Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

H.E. de Swart

Definitions and concepts in the theory of stochastic differential equations

Department of Applied Mathematics Note AM-N8501 Augustus

The Centre for Mathematics and Computer Science is a research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a nonprofit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

DEFINITIONS AND CONCEPTS IN THE THEORY OF STOCHASTIC DIFFERENTIAL EQUATIONS

H.E. DE SWART
Centre for Mathematics and Computer Science, Amsterdam

In this technical note we summarize some definitions and concepts of stochastic processes, which are of importance in the theory of randomly perturbed dynamical systems. The following topics are reviewed: white noise, Chapman-Kolmogorov equation, Fokker-Planck equation, Wiener process, stochastic integrals, stochastic differential equations, and coloured noise.

MATHEMATICAL SUBJECT CLASSIFICATION: 34F05, 60J25.
KEY WORDS: white noise, Markov process, Fokker Planck equation, stochastic differential equation.

NOTE: These investigations were supported by the Netherlands Foundation for the Technical Sciences (S.T.W.), future Technical Science Branch of the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

Report AM-N8501
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

1. INTRODUCTION

Many physical processes can be modelled by sets of nonlinear ordinary differential equations. An intuitive way to study their response to random perturbations is to add stochastic forcing terms to the equations. However, to state the transition from a deterministic- to a stochastic evolution in a mathematical precise way is rather intricate, and requires a thorough knowledge of the theory of stochastic processes.

In this technical note it is not our intention to present a detailed and consistent discussion of stochastic processes; interested readers should consult standard texts on this topic, such as Gardiner (1983), van Kampen (1983) or Ahlberend & Kempe (1984). Merely, we will try to indicate, by means of a global overview, where and how difficulties in the formalism arise. The main reason for presenting this is to show that the introduction of stochastic terms has to be handled with care, and furthermore how certain types of randomly perturbed dynamical systems can be analyzed.

2. DYNAMICAL SYSTEMS WITH STOCHASTIC FORCING

Consider a deterministic dynamical system of the type

$$\frac{dx}{dt} = f_\mu(x,t) \qquad \text{in} \quad R^N, \tag{2.1}$$

where $x = (x^{(1)}, x^{(2)}, \ldots, x^{(N)})$ are state variables and $f_\mu = (f^{(1)}, f^{(2)}, \ldots, f^{(N)})$ the vectorfield with $\mu = (\mu_1, \mu_2, \ldots, \mu_r)$ parameters. From an intuitive point of view the effect of random perturbations on this model may be studied from

$$\frac{dx}{dt} = f_\mu(x,t) + \varepsilon \underset{\sim}{\sigma}(x,t).\eta(t). \tag{2.2}$$

Here ε measures the intensity of the noise and the elements of the $(N \times N)$ diffusion matrix $\underset{\sim}{\sigma}$ are supposed to be known functions of $x(t)$ and t. The N components of the vector $\eta(t)$ are rapidly fluctuating random terms, which are specified by their range of values they can take and assignation of a probability distribution to this range. We will use the conventional axioms of probability theory; for details we refer to Feller (1968, 1971). By specification of the diffusion matrix we can study the effect of randomly varying parameters in the model (internal forcing mechanisms) as well as the effect of external stochastic perturbations on the model.

We first consider some characteristics of the noise terms. We can define their moments, such as the mean values $< \eta(t) >$ and the correlation matrix

$$(2.3) \qquad \underset{\sim}{C}(t,\tau) = < \eta(t)\ \eta(t+\tau) > ,$$

etc., where the operator $< . >$ denotes an ensemble average over a large number M ($M \to \infty$) of realisations of the quantity between brackets. The covariance matrix is defined as $\underset{\sim}{C}(t,t)$, its diagonal elements being the variances.

The noise terms are assumed to be stationary processes, i.e. $\eta(t)$ and $\eta(t+\tau)$ have the same statistics for any τ. In addition they are supposed to have continuous sample paths. As a consequence the moments do not depend explicitely on time, hence $< \eta(t) > = < \eta >$, $\underset{\sim}{C}(t,\tau) = \underset{\sim}{C}(\tau)$, etc.. Furthermore the energy distribution of the noise in the spectral domain is independent of time, and given by the spectral density matrix

$$(2.4) \qquad \underset{\sim}{S}(\omega) = \lim_{T\to\infty} \alpha(\omega,T)\ \alpha^*(\omega,T) ,$$

where

$$(2.5) \qquad \alpha(\omega,T) = \int_{-T}^{T} \eta(t)\ e^{i\omega t}\ dt ,$$

and the asterix denotes a complex conjungation. It can be shown that for certain conditions the spectral density matrix is the Fourier transform of the correlation matrix, i.e.

$$(2.6) \qquad \underset{\sim}{S}(\omega) = \int_{-\infty}^{\infty} \underset{\sim}{C}(\tau)\ e^{i\omega\tau}\ d\tau ,$$

a result known as the Wiener Khinchine theorem , see Reichl (1981).

