DIGITAL SIMULATION OF NONLINEAR STOCHASTIC SYSTEMS

By

JOHN MARK RICHARDSON

Bachelor of Science
Oklahoma State University
Stillwater, Oklahoma
1975

Master of Science
Oklahoma State University
Stillwater, Oklahoma
1977

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF PHILOSOPHY
December, 1980

DIGITAL SIMULATION OF NONLINEAR

STOCHASTIC SYSTEMS

Thesis Approved:

James R. Rowland
Thesis Adviser

Robert J. Mulholland

Craig S. Sims

Marvin S. Keener

Norman N. Durham
Dean of the Graduate College

ACKNOWLEDGMENTS

I owe an enormous debt to my thesis adviser and chairman of my doctoral committee, Dr. James R. Rowland. The time he spent discussing this topic, as well as his friendship, guidance, and insight throughout my graduate education, has been of tremendous value and will always be greatly appreciated. The other members of my committee, Dr. Robert J. Mulholland, Dr. Craig S. Sims, and Dr. Marvin S. Keener, all contributed substantially to my education. I wish to thank them for their willingness to spend their time and share their knowledge with me.

This research was made possible by the graduate research and teaching assistance provided by the School of Electrical Engineering.

I wish to thank Ms. Louise Sumpter for her excellent job of typing this manuscript.

My brother and sisters and their spouses, along with my parents-in-law and brother-in-law, deserve a special thanks for their encouragement throughout my education and it is gratefully extended.

I dedicate this thesis to my parents. Their unfailing encouragement and faith in me have always been of great importance and my love of learning and knowledge is largely a result of their influence. This research effort, the culmination of my formal education, is my thanks for their love and guidance.

To my wife, Kat, whose love and support are absolutely invaluable and whose unending patience and understanding are always recognized, if

seldom acknowledged, I dedicate the fruits of my education and the rest of my life.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I

INTRODUCTION

As scientists continue to explore and discover the interrelationships of man-made and natural phenomena, the need for accurate modeling procedures and analytical techniques becomes more apparent. Because the methods of investigators and the systems which they study are becoming increasingly complex, computers are often essential for gaining insight into physical processes, for solving complicated problems and for analyzing models. Thus, the effects of computer simulation on model performance and on model development are important areas of study.

One current and continually more important subject of research in modeling and analysis is the study of stochastic systems. The modern theory of stochastic systems had its beginnings in the method of least squares, which was developed for parameter estimation of planetary motion from measured data. The modeling and analysis of random dynamic systems using stochastic techniques is very important in systems engineering as evidenced by the use of Kalman filtering, as well as other aspects of stochastic analysis, in such diverse applications as seismic data processing, systems identification and aircraft control. Since stochastic algorithms generally require great amounts of computation, the advent of modern high-speed digital computers, and more recently of microprocessors, has made it feasible to apply stochastic analysis and synthesis to problems where previously it would have been physically or economically impossible.

For linear stochastic systems the theoretical basis for analysis has a fairly firm footing in mathematical rigor and the main problems tend to be computational. However, when the system considered is nonlinear, deeper problems arise. In fact, what is meant when one writes of the solution of a nonlinear stochastic differential equation has not been resolved rigorously. Many theories have been advanced for ending this dilemma, but interpretation of solutions can almost be described as a matter of personal preference rather than one with a complete mathematical justification.

The efficacy of stochastic concepts in systems engineering is no less apparent because of these mathematical controversies, but consequently much care must be used in applying the concepts of stochastic analysis to nonlinear systems. A subject of particular concern is the generation of solutions of nonlinear stochastic systems by digital simulation. Since an accurate computer model of a physical process reflects not only the physics involved, but also the underlying mathematical theory, the computer simulation of nonlinear systems involves discrepancies arising from the disunity of the mathematics. A deeper understanding of the relationships between the mathematics and the modeling of nonlinear random systems is the goal of this research.

History of the Problem

The successful analysis of a physical system is dependent on the choice of a model which reflects adequately the important physical properties of the system. Experimental evidence has indicated that differential equations are accurate mathematical idealizations of reality and, indeed, Newton's motive for developing the theory of differential and integral calculus was as a systematic method for analyzing physical problems. Choosing an appropriate model usually requires insight and experience, however. Among many other problems, if the model is too detailed, a prohibitive amount of computation must be performed, and if the model is too general, results may be meaningless for a specific system.

Another major concern in model development is to construct a model which can be analyzed using proved techniques. If a physical system can be modeled with no random elements, then a fairly complete theory exists for analysis. However, many problems arise in which the random nature of the system must be incorporated into the mathematical model. These systems have necessitated the development of a theory of stochastic differential equations and, more recently, of a stochastic calculus.

One of the first considerations of a comprehensive theory of stochastic systems is the choice of a mathematically tractable model for noise processes which is also a good approximation to physical reality. Many noise processes in nature are approximately Gaussian and approximately stationary, and they have a power spectrum that is essentially flat up to high frequencies. Such processes that decay at high frequencies in a manner that is not known accurately and that has little effect on system solutions at lower frequencies can be modeled as white noise. This noise process, whose mathematical representation has a flat power spectrum corresponding to power present at every frequency, is called "white" in analogy to white light, which has energy at every frequency of the visual spectrum. White noise has values which are independent at any two distinct times and it also has infinite variance. Such a completely nonphysical process provides a meaningful and tractable idealization of physical noise processes and is generally the most useful for modeling stochastic systems.

Wiener [1] [2] [3] developed the first example of a random process, the "Brownian-motion" or "Wiener" process. This stochastic process is historically important because it is the first significant introduction of Lebesgue's theory into probability theory and practically important because it can be interpreted as the integral of white noise. Wiener was also interested in the more general problem of analyzing nonlinear equations with random elements. McKean [4] presents a detailed investigation of differential and integral calculus based upon the Wiener process.

The very important work of Ito [5-11] in the development of a stochastic calculus was motivated primarily by the desire to construct Markov processes whose transition probabilities satisfy particular Kolmogorov equations and to investigate continuity and other properties of sample functions. The Ito stochastic integral is defined only for noise processes which are martingales, but it exists under very general restrictions on the integrand. Stochastic processes resulting from Ito's integral are Markov processes and certain properties of the integral are very useful for the computation of moments resulting from the stochastic integration. Ito also showed that his integral is a martingale of Brownian motion. The extensive theory of Markov processes and martingales, along with the properties mentioned above, explains the popularity of Ito's stochastic integral, especially with mathematicians.

Stratonovich [12] [13] also used stochastic integrals and equations as a means of studying diffusive Markov processes. He proposed a method for defining stochastic integrals which was much like Ito's method, but which had a number of computational advantages. The Stratonovich integral is defined under less general conditions than the Ito integral, but the Stratonovich integral obeys the rules of ordinary calculus, such as integration by parts, whereas the Ito integral does not. Stratonovich integrals are also Markov processes, although the moment properties which hold for the Ito integral do not hold for the Stratonovich integral. Also, estimation is more complicated using Stratonovich integration.

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 [14] 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. He 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 [15-19] made decisive contributions toward unifying the theory of a stochastic calculus. He defined a stochastic integral by a modification of the procedure which Riemann had 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. He removed the discrepancies arising from the Ito and Stratonovich theories by introducing the "doubly stochastic" integral and provided a method for estimating the value of a stochastic integral. McShane laid the foundation for a unified theory of stochastic integration which includes both Lipschitzian and Brownian-motion processes.

Wright [20] considered the digital simulation of stochastic differential equations. He noted correspondences among the various definitions of stochastic integrals and certain well-known numerical integration algorithms. He considered a specific nonlinear stochastic differential equation, solved it numerically using several different numerical integration procedures and then investigated the behavior of the solution at a particular point. This behavior provided preliminary indications that the relationships among the various integral definitions and numerical integration procedures were as conjectured.

Many of the concepts and results mentioned above are formulated precisely in the following section. After a review of the necessary mathematics and a discussion of the general form of the model, various definitions of stochastic integrals are given and some theoretical consequences of these definitions are presented. The next section describes the general approach to the problem, followed by an outline of the remainder of the thesis.

Mathematical Background

In order to establish notational conventions and to provide easy reference, some well-known concepts will be defined. A real-valued function $x(\omega)$, defined on a space Ω, is a random variable if there is a probability measure P defined on sets of Ω and if $\{\omega|x(\omega) \leq \lambda\}$ is P-measurable for every real number λ. The function $F(\lambda) = P\{\omega|x(\omega) \leq \lambda\}$ is the distribution function of $x(\omega)$ and, if $F(\lambda)$ is absolutely continuous, then $f(\lambda) = F'(\lambda)$ is called the density function of $x(\omega)$. The probability of a set Ω_0 conditioned on a set Ω_1 is $P\{\Omega_0|\Omega_1\} = P\{\omega|\omega\varepsilon\Omega_0, \omega\varepsilon\Omega_1\}/P\{\omega|\omega\varepsilon\Omega_1\}$. The distribution of $x(\omega)$ conditioned on the set Ω_0 is defined as $F(\lambda|\Omega_0) = P\{\omega|x(\omega) \leq \lambda, \omega\varepsilon\Omega_0\}/P\{\omega|\omega\varepsilon\Omega_0\}$ and the conditional density of $x(\omega)$ is $f(\lambda|\Omega_0) = F'(\lambda|\Omega_0)$, provided $F(\lambda|\Omega_0)$ is absolutely continuous. For simplicity, the dependence of the random variable x on $\omega \in \Omega$ usually will not be indicated. Then the expected, or mean, value of x is given by $E\{x\} = \int_{-\infty}^{\infty} \lambda f(\lambda)d\lambda$ and the variance is $Var\{x\} = E\{(x - E\{x\})^2\}$. The random variables x and y are said to be independent if $P\{x \leq \lambda, y \leq \gamma\} = P\{x \leq \lambda\}P\{y \leq \gamma\}$.

A stochastic process is a family of random variables $x(t)$, $t \in T$; the set T will be assumed to be a time range. A stochastic process is called Markov if $P\{x(t_2) \leq \lambda|x(t), t \leq t_1\} = P\{x(t_2) \leq \lambda|x(t_1)\}$ for $t_1 < t_2$ and a martingale if $E\{|x(t)|^2\} < \infty$ for all t and if $t_1 < t_2 < \cdots < t_{n+1}$, then $E\{x(t_{n+1})|x(t_1),x(t_2),\ldots,x(t_n)\} = x(t_n)$. Also, if $\{x_n\}$ is a sequence of random variables, x_n converges to x in the mean if $E\{|x_n|^2\} < \infty$, $E\{|x|^2\} < \infty$ and $\lim_{n\to\infty} E\{|x - x_n|^2\} = 0$. This is written $\underset{n\to\infty}{\text{l.i.m.}}\, x_n = x$. If $\lim_{n\to\infty} P\{|x_n - x| \geq \varepsilon\} = 0$ for every $\varepsilon > 0$, x_n converges to x in probability. The L_p-norm of a function $f(t)$ is defined as $||f(t)||_p = (E\{|f(t)|^p\})^{1/p}$.

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

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

are often encountered. Here $\underline{\dot{x}}(t)$ 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 has often been written

$$\dot{\underline{x}}(t) = \underline{f}(\underline{x},t) + \underline{g}(t)\underline{u}(t) \tag{1.1}$$

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

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

$$E\{\underline{u}(t)\} = \underline{0}, \quad E\{\underline{u}(t)\underline{u}^T(t+\tau)\} = Q \cdot \delta(\tau) \tag{1.2}$$

where E denotes expectation and Q is a matrix expressing how the components of $\underline{u}(t)$ are correlated among themselves. Thus the white noise process has infinite variance and independent process values at any two distinct times [21].

In the scalar case of Equation (1.1) with $g(t) = 1$, one is confronted with the integral

$$w(t) = \int_0^t u(s)ds. \tag{1.3}$$

Because of the pathological nature of white noise, it is difficult to interpret Equation (1.3) rigorously. For systems which are linear in noise terms, that is, the noise is additive rather than multiplicative, this point is avoided by simply assuming the absolute convergence of the integral in Equation (1.3) and very useful results obtain, such as covariance analysis [22]. Nonlinear equations require a more critical evaluation of Equation (1.3), 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:

$$w(0) = 0,$$

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

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

This process was studied by Wiener and is often called the Wiener process or Brownian-motion process. Doob [23] and Parzen [24] 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 [25] oscillation property holds, i.e., if $\{a = t_1, t_2, \ldots, t_n = b\}$ is a partition of the interval [a,b] and $\Delta = \max_i |t_i - t_{i-1}|$, then

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

In considering the nonlinear analog of Equation (1.1), the Wiener process turns out to be much more amenable to analysis than white noise. The nonlinear system of equations, written in terms of differentials rather than derivatives, is then given by

$$\underline{dx}(t) = \underline{f}(\underline{x},t) + \underline{g}(\underline{x},t)\underline{dw}(t) \tag{1.4}$$

To find a solution of Equation (1.4), it suffices to display a stochastic process $\underline{x}(t)$ which satisfies

$$\underline{x}(t) = \underline{x}(a) + \int_a^t \underline{f}(\underline{x}(s),s)\, ds + \int_a^t \underline{g}(\underline{x}(s),s)\, \underline{dw}(s). \tag{1.5}$$

If $\underline{w}(t)$ were of bounded variation, there would be no problem in interpreting $\underline{x}(t)$ in Equation (1.5). However, the last integral in Equation

(1.5) 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 $\underline{g}(\underline{x},t)$ and $\underline{w}(t)$.

Stochastic Integral Definitions

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 Equation (1.5) is random, the Ito integral is more useful than the Wiener integral.

Definition 1.1 (Ito Integral):

Let z(t) be a martingale process and suppose there exists a monotone nondecreasing function F such that, if $s < t$, $E\{|z(t) - z(s)|^2\} = F(t) - F(s)$ 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

$$\text{(1)} \quad \int_a^b g(x,s)\, dz(s) = \underset{\Delta \to 0}{\text{l.i.m.}} \sum_{i=1}^{n-1} g(x(t_i),t_i) \cdot (z(t_{i+1}) - z(t_i)) \qquad (1.6)$$

where the series converges in the mean to a random variable denoted by the integral on the left in Equation (1.6).

Doob [23] 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.

When the stochastic process z(t) in Definition 1.1 is the Wiener process, then $F(t) = t$ and the integral hypothesis becomes $\int_{-\infty}^{\infty} E\{|g(x,t)|^2\}dt < \infty$ with probability 1. 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 Equation (1.4) 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 $d\Phi = [\partial\Phi/\partial t + (\partial\Phi/\partial x)f + \frac{1}{2}\ q\ (\partial^2\Phi/\partial x^2)g^2]dt + q(\partial\Phi/\partial x)g\ dw$.

<u>Definition 1.2</u> (Stratonovich Integral):

Let z(t) be a Markov process with $\lim_{h\to 0} E\{(x(t+h) - z(t))/h|z(t) = \xi\} = a(\xi,t)$, $\lim_{h\to 0} E\{(z(t+h) - z(t))^2/h|z(t) = \xi\} = b(\xi,t)$ and $\lim_{h\to 0} E\{|z(t+h) - z(t)| > \delta|z(t) = \xi\} = 0$ with a(z,t) and b(z,t) continuous in both arguments and b(z,t) having continuous partial derivative $\partial b(z,t)/\partial z$. Suppose g(z,t) is continuous in t having continuous partial derivative $\partial g(z,t)/\partial z$ and $\int_{-\infty}^{\infty} E\{g(z,t)a(z,t)\}dt < \infty$ and $\int_{-\infty}^{\infty} E\{|g(z,t)|^2 b(z,t)\}dt < \infty$. Let $\{a = t_o,t_1,\ldots,t_n = b\}$ be a partition of [a,b] and $\Delta = \max_i|t_{i+1} - t_i|$. The Stratonovich integral is defined as

$$(S) \int_a^b g(z,t)\, dz(t) = \underset{\Delta\to 0}{\text{l.i.m.}} \sum_{i=1}^{n-1} g\left(\frac{1}{2}(z(t_{i+1}) + z(t_i)), t_i\right) \cdot (z(t_{i+1}) - z(t_i)). \quad (1.7)$$

Although the Stratonovich integral is only defined for integrands which are functions of the integrator process, Stratonovich [12] showed how to extend the integral to more general situations by defining a multidimensional integral. In particular, if dx(t) and dz(t) are related by a stochastic differential equation of the form of Equation (1.4), then $\int_a^b g(x(t),t)\, dz(t)$ can be defined. If the process z(t) in Definition 1.2 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.

When the hypotheses for the existence of both the Ito and Stratonovich integrals are satisfied, there is a connection between the two theories which was shown by Stratonovich [12] and Wong and Zakai [26]. Their results showed that the solution, in terms of the Stratonovich integral, of the equation

$$dx(t) = f(x,t)dt + g(x,t)dw(t) \quad (1.8)$$

is the same as the Ito solution of the equation

$$dx(t) = f(x,t)dt + g(x,t)dw(t) + \frac{1}{2} g(x,t) \frac{\partial g}{\partial x}(x,t)dt \quad (1.9)$$

where w(t) is the Wiener process. Wong and Zakai proved a further result. If $x_n(t)$ is the solution of the ordinary differential equation obtained from Equation (1.8) by replacing w(t) with $w_n(t)$, where $w_n(t)$ is a continuous, piecewise linear approximation to the Wiener process and $w_n(t)$ converges to w(t), then $x_n(t)$ does not converge to x(t). But $x_n(t)$ does converge in the mean to the solution of Equation (1.9). This result holds for the Ito interpretation of the solution $x(t)$. This is essentially the situation which occurs when a physical process is approximated with white noise as an input.

To illustrate the concepts discussed above, we consider two examples.

Example 1:

Given the nonlinear system of equations

$$\dot{x}_1(t) = x_2(t)u(t) \qquad x_1(0) = 0$$

$$\dot{x}_2(t) = u(t) \qquad x_2(0) = 0$$

where $u(t)$ is white noise input with $\mathrm{Var}\{u(t)\} = q \cdot \delta(t)$, it is seen that $x_2(t) = w(t)$, where $w(t)$ is the Wiener process, and one must evaluate

$$x_1(t) = \int_0^t w(s)\, dw(s).$$

The Stratonovich solution for $x_1(t)$ is obtained by treating $w(t)$ as a smooth function of time and using the rules of ordinary calculus. Thus,

$$x_{1S}(t) = \frac{1}{2} w^2(t).$$

To evaluate the integral in the sense of Ito, one must calculate

$$x_{1I}(t) = \underset{\Delta\to 0}{\mathrm{l.i.m.}} \sum_{i=0}^{n-1} w(t_i)(w(t_{i+1}) - w(t_i))$$

$$= \underset{\Delta\to 0}{\mathrm{l.i.m.}} \sum_{i=0}^{n-1} \frac{1}{2} (w^2(t_{i+1}) - w^2(t_i))$$

$$- \sum_{i=0}^{n-1} \frac{1}{2} (w(t_{i+1}) - w(t_i))^2.$$

From the Levy oscillation property, it follows that

$$x_{1I}(t) = \frac{1}{2} w^2(t) - \frac{1}{2} qt.$$

Example 2:

From the equation

$$\dot{x}(t) = x(t)u(t) \qquad x(0) = 1$$

one obtains the stochastic differential equation

$$dx(t) = x(t)dw(t).$$

The Stratonovich solution is

$$x_S(t) = e^{w(t)}$$

and, from the results of Wong and Zakai, we conclude that this is the same as the Ito solution of

$$dx(t) = x(t)dw(t) + \frac{1}{2}x(t)dt.$$

Using the Ito differential rule with $\Phi(x(t),t) = \ln x(t)$, it is seen that the Ito solution of the original equation is

$$x_I(t) = e^{w(t) - \frac{1}{2}t}.$$

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 $\{t_o, t_1, \ldots, t_n;\ \tau_1, \tau_2, \ldots, \tau_n\}$ where $a = t_o < t_1 < \cdots < t_n = b$ and τ_i is in D for each i and $\tau_i \leq t_i$.

Definition 1.3 (McShane Integral):

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

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

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

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

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

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 also showed that a solution of Equation (1.4) arising from his interpretation of the integral is bounded and continuous in L_2-norm.

The motivation for McShane's work was provided by the discrepancy between solutions of differential equations arising from Lipschitzian inputs and the solutions for inputs satisfying the conditions of Definition 1.3. For Lipschitzian functions, differentials are linearized forms of expressions involving increments. Second-order terms must be included in the stochastic case and this is essentially the source of the problems with stochastic integrals. To handle this problem, McShane defined and proved the existence of "doubly" stochastic integrals of the form $\int g(x,s)dz_1(s)dz_2(s)$ and concluded that the proper stochastic model for systems described by equations of the form $\dot{x}(t) = f(x,t)$ is given by

Equation (1.9) with dt replaced by $(dw)^2$. His interpretation then agrees with the results of Wong and Zakai and he avoids the Ito-Stratonovich controversy by prescribing the form of the stochastic model to be used.

Through the concept of the doubly stochastic integral, McShane has 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. As was noted in Example 1, this 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. Thus the McShane solution of the examples is the same as the Ito solution.

Approach to the Problem

Several objectives were identified as the purpose of this research into the digital simulation of nonlinear stochastic systems. These objectives were: (1) to determine computationally how the Euler and second-order Runge-Kutta methods of numerical integration relate to stochastic integrals, (2) to investigate other numerical schemes in light of the results of the first objective and to show that the discovered relationships are valid, (3) to perform a statistical analysis of the numerical results, (4) to identify the basis of the correspondences between stochastic integrals and their digital simulations and (5) to consider examples which illustrate these concepts. These objectives are described in more detail below.

Euler and Runge-Kutta Integration

The first objective was realized by choosing example problems whose solutions could be determined analytically. These examples were then solved numerically using the Euler and second-order Runge-Kutta integration methods. The digital solutions were then compared to the analytical solutions and from these results, along with comparisons of the details of numerical algorithms with stochastic integral definitions, correspondences were noted for these methods and different stochastic integrals.

Other Numerical Integration Methods

The consequences of using numerical routines which are computationally and conceptually different from the Euler and second-order Runge-Kutta methods were examined. These included a fourth-order Runge-Kutta method and predictor methods. Insights gained from the first objective were utilized here to aid in extending the correspondences. The relationships between numerical algorithms and the corresponding stochastic integrals were formalized in terms of equivalence of moments.

Statistical Analysis

A statistical analysis of the numerical results was performed to verify their validity. Distributions of sample statistics were used to calculate confidence intervals about the true mean and variance of the solutions. All mean values were shown to lie within these intervals, while a small percentage of some variance estimates exceeded the bounds.

Basis for Discovered Correspondences

The discovery of the basis of the relationships between stochastic integral definitions and numerical integration methods was the next objective. This basis was first studied by identifying integral definitions with numerical algorithms which use the same point of functional evaluation of the integrand. The point of evaluation was discovered to be the mechanism which establishes the correspondences and it is manifested most notably in the correlation coefficient function, which was shown to be the unifying concept for the integral-numerical method relationships.

Examples

Instances of nonlinear differential equations with multiplicative noise arise in applications and some examples were studied to illustrate the discovered relationships. These include an optimal nonlinear filtering application, a phase-locked loop example and the estimation of pollution concentration in the air.

Outline of Thesis

The purpose of this research was to gain insight into the effects which modeling and simulation have on the solutions of nonlinear stochastic systems. Preliminary information and objectives have been presented in this chapter. Chapter II is concerned with exploring the connections between stochastic integrals and numerical integration methods. This early investigation involves low-order numerical routines and moment calculations for purposes of comparison. Chapter III continues with higher-order methods and their moments, as well as a broader class of algorithms, and considers the accuracy of solutions and the confidence we may place in them. The theoretical basis of the correspondences is also addressed here. Several applications of nonlinear stochastic systems are presented in Chapter IV. These examples help to illustrate the results obtained earlier. Conclusions and recommendations for further areas of research are discussed in Chapter V.

CHAPTER II

STOCHASTIC INTEGRALS AND NUMERICAL ALGORITHMS

Since few deterministic differential equations of practical interest can be solved analytically, the digital computer has become an invaluable tool for obtaining numerical solutions of these equations. Stochastic differential equations are no easier, and usually much harder, to solve than deterministic equations. The question of immediate concern to one interested in the solution of stochastic equations is the computational procedure employed in obtaining numerical answers.

Numerical methods for solving deterministic differential equations are historically based on area-finding schemes, but the logic of this approach is not apparent for solving stochastic equations. Thus, because of the differences in the definitions of stochastic integrals, the application of deterministic algorithms to stochastic systems cannot automatically be expected to yield consistent results. This chapter presents some consequences of using deterministic numerical integration schemes to solve stochastic equations and investigates the relationships of these results to stochastic integrals.

Preliminary Considerations

Studying the stochastic integral definitions in the previous Chapter, 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 is not true of stochastic integrals. In this case 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 is that the value of the integral is affected by the evaluation point. The examples presented earlier show this discrepancy. 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. That second-order terms do not necessarily vanish is a consequence of the mean square value of the integrator process possibly being on the order of t rather than Δt.

Because of the discrepancies within the theory of stochastic integrals and the differences between it and the deterministic theory, we are 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, this circumstance, and a deeper understanding of its implications, becomes an important topic of study.

Wright [20] provided some additional evidence that care should be taken when solving stochastic equations digitally. His limited experiment involved the behavior of a single point on the trajectory of the solution of a nonlinear stochastic differential equation when different step sizes were used for the numerical algorithm. One sample trajectory was considered and his results indicated that the point of interest tended to converge to a specific value, which was computed theoretically, as the step size decreased. However, the limiting value was different for various numerical integration routines.

