero. It follows that

$$E\{x_i\} = E\{x_0\} \tag{10}$$

for every i.

The mean square value is given by

$$E\{x_{i+1}^2\} = E\{x_i^2\} + 2E\{x_i g(x_i,t_i)(w_{i+1} - w_i)\}$$

$$+ E\{g^2(x_i,t_i)(w_{i+1} - w_i)^2\}$$

$$= E\{x_i^2\} + E\{g^2(x_i,t_i)\}E\{(w_{i+1} - w_i)^2\}$$

$$= E\{x_i^2\} + qE\{g^2(x_i,t_i)\}(t_{i+1} - t_i) \tag{11}$$

which follows from the independence of the noise increment and the integrand and from the properties of the Wiener process. Recalling the identity

$$\mathrm{Var}\{x\} = E\{x^2\} - E^2\{x\} \tag{12}$$

we obtain, using Eq. (10),

$$\mathrm{Var}\{x_{i+1}\} = \mathrm{Var}\{x_i\} + qE\{g^2(x_i,t_i)\}(t_{i+1} - t_i) \ . \tag{13}$$

Numerically, Eq. (13) behaves as the integral of $E\{g^2(x_i,t_i)\}$.

Calculating the mean and variance of a stochastic process arising from an Ito integral may be accomplished by using properties resulting from the Ito definition of a stochastic integral (see Doob[11]). These properties are the following:

$$E\{I\int_a^t g(s)dw(s)\} = 0 \tag{14}$$

$$E\{I\int_a^t g_1(s)dw(s)\ I\int_a^t g_2(s)dw(s)\} = q\int_a^t E\{g_1(s)g_2(s)\}ds \tag{15}$$

where the "I" indicates the integral is to be interpreted in the sense of Ito.

Given the stochastic differential equation

$$dx(t) = g(x,t)dw(t)\ , \tag{16}$$

we have the equivalent integral equation

$$x(t) = x(a) + I\int_a^t g(x,s)dw(s)\ . \tag{17}$$

From Eqs. (14) and (17), the mean value of x(t) is

$$E\{x(t)\} = E\{x(a)\}\ . \tag{18}$$

The variance of x(t) may be computed by noting that the initial condition x(a) is independent of

$$I\int_a^t g(x,s)dw(s)$$

and by using Eqs. (14) and (15) and the Identity (12). Thus

$$E\{x^2(t)\} = E\{x^2(a)\} + 2E\{x(a)\ I\int_a^t g(x,s)dw(s)\}$$

$$+ E\{[I\int_a^t g(x,s)dw(s)]^2\}$$

$$= E\{x^2(a)\} + q\int_a^t E\{g^2(x,s)\}ds \tag{19}$$

and the variance is then

$$Var\{x(t)\} = Var\{x(a)\} + q\int_a^t E\{g^2(x,s)\}ds \ . \qquad (20)$$

The mean value of Euler's method, given by Eq. (10), is the same as the mean value of the Ito integral in Eq. (18). Similarly, Eqs. (13) and (20) indicate that the variances agree also. It thus is concluded that numerical integration by Euler's method corresponds to the Ito integration of stochastic differential equations in the sense that the first two moments coincide.

Runge-Kutta integration methods are somewhat more sophisticated than Euler's method in that they use more than a simple slope for their calculations. They are often used to generate preliminary values for other types of algorithms which are not self-starting. Rather than using the first point in each subinterval of interest as the point of evaluation, as in the Euler method, Runge-Kutta methods use points within a subinterval to generate the solution at the end of the interval. A typical second-order Runge-Kutta method (RK2) is

$$x_{i+1} = x_i + \frac{1}{2}(g(x_i,t_i) + g(x_i + dx,t_i))(w_{i+1} - w_i) \ . \qquad (21)$$