By equation (2.2) we have defined a multivariate stochastic process $X(\eta,t)$, being a continuous function of time t and certain random variables η; it can take on the realisations x. Generally, to give a complete description of such a process requires knowledge of all joint probability densities $\{p(x_\ell,t_\ell;\ldots;\ x_2,t_2;\ x_1,t_1)\}_{\ell=1}^{\infty}$, where $p(x_\ell,t_\ell;\ldots;\ x_2,t_2;\ x_1,t_1)$ denotes the joint probability density for $X_\varepsilon(t)$ to have the values x_1 at t_1, x_2 at t_2,..., and x_ℓ at t_ℓ. An alternative description is

possible with conditional probability densities

$$(2.7)\qquad p(x_n,t_n;\ldots;x_2,t_2;x_1,t_1|\,y_m,\tau_m;\ldots;y_2,\tau_2;y_1,\tau_1) =$$

$$= \frac{p(x_n,t_n;\ldots;x_2,t_2;x_1,t_1;y_m,\tau_m;\ldots;y_2,\tau_2;y_1,\tau_1)}{p(y_m,\tau_m;\ldots;y_2,\tau_2;y_1,\tau_1)} ,$$

which denote the probability densities for $X_\varepsilon(t)$ to have values x_1 at t_1, x_2 at t_2, ... and x_n at t_n, given it had values y_1 at τ_1, y_2 at τ_2,... and y_m at τ_m. From here on an ordering of time is defined as

$$(2.8)\qquad \tau_1 \le \tau_2 \le \ldots \le \tau_m \le t_1 \le t_2 \le \ldots \le t_n .$$

General stochastic processes $\overline{X}(t)$ are difficult to handle, since the evolutions depend on their complete history. Knowledge of the probability densities can only be obtained by performance of a large number of simulations of $\overline{X}(t)$ and evaluation of its statistics. The method is discussed recently by Murthy (1983). However, for the stochastic dynamical system (2.2) it will be shown that if the noise $\eta(t)$ is chosen to be processes $\xi(t)$, with

$$(2.9)\qquad <\xi> = 0 \;; \quad \underset{\sim}{C}(\tau) = \underset{\sim}{I}\,\delta(\tau),$$

$\underset{\sim}{I}$ the unity matrix and $\delta(\tau)$ the Dirac delta function, that probabilistically the evolution of the $X(\xi,t)$ are fully governed by knowledge of its recent condition, and not by any past values. We then have for the conditional probability densities

$$(2.10)\qquad p(x_n,t_n;\ldots;x_2,t_2;x_1,t_1|y_m,\tau_m;\ldots;y_2,\tau_2;y_1,\tau_1) =$$

$$= p(x_n,t_n;\ldots;x_2,t_2;x_1,t_1|y_m,\tau_m).$$

Such processes are called Markov processes and, as will be shown in the next section, their dynamics for the probability densities simplifies considerably.

The proof reads as follows. In order that the stochastic dynamical system is integrable, we must require that

$$(2.11)\qquad \Phi(t) = \int_{t_0}^{t} \underset{\sim}{\sigma}(x,s).\xi(s)ds$$

exists. It seems natural to assume that the $X_\varepsilon(t)$, and hence the $\Phi(t)$, are continuous functions of time. At this point we emphasize that the question whether continuous stochastic processes have continuous sample paths is not a trival one, as will become clear in section 3. However, on this condition we may write

$$(2.12) \qquad \Phi(t) = \lim_{\varepsilon \to 0} \{ \int_{t_0}^{t'-\varepsilon} \underset{\sim}{\sigma}(x(s),s) . \xi(s)ds \} + \int_{t'}^{t} \underset{\sim}{\sigma}(x(s),s) . \xi(s)ds.$$

For any $\varepsilon > 0$, the integrands of both integrals are independent. Hence, by continuity, $\Phi(t')$ and $\Phi(t) - \Phi(t')$ are statistically independent. Furthermore $\Phi(t) - \Phi(t')$ is independent of τ for all $\tau < t'$. This means that probabilistically $\Phi(t)$ is fully determined by knowledge of the value of $\Phi(t')$ and not of any past values. This can also be shown for the function

$$(2.13) \qquad v(t) = \int_{t_0}^{t} f_\mu(x,s)ds.$$

Hence $\Phi(t)$ and $v(t)$, and thus $X_\varepsilon(\xi,t)$, are Markov processes.

Concerning the processes $\xi(t)$ it can be seen from (2.9) that they are uncorrelated. Applying the Wiener Khinchine theorem we obtain for the spectral density matrix

$$(2.14) \qquad \underset{\sim}{S}(\omega) = \underset{\sim}{I} .$$

It appears that the spectra are flat (or white), hence the $\xi(t)$ are referred to as white noise processes. As is clear from (2.14) they have no finite energy contains, hence they are not a suitable parametrization for realistic random processes. Nevertheless, the white noise concept is necessary in order to construct more realistic noise processes. We return to this in the last section.

3. DYNAMICS OF MARKOV PROCESSES

Stochastic processes $\bar{X}(t)$, satisfying the Markov assumption (2.10), are fully governed by a conditional probability density $p(\bar{x},t|y,t')$, since we have by (2.7) that

$$p(\bar{x}_n,t_n;\ldots;\bar{x}_2,t_2;x_1,t_1) = \prod_{i=1}^{N-1} p(\bar{x}_{i+1},t_{i+1}|\bar{x}_i,t_i)\; p(x_1,t_1);$$

(3.1)

$$p(\bar{x}_n,t_n;\ldots;\bar{x}_2,t_2;x_1,t_1|y_m,\tau_m) = \prod_{i=1}^{N-1} p(\bar{x}_{i+1},t_{i+1}|\bar{x}_i,t_i)\; p(\bar{x}_1,t_1|y_m,\tau_m).$$

For processes having a continuous range of values, which are considered here, the one time probability density $p(\bar{x}_2,t_2)$ can be calculated from

$$(3.2) \qquad p(\bar{x}_2,t_2) = \int d\bar{x}_1\; p(\bar{x}_2,t_2|\bar{x}_1,t_1)\; p(x_1,t_1).$$