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

Since we are interested in stochastic integrals, we will restrict attention to scalar equations of the form of Equation (1.4) with $f(x,t) \equiv 0$, i.e., $dx(t) = g(x,t)dw(t)$ with the stochastic process $w(t)$ a Wiener process and $g(x,t)$ a random function. In the deterministic case, we have the equation $\dot{x}(t) = g_1(x,t)$ with $g_1(x,t)$ no longer random. Solving this equation involves computing $\int g_1(x,t)dt$ and in a similar manner we 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 possibly anomalous behavior of individual sample functions from a stochastic process which is not ergodic must be taken into account when simulating a random system. This potential problem may be avoided by employing Monte Carlo simulations rather than studying single solution trajectories. In this method several sample stochastic integrator processes are employed in obtaining an ensemble mean and variance for the digitally generated time solution of the stochastic integral $\int g(x,t)dw(t)$. The behavior of individual time histories is not important; rather the behavior of aggregates of time histories is studied.

In the following sections the generation of input noise processes for use in numerical simulations is discussed and Euler's method of integration is considered with emphasis on its stochastic properties and the results of using this method for solving the examples presented in Chapter I. A second-order Runge-Kutta method is then presented, followed by determination of the mean and variance of the Ito and Stratonovich integrals and a discussion of conclusions which can be drawn from these studies.

Input Noise Generation

When using Monte Carlo methods to generate the mean and variance of random systems, several samples from the noise input process must be simulated digitally. Pseudo-random number generators with appropriate statistics are generally used for the digital generation of these input noise samples. The Monte Carlo solution of a stochastic integral likewise requires a way of obtaining sample functions of the integrator process.

We must therefore have a method of generation for the sample functions from a Wiener process. Any such algorithm must, of course, maintain the salient properties of the theoretical, continuous Wiener process. One procedure employs the concept of integration of white noise, which, as mentioned earlier, is a heuristic way of defining the Wiener process. This may be accomplished digitally by using a pseudo-random number generator to obtain a time history of Gaussianly distributed numbers, adding these numbers sequentially and scaling by a nonlinear time transformation.

Specifically, the value of the generated Wiener process at any time is given by

$$w(t_i) = \sqrt{dt} \sum_{n=1}^{i} u(t_n) \tag{2.1}$$

where t_i denotes the i-th sampling time, dt is the sampling period, which remains fixed and $u(t_i)$ is the i-th zero-mean, uncorrelated, Gaussianly distributed random number with variance q.

Noting that the samples $u(t_i)$ are zero-mean, it is apparent that the mean of the samples $w(t_i)$ is also zero. Computing the variance yields

$$\begin{aligned} Var\{w(t_i)\} &= E\{w^2(t_i)\} \\ &= dt \sum_{n=1}^{i} E\{u^2(t_n)\} \\ &= dt \sum_{n=1}^{i} q \\ &= qt_i . \end{aligned} \tag{2.2}$$

The second step in the above derivation is valid because $E\{u(t_i)u(t_j)\} = 0$ if $i \neq j$ and the third step holds since the weight of each squared random number is q. We also have

$$E\{w(t_i)w(t_j)\} = dt\, E\{ \sum_{n=1}^{i} u(t_n) \sum_{m=1}^{j} u(t_m) \}$$

$$= dt \sum_{n=1}^{k} E\{u^2(t_n)\}, \qquad k = \min(i,j)$$

$$= \begin{cases} qt_i, & i < j \\ qt_j, & j < i \end{cases} . \tag{2.3}$$

It is now an easy calculation to verify that

$$E\{[dw(t_i)]^2\} = E\{[w(t_i) - w(t_{i-1})]^2\}$$

$$= q \cdot dt \tag{2.4}$$

and

$$E\{[w(t_i) - w(t_j)][w(t_k) - w(t_l)]\} = 0 \tag{2.5}$$

if the time intervals do not overlap. The above properties indicate that we now have an acceptable method for digitally generating a Wiener process.

Euler's Method

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 $dx(t) = g(x,t)dw(t)$ 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

$$x_{i+1} = x_i + g(x_i,t_i)(w_{i+1} - w_i) \tag{2.6}$$

where $x_i = x(t_i)$.

From Equation (2.6) we can calculate the statistics of the solution of a stochastic equation which has been solved by Euler's method. Specifically,

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

$$= E\{x_i\} \tag{2.7}$$

since $g(x_i,t_i)$ is independent of the Wiener process increment and the Wiener process has mean value zero. The independence is explained by the fact that

$$E\{u(t_i)u(t_j)\} = 0, \qquad i \neq j.$$

It follows that

$$E\{x_i\} = E\{x_o\} \tag{2.8}$$

for every i.

The mean square value is given by

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

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

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

$$= E\{x_i^2\} + q\, E\{g^2(x_i,t_i)\}(t_{i+1} - t_i) \tag{2.9}$$

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

$$\mathrm{Var}\{x\} = E\{x^2\} - E^2\{x\} \tag{2.10}$$

we obtain, using Equation (2.8),

$$\mathrm{Var}\{x_{i+1}\} = \mathrm{Var}\{x_i\} + q\, E\{g^2(x_i,t_i)\}(t_{i+1} - t_i). \qquad (2.11)$$

Numerically, Equation (2.10) behaves as the integral of $E\{g^2(x_i,t_i)\}$.

The Euler numerical integration of Examples 1 and 2 of Chapter I was performed using a fixed integration step size of approximately 0.002 seconds and 100 sample trajectories of solutions were ensemble-averaged to estimate the mean and variance. The initial condition x(0) and the variance parameter q of the Wiener process were chosen to be unity in both examples. Equation (2.8) indicates that the mean value of the solutions in both cases is also unity. For the equation dx = wdw, Equation (2.11) implies that the variance behaves as the integral of $E\{w_i^2\}$, that is, as the integral of t_i. For dx = xdw, the variance is given by the integral of $E\{x_i^2\}$. Figure 1 presents the simulation results for dx = wdw and Figure 2 shows the corresponding results for the equation dx = xdw.

Runge-Kutta Integration

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 Runge-Kutta method of order two is

$$x_{i+1} = x_i + \frac{1}{2}[g(x_i,t_i) + g(x_i + dx, t_i)](w_{i+1} - w_i). \qquad (2.12)$$

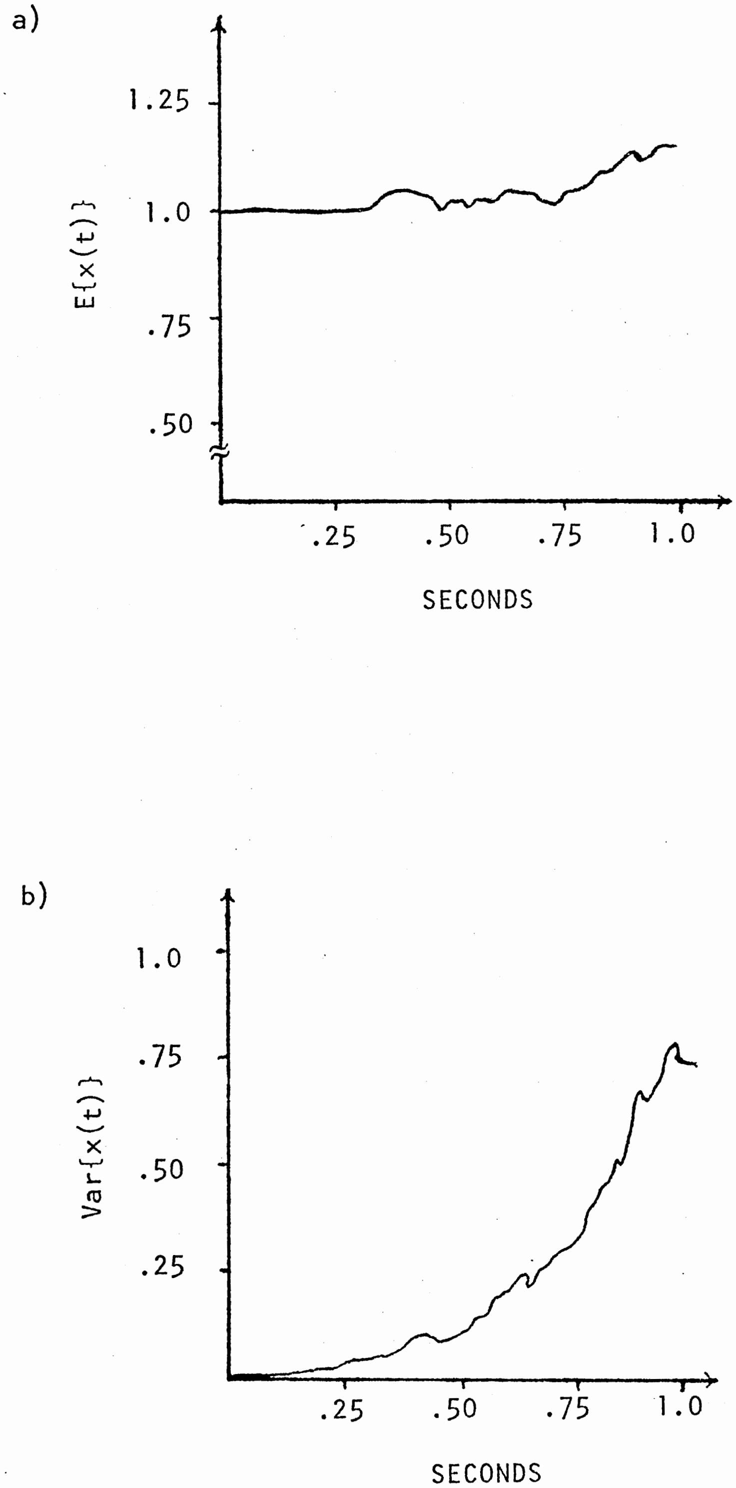


Figure 1. Ensemble-Averaged Mean and Variance, Respectively, for dx = wdw (Euler)

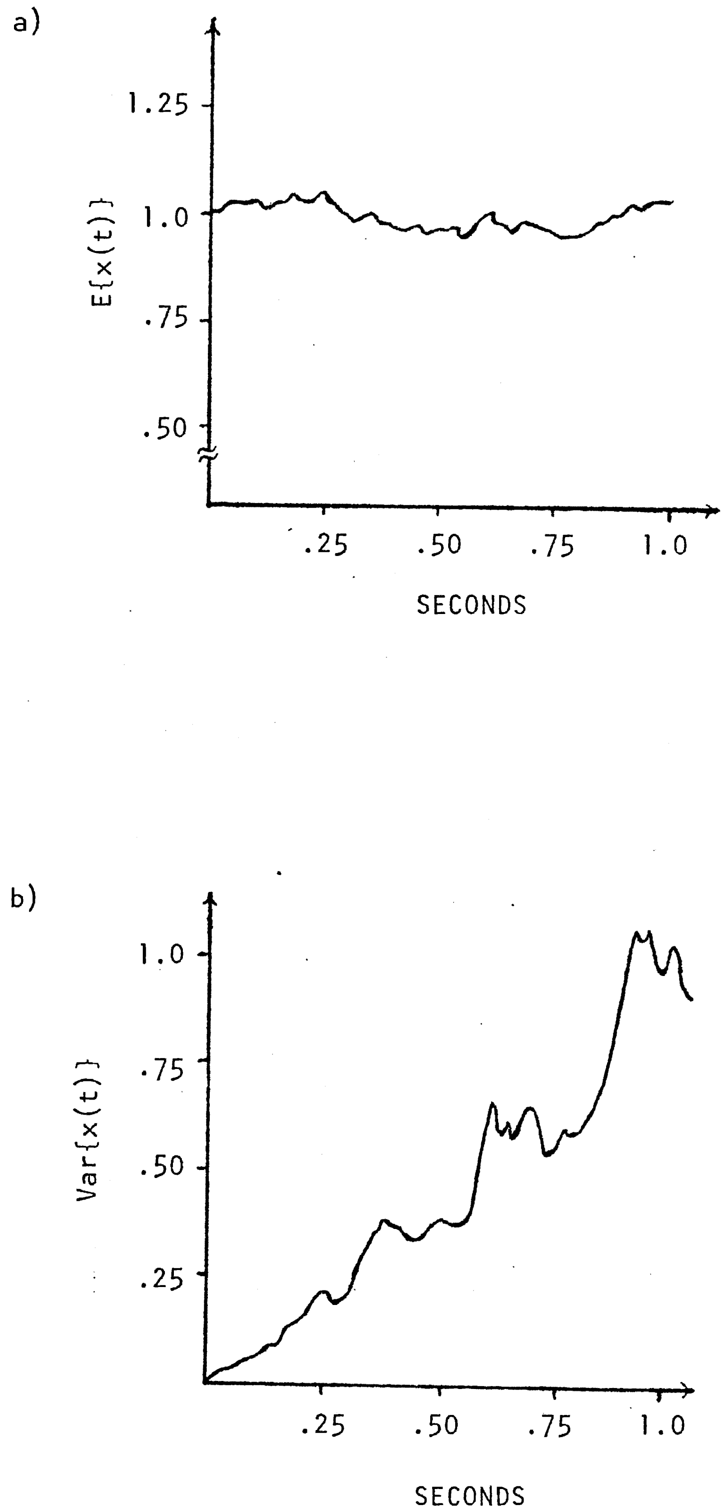


Figure 2. Ensemble-Averaged Mean and Variance, Respectively, for dx = xdw (Euler)

From Equation (2.12) we can determine the statistics of a solution generated by the Runge-Kutta method. The expected value is

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

$$= E\{x_i\} + \frac{1}{2} E\{g(x_i + dx,t_i)(w_{i+1} - w_i)\} \qquad (2.13)$$

From the differentiability of $g(x,t)$, assumed in Definition 1.2, we have

$$\frac{\partial g(x_i,t_i)}{\partial x} = \frac{g(x_i + dx,t_i) - g(x_i,t_i)}{dx} \qquad (2.14)$$

and consequently

$$g(x_i + dx,t_i)(w_{i+1} - w_i) = \left[\frac{\partial g(x_i,t_i)}{\partial x} dx + g(x_i,t_i)\right] \cdot (w_{i+1} - w_i)$$

$$= \frac{\partial g(x_i,t_i)}{\partial x} g(x_i,t_i)(w_{i+1} - w_i)^2 + g(x_i,t_i)(w_{i+1} - w_i) \qquad (2.15)$$

Thus

$$E\{x_{i+1}\} = E\{x_i\} + \frac{1}{2} q\, E\{g(x_i,t_i) \frac{\partial g(x_i,t_i)}{\partial x}\}(t_{i+1} - t_i), \qquad (2.16)$$

since x_i is independent of $(w_{i+1} - w_i)$, and this is the numerical equivalent of ½q times the integral of $E\{g(x_i,t_i)[\partial g(x_i,t_i)/\partial x]\}$.

To find the variance, we first calculate from Equations (2.13) and (2.15)

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

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

(2.17)

To find the mean square value of x_{i+1}, we make use of Equations (2.12) and (2.15) to obtain

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

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

$$= x_i^2 + \frac{1}{4}\ g^2(x_i,t_i)\left[\frac{\partial g(x_i,t_i)}{\partial x}\right]^2 (w_{i+1} - w_i)^4$$

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

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

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

$$+ g^2(x_i,t_i)\ \frac{\partial g(x_i,t_i)}{\partial x}\ (w_{i+1} - w_i)^3 \qquad (2.18)$$

from which it follows that

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

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

$$+ E\{x_i\ g(x_i,t_i)\ \frac{\partial g(x_i,t_i)}{\partial x}\ (w_{i+1} - w_i)\}^2 \qquad (2.19)$$

Equations (2.17) and (2.19) then combine to provide the variance of x_{i+1} as

$$\mathrm{Var}\{x_{i+1}\} = \mathrm{Var}\{x_i\} + q\, E\{g^2(x_i,t_i)\}(t_{i+1} - t_i) + \frac{1}{4}\,\mathrm{Var}\{g(x_i,t_i)\,\frac{\partial g(x_i,t_i)}{\partial x}\,(w_{i+1} - w_i)^2\}. \tag{2.20}$$

Comparison of Equations (2.6) and (2.20) shows that the variance of the Runge-Kutta method is the same as the variance of the Euler method except for an additional term.

The second-order Runge-Kutta numerical integration of the example problems, employing Equation (2.12), was performed using a step size of approximately 0.002 seconds and again 100 sample trajectories of solutions were ensemble-averaged to estimate the mean and variance. The initial condition x(0) and the variance parameter q were again chosen to be unity in both examples. Figure 3 presents the simulation results for the equation dx = wdw and Figure 4 shows the corresponding results for dx = xdw. The Runge-Kutta integration produces a time-varying mean value in both examples, as well as a time-varying variance.

The statistics of these two example problems exhibit very different behavior when obtained through the Runge-Kutta method rather than through Euler's method. Solving deterministic equations with these methods certainly does not produce these discrepancies in solution form, although some difference is observed because of the approximation error inherent in a particular method. One of the most obvious differences in the Euler method and the Runge-Kutta method is that the point of functional evaluation in the algorithm is not the same. This is also the case for

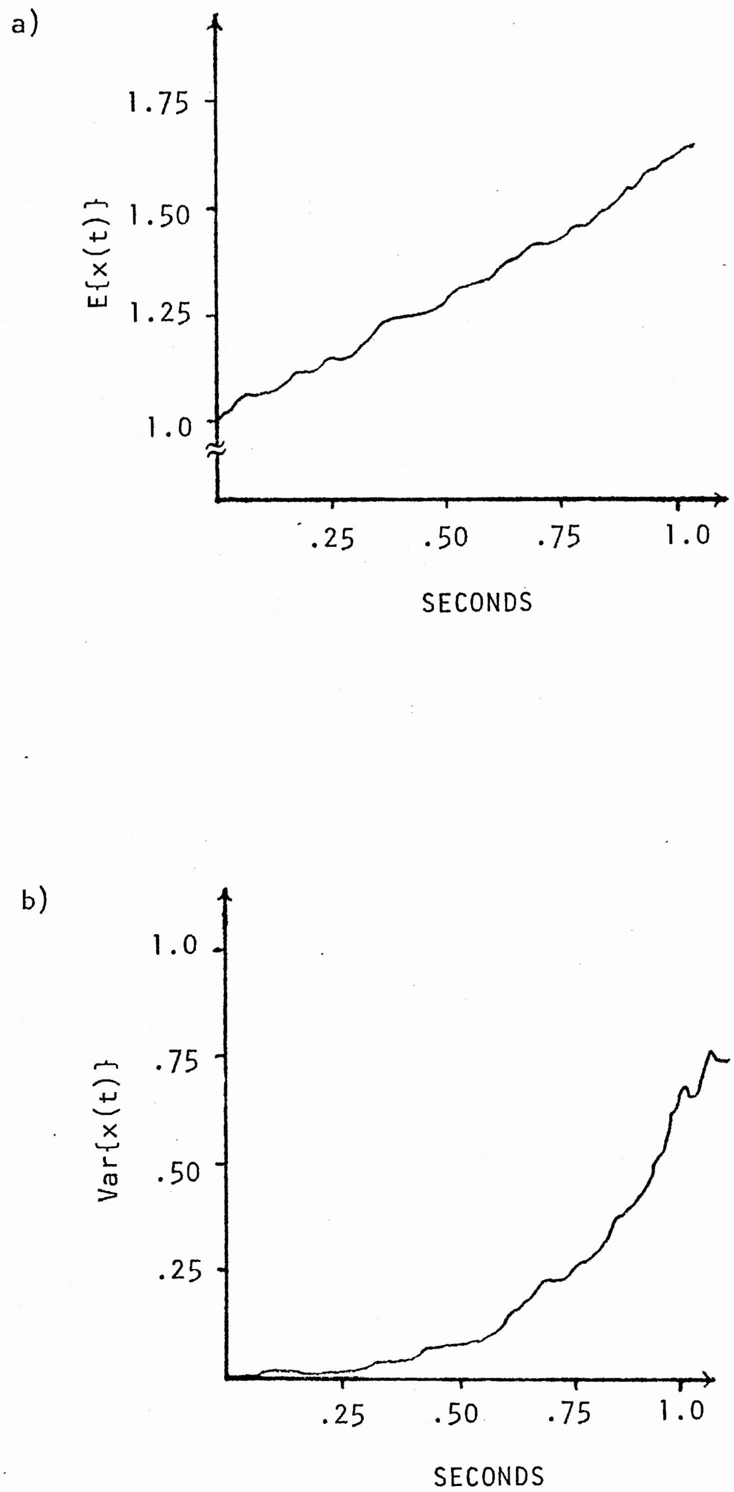


Figure 3. Ensemble-Averaged Mean and Variance, Respectively, for dx = wdw (RK2)

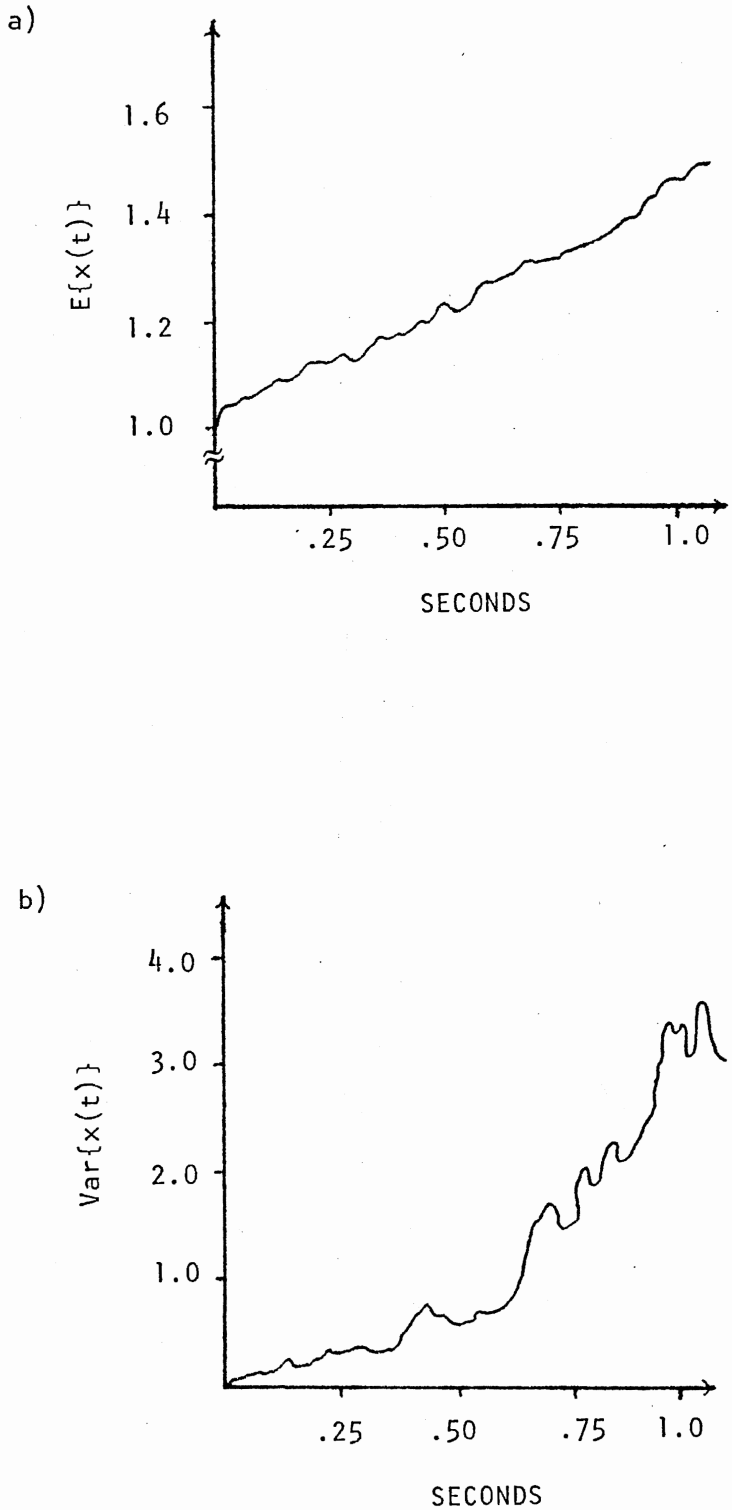


Figure 4. Ensemble-Averaged Mean and Variance, Respectively, for dx = xdw (RK2)

the Ito and Stratonovich definition of the stochastic integral. Further, the evaluation point of the Euler method is the same as that for the Ito integral and the same holds for the Runge-Kutta method and the Stratonovich integral. In order to investigate the effects of these correspondences on the numerical solutions of stochastic equations, we next determine the moments of the Ito and Stratonovich integrals and consider the example problems presented earlier.

Mean and Variance of the Ito Integral

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 [23]). These properties are the following:

$$E\{I \int_a^t g(s)\ dw(s)\} = 0 \tag{2.21}$$

$$E\{I \int_a^t g_1(s)\ dw(s)\ I \int_a^t g_2(s)\ dw(s)\} = q \int_a^t E\{g_1(s)\ g_2(s)\} ds \tag{2.22}$$

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

Given the stochastic differential equation

$$dx(t) = g(x,t)\ dw(t), \tag{2.23}$$

we have the equivalent integral equation

$$x(t) = x(a) + I \int_a^t g(x,t)\ dw(t). \tag{2.24}$$

From Equations (2.21) and (2.24) the mean value of x(t) is

$$E\{x(t)\} = E\{x(a)\}. \tag{2.25}$$

The variance of x(t) may be computed by noting that the initial condition x(a) is independent of $I \int_a^t g(x,t)\, dw(t)$ and by using Equations (2.21) and (2.22) and the identity Equation (2.10). Thus

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

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

$$= E\{x^2(a)\} + q \int_a^t E\{g^2(x,t)\}dt \tag{2.26}$$

and the variance is then

$$\mathrm{Var}\{x(t)\} = \mathrm{Var}\{x(a)\} + q \int_a^t E\{g^2(x,t)\}dt. \tag{2.27}$$

The mean value of Euler's method, given by Equation (2.8), is the same as the mean value of the Ito integral in Equation (2.25). Similarly, Equations (2.11) and (2.27) indicate that the variances agree also. We thus conclude 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. We now consider the Ito solutions of the examples and compare with the numerical results presented earlier.