From Eq. (21), we can determine the statistics of a solution generated by this Runge-Kutta method. The expected value is

$$E\{x_{i+1}\} = E\{x_i\} + \frac{1}{2}E\{(g(x_i,t_i) + g(x_i + dx,t_i))(w_{i+1} - w_i)\}$$

$$= E\{x_i\} + \frac{1}{2}E\{g(x_i + dx,t_i)(w_{i+1} - w_i)\} \ . \qquad (22)$$

From the differentiability of g(x,t), assumed in Definition 2, we have

$$\frac{\partial g(x')}{\partial x} = \frac{g(x_i + dx,t_i) - g(x_i,t_i)}{dx} \ , \ x_i \le x' < x_{i+1} \ , \qquad (23)$$

and, consequently,

$$g(x_i + dx,t_i)(w_{i+1} - w_i) = (\frac{\partial g(x')}{\partial x}dx + g(x_i,t_i))(w_{i+1} - w_i)$$

$$= \frac{\partial g(x')}{\partial x} g(x_i, t_i)(w_{i+1} - w_i)^2$$

$$+ g(x_i, t_i)(w_{i+1} - w_i) \quad . \tag{24}$$

Thus

$$E\{x_{i+1}\} = E\{x_i\} + \tfrac{1}{2} q E\{g(x_i, t_i) \frac{\partial g(x')}{\partial x}\}(t_{i+1} - t_i) \tag{25}$$

since x_i is independent of $(w_{i+1} - w_i)$, and this is the numerical equivalent of $\frac{1}{2}q$ times the integral of $E\{g(x_i, t_i) \frac{\partial g(x')}{\partial x}\}$.

To find the variance, we first calculate from Eqs. (22) and (24)

$$E^2\{x_{i+1}\} = E^2\{x_i\} + E\{x_i\} E\{g(x_i, t_i) \frac{\partial g(x')}{\partial x}(w_{i+1} - w_i)^2\}$$

$$+ \tfrac{1}{4} E^2\{g(x_i, t_i) \frac{\partial g(x')}{\partial x}(w_{i+1} - w_i)^2\} \quad . \tag{26}$$

To find the mean square value of x_{i+1}, we make use of Eqs. (21) and (24) to obtain

$$x_{i+1}^2 = (x_i + \tfrac{1}{2} g(x_i, t_i) \frac{\partial g(x')}{\partial x}(w_{i+1} - w_i)^2 + g(x_i, t_i)(w_{i+1} - w_i))^2$$

$$= x_i^2 + \tfrac{1}{4} g^2(x_i, t_i)(\frac{\partial g(x')}{\partial x})^2 (w_{i+1} - w_i)^4 + g^2(x_i, t_i)(w_{i+1} - w_i)^2$$

$$+ x_i g(x_i, t_i) \frac{\partial g(x')}{\partial x}(w_{i+1} - w_i)^2 + 2 x_i g(x_i, t_i)(w_{i+1} - w_i)$$

$$+ g^2(x_i, t_i) \frac{\partial g(x')}{\partial x}(w_{i+1} - w_i)^3 \tag{27}$$

from which it follows that

$$E\{x_{i+1}^2\} = E\{x_i^2\} + \tfrac{1}{4} E\{g^2(x_i, t_i)(\frac{\partial g(x')}{\partial x})^2 (w_{i+1} - w_i)^4\}$$

$$+ q E\{g^2(x_i, t_i)\}(t_{i+1} - t_i) + E\{x_i g(x_i, t_i)$$

$$\frac{\partial g(x')}{\partial x}(w_{i+1} - w_i)^2\} \ . \tag{28}$$

Eqs. (26) and (28) then combine to provide the variance of x_{i+1} as

$$\mathrm{Var}\{x_{i+1}\} = \mathrm{Var}\{x_i\} + qE\{g^2(x_i,t_i)\}(t_{i+1} - t_i)$$

$$+ \frac{1}{4}\mathrm{Var}\{g(x_i,t_i)\frac{\partial g(x')}{\partial x}(w_{i+1} - w_i)^2\} \ . \tag{29}$$

Noting that the Runge-Kutta algorithms employ functional evaluations within each integration step and that this is also the case for the Stratonovich integral, we now determine the relationship between the moments of these methods and the moments of the Stratonovich integral. Stratonovich[5] introduced the stochastic integral bearing his name and proved the fundamental equality

$$S\int_a^t g(x(s),s)dw(s) = I\int_a^t g(x(s),s)dw(s)$$

$$+ \frac{1}{2}q\int_a^t g(x(s),s)\ \frac{\partial g(x(s),s)}{\partial x}ds \ . \tag{30}$$

Exploiting this relationship between the Stratonovich and Ito integrals allows the computation of the mean resulting from the differential equation, Eq. (16), when the equation is solved in the Stratonovich sense. Thus

$$E\{x(t)\} = E\{x(a)\} + E\{S\int_a^t g(x(s),s)dw(s)\}$$

$$= E\{x(a)\} + E\{I\int_a^t g(x(s),s)dw(s)\}$$

$$+ \frac{1}{2}qE\{\int_a^t g(x(s),s)\frac{\partial g(x')}{\partial x}ds\}$$

$$= E\{x(a)\} + \frac{1}{2}qE\{\int_a^t g(x(s),s)\frac{\partial g(x')}{\partial x}ds\} \ . \tag{31}$$