The conditional probability density $p(\bar{x}_3,t_3|\bar{x}_1,t_1)$, also called the transition probability density, is given by

$$(3.3) \qquad p(\bar{x}_3,t_3|\bar{x}_1,t_1) = \int dx_2\; p(\bar{x}_3,t_3;\bar{x}_2,t_2|\bar{x}_1,t_1).$$

For Markov processes (3.1) may be applied, and (3.3) reduces to

$$(3.4) \qquad p(\bar{x}_3,t_3|\bar{x}_1,t_1) = \int dx_2\; p(\bar{x}_3,t_3|\bar{x}_2,t_2)\; p(\bar{x}_2,t_2|\bar{x}_1,t_1),$$

which is called the Chapman-Kolmogorov equation. However, it is more convenient to use its differential form, which can be derived under rather mild conditions; see Gardiner (1983). It reads

$$(3.5) \qquad \frac{\partial}{\partial t}\, p(\bar{x},t|y,t') = -\nabla.\,[A(\bar{x},t)\; p(\bar{x},t|y,t')] +$$

$$+ \frac{1}{2}\,\nabla\nabla:\,[\underset{\sim}{B}(\bar{x},t)\; p(\bar{x},t|y,t')] +$$

$$+ \int dz\,[W(\bar{x}|z,t)\; p(z,t|y,t') - W(z|\bar{x},t)\; p(\bar{x},t|y,t')],$$

where ∇ denotes the nabla operator with respect to $\bar{x}$, and

$$(3.6)\qquad A(\bar{x},t) = \lim_{\rho\to 0}\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|z-\bar{x}|<\rho} dz(z-\bar{x})\; p(z,t+\Delta t|\bar{x},t)$$

$$= \lim_{\Delta t\to 0}\frac{1}{\Delta t}\langle \bar{X}(t+\Delta t)-\bar{X}(t)\rangle$$

$$(3.7)\qquad \underset{\sim}{B}(\bar{x},t) = \lim_{\rho\to 0}\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|z-\bar{x}|<\rho} dz(z-\bar{x})^2\; p(z,t+\Delta t|\bar{x},t)$$

$$= \lim_{\Delta t\to 0}\frac{1}{\Delta t}\langle [\bar{X}(t+\Delta t)-\bar{X}(t)]^2\rangle$$

$$(3.8)\qquad W(z|\bar{x},t) = \lim_{\Delta t\to 0}\frac{1}{\Delta t}\, p(z,t+\Delta t|\bar{x},t).$$

By definition the initial condition to (3.5) reads

$$(3.9)\qquad p(\bar{x},t_0|y,t_0) = \delta(\bar{x}-y).$$

The function $W(z|\bar{x},t)$ is the conditional probability density per unit time that the system changes from state $\bar{x}$ to z, and $A(\bar{x},t)$ and $\underset{\sim}{B}(\bar{x},t)$ are the first- and second order evolution moment respectively. It appears that higher order moments are at least $\mathcal{O}(\rho)$, so they need not to be considered in the limit $\rho \to 0$.

To give a further interpretation of the functions in (3.6) - (3.8) we consider the condition for Markov processes to have continuous sample paths.
It reads

$$(3.10)\qquad \lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|z-x|>\rho} dz\; p(z,t+\Delta t|\bar{x},t) = 0,$$

which means that the probability for the final position z to be finitely different from $\bar{x}$ tends to zero faster than Δt in the limit $\Delta t \to 0$. From (3.8) and (3.10) it follows that $W(z|\bar{x},t)$ is associated with the discontinuity properties of the processes, while the evolution moments are connected with a continuous motion. For $A(\bar{x},t) = 0$, $\underset{\sim}{B}(\bar{x},t) = 0$ equation (3.5) reduces to the Master equation, which describes a discontinuous jump process. For Markov processes having continuous sample paths (3.5) reduces to the Fokker-Planck equation

$$\frac{\partial}{\partial t}\, p(\bar{x},t|y,t') = -\nabla.\,[A(\bar{x},t)\; p(\bar{x},t|y,t')] + \tag{3.11}$$
$$+\frac{1}{2}\,\nabla\nabla:\,[\underset{\sim}{B}(\bar{x},t)\; p(\bar{x},t|y,t')]\,.$$

Up to now we have considered the time evolution of $p(\bar{x},t|y,t')$ with respect to the final variables $\bar{x}$. In the same way an evolution equation can be derived with respect to the initial variables y. This results in the differential form of the backward Chapman-Kolmogorov equation:

$$\frac{\partial}{\partial t'}\, p(\bar{x},t|y,t') = -A(y,t').\nabla_y\, p(\bar{x},t|y,t') + \tag{3.12}$$
$$-\frac{1}{2}\,\underset{\sim}{B}(y,t'):\,\nabla_y\nabla_y\, p(\bar{x},t|y,t') +$$
$$+\int dz\; W(z|y,t')\;[p(\bar{x},t|y,t') - p(\bar{x},t|z,t')],$$

where ∇_y is the nabla operator with respect to y, and (3.9) must now be interpreted as a final condition. For continuous Markov processes the integral contribution on the right hand side identically vanishes.

The backward equation has a better mathematical foundation than the forward equation (Feller, 1968). Nevertheless, both equations are useful; the forward equation gives more directly the values of measurable quantities as functions of time. The backward equation is mostly applicated in the study of first passage- or exit problems, see Gardiner (1983).