<u>Example 1</u>:

Given the equation $dx(t) = w(t)\, dw(t)$ with initial condition $x(0) = 1$, the mean value is easily seen to be

$$E\{x(t)\} = 1 \tag{2.28}$$

and the variance, since the initial condition is given, is computed as follows:

$$\mathrm{Var}\{x(t)\} = q \int_0^t E\{w^2(\tau)\}d\tau$$

$$= q \int_0^t q\tau \ d\tau$$

$$= \frac{1}{2} q^2 \ t^2. \qquad (2.29)$$

With $q = 1$, $\mathrm{Var}\{x(t)\} = \frac{1}{2} t^2$. Figure 5 shows the simulation results for Euler's method as obtained earlier, along with the theoretical results from the Ito integral indicated with dashed lines.

Example 2:

Given the equation $dx(t) = x(t)\ dw(t)$ with $x(0) = 1$, the mean value is

$$E\{x(t)\} = 1 \qquad (2.30)$$

and the variance is given by

$$E\{x^2(t)\} - E^2\{x(t)\} = q \int_0^t E\{x^2(\tau)\}d\tau$$

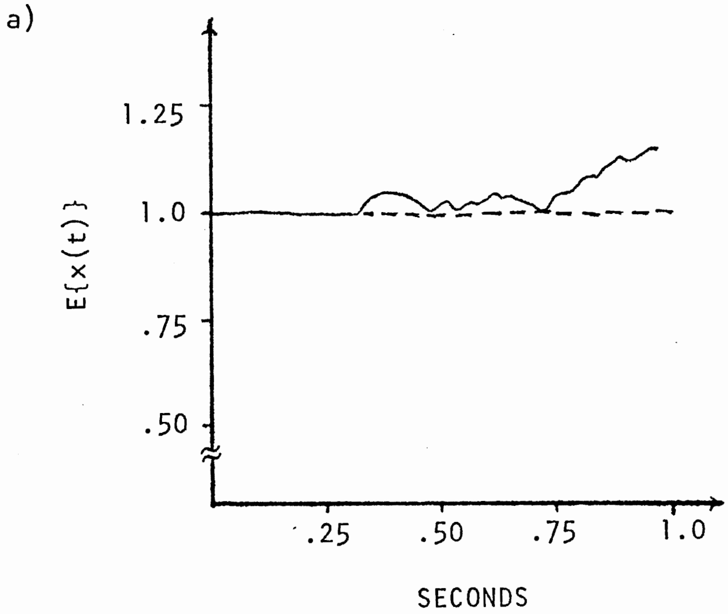
which implies

$$E\{x^2(t)\} - 1 = q \int_0^t E\{x^2(\tau)\}d\tau$$

and it is easily seen that the variance is an exponential function, that is,

$$\mathrm{Var}\{x(t)\} = e^{qt} - 1. \qquad (2.31)$$

With $q = 1$, $\mathrm{Var}\{x(t)\} = e^t - 1$. Figure 6 presents the simulation results for this example, again with the theoretical results indicated by dashed lines.



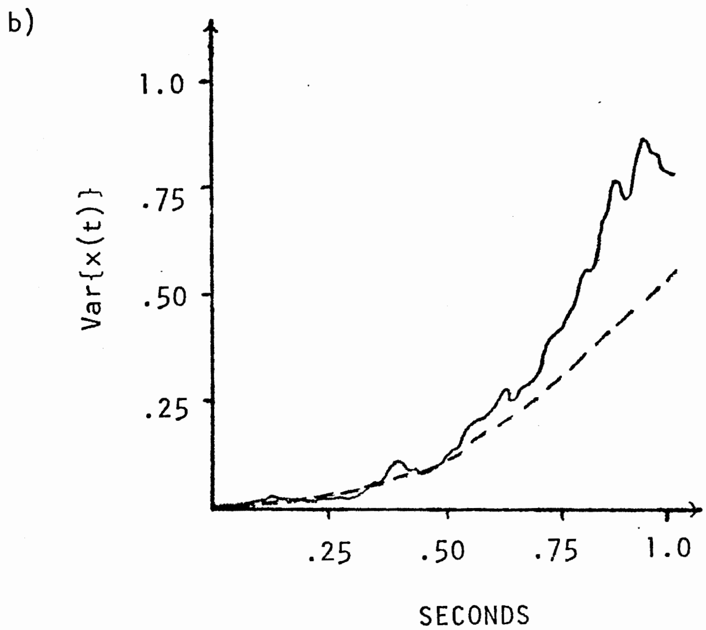


Figure 5. Theoretical and Ensemble-Averaged Mean and Variance, Respectively, for dx = wdw (Euler)

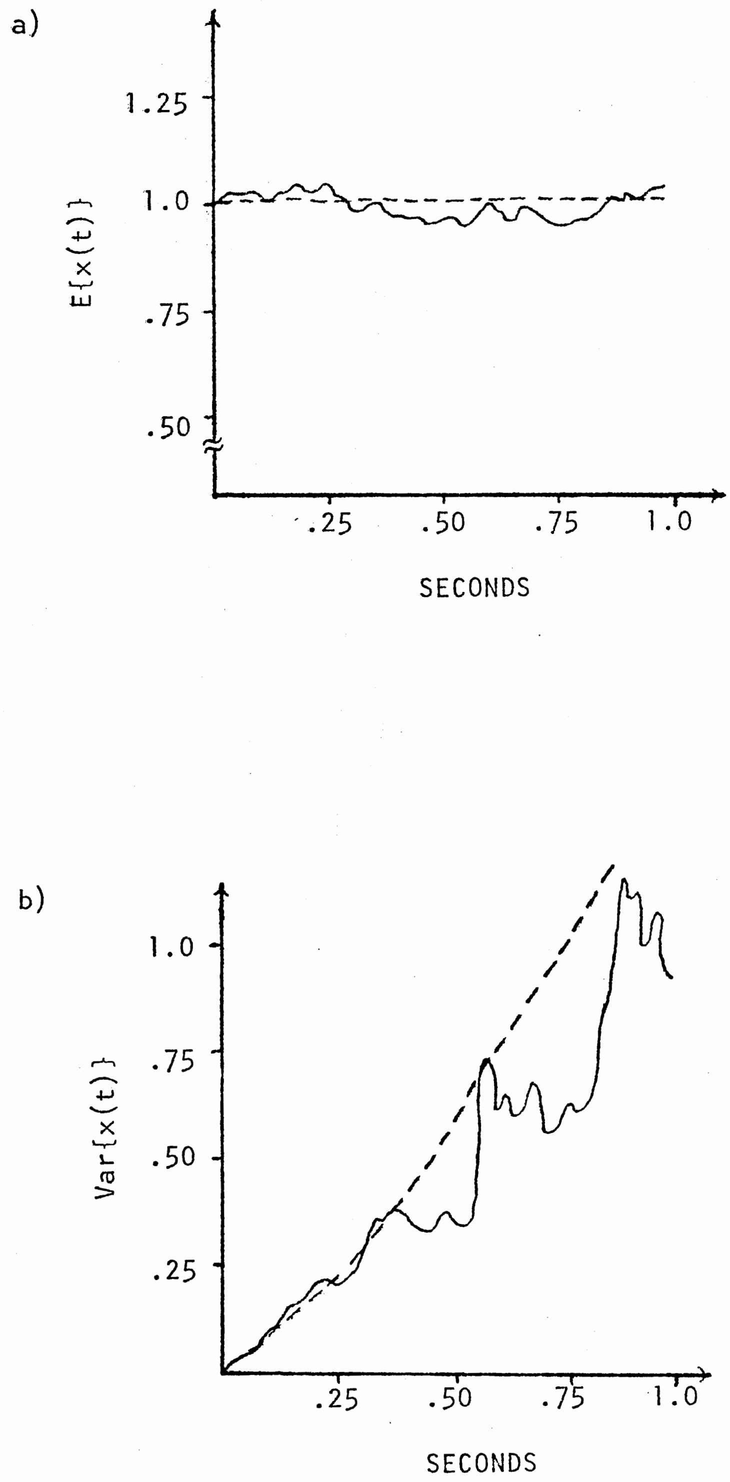


Figure 6. Theoretical and Ensemble-Averaged Mean and Variance, Respectively, for dx = xdw (Euler)

Mean and Variance of the Stratonovich Integral

Stratonovich [12] introduced the stochastic integral bearing his name and proved the fundamental equality

$$S \int_a^t g(w(t),t)\, dw(t) = I \int_a^t g(w(t),t)\, dw(t) + \frac{1}{2} q \int_a^t \frac{\partial g(w(t),t)}{\partial w}\, dt. \tag{2.32}$$

He then showed how to extend this theory for the case in which $g(\cdot,t)$ is not necessarily a function of the integrator process. In this case, the relationship between the Stratonovich and Ito integrals is given by

$$S \int_a^t g(x(t),t)\, dw(t) = I \int_a^t g(x(t),t)\, dw(t) + \frac{1}{2} q \int_a^t g(x(t),t) \frac{\partial g(x(t),t)}{\partial x}\, dt. \tag{2.33}$$

Exploiting this relationship between the Stratonovich and Ito integrals allows the computation of the mean resulting from the differential equation (2.23) when the equation is solved in the Stratonovich sense. Thus

$$\begin{aligned} E\{x(t)\} &= E\{x(a)\} + E\{S \int_a^t g(x(t),t)\, dw(t)\} \\ &= E\{x(a)\} + \frac{1}{2} q\, E\{ \int_a^t g(x(t),t) \frac{\partial g(x(t),t)}{\partial x}\, dt\}. \end{aligned} \tag{2.34}$$

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

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

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

$$+ \frac{1}{2} q \int_a^t g(x(t),t) \frac{\partial g(x(t),t)}{\partial x} dt]\}$$

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

$$+ \frac{1}{2} q \int_a^t g(x(t),t) \frac{\partial g(x(t),t)}{\partial x} dt]^2\} \qquad (2.35)$$

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

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

$$+ \frac{1}{4} q^2 [E\{[\int_a^t g(x(t),t) \frac{\partial g(x(t),t)}{\partial x} dt]^2\}$$

$$- E^2\{\int_a^t g(x(t),t) \frac{\partial g(x(t),t)}{\partial x} dt\}]$$

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

$$\cdot \frac{\partial g(x(t),t)}{\partial x} dt\}.$$

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

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

$$+ \frac{1}{4} q^2\, \mathrm{Var}\{\int_a^t g(x(t),t) \frac{\partial g(x(t),t)}{\partial x} dt\}. \qquad (2.36)$$

By comparing the mean and variance of the Ito solution with these same statistics of the Stratonovich solution, it is seen that the Stratonovich results are the same as the Ito results except for the additional term involving $g(x(t),t)(\partial g(x(t),t/\partial x)$. In light of Equation

(2.33), this is not unexpected. It should be noted that the variance of the Stratonovich solution coincides with the Ito variance if $g(x(t),t) \cdot (\partial g(x(t),t)/\partial x)$ is not a random function.

The mean value of the Runge-Kutta method, given by Equation (2.16), is now seen to be the numerical equivalent of the mean value of the Stratonovich integral, given by Equation (2.34). Comparison of Equations (2.20) and (2.36) indicates that the variances of the Runge-Kutta method and the Stratonovich integral also coincide. We thus conclude 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. We now consider the Stratonovich solutions of the examples and compare these with the results from the Runge-Kutta integration method.

<u>Example 1</u>:

To determine the mean, in the Stratonovich sense, of the equation $dx(t) = w(t)\ dw(t)$ with initial condition $x(0) = 1$, we first calculate $\partial g(t)/\partial w = 1$. Then the expected value of Equation (2.32) yields

$$E\{x(t)\} = 1 + \frac{1}{2} qt. \tag{2.37}$$

To determine the variance, first note that $\partial g(t)/\partial w$ is not a random function. Then an analysis of Equation (2.32), performed in the same way as was done for Equation (2.33), shows that the Stratonovich variance is the same as the Ito variance. Thus

$$\mathrm{Var}\{x(t)\} = \frac{1}{2} q^2\ t^2. \tag{2.38}$$

With $q = 1$, again $\mathrm{Var}\{x(t)\} = \frac{1}{2} t^2$. Figure 7 provides the simulation

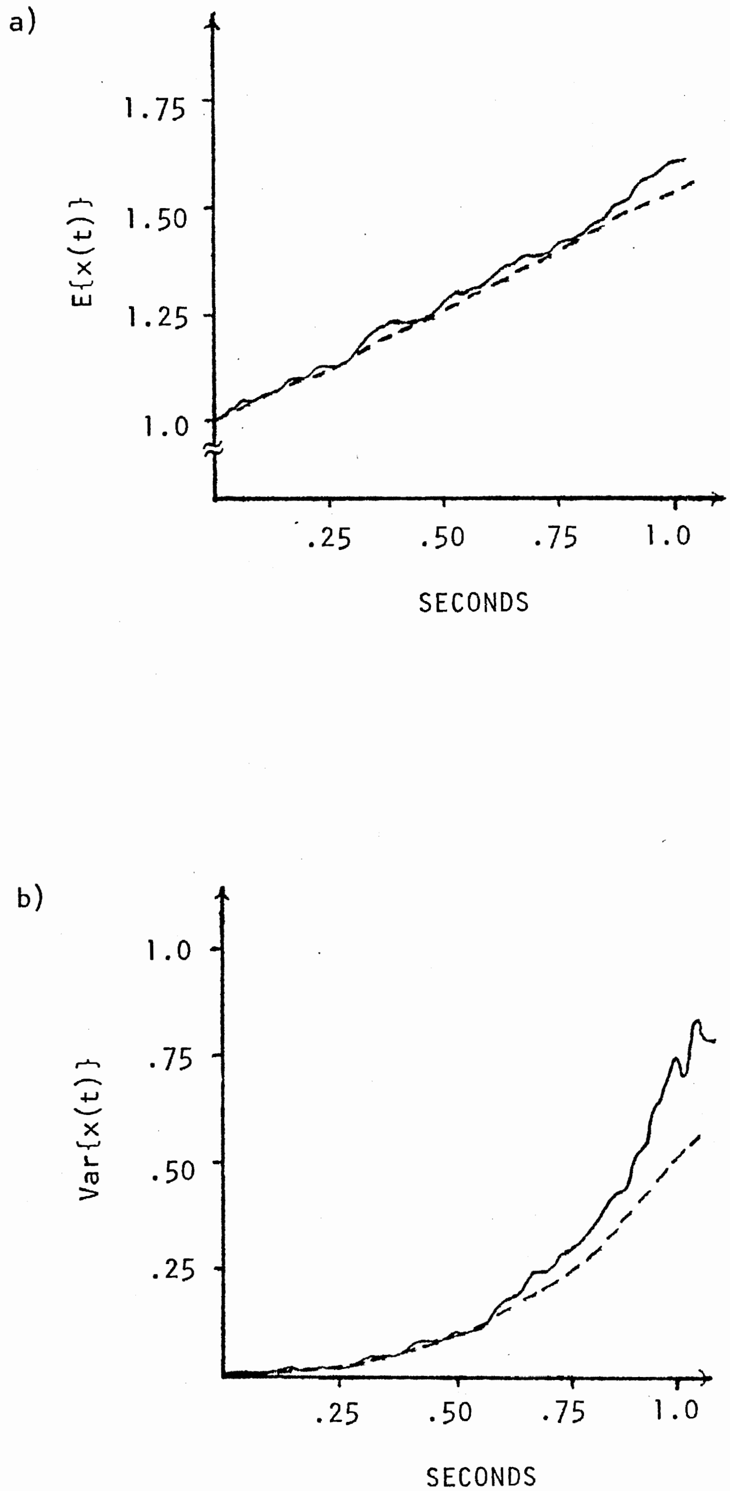


Figure 7. Theoretical and Ensemble-Averaged Mean and Variance, Respectively, for dx = wdw (RK2)

results for the Runge-Kutta method as obtained earlier, along with these theoretical results indicated by the dashed lines.

Example 2:

For the equation $dx(t) = x(t)\,dw(t)$ with $x(0) = 1$, we calculate that $\partial g(x(t),t)/\partial x = 1$ and thus obtain from Equation (2.34)

$$E\{x(t)\} = 1 + \frac{1}{2} q \int_0^t E\{x(\tau)\}d\tau$$

from which it follows that

$$E\{x(t)\} = e^{\frac{1}{2}qt}. \tag{2.39}$$

With $q = 1$, the mean value becomes $e^{\frac{1}{2}t}$.

To determine the variance of $x(t)$, recall from Chapter I that $x(t) = e^{w(t)}$ when integrating $dx(t) = x(t)\,dw(t)$ in the Stratonovich sense. Equation (2.39) shows that

$$E\{e^{w(t)}\} = e^{\frac{1}{2}qt} \tag{2.40}$$

and an easy calculation shows that, in general,

$$E\{e^{\alpha w(t)}\} = e^{\frac{1}{2}\alpha^2 qt}. \tag{2.41}$$

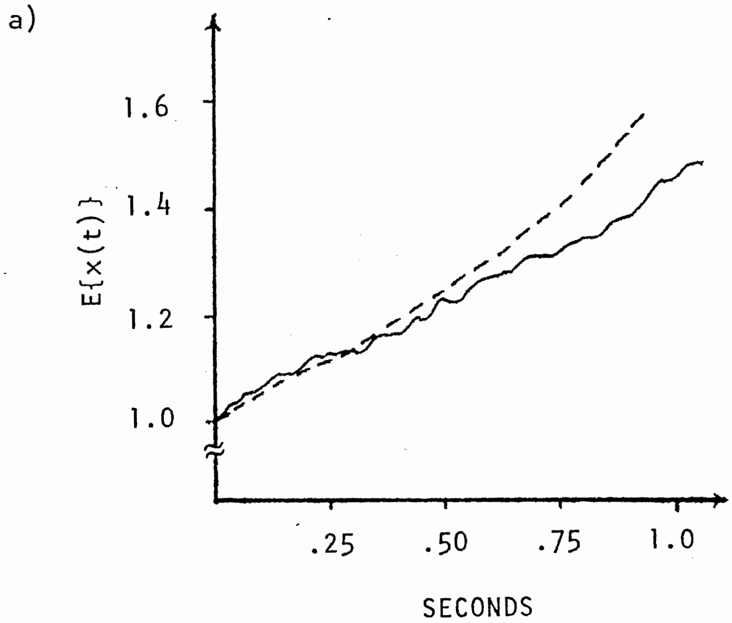
Then the variance is given by

$$\begin{aligned} Var\{x(t)\} &= E\{e^{2w(t)}\} - E^2\{e^{w(t)}\} \\ &= e^{2qt} - e^{qt} \end{aligned}$$

and, when $q = 1$, we have

$$Var\{x(t)\} = e^{2t} - e^{t}. \tag{2.42}$$

Figure 8 shows the simulation results for this example with the calculated mean and variance indicated by the dashed lines.



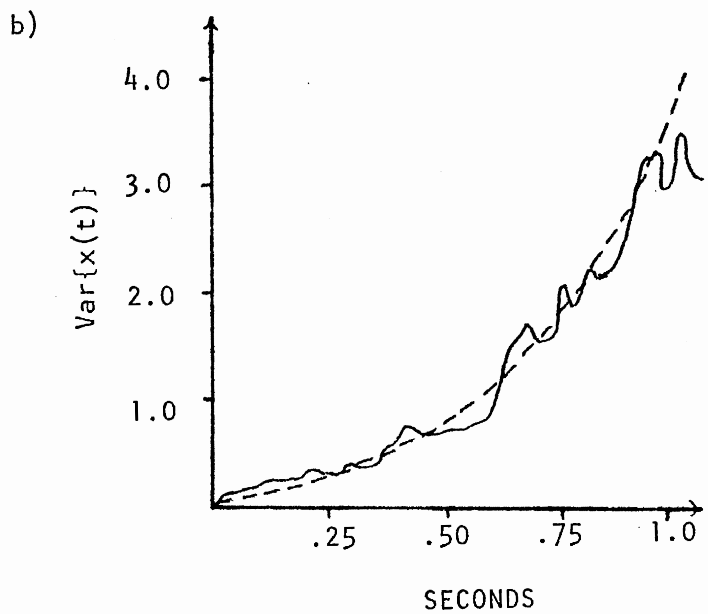


Figure 8. Theoretical and Ensemble-Averaged Mean and Variance, Respectively, for dx = xdw (RK2)

Summary

This chapter dealt with employing Euler and Runge-Kutta methods of numerical integration for solving stochastic differential equations. First and second moments of these results were calculated and Monte Carlo simulations of two examples were performed. The similarity of the Euler method and the Ito stochastic integral were noted and, upon determination of the moments of the Ito integral, the numerical equivalence of these two concepts was shown. Analogous results were given for the equivalence of a second-order Runge-Kutta method and the Stratonovich stochastic integral. Theoretical moments were then compared to the moments obtained from the numerical simulation.

The results of this chapter indicate the importance which the point of functional evaluation of the integrand plays in the theory and practical implementation of stochastic integration. These conculsions lead naturally to consideration of the relationships among stochastic integrals and other, more sophisticated, conceptually different numerical integration schemes. The next chapter discusses these ideas and their consequences and also addresses the question of accuracy of numerical routines.

CHAPTER III

ERROR BOUNDS AND SYSTEM CORRELATION

There are many different types of numerical integration routines which have been developed for solving systems of differential equations. When applied to deterministic systems these methods produce consistent results, but, as was shown in Chapter II, this consistency does not carry over to the stochastic case. This chapter begins with a look at broad classes of numerical algorithms and how these correspond to stochastic integration.

The question of accuracy is always important in the numerical solution of equations. The algorithm itself gives rise to errors and another cause of uncertainty in the analysis of stochastic systems is the use of ensemble statistics. These sources of errors and their consequences in terms of reliability of solutions is also discussed in this chapter.

Theoretical differences in stochastic integrals manifest themselves as profound discrepancies in the actual solution process. This anomalous behavior in the theory, which is also apparent in digitally generated solutions of these integrals, raises the question of what mechanism is responsible for the differences. It is shown that the correlation is the key to the relationships discovered earlier. The final subject of this chapter is thus a discussion of the correlation coefficient function as the unifying concept in the study of stochastic integration and digital simulations.

Numerical Integration Methods

Numerical integration methods have been developed in many different contexts and with many different purposes in mind. However, convergence properties, error characteristics and computation requirements are the most important considerations in their use. Complexity also varies greatly among the various methods with the accuracy of computed solutions tending to be greater with the increased complexity, as would be expected. The methods can be broadly classified as being single-step or multi-step, with single-step methods generally involving more computation and multi-step methods requiring more storage space.

Single-step methods do not require any functional evaluations before the current interval of interest. For this reason, these methods are often used to provide starting values for multi-step formulas. Several computations of the derivative within each interval may be required and, consequently, these methods can be quite time-consuming. Error calculations are somewhat difficult, although the accuracy of solutions is good. Incorporating a variable step size into these algorithms is also comparatively easy. Runge-Kutta formulas are the most widely used of the single-step methods.

Multi-step methods require a past history of derivative values for their use. This implies that storage capacity be slightly greater than for single-step methods. Changing the step size is complicated and these methods must be primed with values from a single-step method. However, error calculations are not difficult and computation time is reduced since several derivative calculations are not necessary for each computed solution value. Predictor methods are typical of multi-step integration formulas.

In considering the use of numerical integration routines to solve stochastic differential equations, another classification method suggests itself. This is based on the point of functional evaluation of the numerical algorithm. Suppose the current interval is $[t_i, t_{i+1}]$. Then the point of functional evaluation may lie before the current interval, say t_{i-1}, or within the current interval, say $t_{i+\frac{1}{2}}$. In this classification scheme, the evaluation point t_i may belong to either class of evaluation points. The point t_{i+1} obviously must be included within the current interval.

The results of the previous chapter indicate some possible consequences of this grouping of numerical methods. Recall that the Ito stochastic integral definition requires that t_i be the point of functional evaluation. Euler's numerical integration method also uses t_i as the point of evaluation for its computations. The moments of Euler's method, which are presented in Chapter II, show that this Ito-Euler correspondence based on the point of functional evaluation is a valid classification scheme and the numerical results presented in Figures 5 and 6 bear out this correspondence also.

The second-order Runge-Kutta method discussed earlier requires evaluations at t_i and t_{i+1}, that is, within the current interval. The Stratonovich definition of the stochastic integral also specifies an evaluation point within the current interval. Moments of the second-order Runge-Kutta method, calculated in Chapter II, substantiate the correspondence of the Stratonovich integral with numerical routines which require functional evaluations within the current interval. The numerical results in Figures 7 and 8 corroborate this relationship.

Fourth-order Runge-Kutta methods are more accurate and more widely used than second-order methods. However, they still require evaluations within the current interval and thus will be grouped with the Stratonovich integral. In general, single-step integration formulas introduce correlation between the integrand and the noise process by calculating successive integrand values from earlier ones within the same interval. The Stratonovich integral also introduces correlation between the integrand and the noise process and this correlation is the basis of the relationship between the Stratonovich integral and single-step numerical formulas. The next section extends this Runge-Kutta-Stratonovich relationship to the important class of fourth-order Runge-Kutta methods.