In a similar manner, the variance of the Stratonovich solution may be found.

$$E\{x^2(t)\} = E\{[x(a) + S\int_a^t g(x(s),s)dw(s)]^2\}$$

$$= E\{x^2(a)\} + 2E\{x(a)[I\int_a^t g(x(s),s)dw(s)$$

$$+ \frac{1}{2}q\int_a^t g(x(s),s)\frac{\partial g(x')}{\partial x}ds]\}$$

$$+ E\{[I\int_a^t g(x(s),s)dw(s)$$

$$+ \frac{1}{2}q\int_a^t g(x(s),s)\frac{\partial g(x')}{\partial x}ds]^2\} \qquad (32)$$

After subtracting the square of the mean and performing some algebraic manipulation, we obtain

$$Var\{x(t)\} = E\{x^2(a)\} - E^2\{x(a)\} + q\int_a^t E\{g^2(x(s),s)\}ds$$

$$+ \frac{1}{4}q^2 E\{[\int_a^t g(x(s),s)\frac{\partial g(x')}{\partial x}ds]^2\}$$

$$- E^2\{\int_a^t g(x(s),s)\frac{\partial g(x')}{\partial x}ds\}]$$

$$+ qE\{I\int_a^t q(x(s),s)dw(s)\int_a^t g(x(s),s)\frac{\partial g(x')}{\partial x}ds\} .$$

The last term vanishes, however, since the integrals are independent. We now have the result

$$Var\{x(t)\} = Var\{x(a)\} + q\int_a^t E\{g^2(x(s),s)\}ds$$

$$+ \frac{1}{4}q^2 \mathrm{Var}\{\int_a^t g(x(s),s)\frac{\partial g(x')}{\partial x}ds\} \ . \tag{33}$$

The mean value of the Runge-Kutta method, given by Eq. (25), is now seen to be the numerical equivalent of the mean value of the Stratonovich integral, given by Eq. (31). Comparison of Eqs. (29) and (33) indicates that the variances of the Runge-Kutta method and the Stratonovich integral also coincide. It thus is concluded that numerical integration by this Runge-Kutta method corresponds to Stratonovich integration of stochastic differential equations in the sense that the first two moments are identical.

Predictor methods are another type of numerical integration algorithm. A k-th order predictor estimates the value of x_{i+1} from the previous values $x_i, x_{i-1}, \ldots, x_{i-k+1}$. Predictor methods are thus multistep and require starting values. These methods do not introduce correlation between the integrand and the noise process because different noise increments are used for each calculated value of x_i. Functional evaluations prior to the current interval are required for these methods. This fact suggests the correspondence of predictor methods with the McShane definition of the stochastic integral.

The rationale behind the development of the McShane integral was the construction of a theory of stochastic systems which would also provide results consistent with deterministic systems. This objective was achieved without introducing profound differences in the existing stochastic theories. Consequently, solutions of McShane integrals arising from practical applications agree with results obtained from Ito's theory. More precisely, if the integrator process is a martingale and the integrand is bounded and continuous in the L_2 sense, then the Ito and McShane integrals agree. The Wiener process is a martingale and the class of L_2-bounded and L_2-continuous functions is general enough to include systems of practical interest. Thus, the McShane and Ito theoretical solutions for those systems agree. More generally, the McShane integral agrees with the Ito integral when the hypotheses for the existence of both are satisfied, and the McShane integral agrees with the Riemann integral in the case of Lipschitzian inputs. These facts lead to the conclusion that the predictor methods should correspond to the Ito stochastic integral.

The Adams-Bashforth second-order predictor method is given by the formula

$$x_{i+1} = x_i + \frac{1}{2}(3g(x_i) - g(x_{i-1}))(w_{i+1} - w_i) \ . \tag{34}$$

The mean value of x_{i+1} from Eq. (34) is given by

$$E\{x_{i+1}\} = E\{x_i\} + \frac{3}{2}E\{g(x_i)(w_{i+1} - w_i)\}$$

$$- \frac{1}{2}E\{g(x_{i-1})(w_{i+1} - w_i)\} = E\{x_i\} \tag{35}$$

and the mean value equivalence with the Ito integral holds.