A special class of processes described by a Fokker-Planck equation are the $X(\eta,t)$ governed by the stochastic dynamical system (2.2). The question is how the evolution moments $A(x,t)$ and $\underset{\sim}{B}(x,t)$ can be related to this system. As a first answer we consider the deterministic part of the evolution, i.e. in fact the dynamical system (1.1). We then have

$$p(z,t+\Delta t|x,t) = \delta(z-x-f_\mu(x,t)\;\Delta t)\;;\;\Delta t \to 0, \tag{3.13}$$

and from (3.6) and (3.7) it follows that

$$A(x,t) = f_\mu(x,t);\quad \underset{\sim}{B}(x,t) = 0. \tag{3.14}$$

Obviously the so-called driftvector $A(x,t)$ describes the deterministic evolution of the system. The reduced Fokker-Planck equation corresponding to (3.14) is known as the Liouville equation.

Next consider the stochastic evolution, as described by the $\Phi(t)$ defined in (2.11).

The evolution moments for the processes appear to be

$$(3.15) \qquad A(x,t) = 0; \quad \underset{\sim}{B}(x,t) = \underset{\sim}{\sigma}(x,t).\underset{\sim}{\sigma}^{T}(x,t),$$

where $\underset{\sim}{\sigma}^{T}$ is the adjungated of $\underset{\sim}{\sigma}$. Obviously the second evolution moment is connected with stochastic motion.

Combining the results we have that the dynamics of the stochastic dynamical system (2.2) is described by the Fokker-Planck equation

$$(3.16) \qquad \frac{\partial}{\partial t}\, p(x,t|y,t') = -\nabla.[f_{\mu}(x,t)\; p(x,t|y,t')] + \\ + \frac{1}{2}\nabla\nabla: [\,\underset{\sim}{\sigma}(x,t).\underset{\sim}{\sigma}^{T}(x,t)\; p(x,t|y,t')].$$

In the next section some implications of this result will be derived.

4. RELATION BETWEEN WHITE NOISE AND WIENER PROCESSES: DYNAMICAL IMPLICATIONS

In this section we consider the time-integrated white noise processes

$$(4.1) \qquad W(t) = \int_{t_0}^{t} \xi(s)ds.$$

From the results of the previous section it follows that the corresponding Fokker-Planck equation reads

$$(4.2) \qquad \begin{aligned} &\frac{\partial}{\partial t}\, p(w,t|w_0,t_0) = \frac{1}{2}\nabla_w^2\, p(w,t|w_0,t_0), \\ &p(w,t_0|w_0,t_0) = \delta(w - w_0). \end{aligned}$$

By definition, stochastic processes with a conditional probability density being a solution of (4.2) are called Wiener processes.

The solution of the diffusion problem (4.2) can be found by standard methods, as e.g. discussed in Morse & Feshbach (1953). The result is

$$(4.3) \qquad p(w,t|w_0,t_0) = \frac{1}{[2\pi(t - t_0)]^{N/2}} \exp\left[\frac{-(w - w_0)^2}{2(t - t_0)}\right],$$

which is a multivariate Gaussian distribution with mean w_0. In figure 1 the one-dimensional solution is shown at different times $(t - t_0)$.

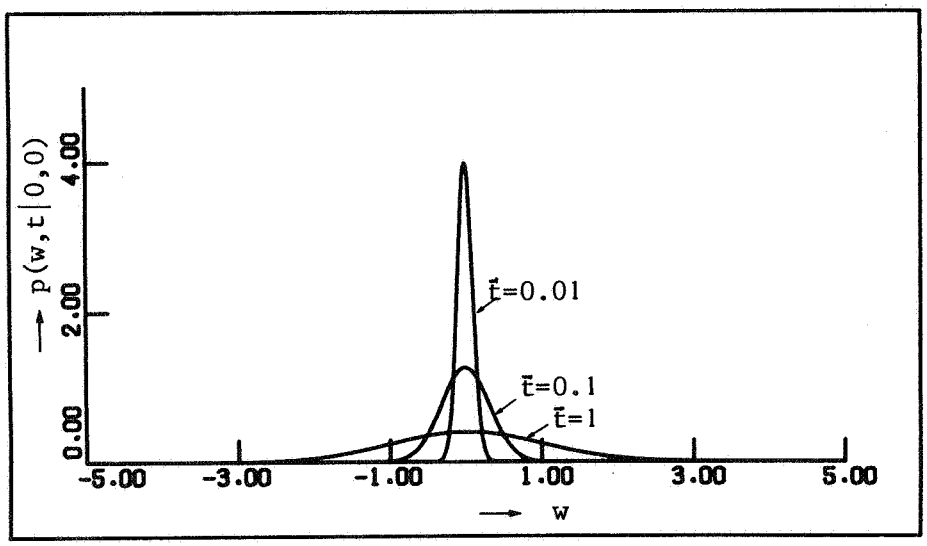

Figure 1. Conditional probability density of a Wiener process at different times $t - t_0 = \tau$.