Further Runge-Kutta Results

Consider the fourth-order Runge-Kutta method given by

$$x_{i+1} = x_i + \frac{1}{6}\left[g(x_i) + 2g(x_{i1}) + 2g(x_{i2}) + g(x_{i3})\right] \cdot (w_{i+1} - w_i) \tag{3.1}$$

where

$$x_{i1} = x_i + \frac{1}{2}\, g(x_i)(w_{i+1} - w_i)$$

$$x_{i2} = x_i + \frac{1}{2}\, g(x_{i1})(w_{i+1} - w_i)$$

$$x_{i3} = x_i + g(x_{i2})(w_{i+1} - w_i).$$

Three expressions for the derivative of $g(x)$ may be written to aid in the calculation of moments of Equation (3.1):

$$\frac{\partial g_1}{\partial x} = \frac{g(x_i + \frac{1}{2}dx) - g(x_i)}{\frac{1}{2}dx},$$

$$dx = g(x_i)(w_{i+1} - w_i) \qquad (3.2a)$$

$$\frac{\partial g_2}{\partial x} = \frac{g(x_i + \frac{1}{2}dx_1) - g(x_i)}{\frac{1}{2}dx_1},$$

$$dx_1 = g(x_i + \frac{1}{2}dx)(w_{i+1} - w_i) \qquad (3.2b)$$

$$\frac{\partial g_3}{\partial x} = \frac{g(x_i + dx_2) - g(x_i)}{dx_2},$$

$$dx_2 = g(x_i + \frac{1}{2}dx_1)(w_{i+1} - w_i). \qquad (3.2c)$$

Substituting Equation (3.2) into Equation (3.1) yields

$$\begin{aligned} x_{i+1} = {} & x_i + \frac{1}{6} g(x_i)(w_{i+1} - w_i) \\ & + \frac{1}{6} g(x_i) \frac{\partial g_1}{\partial x} (w_{i+1} - w_i)^2 + \frac{1}{3} g(x_i)(w_{i+1} - w_i) \\ & + \frac{1}{6} g(x_i + \frac{1}{2}dx) \frac{\partial g_2}{\partial x} (w_{i+1} - w_i)^2 \\ & + \frac{1}{3} g(x_i)(w_{i+1} - w_i) \\ & + \frac{1}{6} g(x_i + \frac{1}{2}dx_1) \frac{\partial g_3}{\partial x} (w_{i+1} - w_i)^2 \\ & + \frac{1}{6} g(x_i)(w_{i+1} - w_i) \end{aligned} \qquad (3.3)$$

and upon rewriting, we obtain

$$x_{i+1} = x_i + g(x_i)(w_{i+1} - w_i) + \frac{1}{6}[g(x_i)\frac{\partial g_1}{\partial x} + g(x_i + \frac{1}{2}dx)\frac{\partial g_2}{\partial x} + g(x_i + \frac{1}{2}dx_1)\frac{\partial g_3}{\partial x}] \cdot (w_{i+1} - w_i)^2 \quad (3.4)$$

and consequently

$$E\{x_{i+1}\} = E\{x_i\} + \frac{1}{6}q\, E\{g(x_i)\frac{\partial g_1}{\partial x} + g(x_i + \frac{1}{2}dx)\frac{\partial g_2}{\partial x} + g(x_i + \frac{1}{2}dx_1)\frac{\partial g_3}{\partial x}\}(t_{i+1} - t_i) \quad (3.5)$$

since all functions $g(\cdot)$ are functions of x_i and hence independent of the noise process. Noting the regularity conditions on $g(x)$ and $\partial g(x)/\partial x$ which are required in the Stratonovich definition, we can see that the equations in Equations (3.2) are different approximations of the same quantity for small step sizes. We then calculate the mean value associated with the fourth-order Runge-Kutta method as

$$E\{x_{i+1}\} = E\{x_i\} + \frac{1}{2}q\, E\{g(x_i)\frac{\partial g(x_i)}{\partial x}\}(t_{i+1} - t_i) \quad (3.6)$$

and the mean value equivalence is established upon comparison with Equation (2.34).

To determine the variance of the fourth-order Runge-Kutta method, we first let $g_{ij} = g(x_{ij})$, $j = 1, 2, 3$, and $g_i = g(x_i)$ and calculate

$$x_{i+1}^2 = x_i^2 + \frac{1}{3}x_i(g_i + 2g_{i1} + 2g_{i2} + g_{i3})dw$$

$$+ \frac{1}{36} (g_i^2 + 4g_{i1}^2 + 4g_{i2}^2 + g_{i3}^2 + 4g_i g_{i1} + 4g_i g_{i2}$$

$$+ 2g_i g_{i3} + 8g_{i1} g_{i2} + 4g_{i1} g_{i3} + 4g_{i2} g_{i3}) . \qquad (3.7)$$

From Equations (3.2), we see that

$$g_{i1} = \frac{1}{2} g_i \frac{\partial g_1}{\partial x} dw + g_i \qquad (3.8a)$$

$$g_{i2} = \frac{1}{2} g_{i1} \frac{\partial g_2}{\partial x} dw + g_i \qquad (3.8b)$$

$$g_{i3} = g_{i2} \frac{\partial g_3}{\partial x} dw + g_i . \qquad (3.8c)$$

Substituting Equations (3.8) into Equation (3.7) and taking the expected value gives

$$E\{x_{i+1}^2\} = E\{x_i^2\} + \frac{1}{3} E\{x_i (g_i \frac{\partial g_1}{\partial x} + g_{i1} \frac{\partial g_2}{\partial x} + g_{i2} \frac{\partial g_3}{\partial x}) dw^2\}$$

$$+ \frac{1}{36} E\{36 g_i^2 dw^2\} + \frac{1}{36} E\{(g_i^2 \frac{\partial g_1^2}{\partial x} + g_{i1}^2 \frac{\partial g_2^2}{\partial x} + g_{i2}^2 \frac{\partial g_3^2}{\partial x}$$

$$+ 2g_i g_{i1} \frac{\partial g_1}{\partial x} \frac{\partial g_2}{\partial x} + 2g_i g_{i2} \frac{\partial g_1}{\partial x} \frac{\partial g_2}{\partial x}$$

$$+ 2g_{i1} g_{i2} \frac{\partial g_2}{\partial x} \frac{\partial g_3}{\partial x}) dw^4\}. \qquad (3.9)$$

Again noting that the Equations (3.2) are approximations of the same value, we obtain

$$E\{x_{i+1}^2\} = E\{x_i^2\} + E\{x_i g_i \frac{\partial g_1}{\partial x} dw^2\} + E\{g_i^2 dw^2\}$$

$$+ \frac{1}{4} E\{g_i^2 \frac{\partial g_1^2}{\partial x} dw^4\}. \qquad (3.10)$$

Subtracting the square of Equation (3.5) from Equation (3.10) shows that the variance of the fourth-order Runge-Kutta method is given by

$$\mathrm{Var}\{x_{i+1}\} = \mathrm{Var}\{x_i\} + E\{g_i^2\, dw^2\} + \frac{1}{4}\,\mathrm{Var}\{g_i \frac{\partial g_1}{\partial x}\, dw^2\}. \tag{3.11}$$

These results are the same as the second order Runge-Kutta results, Equations (2.16) and (2.20), and extend the Runge-Kutta-Stratonovich correspondence to the fourth-order case.

The fourth-order Runge-Kutta numerical integration of the example problems, using Equation (3.1), was performed using a step size of approximately 0.002 seconds and an ensemble of 100 sample trajectories. The initial value x(0) and the variance parameter q were chosen to be unity. Figure 9 presents the simulation results for the equation dx = wdw, with the solutions from the Stratonovich integral indicated by dashed lines. Figure 10 shows the corresponding results for the equation dx = xdw. Once again, these numerical results behave as expected for the Runge-Kutta-Stratonovich relationship.

Predictor Methods

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}, ..., x_{i-k+1}. Predictor methods are thus multi-step 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

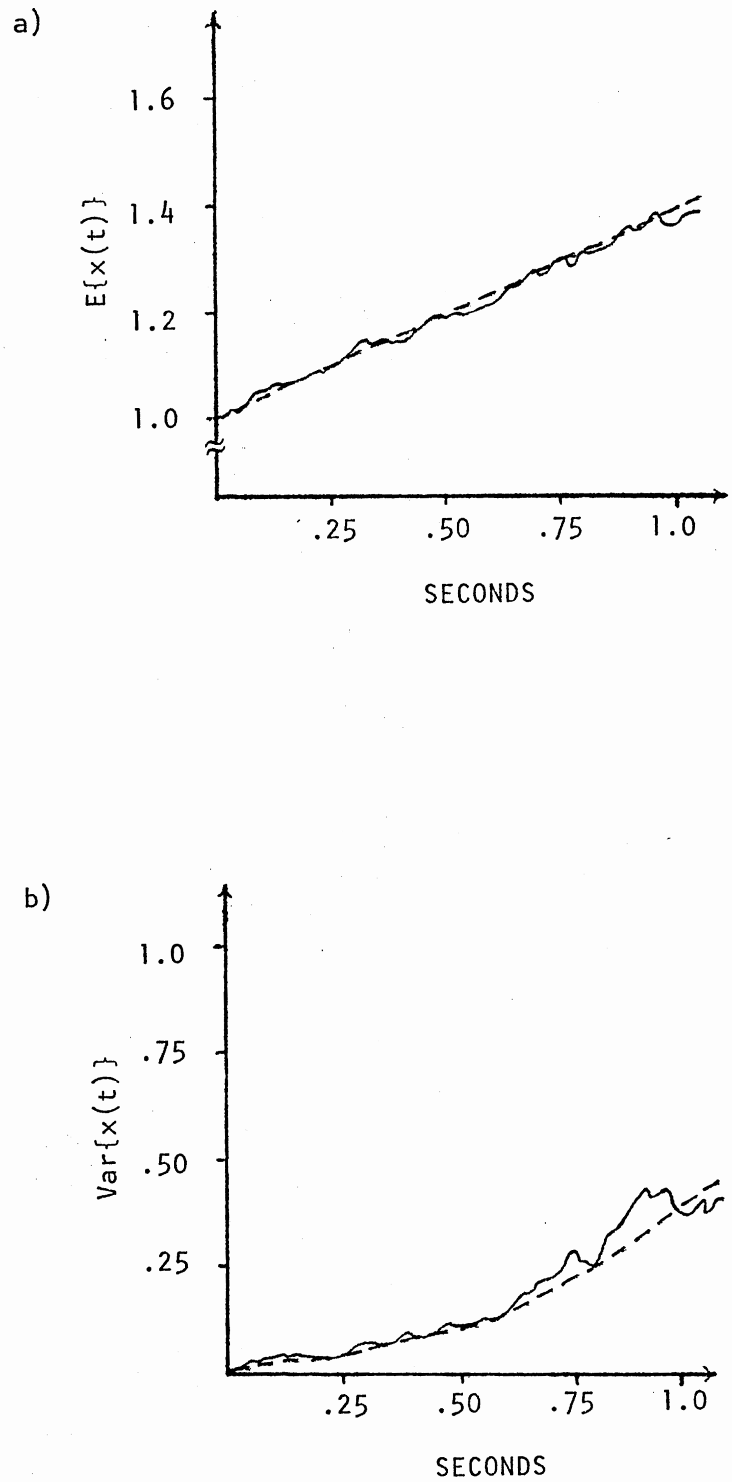
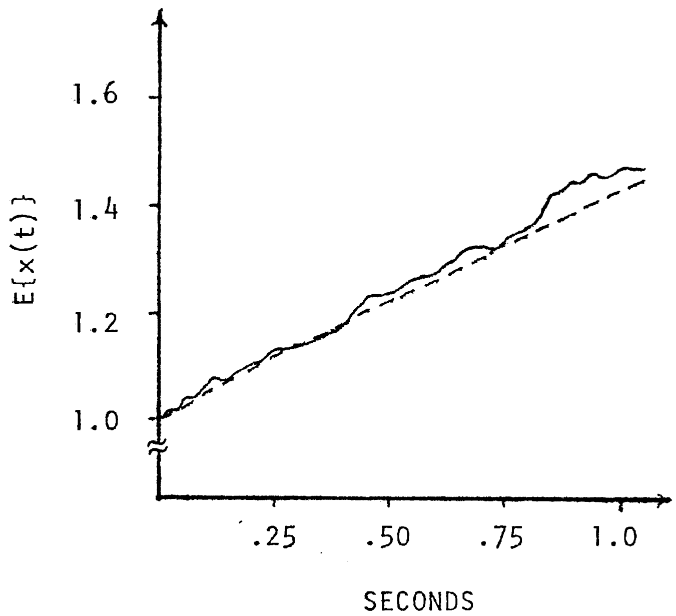


Figure 9. Theoretical and Ensemble-Averaged Mean and Variance, Respectively, for dx = wdw (RK4)

a)



b)

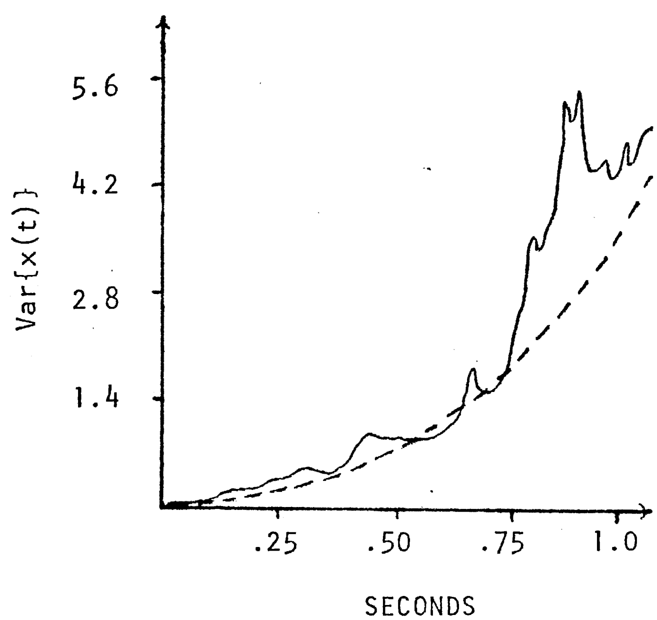


Figure 10. Theoretical and Ensemble-Averaged Mean and Variance, Respectively, for dx = cdw (RK4)

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 it 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

$$x_{i+1} = x_i + \frac{1}{2}(3g(x_i) - g(x_{i-1}))(w_{i+1} - w_i). \tag{3.12}$$

The mean value of x_{i+1} from Equation (3.12) is given by

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

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

$$= E\{x_i\} \tag{3.13}$$

and the mean value equivalence with the Ito integral holds.

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

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

$$+ \frac{1}{4}[3g(x_i) - g(x_{i-1})]^2(w_{i+1} - w_i)^2 \tag{3.14}$$

and then obtain

$$E\{x_{i+1}^2\} = E\{x_i^2\} + \frac{1}{4}E\{[3g(x_i) - g(x_{i-1})]^2(w_{i+1} - w_i)^2\} \tag{3.15}$$

and, from the independence of the noise increment and the other expressions in the second term on the right in Equation (3.15), we find

$$E\{x_{i+1}^2\} = E\{x_i^2\} + \frac{1}{4} q\, E\{[3g(x_i) - g(x_{i-1})]^2\}dt. \tag{3.16}$$

Noting that

$$\frac{\partial g(x_i)}{\partial x} = \frac{g(x_i) - g(x_{i-1})}{x_i - x_{i-1}} \tag{3.17}$$

we find that

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

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

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

$$+ 4g^2(x_{i-1}). \tag{3.18}$$

Equation (3.16) then becomes

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

$$+ 3\,q\,E\{g^2(x_{i-1}) \frac{\partial g(x_i)}{\partial x} (w_i - w_{i-1})\}dt$$

$$+ q \; E\{g^2(x_{i-1})\}dt. \tag{3.19}$$

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

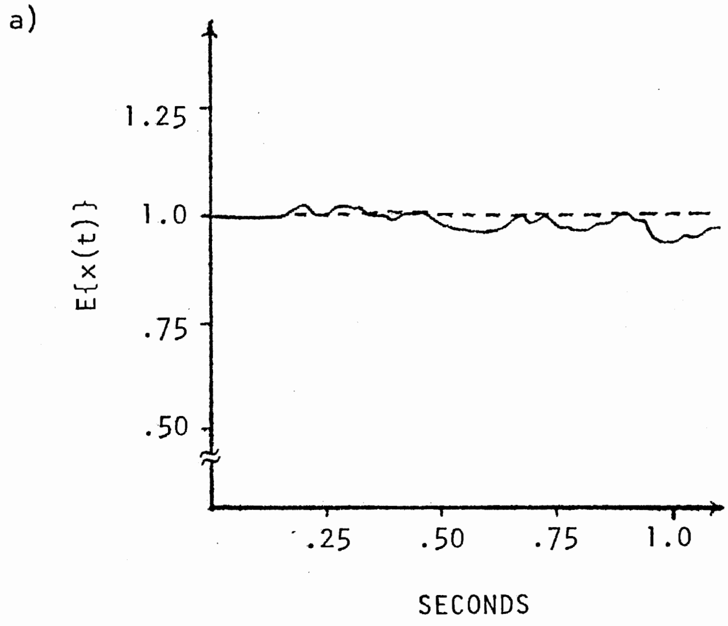
$$Var\{x_{i+1}\} = Var\{x_i\} + q \; E\{g^2(x_{i-1})\}dt \tag{3.20}$$

which is of the same form as Equation (2.27), the variance of the Ito integral.

The numerical solution of Example 1 using the second-order Adams predictor method was obtained with a fixed integration step size of approximately 0.002 seconds and 100 sample trajectories were ensemble-averaged to provide an estimate of the mean and variance. The initial condition was again chosen to be unity, as was q. Figure 11 presents these results, with the mean and variance of the Ito solution indicated by the dashed lines. Figure 12 presents the corresponding results for Example 2, which was simulated as described above. These figures show good agreement for the Ito integral and the second-order Adams predictor method.

The fourth-order Adams-Bashforth predictor method, given by

$$x_{i+1} = x_i + \frac{1}{24} (55g(x_i) - 59g(x_{i-1}) + 37g(x_{i-2})$$



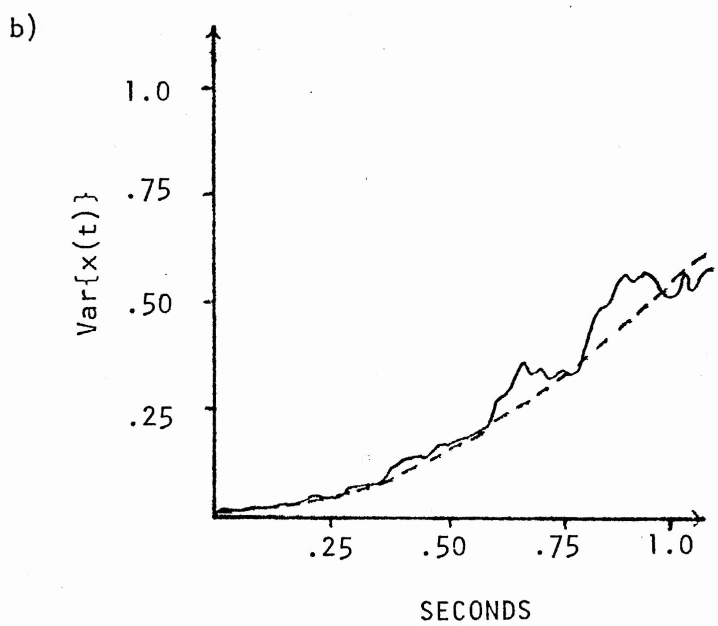
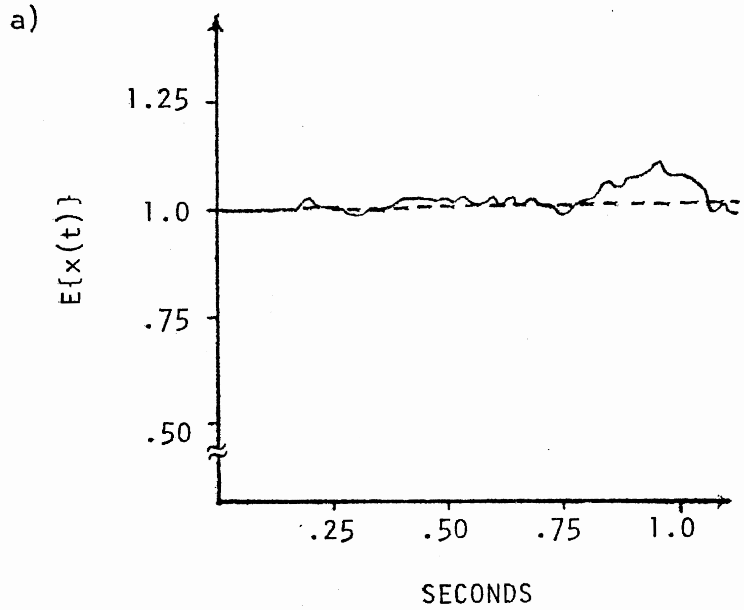


Figure 11. Theoretical and Ensemble-Averaged Mean and Variance, Respectively, for dx = wdw (AB2)



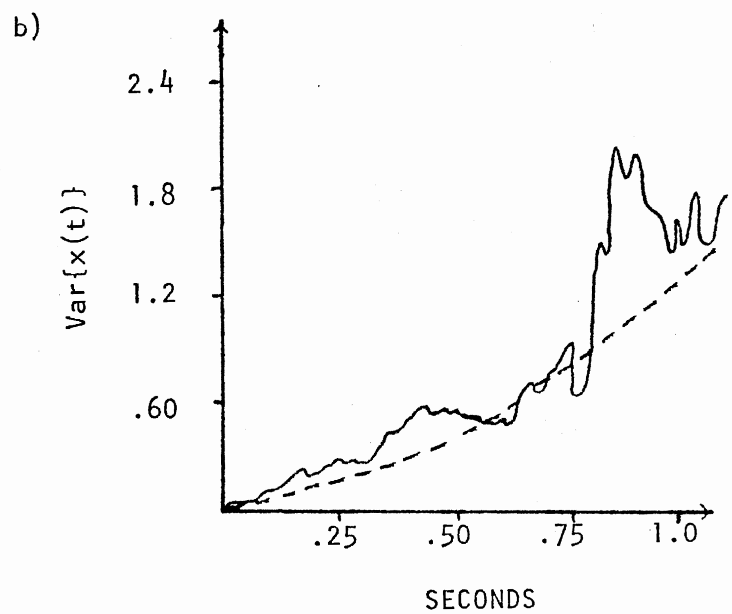


Figure 12. Theoretical and Ensemble-Averaged Mean and Variance, Respectively, for dx = xdw (AB2)

$$- 9g(x_{i-3}))(w_{i+1} - w_i), \tag{3.21}$$

is easily seen to have the mean value

$$E\{x_{i+1}\} = E\{x_i\}, \tag{3.22}$$

since the noise increment is independent of the expression in parentheses in Equation (3.21).

To determine the variance of this fourth-order predictor, we define two expressions similar to Equation (3.17). Namely,

$$\frac{\partial g_1}{\partial x} = \frac{g(x_i) - g(x_{i-1})}{x_i - x_{i-1}} \tag{3.23a}$$

$$\frac{\partial g_2}{\partial x} = \frac{g(x_{i-2}) - g(x_{i-3})}{x_{i-2} - x_{i-3}} . \tag{3.23b}$$

Now

$$x_{i+1}^2 = x_i^2 + \frac{1}{24} x_i [55g(x_i) - 59g(x_{i-1}) + 37g(x_{i-2})$$

$$- 9g(x_{i-3})](w_{i+1} - w_i) + \frac{1}{(24)^2} [55g(x_i) - 59g(x_{i-1})$$

$$+ 37g(x_{i-2}) - 9g(x_{i-3})]^2 (w_{i+1} - w_i)^2$$

and taking the expected value gives

$$E\{x_{i+1}^2\} = E\{x_i^2\} + \frac{q}{(24)^2} E\{[55g(x_i) - 59g(x_{i-1}) + 37g(x_{i-2})$$

$$- 9g(x_{i-3})]^2\}(t_{i+1} - t_i) . \tag{3.24}$$

But, by using Equation (3.23), we can see that

$$[55g(x_i) - 59g(x_{i-1}) + 37g(x_{i-2}) - 9g(x_{i-3})]^2 =$$

$$[55g(x_{i-1}) \frac{\partial g_1}{\partial x} (w_i - w_{i-1}) - 4g(x_{i-1})$$

$$+ 37g(x_{i-3}) \frac{\partial g_2}{\partial x} (w_{i-2} - w_{i-3}) + 28g(x_{i-3})]^2$$

$$= (55)^2 g^2(x_{i-1}) (\frac{\partial g_1}{\partial x})^2 (w_i - w_{i-1})^2 + 16g^2(x_{i-1})$$

$$+ (37)^2 g^2(x_{i-3}) (\frac{\partial g_2}{\partial x})^2 (w_{i-2} - w_{i-3})^2 + (28)^2 g^2(x_{i-3})$$

$$- (8)(55) g^2(x_{i-1}) \frac{\partial g_1}{\partial x} (w_i - w_{i-1})$$

$$+ (37)(110) g(x_{i-1}) g(x_{i-3}) \frac{\partial g_1}{\partial x} \frac{\partial g_2}{\partial x} (w_i - w_{i-1})(w_{i-2} - w_{i-3})$$

$$+ (28)(110) g(x_{i-1}) g(x_{i-3}) \frac{\partial g_1}{\partial x} (w_i - w_{i-1})$$

$$- (8)(37) g(x_{i-1}) g(x_{i-3}) \frac{\partial g_2}{\partial x} (w_{i-2} - w_{i-3})$$

$$- (8)(28) g(x_{i-1}) g(x_{i-3})$$

$$+ (37)(56) g^2(x_{i-3}) \frac{\partial g_2}{\partial x} (w_{i-2} - w_{i-3}). \qquad (3.25)$$

On calculating the expected value of the above expression, one can see that the terms involving noise increments lead, after substitution into Equation (3.24), to terms of higher order in dt and consequently they become negligible for small step sizes. As in the definition of the McShane integral, we consider a belated partition, in which $\tau_i = t_{i-1}$. Then the remaining terms in Equation (3.25), those containing functions of the form $g^2(\cdot)$, are approximately the same as the term $(24)^2 g^2(x_{i-1})$ and the variance of the fourth-order Adams-Bashforth method is given by

$$\mathrm{Var}\{x_{i+1}\} = \mathrm{Var}\{x_i\} + q\, E\{g^2(x_{i-1})\}(t_{i+1} - t_i) \qquad (3.26)$$

which again is of the same form as the variance of the Ito integral.