To analyze the variance of x_{i+1}, we first calculate

$$x_{i+1}^2 = x_i^2 + x_i(3g(x_i) - g(x_{i-1}))(w_{i+1} - w_i)$$

$$+ \frac{1}{4}[3g(x_i) - g(x_{i-1})]^2(w_{i+1} - w_i)^2 \tag{36}$$

and then obtain

$$E\{x_{i+1}^2\} = E\{x_i^2\} + \frac{1}{4}E\{[3g(x_i) - g(x_{i-1})]^2(w_{i+1} - w_i)^2\} \tag{37}$$

and, from the independence of the noise increment and the other expressions in the second term on the right in Eq. (37), we find

$$E\{x_{i+1}^2\} = E\{x_i^2\} + \frac{1}{4}qE\{[3g(x_i) - g(x_{i-1})]^2\}dt \ . \tag{38}$$

Noting that

$$\frac{\partial g(x')}{\partial x} = \frac{g(x_i) - g(x_{i-1})}{x_i - x_{i-1}} \ , \quad x_{i-1} \le x' < x_i \tag{39}$$

we find that

$$[3g(x_i) - g(x_{i-1})]^2 = [3\frac{\partial g(x')}{\partial x}(x_i - x_{i-1}) + 2g(x_{i-1})]^2$$

$$= 9(\frac{\partial g(x')}{\partial x})^2(x_i - x_{i-1})^2$$

$$+ 12g(x_{i-1})\frac{\partial g(x')}{\partial x}(x_i - x_{i-1})$$

$$+ 4g^2(x_{i-1}) \quad . \tag{40}$$

Equation (38) then becomes

$$E\{x_{i+1}^2\} = E\{x_i^2\} + \frac{9}{4}qE\{g^2(x_{i-1})(\frac{\partial g(x')}{\partial x})^2(w_i - w_{i-1})^2\}dt$$

$$+ 3qE\{g^2(x_{i-1})\frac{\partial g(x')}{\partial x}(w_i - w_{i-1})\}dt$$

$$+ qE\{g^2(x_{i-1})\}dt \quad . \tag{41}$$

It is easy to see that the second and third terms on the right vanish since they are of order higher than one in dt. We thus have the result that

$$\mathrm{Var}\{x_{i+1}\} = \mathrm{Var}\{x_i\} + qE\{g^2(x_{i-1})\}dt \tag{42}$$

which is of the same form as Eq. (20), the variance of the Ito integral.

These results show that the correspondences among stochastic integral definitions and numerical integration algorithms are not merely intuitive concepts. Instead, the Ito-Euler-predictor association and the Stratonovich-Runge-Kutta association are valid classifications based on the equivalence of first and second moments. In light of these results, it is not surprising that the digital simulation of a nonlinear stochastic system should give rise to different solutions based on the method of numerical integration. Indeed, this difference of solutions will appear digitally whenever the Ito and Stratonovich integrals provide different solutions theoretically.[13] The next section illustrates these results through the simulation of a nonlinear filter.

Example

A broad area of general interest in the field of stochastic systems is filtering theory. The form of the optimal filter in the case of linear stochastic systems with white Gaussian noise inputs is widely known, but, in the nonlinear case, no such generally applicable optimal results have been found. The estimation of the state of a physical system, based upon data corrupted by noise, is easily accomplished if the probability distribution of the system state, conditioned on the measurement data, is known for all times. The problem thus becomes that of describing the time history of this distribution and the specification of the structure of the filter whose output is this distribution when the input is the given input measurement function.

Stochastic differential equations have been used in the analysis of this optimal nonlinear filtering problem. The study of the evolution of the probability distribution of the system state by means of stochastic differential equations was initiated by Stratononvich.[6] In these equations, the observed noisy input time function is the forcing term. The result of these studies has been the specification of the probability distribution in terms of a nonlinear stochastic differential equation.

Consider the observation process defined by

$$dy(t) = \frac{\alpha}{\beta^2}\,dt + \frac{1}{\beta}dw(t) \tag{43}$$

where α and β are constants. The optimal estimate for the posterior probability distribution of the observed process was derived by Wonham[14] and is given by the stochastic equation

$$dx(t) = -\beta^2 x(t)(1 - x^2(t))dt - \alpha(1 - x^2(t))dt + \beta(1 - x^2(t))dw \tag{44}$$

Equation (44) defines the structure of an ideal filter which generates the optimal estimate of the posterior distribution from the observed input function.