We will first derive an important property for W(t). By (3.1) and (4.2) we have

$$(4.4)\qquad p(w_n,t_n;w_{n-1},t_{n-1};\ldots;w_0,t_0) =$$

$$= \left\{\prod_{i=0}^{N-1} \frac{1}{[2\pi(t_{i+1} - t_i)]^{N/2}} \exp\left[\frac{-(w_{i+1} - w_i)^2}{2(t_{i+1} - t_i)}\right]\right\} \cdot p(w_0,t_0).$$

Defining the increments

$$(4.5)\qquad \Delta W_i = W(t_{i+1}) - W(t_i),$$

we find that the joint probability density for the ΔW_i is

$$(4.6)\qquad p(\Delta w_n,t_n;\Delta w_{n-1},t_{n-1};\ldots;\Delta w_2,t_2;\Delta w_1,t_1;\Delta w_0,t_0) =$$

$$= \left\{\prod_{i=1}^{N} \frac{1}{[2\pi\,\Delta t_i]^{N/2}} \exp\left[\frac{-\Delta w_i^2}{2\Delta t_i}\right]\right\} p(w_0,t_0).$$

Using definition (2.7) for the conditional probability densities and the Markov condition (2.10) we can show that the conditional probability densities $p(\Delta w_i,\Delta t_i|\Delta w_j,\Delta t_j)$ are independent of Δw_j and Δt_j for all i and $j < i$, and furthermore $p(\Delta w_i,t_i|w_0,t_0)$ does not depend on w_0 and t_0 for all i. This means that the increments ΔW_i are statistically independent of each other and of $W(t_0)$.

Next consider the correlation matrix for the Wiener process. It is written

$$\underset{\sim}{C}(t,\tau) = \langle [W(t) - W(t+\tau)]W(t+\tau)\rangle + \langle W(t+\tau)W(t+\tau)\rangle; \quad \tau \leq 0,$$

(4.7)

$$\underset{\sim}{C}(t,\tau) = \langle [W(t+\tau) - W(t)]W(t)\rangle + \langle W(t)W(t)\rangle; \quad \tau \geq 0.$$

The first matrices on the right hand sides of equations (4.7) are zero, because of the independence of the increments. The remaining covariance matrices can be calculated from an alternative definition of the ensemble averaging operator, i.e.

$$(4.8) \qquad \langle W(s)W(s)\rangle = \int dwdw'.ww' \; p(w,s;w',s|w_0,t_0).$$

By means of (3.1) and (4.3) the result becomes

$$(4.9) \qquad \langle W(s)W(s)\rangle = w_0 w_0 + s \underset{\sim}{I},$$

and hence the autocorrelation matrix of the Wiener process reads

$$(4.10) \qquad \underset{\sim}{C}(t,\tau) = \min(t,t+\tau)\, \underset{\sim}{I} + w_0 w_0.$$

Obviously the variances of individual realisations of the Wiener process increase linearly with time, hence the sample paths may vary indefinitely. This can be seen from figure 2, which shows some realisations of a one-dimensional Wiener process.

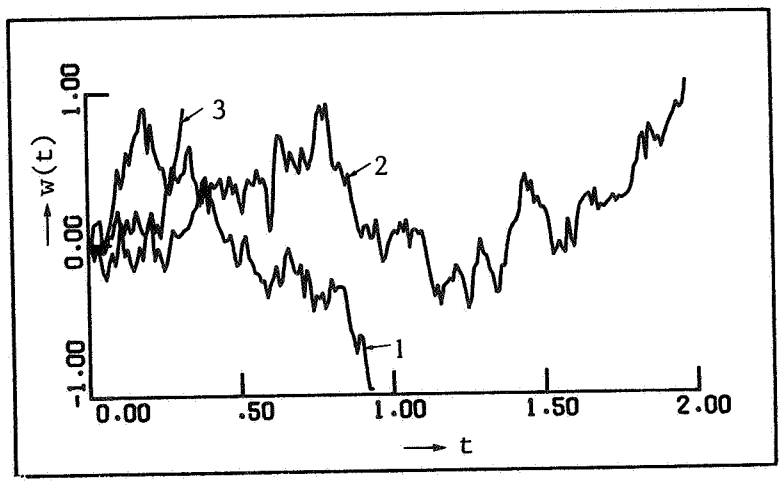

Figure 2. Three simulated sample paths of the Wiener process, illustrating their great variability

Furthermore the sample paths are continuous but nowhere differentiable, as is seen from the probability that the finite time-difference quotient of W(t) is larger than a certain number γ. By (4.3) this is

$$(4.11)\qquad P\left(\left|\frac{W(t+\Delta t)-W(t)}{\Delta t}\right|>\gamma\right)=2\int_{\gamma\Delta t}^{\infty}dw\,\frac{1}{(2\pi\,\Delta t)^{N/2}}\exp\left[\frac{-(w-w_0)^2}{2\Delta t}\right],$$

which tends to 1 as $\Delta t \to 0$. Hence the derivative at any point of the sample path is almost certainly infinite. This implies that white noise can only be properly interpreted in its time-integrated form.

As a consequence the stochastic dynamical system (2.2) is not correctly formulated. This objection is met by integrating the equations once:

$$(4.12)\qquad x(t)-x(t_0)=\int_{t_0}^{t}f_\mu(x,s)ds+\varepsilon\int_{t_0}^{t}\underset{\sim}{g}(x,s).\xi(s)ds.$$

By writing

$$(4.13)\qquad dW(t)=W(t+dt)-W(t)=\xi(t)dt,$$

we can redefine (4.12) as

$$(4.14)\qquad x(t)-x(t_0)=\int_{t_0}^{t}f_\mu(x,s)ds+\varepsilon\int_{t_0}^{t}\underset{\sim}{g}(x,s).dW(s).$$

The second term on the right hand side, being some kind of a stochastic integral with respect to the W(t), will be defined in the next section.