The numerical solution of Example 1 using the fourth-order Adams-Bashforth predictor method was obtained employing a step size of 0.002 seconds and again 100 sample trajectories were ensemble-averaged. The initial condition $x(0)$ was unity and $q = 1$ also. Figure 13 presents these results with the mean and variance of the Ito solution given by the dashed lines. Figure 14 gives the corresponding results for Example 2, which was simulated as described above. These figures show good agreement between the Ito integral and the fourth-order Adams-Bashforth method.

We now see that the connection between numerical integration methods and stochastic integrals lies in the fact that the point of functional evaluation determines whether or not the integrand is uncorrelated with the noise. Consequently, the point t_i in the interval $[t_i, t_{i+1}]$ is considered to be outside the interval, since no correlation results from this as an evaluation point. Single-step formulas, which evaluate the integrand within the current interval and thereby introduce correlation, correspond to the Stratonovich stochastic integral. Multi-step formulas, which require functional evaluations before the current interval, allow the integrand to remain uncorrelated with the noise input and thus correspond to the Ito stochastic integral.

These ideas are discussed again in later parts of this chapter. The next section is concerned with the statistical analysis of errors resulting from the simulation of the examples.

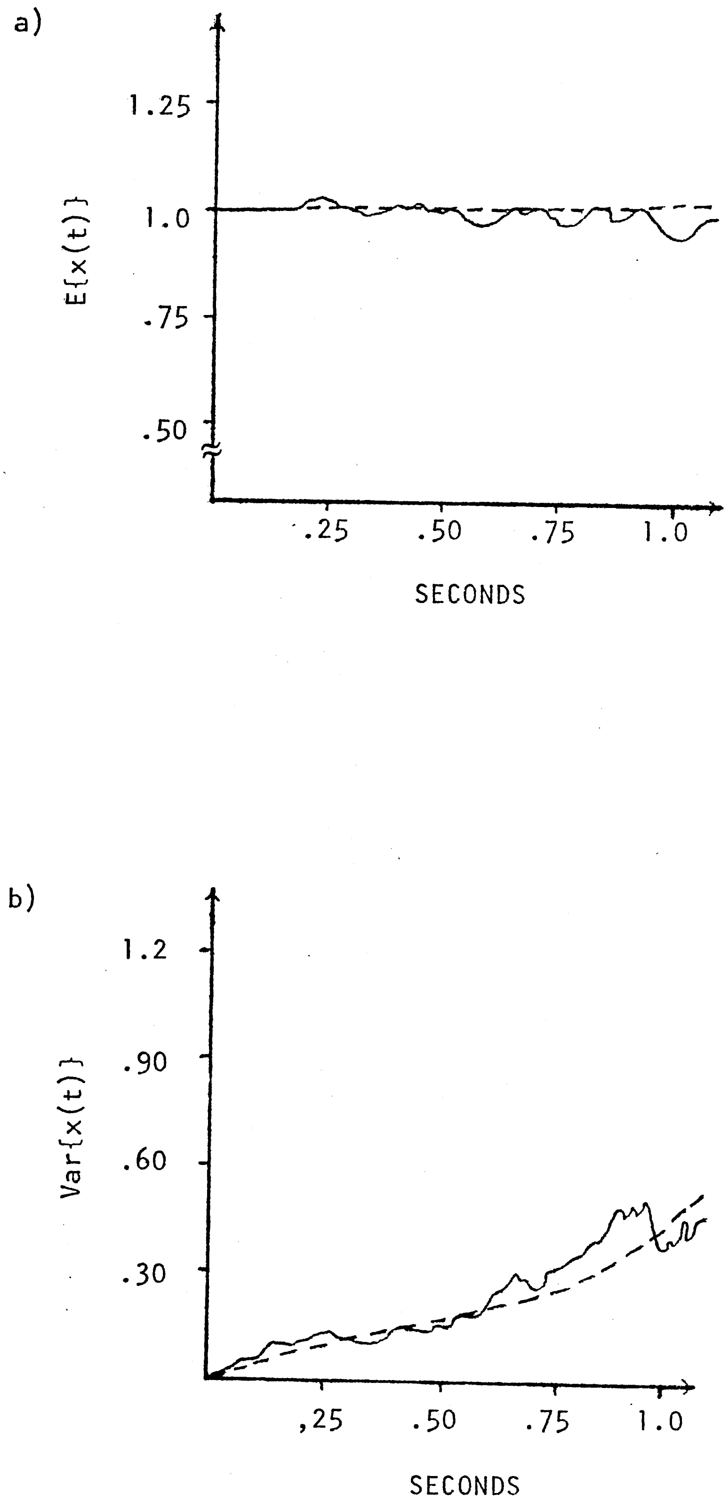
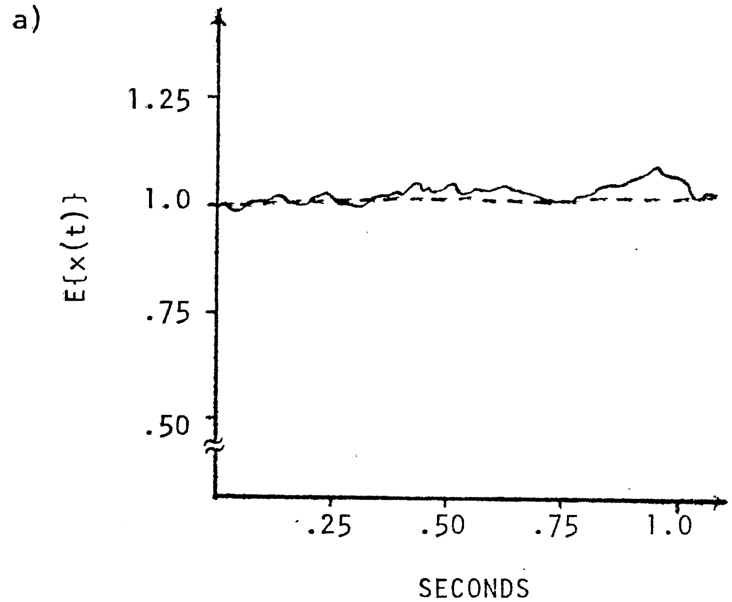


Figure 13. Theoretical and Ensemble-Averaged Mean and Variance, Respectively, for dx = wdw (AB4)



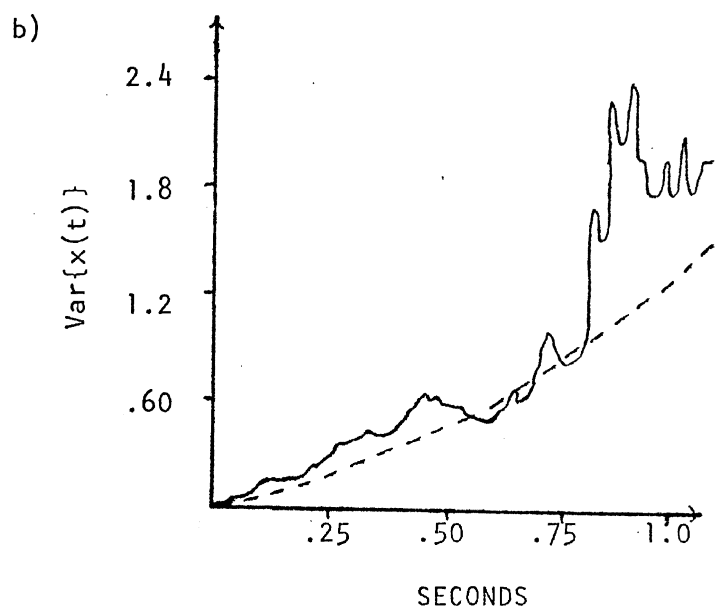


Figure 14. Theoretical and Ensemble-Averaged Mean and Variance, Respectively, for dx = xdw (AB4)

Statistical Errors

The purpose of using a Monte Carlo simulation to analyze a system is to average out the effects of the noise input on the output characteristics. Little, if any, meaningful information may be gained from considering a single solution trajectory, unless it is known that the resulting stochastic process is ergodic, which is usually not the case. The use of aggregates of sample solutions necessitates a statistical study of the results, however, since these results arise from a finite number of samples. We need to find the reliability with which these ensemble-averaged statistics reflect the true behavior of the system. For this reason, we want to place confidence intervals around the theoretical solutions. This will enable us to make probability statements concerning the reliability of the ensemble-averaged solutions.

Consider first the mean value of the ensemble of solution trajectories. Let x_{ij} denote the computed value of $x(t_i)$ in the j-th sample and let x_i denote the average of these values, that is,

$$\overline{x}_i = \frac{1}{N} \sum_{j=1}^{N} x_{ij}. \tag{3.27}$$

The values $\overline{x}_i$ are the ensemble mean values plotted in the figures presented earlier. Bendat and Piersol [27] show that the distribution of the sample mean $\overline{x}_i$, under mild conditions, approaches a normal distribution regardless of how the original variable x_i is distributed. We may then calculate confidence levels from the following:

$$\Pr[(\mu_{x_i} - \frac{\sigma_{x_i} Z_\alpha}{\sqrt{N}}) < \overline{x}_i < (\mu_{x_i} + \frac{\sigma_{x_i} Z_\alpha}{\sqrt{N}})] = 1 - 2\alpha, \tag{3.28}$$

where μ_{x_i} is the known mean of the random variable x_i, σ_{x_i} is the known standard deviation of the random variable x_i and z_α is the 100α percentage point of the normal distribution, that is, $\Pr(z_\alpha) = 1 - \alpha$. For this case of a normally distributed random variable, $z_\alpha = 1$ gives the 1σ band about the true mean μ_{x_i}. We now have

$$\Pr[(\mu_{x_i} - \frac{\sigma_{x_i}}{\sqrt{N}}) < \bar{x}_i < (\mu_{x_i} + \frac{\sigma_{x_i}}{\sqrt{N}})] = 0.6826 \qquad (3.29)$$

and conclude that approximately 68 percent of sample means $\bar{x}_i$ will fall within the calculated interval. Alternatively, we may conclude that for the case in which a time history of sample means is studied, no more than 32 percent of the samples should lie outside of the interval. Similar statements may be made concerning the cases $z_\alpha = 2$ and $z_\alpha = 3$ to determine 2σ and 3σ bands about the true mean.

To illustrate, consider the results of the Euler numerical integration of the equation $dx = wdw$ at the point $t_i = 1$ second. As calculated in Chapter I for the Ito integral, $\mu_{x_i} = 1$ and $\sigma^2_{x_i} = \frac{1}{2}t_i = \frac{1}{2}$, from which it follows that $\sigma_{x_i} = 0.707$. Also, $N = 100$. Equation (3.28) becomes, with $z_\alpha = 2$,

$$\Pr[0.8586 < \bar{x}_i < 1.1414] = 0.9544.$$

Similar confidence intervals about the true mean may be computed for each point $t_i \in [0,1]$ for the example problems.

The unbiased estimator for the sample variance of a point x_i was computed as

$$s_i^2 = \frac{1}{N - 1} \sum_{j=1}^{N} (x_{ij} - \bar{x}_i)^2. \qquad (3.30)$$

This estimate of the variance is not normally distributed, as was the mean value estimator, but rather chi-square distributed with N - 1 degrees of freedom. It can be shown [27] that the sampling distribution of the sample variance is given by

$$s_i^2 = \frac{\chi^2_{N-1}\ \sigma^2_{x_i}}{N-1}$$

where χ^2_{N-1} is chi-square distributed. We can make the statement that

$$\Pr\left[\frac{\chi^2_{N-1;1-\alpha}\ \sigma^2_{x_i}}{N-1} < s_i^2 < \frac{\chi^2_{N-1;\alpha}\ \sigma^2_{x_i}}{N-1}\right] = 1 - 2\alpha \tag{3.31}$$

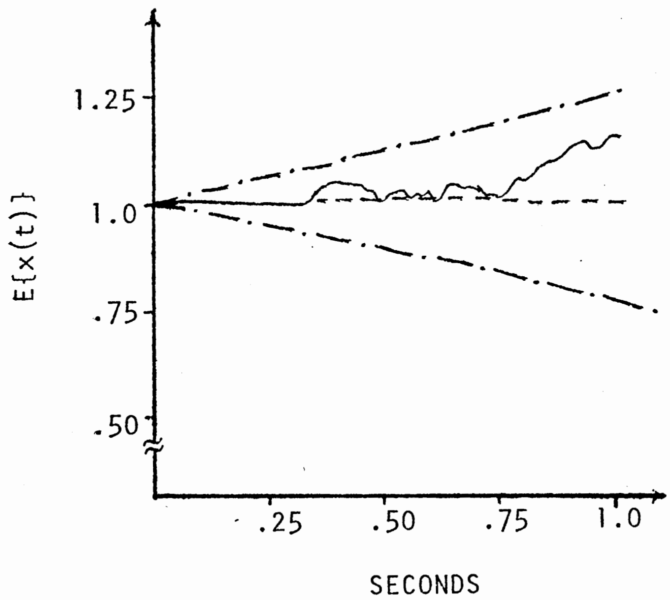
where $\chi^2_{n;\alpha}$ is the 100α percentage point for the chi-square distribution with n degrees of freedom, that is, $\Pr(\chi^2_{n;\alpha}) = 1 - \alpha$. Confidence intervals, corresponding to σ bounds in the case of a normal distribution, may be computed from a table of percentage points for the chi-square distribution.

Again using the Euler integration results for dx = wdw as an illustration, we calculate the 95 percent confidence interval about the variance. We find that $\alpha = 0.025$ and $\chi^2_{99;0.025} = 128.43$ and $\chi^2_{99;0.975} = 73.35$. Since $N - 1 = 99$ and $\sigma^2_{x_i}(1) = 1/2$, Equation (3.31) becomes

$$\Pr[0.3705 < s_i^2 < 0.6486] = 0.95.$$

Figures 15 through 24 show the 99 percent confidence intervals for each numerical integration method for both example problems. The theoretical and ensemble-averaged solutions are indicated as before with the error bands given by the dashed and dotted lines. All mean values

a)



b)

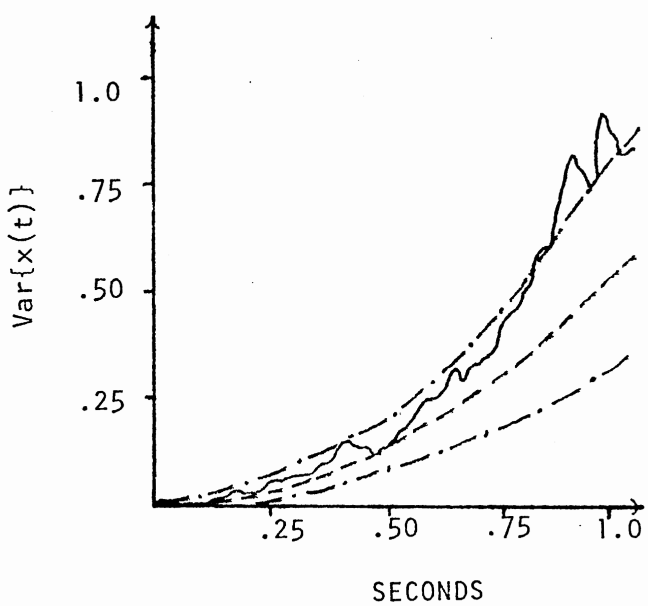


Figure 15. Theoretical and Ensemble-Averaged Mean and Variance, with Error Bands, for dx = wdw (Euler)

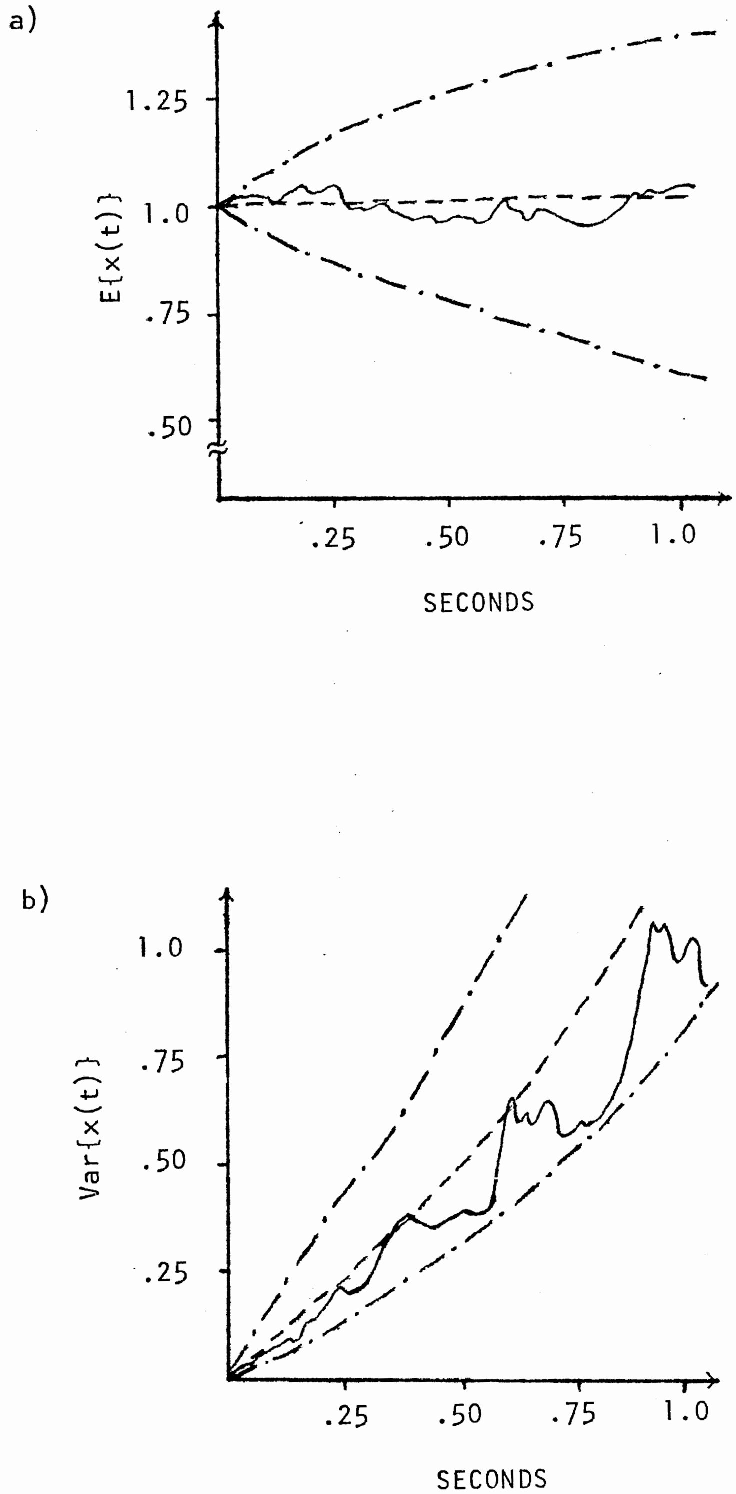


Figure 16. Theoretical and Ensemble-Averaged Mean and Variance, with Error Bands, for dx = xdw (Euler)

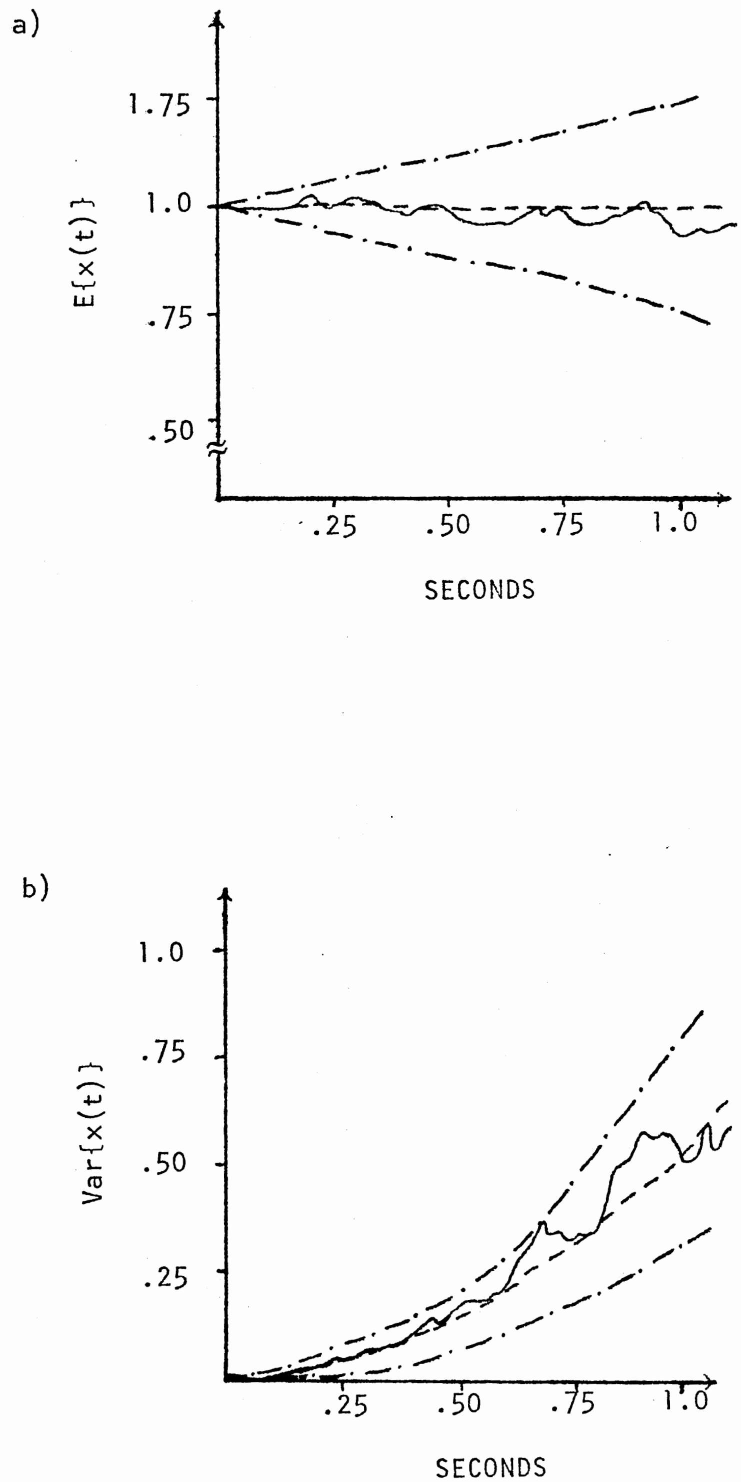
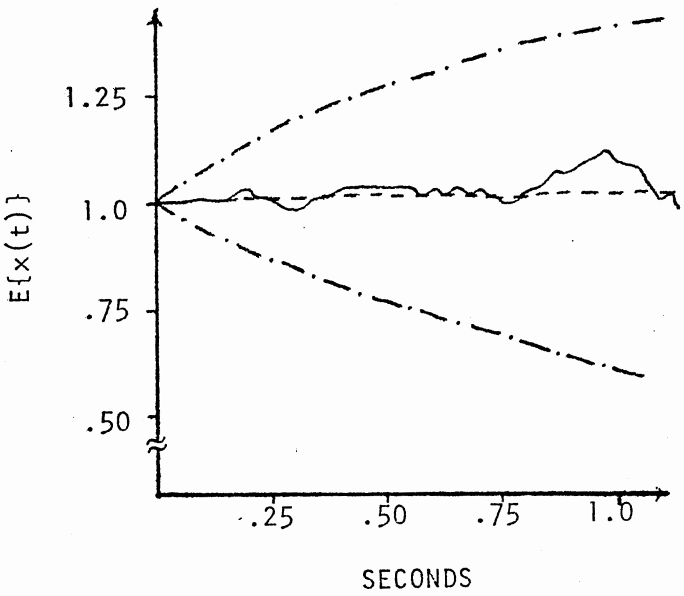


Figure 17. Theoretical and Ensemble-Averaged Mean and Variance, with Error Bands, for dx = wdw (AB2)

a)



b)

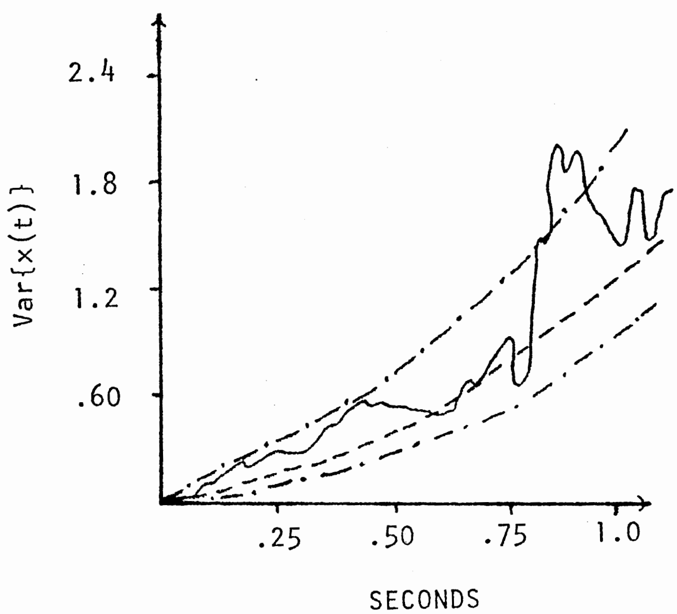


Figure 18. Theoretical and Ensemble-Averaged Mean and Variance, with Error Bands, for dx = xdw (AB2)