The simulation of this example was performed with initial condition x(0) = 0.0. The integration step size was chosen to be approximately 0.002 second, and 100 sample runs were ensemble-averaged to provide the results. The parameters α and β were given the same value, α = β = -2.0, and the variance parameter was chosen to be unity. The solid lines of Figures 1 and 2 give the simulation results for the RK4 and Euler numerical integration methods, respectively. The AB2 method resulted in the same mean and variance as the Euler method. The mean value generated by the RK4 method is different, however. There is about a 30% difference in the mean value after 1 second. The variance estimates for all the methods appear to achieve a steady-state value of approximately one-fourth. The Euler method overshoots this value somewhat and damps out rather slowly, while the RK4 method achieves the value quickly and then exhibits small random perturbations.

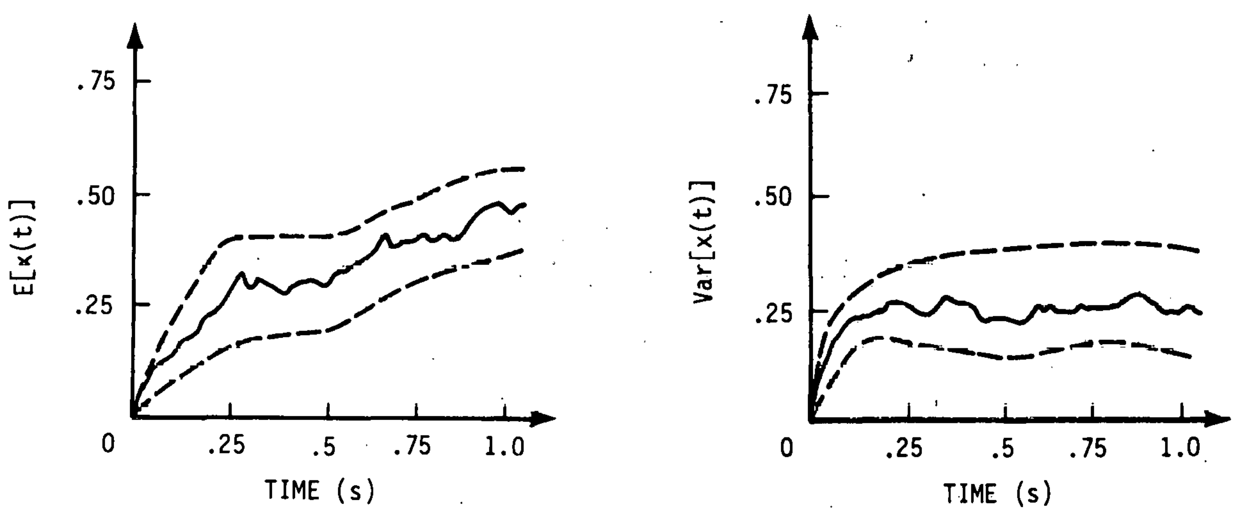

Figure 1. Ensemble-Averaged Mean and Variance (RK4)

Confidence intervals at a significance level of 95% are shown for the simulation results by dashed lines. From approximately t = 0.4 second to t = 1 second, the mean generated by the Runge-Kutta method lies outside the confidence interval for the Euler method and the Euler mean value lies outside the Runge-Kutta confidence interval for the same time period. For times near 1 second, the confidence intervals do not overlap. These results show that the generated mean values are in fact different time functions and not merely different approximations to the same one. The variance estimates and associated confidence intervals exhibit the same type behavior, although not to the same extent. About