5. STOCHASTIC INTEGRALS AND STOCHASTIC DIFFERENTIAL EQUATIONS

Up to now we have found that a white noise-forced dynamical system of the type (2.2) is a Markov process. Then its dynamics are governed by a Fokker-Planck equation, which in this case becomes equation (3.16). Furthermore it appeared that the evolution of the system is only properly described by the integral equation (4.14). In order to interprete this result we must define the stochastic integral $\int_{t_0}^{t}\underset{\sim}{g}(x,s)\,dW(s)$. The general procedure to do so is to divide the interval $[t_0,t]$ into n subintervals by means of partitioning points, such that $t_0\le t_1\le\ldots\le t_{n-1}\le t_n\equiv t$. Next we introduce intermediate points τ_i, where $t_{i-1}\le\tau_i\le t_i$; see figure 3.

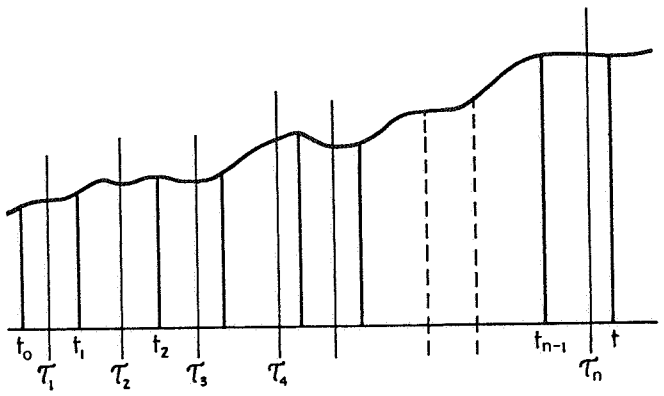

Figure 3. Partioning of the time interval used in the definition of stochastic integration.

The stochastic integral is defined as a kind of a Riemann-Stieltjes integral, viz. as a limit of the partial sums

$$(5.1) \qquad R_n = \sum_{i=1}^{n} \underset{\sim}{g}(x,\tau_i).[W(t_i) - W(t_{i-1})].$$

Generally the integral depends on the particular choice of the τ_i. Two well-known concepts are the Ito stochastic integral

$$(5.2) \qquad I \int_{t_0}^{t} \underset{\sim}{g}(x,s).dW(s) = \underset{n\to\infty}{\text{ms-lim}} \sum_{i=1}^{n} \underset{\sim}{g}(x,t_{i-1}).[W(t_i) - W(t_{i-1})]$$

and the Stratonovich integral

$$(5.3) \qquad S \int_{t_0}^{t} \underset{\sim}{g}(x,s).dW(s) = \underset{n\to\infty}{\text{ms-lim}} \sum_{i=1}^{n} \underset{\sim}{g}(x,\tfrac{1}{2}(t_{i-1}+t_i)).[W(t_i) - W(t_{i-1})].$$

The notation ms-lim denotes a mean square limit in the sense that the variance of the difference between R_n and the result of the integral tends to zero for $n \to \infty$. The difference in interpretation of the two integrals will be more clear after the introduction of nonanticipating functions.

Functions G(t) are called nonanticipating of t if for all $t < s$, G(t) is statistically independent of $W(s) - W(t)$. This definition states that the values G(t) at time t are not effected by the evolution of the Wiener processes W(s) for $s > t$. For example, W(t) themselves are examples of such functions.

A reason to consider nonanticipating functions specifically is that many results can be derived, being valid only for such functions. Furthermore, in physical applications, where always a causality is assumed in the sense that the unknown future cannot effect the present, functions describing stochastic processes are nonanticipating. Finally, they are involved in the concept of stochastic differential equations, as will become clear.

For nonanticipating functions G(t) the following identities hold:

$$(5.4) \qquad \int_{t_0}^{t} G(s)\, dW(s)\, dW(s) = \int_{t_0}^{t} G(s)\, \underset{\sim}{I}\, ds,$$

$$(5.5) \qquad \int_{t_0}^{t} G(s)\, [dW(s)]^{m+2} = 0, \quad m > 0$$

$$(5.6) \qquad \int_{t_0}^{t} G(s)\, dW(s)\, ds = 0,$$

as long as $G(\tau_i)$ and $W(t_i)$ are statistically independent. As a consequence the integrals in (5.4) - (5.6) must be Ito stochastic integrals. A proof can be found in Gardiner (1983). The method resulting in these relations is known as Ito calculus. Roughly spoken (5.4) - (5.6) show that the dW(s) can be interpreted as differentials of order $\frac{1}{2}$, while ds is of order 1.

According to the definitions (5.2) and (5.3) we can associate two types of stochastic differential equations to the integral equation (4.14). A stochastic process $X_\varepsilon(t)$ is the solution of Ito stochastic differential equation

$$(5.7) \qquad dx(t) = f_\mu(x,t)\, dt + \varepsilon\, \underset{\sim}{g}(x,t).dW(t)$$

if for all t and t_0, (4.14) holds with the Ito interpretation of the stochastic integral. In a similar way $X_\varepsilon(t)$ satisfies a Stratonovich stochastic differential equation

(5.8) $\quad (S)\ dx(t) = f_\mu(x,t)\ dt + \varepsilon\ \underset{\sim}{g}\ (x,t).dW(t)$

if in (4.14) the stochastic integral is considered as a Stratonovich integral.

From (5.4) - (5.6) we may conclude that the Ito formulation is preferable from a mathematical point of view, and it will be used from now on. A discussion about this choice is presented at the end of section 7.

Using Lipschitz continuity conditions for $f_\mu(x,t)$ and $\underset{\sim}{g}(x,t)$ Arnold (1974) has proved that there exists a unique nonanticipating solution of (5.7) on a finite time interval, which is bounded if $f_\mu(x,t)$ and $\underset{\sim}{g}(x,t)$ satisfy certain growth conditions; see also Schuss (1980).