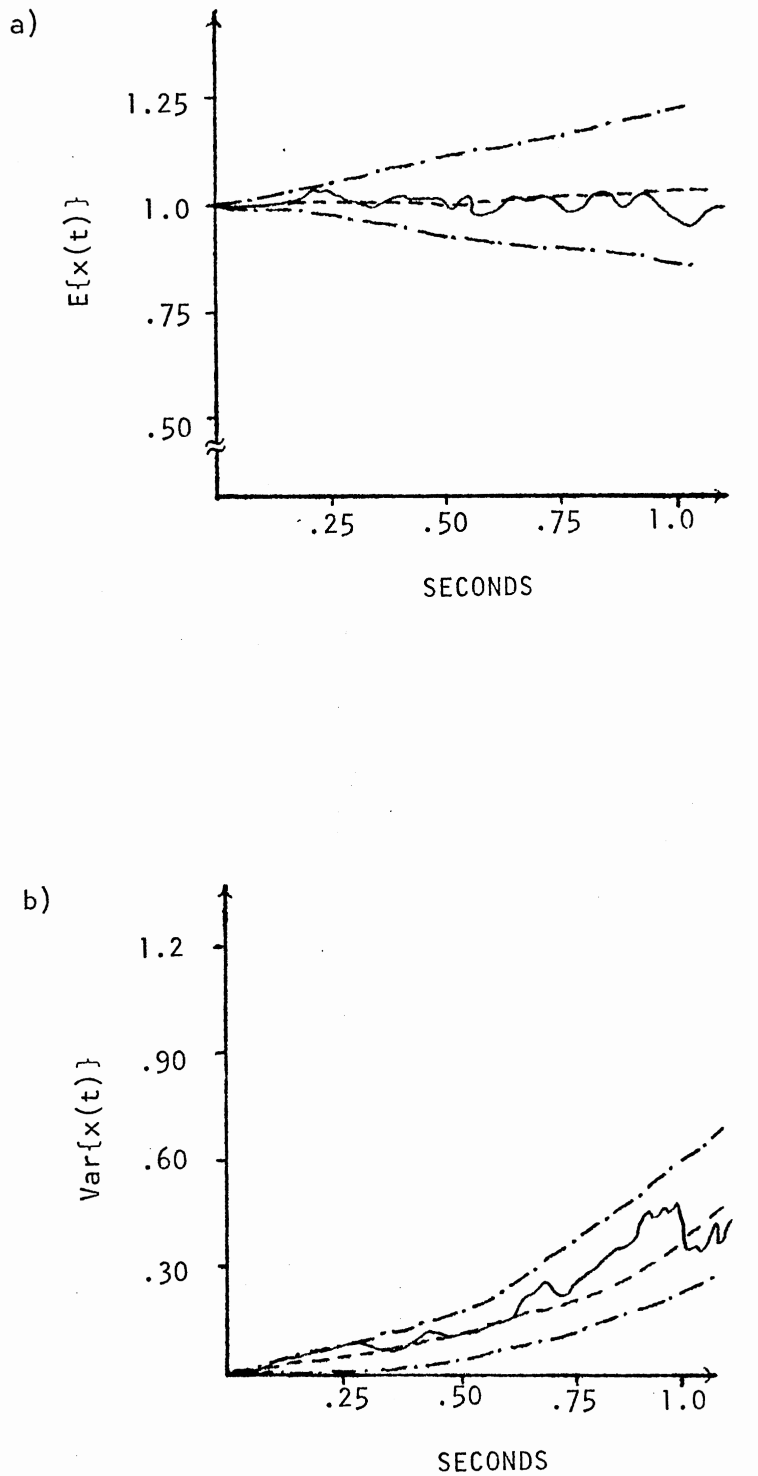
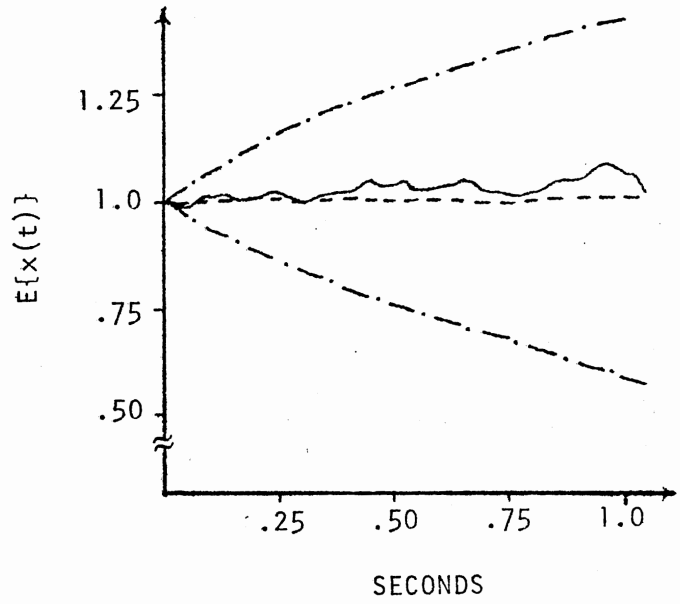


Figure 19. Theoretical and Ensemble-Averaged Mean and Variance, with Error Bands, for dx = wdw (AB4)

a)



b)

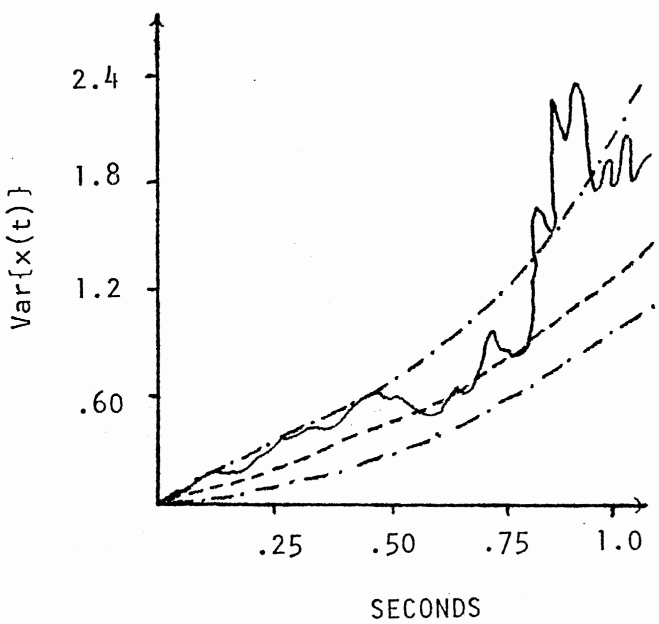


Figure 20. Theoretical and Ensemble-Averaged Mean and Variance, with Error Bands, for dx = cdw (AB4)

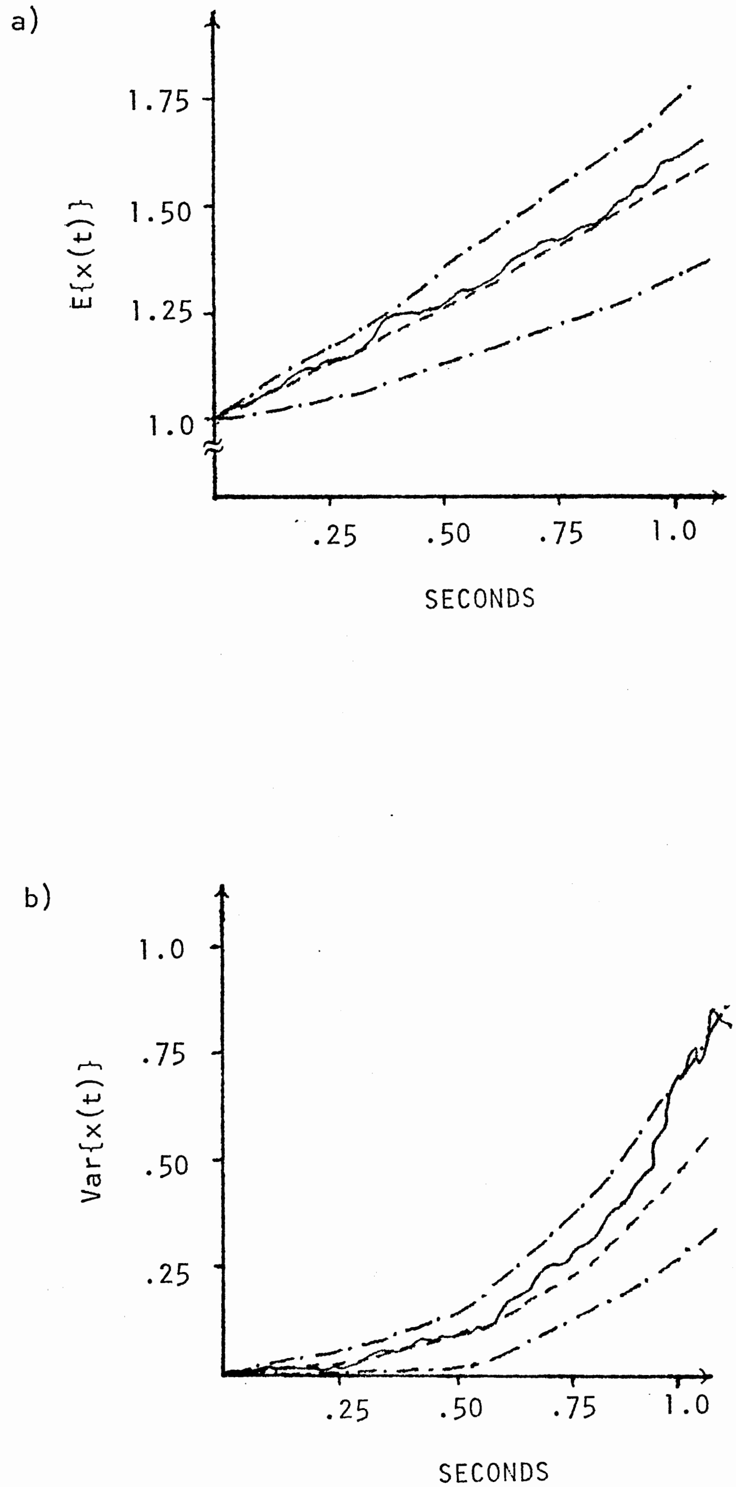
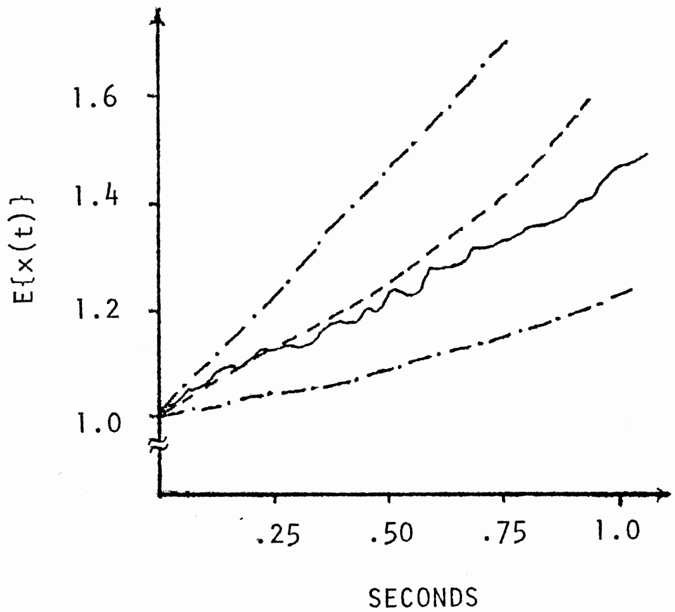


Figure 21. Theoretical and Ensemble-Averaged Mean and Variance, with Error Bands, for dx = wdw (RK2)

a)



b)

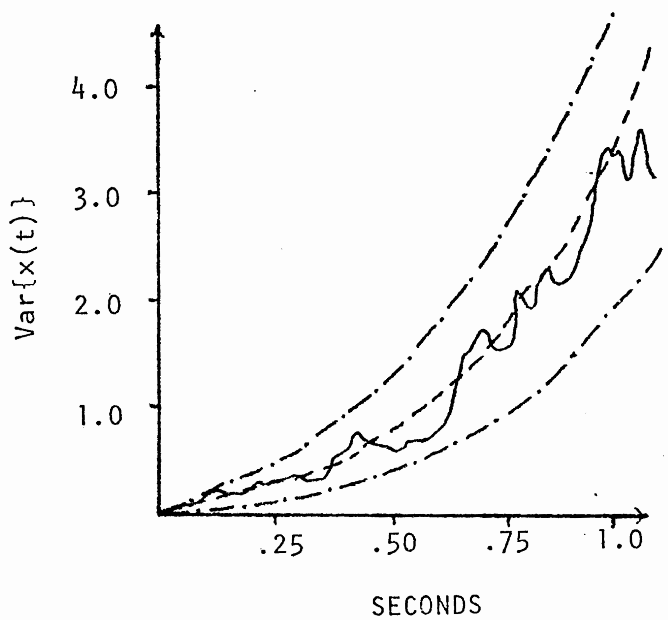
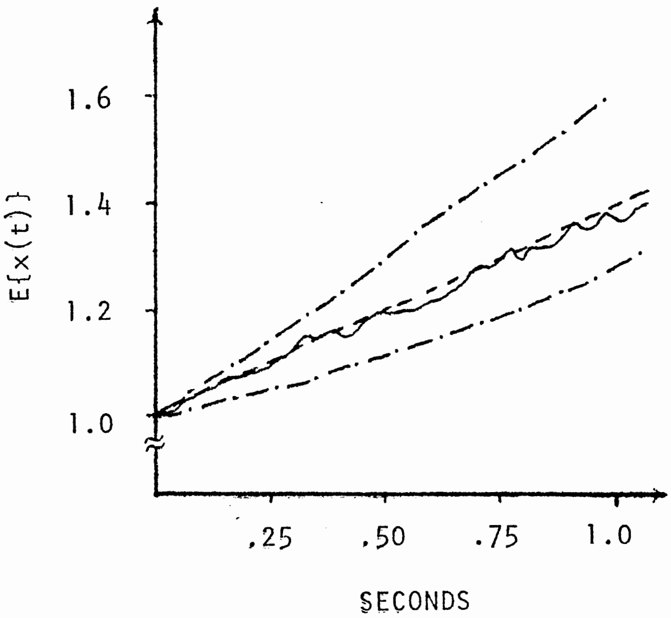


Figure 22. Theoretical and Ensemble-Averaged Mean and Variance, with Error Bands, for dx = xdw (RK2)

a)



b)

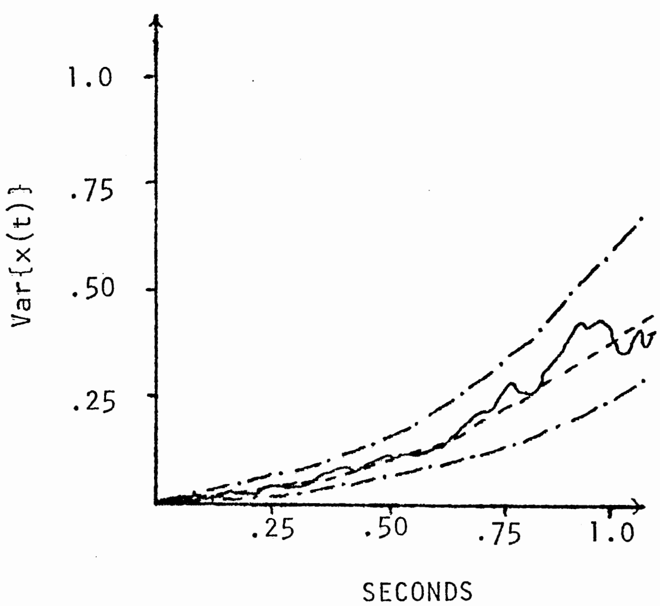


Figure 23. Theoretical and Ensemble-Averaged Mean and Variance, with Error Bands, for dx = wdw (RK4)

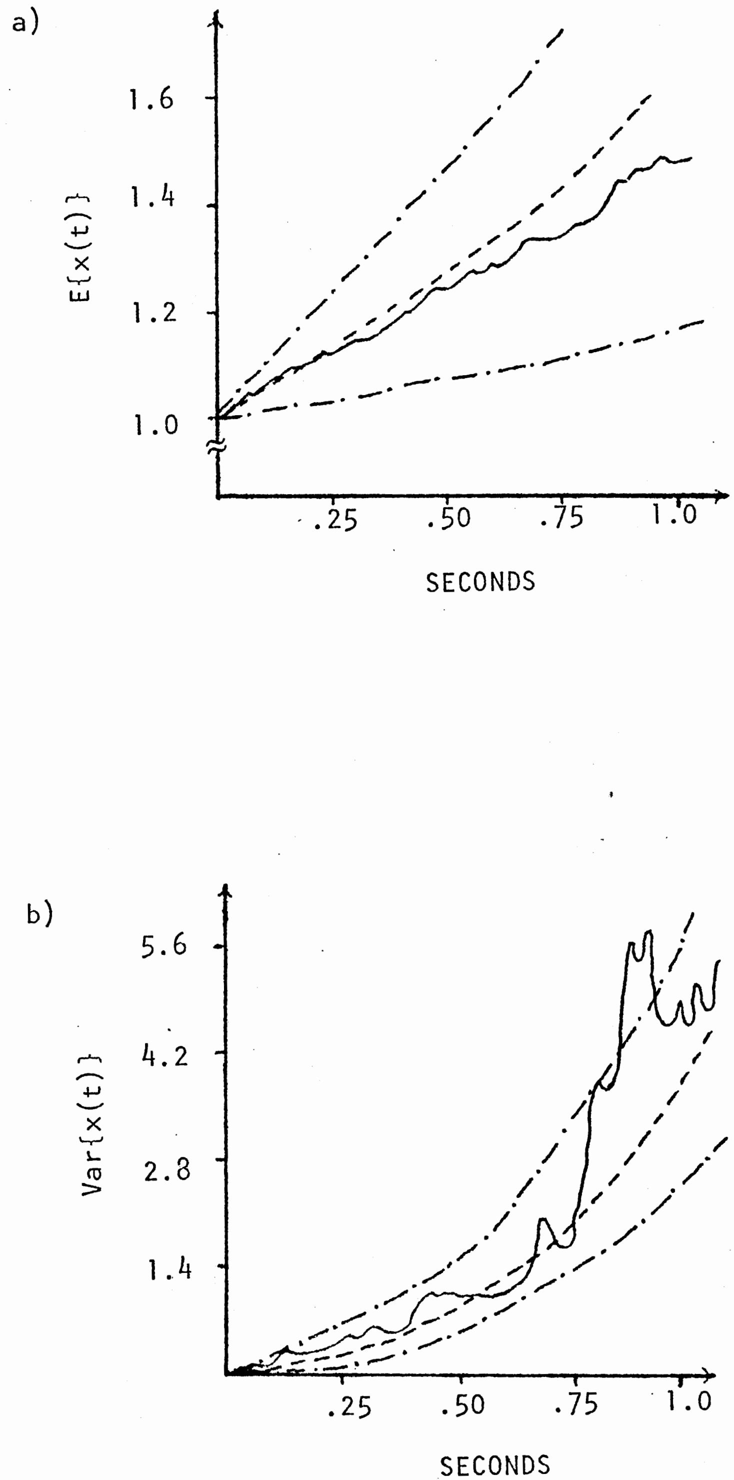


Figure 24. Theoretical and Ensemble-Averaged Mean and Variance, with Error Bands, for dx = xdw (RK4)

lie within the calculated confidence intervals while some portions of trajectories for the variances lie outside the bounds. That small segments of the variance trajectories extend beyond the confidence limits is explained by the fact that these calculated intervals do not contain information on uncertainty arising from numerical algorithm errors such as truncation and roundoff. Consideration of these errors would provide somewhat wider intervals at the same confidence level.

Having attained these results on the relationships between digital simulation and stochastic integral definitions, the remainder of the Chapter discusses the unifying idea in this area, the concept of correlation.

System Correlation

As indicated earlier, the unifying concept in the simulation of nonlinear stochastic systems is the correlation between the noise input and the system dynamics. No correlation is present in the Ito integral, but the Stratonovich definition introduces correlation by evaluating the integrand at the midpoint of each interval of the partition. Digitally, correlation is introduced by repeated use of each noise increment in the evaluation of the integrand.

The amount of correlation between the integrand and the noise increment is represented by the correlation coefficient function, given by

$$\rho(t) = \frac{E\{gdw\} - E\{g\}E\{dw\}}{\sigma_g \, \sigma_{dw}} . \tag{3.32}$$

Since the mean increment of the Wiener process is zero, the correlation coefficient function becomes

$$\rho(t) = \frac{E\{g dw\}}{\sigma_g \sigma_{dw}} . \tag{3.33}$$

That the correlation is a consistent measure of system characteristics, which carries over from theory to digital simulation, may be shown as follows. First, Ito requires the noise increment and the integrand be uncorrelated, so $\rho(t) = 0$. For Euler's method, we have

$$\rho(t_i) = \frac{E\{g(x_i)(w_{i+1} - w_i)\}}{\sigma_{g_i} \sigma_{dw_i}}$$

$$= 0 \tag{3.34}$$

It follows similarly that the second- and fourth-order Adams-Bashforth predictors also have correlation coefficient functions identically zero.

Correlation in the Stratonovich integral may be found by utilizing the relationship in Equation (2.33). In differential equation form, Equation (2.33) states that the Stratonovich solution of the equation

$$dx(t) = g(x,t)\ dw(t)$$

is the same as the Ito solution of

$$dx(t) = g(x,t)\ dw(t) + \frac{1}{2} q\ g(x,t) \frac{\partial g}{\partial x}(x,t)\ dt .$$

We note that

$$E\{g dw\} = \frac{1}{2} q\ E\{g \frac{\partial g}{\partial x}\} dt \tag{3.35}$$

in terms of the Stratonovich definition. We then determine the Stratonovich correlation coefficient function to be

$$\rho(t) = \frac{\frac{1}{2} q\ E\{g \frac{\partial g}{\partial x}\}\ dt}{\sigma_g \sigma_{dw}}$$

$$= \frac{1}{2} \sqrt{q\,dt}\; \frac{E\{g \frac{\partial g}{\partial x}\}}{\sigma_g} \tag{3.36}$$

since $\sigma_{dw} = \sqrt{qdt}$.

For the second-order Runge-Kutta method, the correlation between the integrand and the noise increment is found by first calculating

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

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

$$= \frac{1}{2} q\; E\{g(x_i) \frac{\partial g}{\partial x} (x')\}dt, \tag{3.37}$$

as in Equations (2.15) and (2.16). The correlation coefficient function is thus

$$\rho(t_i) = \frac{\frac{1}{2} q\; E\{g(x_i) \frac{\partial g}{\partial x} (x')\}dt}{\sigma_{g_i}\, \sigma_{dw_i}}$$

$$= \frac{1}{2} \sqrt{q\,dt}\; \frac{E\{g(x_i) \frac{\partial g}{\partial x} (x')\}}{\sigma_{g_i}} \tag{3.38}$$

which is the same as Equation (3.36). The correlation of the fourth-order Runge-Kutta method is determined in a similar manner and has the same value as Equation (3.38).

These results illustrate the critical importance of the evaluation point of the integrand in the theory and simulation of stochastic differential equations. The location of this evaluation point in each interval of the partition is the determining factor in the value of the integral and its effect is manifested through the concept of the correlation coefficient function. This effect is seen in the theory of stochastic integration as well as in the digital simulation of stochastic systems and can be seen to be the unifying principle between these areas and also the basis for the Ito-Euler-predictor relationship and the Stratonovich-Runge-Kutta relationship.

Since $0 \leq |\rho(t)| \leq 1$, the Ito theory has the least correlation possible in stochastic integration. The questions then arise of what is the maximum correlation possible between the integrand and noise and of what role the Stratonovich theory plays in this correlation analysis. A question related to this last idea is what effect the point of evaluation within the interval has on the correlation in the Stratonovich integral. It is notable that when the midpoint is used, as in the Stratonovich definition of the integral, the coefficient in the correlation function is one-half.

In general,

$$\frac{\partial g}{\partial x}(x') = \frac{g(ax_i + bx_{i+1}) - g(x_i)}{ax_i + bx_{i+1} - x_i}, \quad x_i \leq x' \leq ax_i + bx_{i+1} \tag{3.39}$$

and then

$$b\frac{\partial g}{\partial x}(x')(x_{i+1} - x_i) = g(ax_i + bx_{i+1}) - g(x_i) \tag{3.40}$$

where $0 \leq a \leq 1$ and $0 \leq b \leq 1$ and $a + b = 1$. Multiplying both sides of Equation (3.40) by the noise increment and calculating the correlation function gives

$$\rho(t) = b\sqrt{q\,dt}\,\frac{E\{g(x_i)\frac{\partial g}{\partial x}(x')\}}{\sigma_g}. \tag{3.41}$$

The case $b = o$ corresponds to the Ito integral and the Euler and predictor methods of numerical integration as can be seen because $\rho(t) = 0$ and the integrand is evaluated only at $x(t_i)$. The Stratonovich theory and Runge-Kutta integration methods result when $a = b = 1/2$. With the evaluation point constrained to be within the interval, the maximum correlation will occur when $b = 1$. Indeed, this instance provides the maximum correlation even if points after the current interval are used as evaluation points.