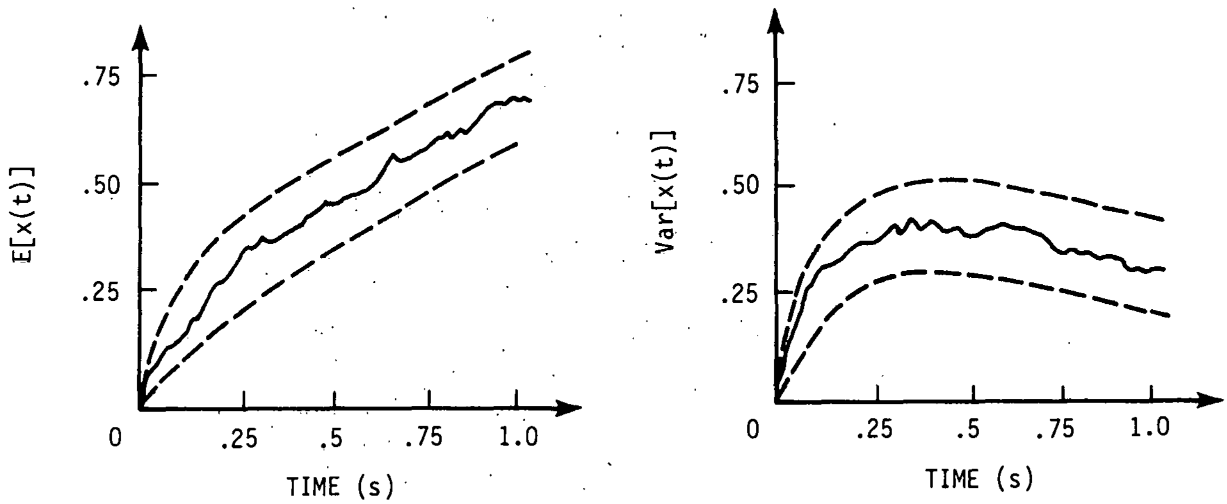

Figure 2. Ensemble-Averaged Mean and Variance (Euler)

40% to 50% of the variance trajectories lies outside the confidence intervals associated with the different type of numerical method.

Hence, in evaluating the performance of the filter, the effect of the numerical integration algorithm must be accounted for. The results, and conclusions, of an analysis of the filter dynamics would seem to be somewhat arbitrary, to the extent that they ignore this algorithm dependence.

Conclusions

It has been shown that digitally generated solutions of nonlinear stochastic systems are not unique. Correspondences among stochastic integral definitions and numerical integration algorithms based on the point of evaluation of the integrand have been validated in the sense of equivalence of first and second moments. Thus digital simulations of the same system result in different solutions whenever there is a difference in the Ito and Stratonovich solutions. This divergence has been illustrated through the simulation of an optimal nonlinear filter.

References

[1] N. Wiener, Nonlinear Problems in Random Theory (New York: Wiley, 1958).

[2] _____, "Differential-Space," J Math Phys, 2, 1923.

[3] K. Ito, "Stochastic Integral," Proc Imperial Academy, vol 20 (Tokyo: 1954).

[4] _____, Lectures on Stochastic Processes (notes by K. M. Rao) given at the Tata Institute for Fundamental Research, Bombay, India, 1961.

[5] R. L. Stratonovich, "A New Representation for Stochastic Integrals and Equations," J SIAM Control, 4:2, 1966.

[6] _____, "Conditional Markov Processes," Theor Probability Appl, V, 1960.

[7] R. E. Mortensen, "Mathematical Problems of Modeling Stochastic Nonlinear Dynamic Systems," J Stat Phys, 1:2, 1969.

[8] E. J. McShane, Stochastic Calculus and Stochastic Models (New York: Academic Press, 1974).

[9] D. J. Wright, "The Digital Simulation of Stochastic Differential Equations," IEEE Trans on Aut Contr, February, 1974.

[10] J. R. Rowland and W. H. Holmes, "A Direct Covariance Algorithm for Computer-Aided Statistical Electronics Circuit Design," Int J Elec, 37:1, 1974.

[11] J. L. Doob, Stochastic Processes (New York: Wiley, 1953).

[12] E. Parzen, Stochastic Processes (San Francisco: Holden-Day, 1962).

[13] J. M. Richardson, "Digital Simulation of Nonlinear Stochastic Systems" (PhD thesis, Oklahoma State University, 1980).

[14] W. M. Wonham, "Some Applications of Stochastic Differential Equations to Optimal Nonlinear Filtering," J SIAM Control, 2:3, 1964.

Distribution:
TID-4500 UC-32 (191 Copies)

James R. Rowland (5)
303 Engineering South
Oklahoma State University
Stillwater, OK 74074

4410 D. J. McCloskey
4416 L. D. Chapman
4416 J. M. Richardson (5)
4416 S. L. K. Rountree
5642 L. F. Shampine
8214 M. A. Pound
3141 L J Erickson (5)
3151 W. L. Garner (3)
For DOE/TIC (Unlimited Distribution)

Org.	Bldg.	Name	Rec'd by *	Org.	Bldg.	Name	Rec'd by