6. NUMERICAL SIMULATIONS OF STOCHASTIC DIFFERENTIAL EQUATIONS

To obtain information about the solutions of the Ito stochastic differential equation (5.7) one can try to solve the corresponding Fokker-Planck equation (3.16), however in general this is a complicated problem. An alternative way consists of the statistical analysis of a large number of realizations of the process. Employing a finite difference scheme of equation (5.7), viz.

(6.1) $\quad x(t+\Delta t) = x(t) + f_\mu(x(t),t)\ \Delta t + \varepsilon \underset{\sim}{g}(x(t),t).\Delta W(t),\ ;\ \Delta t \to 0,$

we make for simulations the substitution

(6.2) $\quad \Delta W(t) = A(\Delta t)\ G.$

Here the components of G are random generators and $A(\Delta t)$ is a qualified function of the time stap Δt. For G we take mutual independent Gaussian generators with zero mean and standard deviation 1, which is a direct consequence of the characteristics of the Wiener process as discussed in section 4. The function $A(\Delta t)$ is determined by the condition that in the limit $\Delta t \to 0$, ΔW and $A(\Delta t)G$ result in equal variances of the increment Δx, i.e.

(6.3) $\quad \langle \Delta W(t)\ \Delta W(t) \rangle = A^2(\Delta t)\ \langle GG \rangle;\quad \Delta t \to 0.$

By (5.4) we have that the left hand side of (6.3) equals Δt, and using the generator characteristics we obtain

$$A(\Delta t) = \sqrt{\Delta t}. \tag{6.4}$$

Hence the numerical scheme for the stochastic differential equation (5.7) reads

$$x(t+\Delta t) = x(t) + f_{\mu}(x(t),t)\,\Delta t + \varepsilon \underset{\sim}{\sigma}(x(t),t).G\,\sqrt{\Delta t}. \tag{6.5}$$

Finally we remark that there are two conditions on the time step. First,

$$\Delta t \ll 1, \tag{6.6}$$

in order to provide instabilities due to the deterministic integration. Secondly, in the limit $\varepsilon \to 0$ we expect the variance of the difference between the stochastic- and deterministic trajectories to be zero. Therefore we must also require

$$\Delta t = o(\varepsilon); \quad \varepsilon \to 0. \tag{6.7}$$

7. SYSTEMS PERTURBED BY COLOURED NOISE AND THE WHITE NOISE LIMIT

Up to now we have assumed the stochastic forcing terms to be white noise processes, which are fully uncorrelated. It was found in section 2 that the corresponding power spectra are flat, consequently white noise cannot have a finite energy contains. Furthermore, in section 4 it was concluded that the process can only be interpreted in its time-integrated form. For these reasons white noise is a poor representation for fluctuations in physical systems.

The motivation of the present section is therefore to construct a more realistic parametrerization of a random signal. This can be done by means of the previously obtained results, i.e. we will construct a stationary Markov process $\zeta(t)$ from the Ito stochastic differential equation

$$dz(t) = a(z)\,dt + \underset{\sim}{b}(z).dW(t), \tag{7.1}$$

by suitable choices of $a(z)$ and $\underset{\sim}{b}(z)$. Next the $\zeta(t)$ will be considered as the new stochastic forcing terms.

A stationary Markov process can be generated by putting

(7.2) $\quad a(z) = -\alpha z; \quad \underset{\sim}{b}(z) = \beta \underset{\sim}{I}$

in (7.1) with $\alpha > 0$ and β constants. The solution of (7.1) for this choice can be found by means of the variation of parameter-method and Ito calculus. It reads

$$(7.3) \qquad z(t) = z(t_0)\, e^{-\alpha(t-t_0)} + \beta \int_{t_0}^{t} e^{-\alpha(t-s)}\, dW(s),$$

and the process is called a multi-variate Ornstein-Uhlenbeck process.

The correlation matrix of this process is obtained by application of the correlation formula

$$(7.4) \qquad \left\langle \int_{t_0}^{t} G(s)\, dW(s) \int_{t_0}^{t'} H(s)\, dW(s) \right\rangle = \int_{t_0}^{\min(t,t')} \langle G(s)\, H(s) \rangle\, ds,$$

being valid for arbitrary nonanticipating functions G(t) and H(t). A proof of (7.4) is given by Gardiner (1983). The result is

$$(7.5) \qquad \underset{\sim}{C}(t,\tau) = z(t_0)\, z(t_0)\, e^{-2\alpha(t-t_0)}\, e^{-\alpha|\tau|} + \frac{\beta^2}{2\alpha}\, e^{-\alpha|\tau|}\, \underset{\sim}{I}.$$

Obviously α^{-1} is a measure of the correlation time of the $z(t)$, and hence of the memory of the processes.

In the limit $(t-t_0) \to \infty$ the correlations do no longer depend explicitly on time, which means that

$$(7.6) \qquad \zeta(t) = \beta \int_{-\infty}^{t} e^{-\alpha(t-s)}\, dW(s)$$

is a stationary multi-variate Markov process. Introduction of the $\zeta(t)$ as stochastic forcing terms in equation (2.2) the stochastic dynamical system can be formulated as

$$(7.7) \qquad \begin{aligned} d\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} &= \begin{pmatrix} f_\mu(x,t) + \underset{\sim}{g}(x,t) \cdot y \\ -\alpha y \end{pmatrix} dt + \varepsilon \begin{pmatrix} 0 & 0 \\ 0 & \beta \underset{\sim}{I} \end{pmatrix} \cdot dW(t), \\ y(t) &= \varepsilon \zeta(t), \end{aligned}$$

which is a system of 2N Ito stochastic differential equations. This system is of the same type as (5.7), hence a Fokker-Planck equation can be found in the same way as has previously been done.