To see this, consider using values n steps beyond the current interval as evaluation points. Then

$$\frac{\partial g}{\partial x}(x') = \frac{g(ax_i + bx_{i+n}) - g(x_i)}{ax_i + bx_{i+n} - x_i}, \quad x_i \leq x' \leq ax_i + bx_{i+n} \tag{3.42}$$

and

$$g(ax_i + bx_{i+n}) = b\,\frac{\partial g}{\partial x}(x')(x_{i+n} - x_i) - g(x_i).$$

Since $g(x_i)$ is uncorrelated with noise increments beginning at t_i, the first term on the right is the quantity of interest. If we consider that the noise increment is $w_{i+1} - w_i$, we obtain

$$\begin{aligned} g(ax_i + bx_{i+n})(w_{i+1} - w_i) \\ &= b\,g(x_i)\,\frac{\partial g}{\partial x}(x')(w_{i+n} - w_i)(w_{i+1} - w_i) \\ &= b\,g(x_i)\,\frac{\partial g}{\partial x}(x')[(w_{i+n} - w_{i+n-1}) \\ &\quad + (w_{i+n-1} - w_{i+n-2}) + \ldots + (w_{i+1} - w_i)] \\ &\quad \cdot (w_{i+1} - w_i). \end{aligned} \tag{3.43}$$

It then follows that

$$\rho(t) = b \sqrt{q\,dt}\ \frac{E\{g(x_i)\ \frac{\partial g}{\partial x}\ (x')\}}{\sigma_g}\ , \tag{3.44}$$

since all noise increments are uncorrelated except the last in the brackets in Equation (3.43).

Alternatively, we could consider that the multiplicative noise increment is $w_{i+n} - w_i$ instead of $w_{i+1} - w_i$. Then we would obtain

$$g(ax_i + bx_{i+n})(w_{i+n} - w_i) = b\ g(x_i)\ \frac{\partial g}{\partial x}\ (x')(w_{i+n} - w_i)^2 \tag{3.45}$$

and

$$E\{g(ax_i + bx_{i+n})(w_{i+n} - w_i)\} = b\,q\,E\{g(x_i)\frac{\partial g}{\partial x}(x')\}n\ dt. \tag{3.46}$$

Also

$$\sigma_{dw} = \sqrt{E\{(w_{i+n} - w_i)^2\}}$$

$$= \sqrt{n\,q\,dt}$$

and we obtain

$$\rho(t) = b \sqrt{n\,q\,dt}\ \frac{E\{g(x_i)\ \frac{\partial g}{\partial x}\ (x')\}}{\sigma_g}\ . \tag{3.47}$$

This method introduces more correlation by a factor of $\sqrt{n}$. However, this technique essentially utilizes a step size of ndt rather than dt and hence cannot be compared with the results in Equations (3.36) or (3.44).

We see then that the system correlation is a consistent and meaningful way of comparing and contrasting stochastic integrals and numerical integration methods. The general form of the correlation function, given by Equation (3.41), may be used for these purposes. The Ito-Euler-predictor grouping provides the limiting case of no correlation between integrand and noise. The Stratonovich-Runge-Kutta association is an intermediate case corresponding to evaluation of the integrand at the midpoint of each interval. The maximum correlation between integrand and noise occurs if the integrand is evaluated at $x(t_{i+1})$, which value is not known exactly when using a digital algorithm. The point of functional evaluation within each interval therefore determines the numerical value of the correlation function, although the general form is specified.

Summary

The results in Chapter II were extended in this chapter to other numerical integration methods. Moments of a fourth-order Runge-Kutta method were determined to extend the Stratonovich-Runge-Kutta relationship to this important type of integration algorithm. Predictor methods were shown to correspond to the Ito integral in the practical case of bounded and continuous integrands.

Confidence intervals were calculated about the ensemble-averaged mean and variance for the example problems and each type of numerical algorithm. The mean values were completely within the calculated intervals with the variance trajectories outside the intervals a small portion of the time. It was shown that the unifying concept in the relationship between integration methods and stochastic integrals is the correlation between the integrand and noise. The general form of this correlation was found and it was shown that the Ito and Stratonovich integrals correspond to specific cases of the correlation function.

CHAPTER IV

EXAMPLES

There are many examples of nonlinear stochastic systems in man-made processes and in nature. Estimation and control of systems are broad areas of interest which sometimes necessitate the use of nonlinear stochastic differential equations. The structure of the optimal filter for obtaining state estimates provides one important example. Problems in communication theory involving phase-locked loops provide a wide range of applications for the concepts discussed earlier. This area is particularly interesting because of the multitude of uses of phase-locked loops. Estimating the concentration of pollutants in the air is another problem which involves nonlinear stochastic theory.

Many of the nonlinear stochastic problems have been studied primarily through linear approximations or through neglecting the multiplicative noise terms and its consequences. Some examples are considered in this chapter which illustrate the effects of digital simulation on these systems.

Optimal Nonlinear Filtering

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. Estimation of the state of a physical system, based on 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 Stratonovich [13], who also studied the implications for stochastic control problems [28]. 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

$$dy(t) = \frac{\alpha}{\beta^2} dt + \frac{1}{\beta} dw(t) \tag{4.1}$$

where α and β are constants. The optimal estimate for the posterior probability distribution of the observed process was derived by Wonham [29] and is given by the stochastic equation

$$\begin{aligned} dx(t) = & -\beta^2 x(t)(1 - x^2(t))dt - \alpha(1 - x^2(t))dt \\ & + \beta(1 - x^2(t))dw \end{aligned} \tag{4.2}$$

Equation (4.2) defines the structure of an ideal filter which generates the optimal estimate of the posterior distribution from the observed input function. The form of the optimal filter is given in Figure 25.

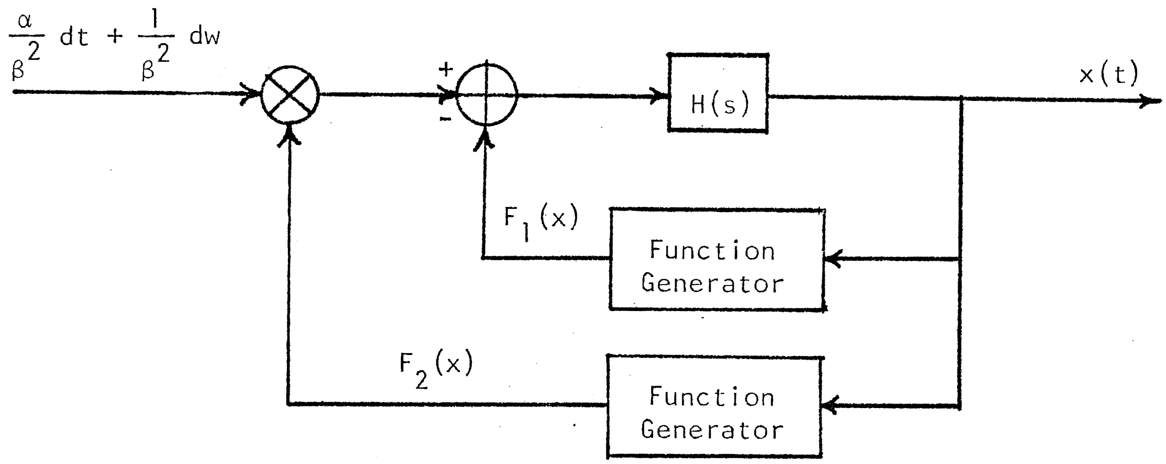


Figure 25. Optimal Nonlinear Filter

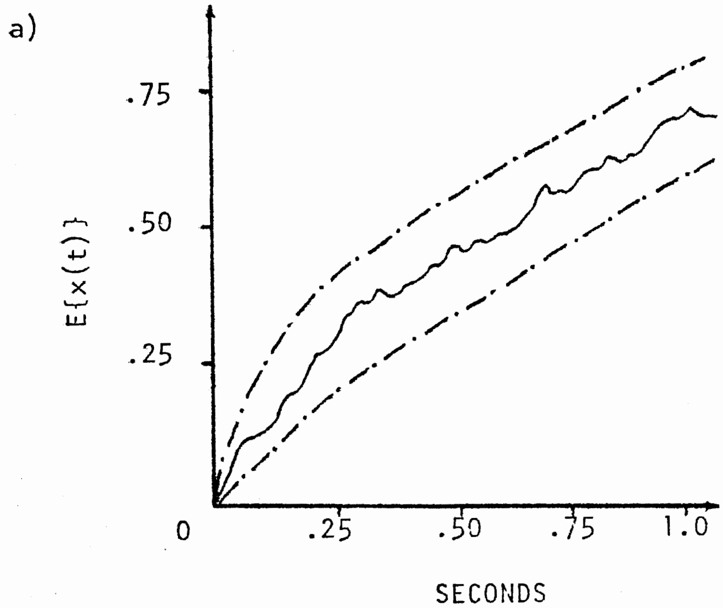
The output of the function generators is

$$F_1(x) = \beta^2 x(1 - x^2)$$

$$F_2(x) = \beta^2 (1 - x^2)$$

and the transfer function $H(s) = 1/s$.

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 seconds and 100 sample runs were ensemble-averaged to provide the results. The parameters α and β were given the same value, $\alpha = \beta = -2.0$ and the variance parameter was chosen to be unity. Figures 26, 27 and 28 give the simulation results for the Euler, AB2 and RK4 numerical integration methods, respectively.



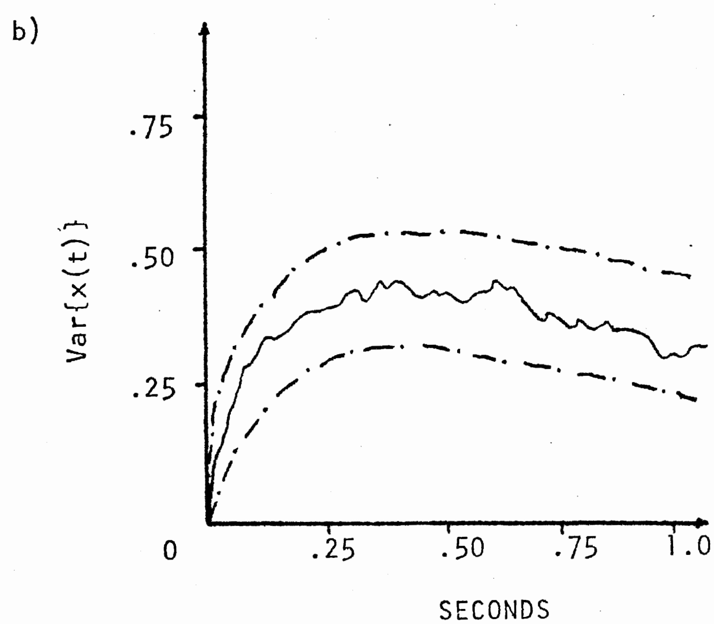


Figure 26. Ensemble-Averaged Mean and Variance for Optimal Nonlinear Filter (Euler)

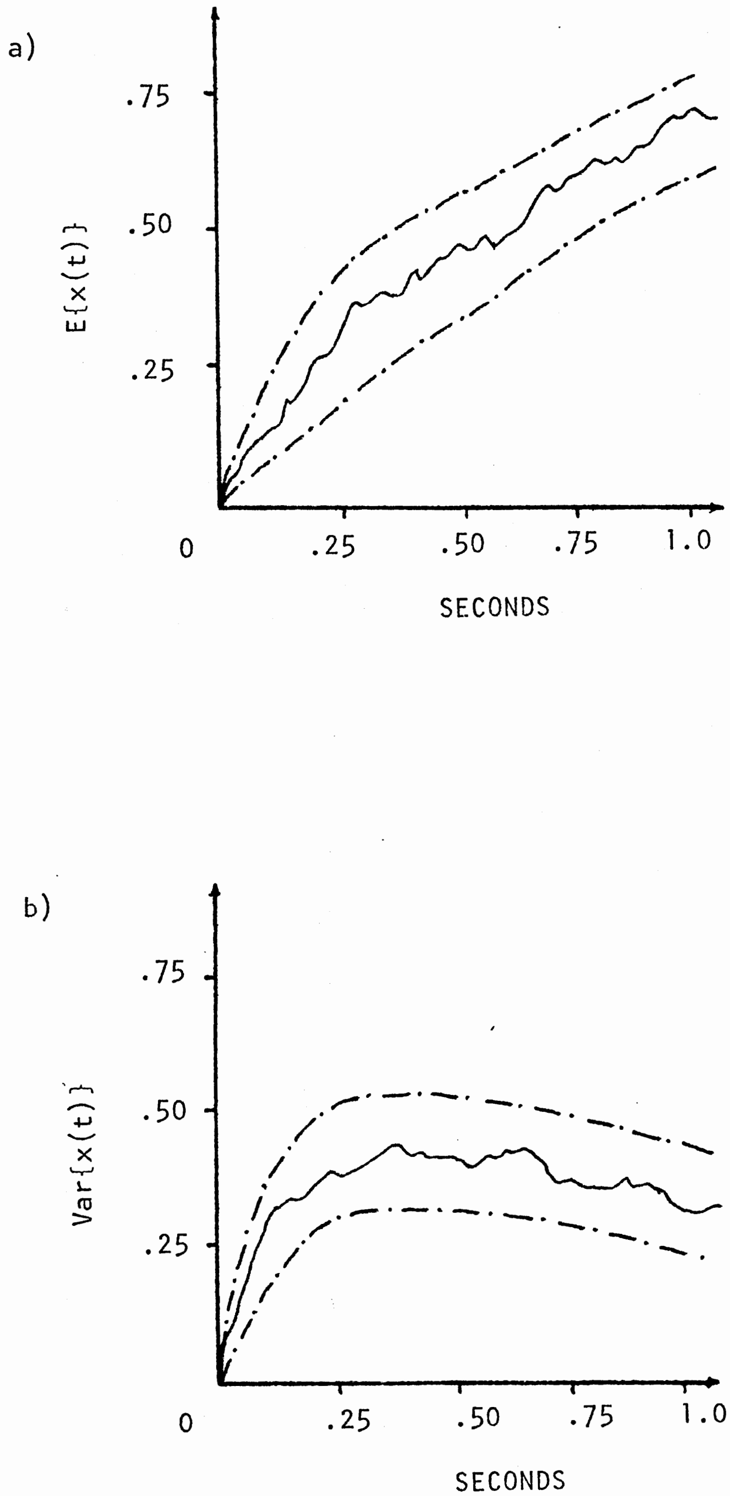
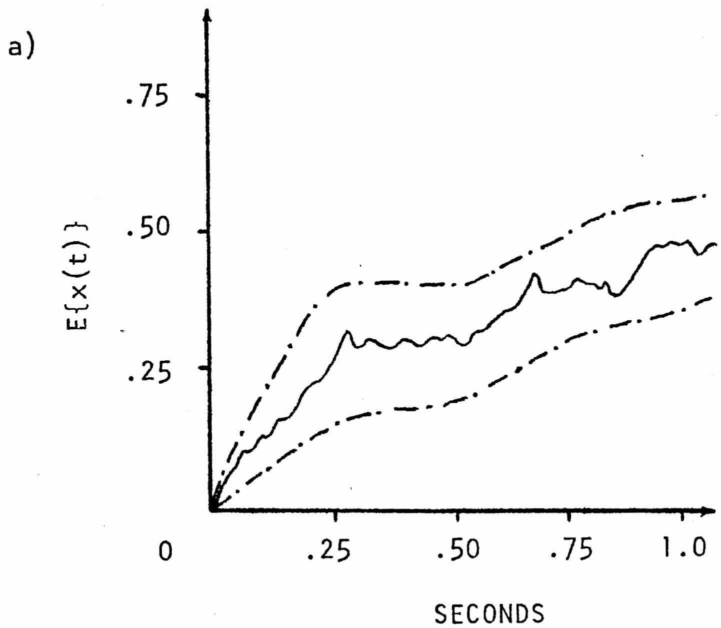


Figure 27. Ensemble-Averaged Mean and Variance for Optimal Nonlinear Filter (AB2)



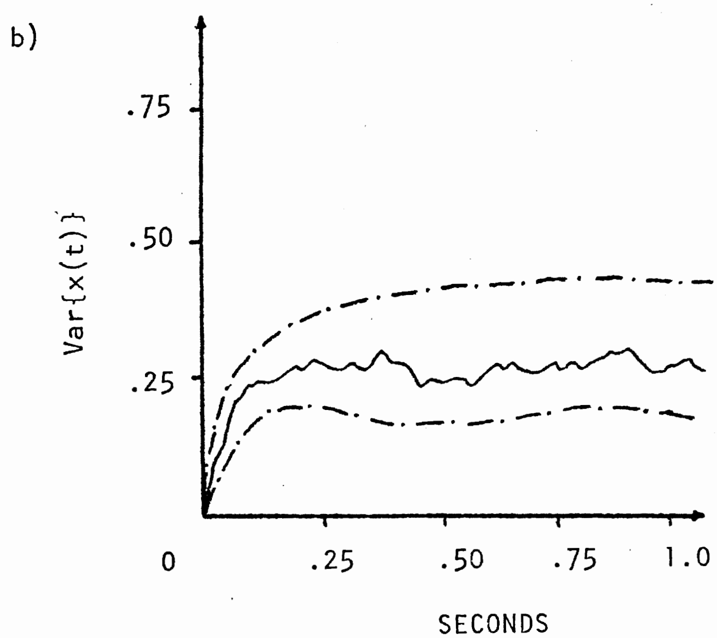


Figure 28. Ensemble-Averaged Mean and Variance for Optimal Nonlinear Filter (RK4)

The mean and variance estimates of the Euler and AB2 methods are nearly identical, as expected. The mean value generated by the RK4 method is different, however. There is about a 30% difference in the mean value after one second. The variance estimates for all the methods appear to achieve a steady-state value of approximately one-fourth. The Euler and AB2 methods overshoot this value somewhat and damp out rather slowly, while the RK4 method achieves the value quickly and then exhibits small random perturbations.

Confidence intervals at a significance level of 95 percent are shown for the simulation results. As is the case for all examples in this chapter, the theoretical mean and variance are not known, and the confidence intervals are calculated around the ensemble-averaged solutions. 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 and AB2 methods and the Euler and AB2 mean values lie outside the Runge-Kutta confidence interval for the same time period. For times near one 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 40 percent to 50 percent 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.

Phase-Locked Loops

The phase-lock principle is a very powerful and general tool in the analysis and design of systems in which one of the requirements is the acquisition and tracking of an input signal. The basic configuration of a phase-locked loop (PLL) is shown in Figure 29. The input signal to be tracked contains background noise, and nonlinearities in the input as well as the voltage-controlled oscillator (VCO) indicate the need of nonlinear stochastic analysis in the study of these systems. The acquisition signal aids in driving the VCO to obtain a lock on the input signal.

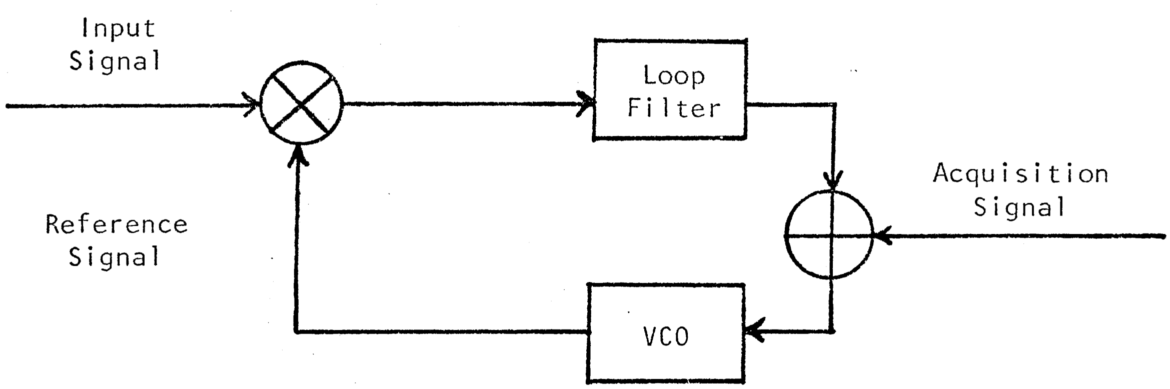


Figure 29. Phase-Locked Loop

Phase-locked loop concepts were first studied about 50 years ago by de Bellescize [30], who was interested primarily in the subject of synchronous reception of radio signals. Practical limitations of the time, however, prevented the perfect synchronization of the carrier frequency and the oscillator frequency in the receiver and PLLs were not developed until later. Subsequently it has been discovered that the general closed-loop configuration of the PLL can be applied to a very general class of not only communication and control systems, but also to certain natural systems.

There are many practical utilizations of the basic PLL configuration in Figure 29. Applications to communications include the demodulation of analog or digital systems [31] [32], reference extraction for linear demodulation and amplitude detection [33], carrier tracking [34] and as a synchronizer for various systems [35] [36]. Other applications include frequency synthesis, multiplication and division, electric power generation and the study of circadian rhythms by physiologists and biophysicists. These last are concerned with the electrical rhythms of the brain as well as the synchronous activity of the heart.

Consider the problem arising in communication theory of the demodulation of an angle-modulated signal. Figure 30 depicts a demodulation system for this type of application. The output of the linear filter is given by

$$x(t) = \sqrt{2} \int_{-\infty}^{t} \cos(\omega\,\tau + x(\tau))\; u(\tau)\; f(t - \tau)\; d\tau. \qquad (4.3)$$

If the angle-tracking linear filter is given by K/s, then Figure 30 represents a second-order phase-locked loop. Let the input u(t) be the angle-modulated signal

$$u(t) = \sqrt{2}\,\sin(\omega\,\tau + m(t)) \qquad (4.4)$$

where m(t) is zero-mean Gaussian white noise. The resulting filter output is then given by the integral equation

$$x(t) = 2K \int_{-\infty}^{t} \sin\,(\omega\tau + m(\tau))\;\cos(\omega\tau + x(\tau))\; d\tau. \qquad (4.5)$$

The equivalent stochastic differential equation may be written in the form

$$dx(t) = 2K \sin(\omega t + m(t)) \cos(\omega t + x(t))\, dt \qquad (4.6)$$

in which the noise term enters multiplicatively and nonlinearly through the sine function.

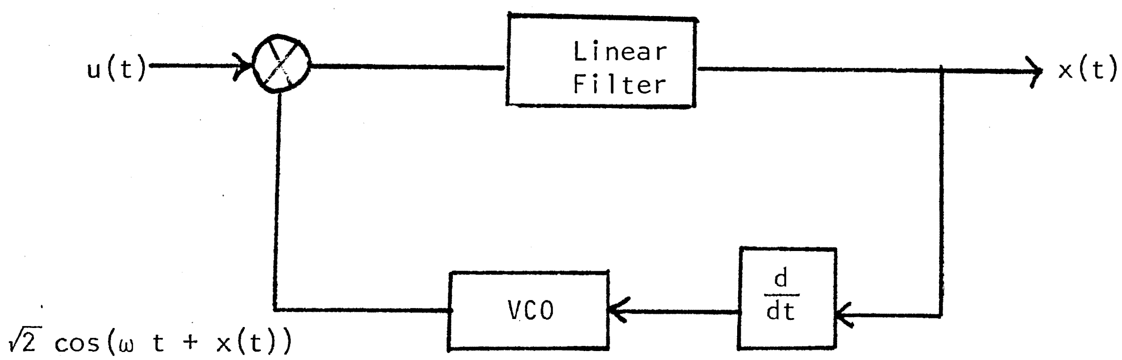


Figure 30. Demodulator for Angle-Modulated Signal

The output x(t) of the PLL is the estimate of the input angle modulation, in this case white Gaussian noise. The simulation was performed with parameters $K = 1/2$, $\omega = 50$, and $x(0) = 1/4$ and the variance of the noise input equal to unity. The output for the RK4 simulation is given in Figure 31. The Euler and AB2 methods gave nearly identical results.

Several interesting observations can be made from the results of this simulation. The mean value tracks the input angle modulation from the initial error to the mean of zero, as expected. However, the variance estimate remains vanishingly small and, perhaps the most surprising occurrence, the numerical algorithm used had no effect on the outcome of

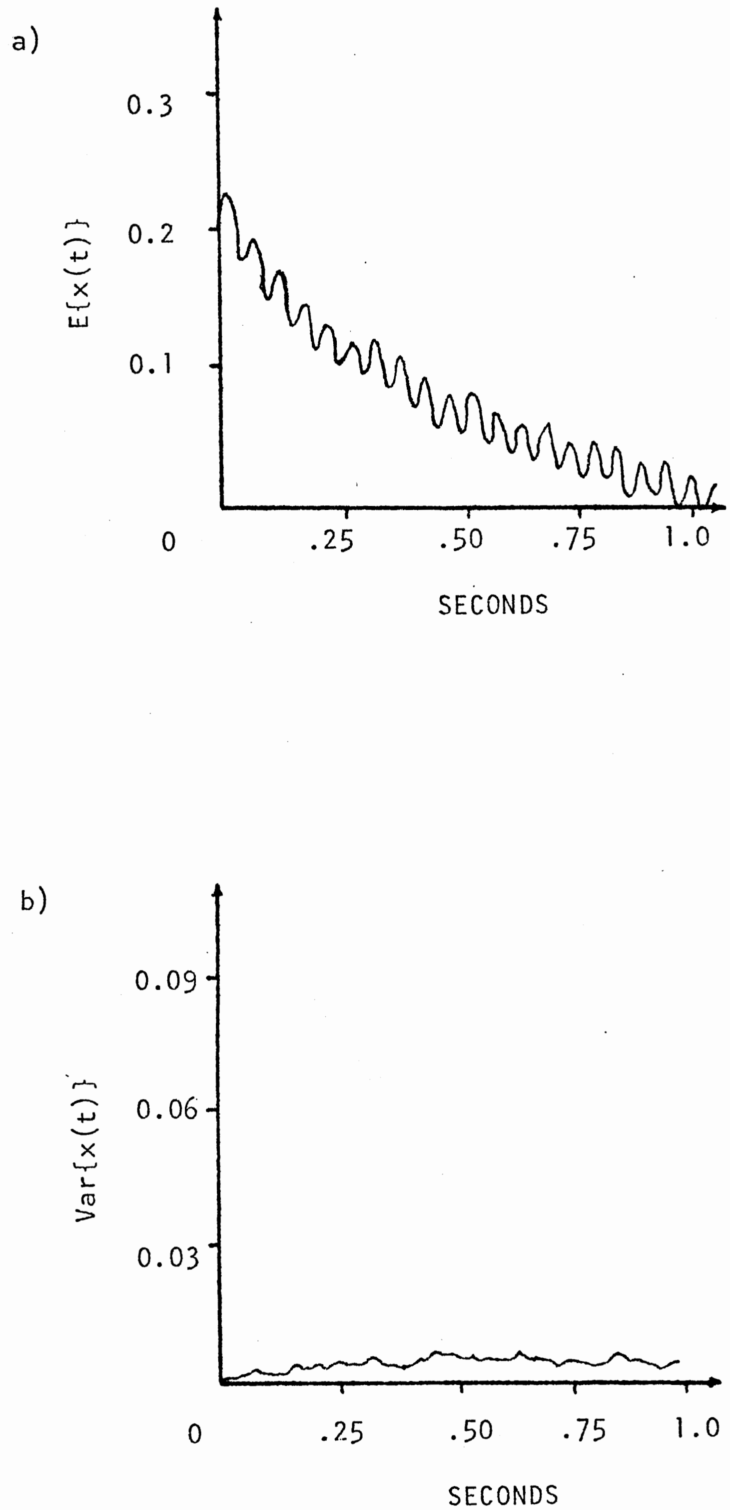


Figure 31. Ensemble-Averaged Mean and Variance for Angle-Demodulation (RK4)

the simulation. These unexpected results have a common origin. In the digital simulation, the white noise input is modeled using a random number generator and these numbers are used directly in the calculation of the sine function. From the series expansion of the sine and cosine, it is easily seen that Equation (4.6) results in terms containing $u^n(t_i)dt$, where n is a positive integer and $u(t_i)$ is a digitally generated random number. The statistics of such products are not the same as those required for the simulation of a Wiener process, as discussed in Chapter II. Without the correlation properties and time-dependent behavior associated with this process, the simulation results will not exhibit the range of behavior expected.

This example once again calls attention to the proper use of random numbers in a digital simulation. The technique in Chapter II is effective in instances when the Wiener process enters explicitly. It is not known at present how to best model the process when it is an implicit function of the integrand.

Concentration of Air Pollution

Air pollution has become a fact of life in certain industrialized areas throughout the world. The monitoring of pollution levels in an attempt to maintain them at a safe maximum has thus become more important and many mathematical models have been developed for the study of the diffusion of pollutants in the atmosphere. One class of models which has been widely used is the steady-state Gaussian plume model based on the diffusion equation. These models have inherent accuracy limits, however, as pointed out by Desalu, Gould and Schweppe [38], who developed a stochastic model for air pollution.

Their model involves partial differential equations which, when discretized, assume the form of nonlinear differential equations with multiplicative noise inputs. This advection-diffusion model accounts for the continuous fluctuation in meteorological conditions and incomplete knowledge of the true system model.

For practical implementation, the advection-diffusion equation is discretized into grid cells with the output of the model being the average pollution concentration within each cell. The equation can then be written as

$$A\,dx(t) = (2 - A + B)x(t)dt + C\,dt + D\,Q\,dt + E\,x(t)dw(t) + F\,q\,dt$$

where A, D, E, F are constants, B is a constant related to the gradients of wind velocity and pollution concentration, C is the pollution concentration outside the grid cell, Q and q are the deterministic and stochastic components of the pollution source and u(t) is Gaussianly distributed white noise. The multiplicative noise term Ex(t)u(t) introduces the need for nonlinear stochastic analysis.

Let A = 2 and B = 1 and assume that the stochastic component of the pollution source is zero, i.e., q = 0. Assume that the initial pollution concentration outside the grid cell is 4 gm/m^3 and that the deterministic component of the pollution source is 0.02 gm/hr with D = 1. With the initial pollution concentration within the grid cell taken to be zero, we simulate the equation

$$dx(t) = \frac{1}{2}x(t)dt + x(t)dw(t) + 2.01dt.$$

Figures 32, 33 and 34 give the results of this simulation.

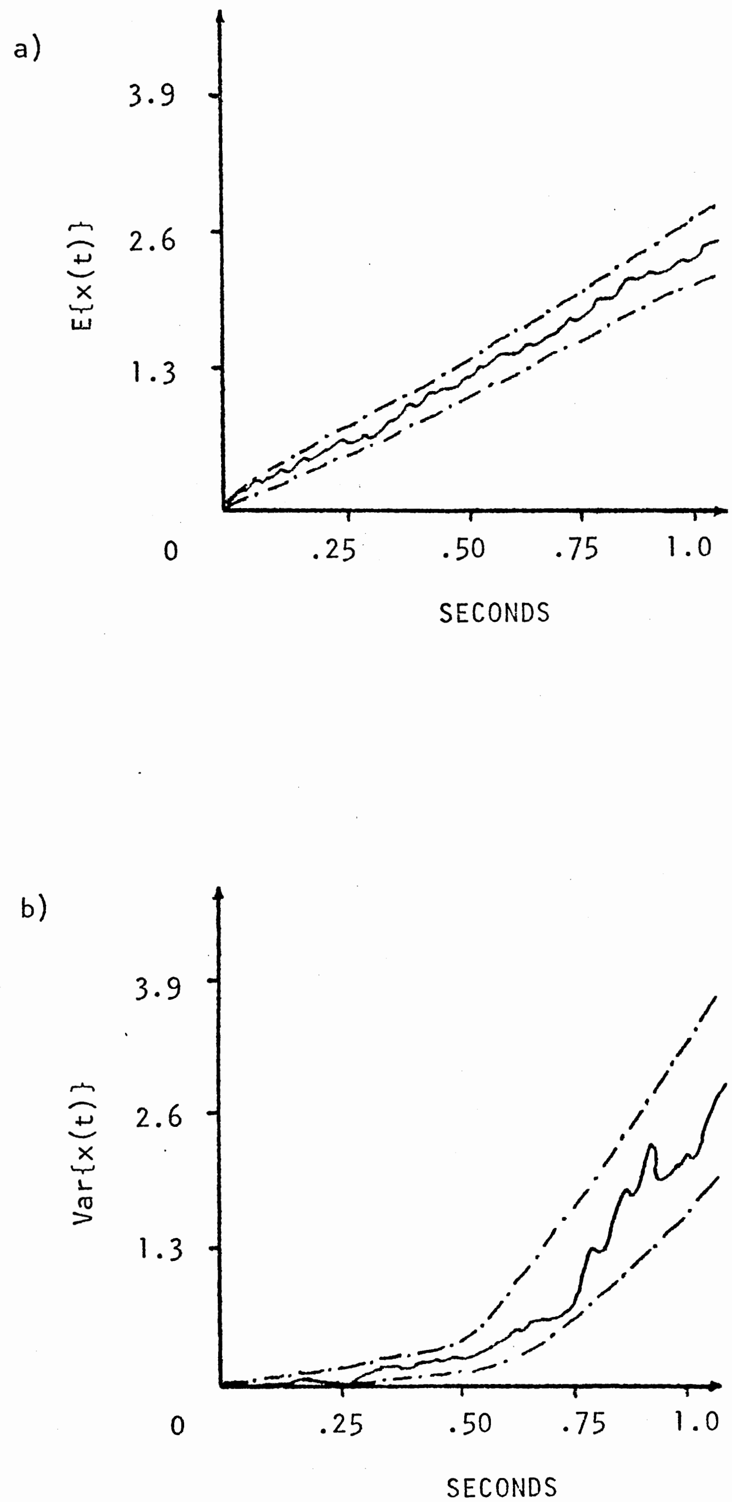


Figure 32. Ensemble-Averaged Mean and Variance for Pollution Concentration (Euler)

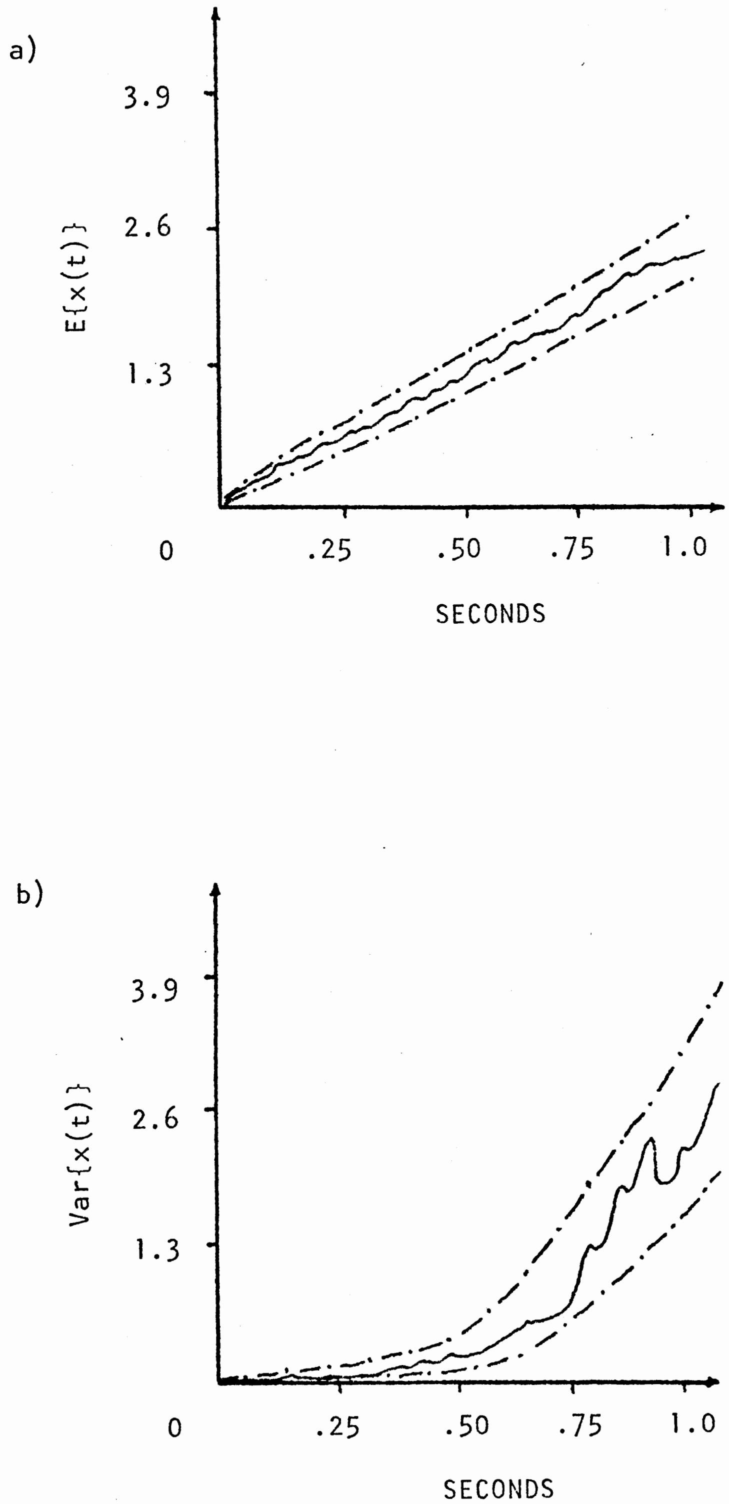
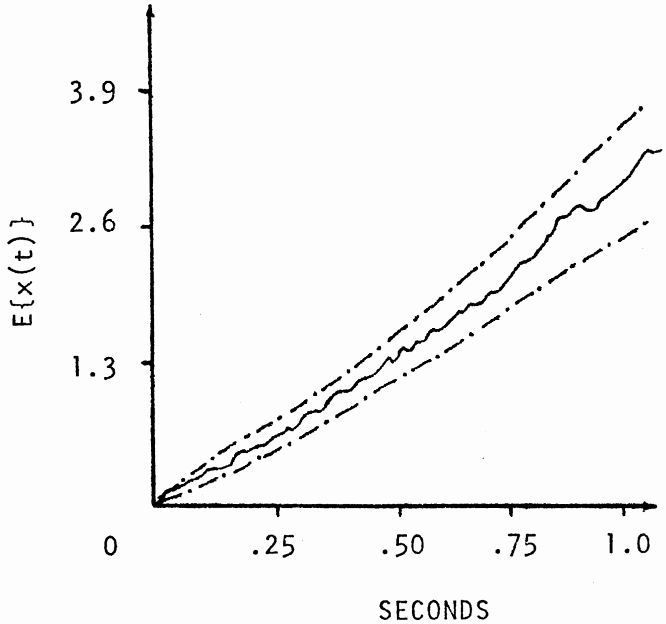


Figure 33. Ensemble-Averaged Mean and Vari- and for Pollution Concentration (AB2)

a)



b)

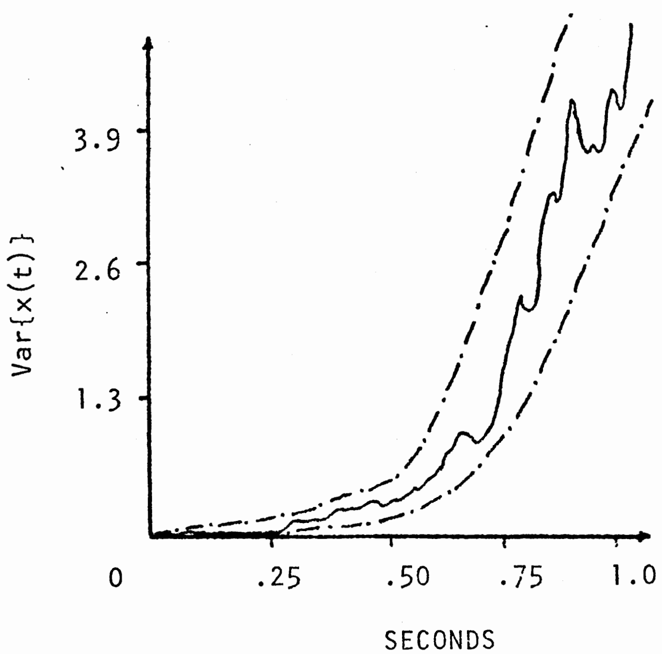


Figure 34. Ensemble-Averaged Mean and Variance for Pollution Concentration (RK4)

The mean and variance of the Euler and AB2 methods are virtually identical, with the mean pollution concentration after one second approximately 2.5 gm/m^3. The results of the RK4 simulation are rather different, with a mean concentration of 3.5 gm/m^3. The variance from the RK4 method is about twice as large as for the Euler and AB2 methods. Once again we see the influence that digital simulations can have on the analysis of system behavior.

Summary

Conclusions and insights gained from digital simulations of dynamic systems are invaluable to system designers and policy makers. However, care must be taken in interpreting results from simulations, especially in systems including multiplicative noise components. Errors arising because of truncation, roundoff, statistical anomalies and inaccuracies in the system model must be considered and, as illustrated in this chapter, possible influences of the particular numerical algorithm must also be noted. The numerical method can have novel effects on system outputs which could result in faulty conclusions for a system in which the nonlinear noise input is important.

CHAPTER V

CONCLUSIONS AND RECOMMENDATIONS

Conclusions

The mechanisms governing the relationships among theoretical stochastic integration results and numerical integration methods have been elucidated and illustrated in this thesis. The major contributions are the establishment of the moment equivalence of broad classes of numerical algorithms with various interpretations of stochastic integrals and the discovery of the unifying concept in these relationships, the correlation coefficient function.

Familiarity with various definitions of stochastic integrals and with numerical integration methods, as well as some early simulation work on systems corrupted by multiplicative noise, indicated that digital solutions of these systems varied with the numerical algorithm used. After developing an acceptable method for generating samples from a Wiener process, the mean and variance of the Euler and second-order Runge-Kutta methods were found. From comparison with the moments of the Ito and Stratonovich stochastic integrals, it was shown that the Euler method gives rise to the Ito solution and that the Runge-Kutta method produces the Stratonovich solution. Actual simulations of example problems showed good agreement with these results.

These results were then extended to the important case of the fourth-order Runge-Kutta method, in which it was shown that the Stratonovich-

Runge-Kutta relationship still held. The class of predictor integration methods were then discussed, and it was shown that the second-order and fourth-order Adams-Bashforth predictors correspond to the Ito stochastic integral. Digital simulations of examples were performed for these methods also and again showed a good agreement with the above results. Quantification of the agreement between simulated solutions and theoretical solutions was obtained by the derivation of confidence intervals about the exact solutions. All mean value estimates were within the calculated confidence limits, while a small percentage of some variance estimates lay outside these limits.

The study of the aforementioned relationships and their implications led to the realization that the point of functional evaluation of the stochastic integrand within each subinterval of a partition plays a determining role in the solution process. It was shown that this effect is evident in the correlation introduced between the integrand the the noise input. The correlation coefficient function was shown to provide a generalized method for analyzing the connections between stochastic integrals and numerical methods, with the particular functional evaluation point determining the amount of correlation. It was also demonstrated that the correlation function of the Stratonovich integral provides the general form of the correlation with the particular value determined by the evaluation point, which effect is manifested in the coefficient to the correlation function. The Ito and Stratonovich correlation functions were shown to be special cases of this generalized correlation coefficient function.

Several examples involving multiplicative noise have been discussed in light of the results obtained here. Some of the effects of digital simulation on these systems have been illustrated. Many ideas for further research have occurred thoughout this work and some of them are mentioned in the next section.

Recommendations for Further Research

Within the context of this thesis, there have arisen many possibilities for additional work which would be of interest to those involved in simulation research, as well as the broad area of nonlinear stochastic systems. An analysis of errors arising from truncation and roundoff in numerical algorithms would be of benefit. Classical error analysis is not appropriate since total derivatives are utilized, while differentials must be employed in the nonlinear stochastic case.

The development of numerical integration algorithms designed specifically for stochastic differential equations is another area of possible research. Some work has been done for Ito stochastic differential equations and Euler and Runge-Kutta numerical methods, considering mean-square convergence of solutions and convergence of distributions. Predictor methods could be profitably studied and also the implications for solving a stochastic equation interpreted in the sense of Stratonovich. The convergence of solutions in the sense of absolute error could also be addressed.

The discretization step is a critical factor in the stability and convergence of numerical integration methods to correct solutions. The step size enters into stochastic algorithms not only as in the deterministic case, but also through the variance of the noise input. An analytical study of the effect of the integration step size on the stability characteristics and convergence properties of numerical algorithms would aid in the development of stochastic numerical integration algorithms. Limited work by the author has indicated that the mean value is especially sensitive to changes in the integration step size.

It is well known that the number of sample runs in a Monte Carlo simulation plays a crucial role in the accuracy of the results. This effect is seen in the calculation of error bands for ensemble-averaged statistics. Variance estimates are much more sensitive to the number of runs than the mean value estimates, which effect can be seen from the form of the confidence intervals for these estimates. An investigation of the quantitative effects of the number of sample runs on solution accuracy would aid in determining computational requirements for particular applications.

SELECTED BIBLIOGRAPHY

[1] Wiener, N. Extrapolation, Interpolation and Smoothing of Stationary Time Series. M.I.T. Press, Cambridge, Mass., 1949.

[2] ________. "Differential-Space." J. Math. Phys., 2, 131-174, 1923.

[3] ________. Nonlinear Problems in Random Theory. Wiley, New York, 1958.

[4] McKean, H. P. Stochastic Integrals. Academic Press, New York, 1969.

[5] Ito, K. "Stochastic Integral." Proc. Imperial Academy, Tokyo, 20, 519-524, 1954.

[6] ________. "On a Stochastic Integral Equation." Proc. Japan Academy, 22, 33-35, 1946.

[7] ________. "Multiple Wiener Integrals." J. Math. Soc. Japan, 3, 68-79, 1951.

[8] ________. "On Stochastic Differential Equations." Mem. Amer. Math. Soc., 4, 51-89, 1951.

[9] ________. "On a Formula Concerning Stochastic Differentials." Nagoya Math. J., 3, 55-65, 1951.

[10] ________. Lectures on Stochastic Processes (notes by K. M. Rao). Tata Institute for Fundamental Research, Bombay, India, 1961.

[11] ________ and H. P. McKean. Diffusion Processes and Their Sample Paths. Academic Press, New York, 1964.

[12] Stratonovich, R. L. "A New Representation for Stochastic Integrals and Equations." J. SIAM Control, 4, 362-371, 1966.

[13] ________. "Conditional Markov Processes." Theor. Probability Appl., 5, 156-195, 1960.

[14] Mortensen, R. E. "Mathematical Problems of Modeling Stochastic Nonlinear Dynamic Systems." J. Stat. Phys., 1, 271-296, 1969.

[15] McShane, E. J. "A Riemann-Type Integral that Includes Lebesgue-Stieltjes, Bochner and Stochastic Integrals." Mem. Amer. Math. Soc., No. 88, 1969.

[16] __________. "Stochastic Integrals and Stochastic Functional Equations." SIAM J. Applied Math., 17, 287-306, 1969.

[17] __________. "Towards a Stochastic Calculus." Proc. Nat. Acad. Sci. U. S., 63, Part I, 275-280; Part II, 1084-1087, 1969.

[18] __________. "Stochastic Differential Equations and Models of Random Processes." Proc. 6th Berkeley Symp., 1971.

[19] __________. Stochastic Calculus and Stochastic Models. Academic Press, New York, 1974.

[20] Wright, D. J. "The Digital Simulation of Stochastic Differential Equations." IEEE Trans. on Aut. Cont., 19, 75-76, 1974.

[21] Kailath, T. and P. Frost. "Mathematical Modeling of Stochastic Processes." Stochastic Problems in Control Symp., Joint Automatic Control Conference, 1968.

[22] Rowland, J. R. and W. H. Holmes. "A Direct Covariance Algorithm for Computer-Aided Statistical Electronics Circuit Design." Int. J. Elec., 37, 65-77, 1974.

[23] Doob, J. L. Stochastic Processes. Wiley, New York, 1953.

[24] Parzen, E. Stochastic Processes. Holden-Day, San Francisco, 1962.

[25] Levy, P. Processus Stochastiques et Mouvement Brownien. Gauthier-Villars, Paris, 1948.

[26] Wong, E. and M. Zakai. "On the Relation Between Ordinary and Stochastic Differential Equations." Int. J. Engr. Sci., 3, 213-229, 1965.

[27] Bendat, J. S. and A. G. Piersol. Random Data: Analysis and Measurement Procedures. Wiley-Interscience, New York, 1971.

[28] Kolosov, G. E. and R. L. Stratonovich. "A Problem of Synthesis of an Optimal Regulator by Methods of Dynamic Programming." Avtomat. i Telemeh., 24, 1165-1173, 1963.

[29] Wonham, W. J. "Some Applications of Stochastic Differential Equations to Optimal Nonlinear Filtering." J. SIAM Cont., 2, 347-369, 1964.

[30] De Bellescize, H. "La Reception Synchrone." Onde(Electronics), 11, 225-240, 1932.

[31] Lindsey, W. C. Synchronization Systems in Communications and Control. Prentice-Hall, Englewood Cliffs, New Jersey, 1972.

[32] __________ and M. K. Simon. Telecommunications Systems Engineering. Prentice-Hall, Englewood Cliffs, New Jersey, 1972.

[33] Garodnick, J., J. Greco, and D. L. Schilling. "An All Digital Phase-Locked Loop for FM Demodulation." IEEE International Conference on Communications, Montreal, Canada, June, 1971.

[34] Viterbi, A. J. Principles of Coherent Communications. McGraw-Hill, New York, 1966.

[35] Anderson, T. O. and W. C. Lindsey. "Digital Transition Tracking Loops." Proceedings of International Telemetering Conference, Los Angeles, Calif., October, 1968.

[36] Nozaka, K. "Carrier Synchronization Techniques of PSK-Modem for TDMA Systems." Proceedings of INTELSAT/IEEE Digital Satellite Communication Conference, London, England, November, 1969.

[37] Bennett, W. R. Electrical Noise. McGraw-Hill, New York, 1960.

[38] Desalu, A. A., L. A. Gould, and F. C. Schweppe. "Dynamic Estimation of Air Pollution." IEEE Trans. on Aut. Cont., 19, 904-910, 1974.

VITA

John Mark Richardson

Candidate for the Degree of

Doctor of Philosophy

Thesis: DIGITAL SIMULATION OF NONLINEAR STOCHASTIC SYSTEMS

Major Field: Electrical Engineering

Biographical:

Personal Data: Born in Duncan, Oklahoma, April 27, 1954, the son of Mr. and Mrs. Cecil R. Richardson, Jr.

Education: Graduated from Duncan High School, Duncan, Oklahoma, in May, 1972; received the Bachelor of Science degree in Mathematics from Oklahoma State University, Stillwater, Oklahoma, in December, 1975; received the Master of Science degree in Electrical Engineering from Oklahoma State University in May, 1977; completed requirements for the Doctor of Philosophy degree in Electrical Engineering at Oklahoma State University in December, 1980.

Professional Experience: Undergraduate Research Assistant, Electrical Engineering, Oklahoma State University, June, 1975, to December, 1975; Graduate Research Assistant, Electrical Engineering, Oklahoma State University, January, 1976, to May, 1977, and June, 1979, to present; Summer Engineer, Halliburton Company, Duncan, Oklahoma, summer of 1977; Graduate Teaching Assistant, Electrical Engineering, Oklahoma State University, September, 1977, to May, 1978, and September, 1978, to May, 1979; Summer Research Engineer, Sandia National Laboratories, Albuquerque, New Mexico, summer of 1978.

Professional Organizations: Member of Institute of Electrical and Electronics Engineers, Pi Mu Epsilon and Eta Kappa Nu; Associate Member of Sigma Xi.