A relation between α and β follows from the condition that the white noise forcing in (5.7) and the $\zeta(t)$ forcing terms in (7.7) result in equal variances of the increments dx, i.e.

$$(7.8) \qquad \langle dW(t)\, dW(t)\rangle = \langle dR(t)\, dR(t)\rangle$$

must hold. Here W(t) is a multi-variate Wiener process and

$$(7.9) \qquad R(t) = \int_{t_0}^{t} \zeta(s)ds.$$

Using (5.4) we have that the left hand side of (7.8) equals dt, and from (7.6) it follows that

$$(7.10) \qquad \langle dR(t)\, dR(t)\rangle = \langle [R(t+dt) - R(t)]\, [R(t+dt) - R(t)]\rangle = \frac{\beta^2}{\alpha^2}\, dt,$$

hence by (7.8) we have

$$(7.11) \qquad \beta^2 = \alpha^2.$$

For these parameter values the spectral density matrix of the $\zeta(t)$ is calculated by application of the Wiener-Khinchine theorem (2.6) to the stationary form of the correlation matrix in (7.5). The result is

$$(7.12) \qquad \underset{\sim}{S}(\omega) = \frac{\alpha^2}{\alpha^2 + \omega^2}\, \underset{\sim}{I}.$$

Some examples of energy spectra are shown in figure 4 for different values of α.

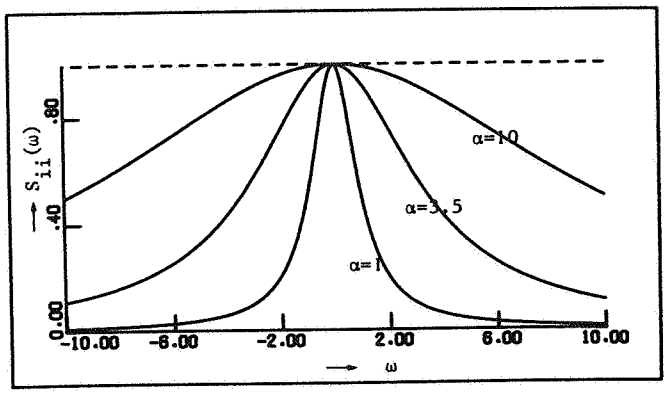

Figure 4. Energy spectrum of a process $\zeta_i(t)$ for different values of α.

Since their spectra resemble those of coloured light the $\zeta(t)$ are called coloured noise processes (for large values of α pink- or red noise). In the limit $\alpha \to \infty$, $\underset{\sim}{S}(\omega)$ reduces to the unity matrix and the $\zeta(t)$ become white noise processes.

It can be shown (Gardiner, 1983) that in this limit the $X(\eta,t)$ in (2-2), with $\eta(t) = \zeta(t)$, are governed by the Stratonovich stochastic differential equations (5.8). Thus it follows that the Ito formalism, although preferable from a mathematical point of view, is not always the most natural choice physically. This is no actual problem, since it can be shown that in the white noise limit the two types of stochastic intervals can be related to each other. It then appears that (5.8) can be identified with the Ito stochastic differential equation

$$(7.13) \qquad dx(t) = [f_\mu(x,t) + \tfrac{1}{2}\underset{\sim}{g}(x,t) : \nabla\sigma(x,t)]\, dt + \underset{\sim}{g}(x,t)\, dW(t),$$

Gardiner (1983). A comparison between the two concepts of stochastic differential equations, including specific examples, is given by Vasudevan (1983). For more details including applications, we refer to most textbooks on stochastic processes and stochastic differential equations, e.g. Schuss (1980).

REFERENCES

AHLBEREND, N. & V. KEMPE (1984). *Analyse stochastische Systeme: nichtlineare dynamische Systeme*. R. Oldenbourg Verlag.

ARNOLD, L. (1974). *Stochastic differential equations*. John Wiley & Sons.

FELLER, W. (1968). *An introduction to probability theory and its applications*, 3rd ed., vol. I. John Wiley & Sons.

FELLER, W. (1971). *An introduction to probability theory and its applications*, 2nd ed., vol. II. John Wiley & Sons.

GARDINER, C.W. (1983). *Handbook of stochastic methods for physics, chemistry and the natural sciences*. Springer Verlag.

KAMPEN, N.G. van (1983). *Stochastic processes in physics and chemistry*, 2nd ed.. North Holland.

MORSE, P.M. & H. FESHBACH (1953). *Methods of theoretical physics*, part 1. Mc Graw Hill.

MURTHY, K.P.N. (1983). *Monte Carlo methods: an introduction*. In: G.S. Agarwal & S. Dattagupta (eds). *Stochastic processes: formalism and applications*. Lect. Notes Ph. 184, Springer Verlag, pp. 116 - 121.

REICHL, L.E. (1981). *A modern course in statistical physics.* Edward Arnold (Publ.) Ltd.

SCHUSS, Z. (1989). *Theory and applications of stochastic differential equations.* John Wiley & Sons.

VASUDEVAN, R. (1983). *Stochastic differential equations.* In: G.S. Agarwal & S. Dattagupta (eds). *Stochastic processes: formalism and applications.* Lect. Notes Ph. 184, Springer Verlag, pp. 37 - 53.