Journal of Statistical Physics, Vol. 42, Nos. 5/6, 1986

A Variational Approach to Stochastic Nonlinear Problems

R. Phythian[1] **and W. D. Curtis**[1]

Received April 12, 1985; revision received October 8, 1985

A variational principle is formulated which enables the mean value and higher moments of the solution of a stochastic nonlinear differential equation to be expressed as stationary values of certain quantities. Approximations are generated by using suitable trial functions in this variational principle and some of these are investigated numerically for the case of a Bernoulli oscillator driven by white noise. Comparison with exact data available for this system shows that the variational approach to such problems can be quite effective.

KEY WORDS: Stochastic nonlinear equations; moments; variational principle; trial functions; reserved variables; statistical linearization.

1. INTRODUCTION

The mathematical modeling of a wide variety of systems in science and engineering gives rise to stochastic nonlinear equations, and the development of reliable and tractable procedures for the approximate solution of these is regarded as a matter of great importance. The stochastic element in these equations often takes the form of a Gaussian random force which is introduced in an ad hoc manner to simulate the effect of those degrees of freedom of the system which are not explicitly taken into account in the equations of motion and whose behavior is too complicated to be analyzed in detail. Dissipative terms also usually appear as a result of this "modeling away" of degrees of freedom. As a prototype of such equations we shall consider here the Bernoulli oscillator driven by white noise, the equation for which may be written

$$\dot{X}(\tau) + \beta X(\tau) + \lambda X^3(\tau) = e(\tau) \tag{1}$$

[1] Department of Physics, University College of Swansea, Singleton Park, Swansea, SA2 8PP, United Kingdom.

0022-4715/86/0300-1019$05.00/0

where

$$\langle e(\tau)\rangle = 0$$

$$\langle e(\tau)\, e(\tau')\rangle = \delta(\tau - \tau')$$

with expectation values denoted by angular brackets. The oscillator is assumed to evolve from a sharp initial state given by $X(0) = x_0$ and eventually attains a statistically stationary state as $\tau \to \infty$. Accurate values of the time-dependent moments $\langle X(\tau)\rangle$,..., $\langle X^4(\tau)\rangle$ have been calculated by Eaves and Reinhardt[1] for particular values of β, λ, and x_0 and their data will be used in assessing the accuracy of the approximations considered here. The ultimate object of course is to develop methods which can be applied to systems with many degrees of freedom without the computation involved becoming prohibitively lengthy.

If we exclude methods based on the Fokker–Planck equation, which appear to be impracticable for systems with several degrees of freedom, most existing approximation procedures for such equations may be conveniently divided into two categories which will be referred to as "decoupling" and "direct" methods. Decoupling methods[2] seek to derive approximate relationships between averaged quantities, such as correlation and response functions or reduced probability density functions, by truncating the infinite hierarchy of equations which they satisfy. This requires the introduction of some sort of decoupling procedure and a variety of methods, not all convincing, exists for this purpose. Although some simple approximations obtained in this way, such as the direct interaction approximation, have proved quite effective in certain circumstances, more elaborate decouplings have so far failed to produce any useful approximations. We shall not be concerned here with this sort of procedure.

Direct methods on the other hand seek to approximate the stochastic quantities themselves, such as the solution $X(\tau)$ of Eq. (1). The averaged quantities such as moments and correlation functions are then calculated by performing the averaging over the realizations of the random force $e(\tau)$. Some examples of direct methods are (i) random Taylor series[3] in which $X(\tau)$ is approximated by the first few terms of a power series in τ; (ii) perturbation theory,[2] where $X(\tau)$ is approximated by a truncated power series in the nonlinearity parameter λ; (iii) the Wiener–Hermite method,[4] which may be obtained by approximating $X(\tau)$ by a functional polynomial in e, this being chosen so that the approximation is good in some average sense; this method includes statistical linearization[1] as a particular case; (iv) direct numerical simulation.[5]

The last two of these methods will be of most relevance to the work

reported here. An advantage of direct methods is that realizability conditions are not violated, a virtue which seems difficult to attain with decoupling schemes.

In the first three of the direct approximation procedures mentioned above, the averaging over realizations of the random force e can be carried out analytically using the familiar properties of Gaussian random functions which enable higher-order correlation functions to be expressed in terms of the second-order one $\langle e(\tau)\, e(\tau')\rangle$. With direct numerical simulation this is of course impossible and some way must be found of representing e approximately in terms of a finite number of random variables which can be averaged over numerically. For example, if we imagine the random force e to be approximated by a finite sum of terms of the form

$$e(\tau) \approx \sum_{j=1}^{N} \phi_j \xi_j(\tau)$$

where $\xi_j(\tau)$ form an orthonormal set of nonrandom functions, then the averaging would be accomplished by multiple integration over the ϕ_j with a Gaussian weight function

$$\frac{1}{(2\pi)^{N/2}}\, e^{-(1/2)\phi^2}$$

where ϕ^2 denotes $\sum_{j=1}^{N} \phi_j^2$. Clearly white noise has to be treated as a limiting case of true random functions. If the functions $\xi_j(\tau)$ can be chosen so that a small number of terms in the sum provides an adequate approximation for e, then the integration over the ϕ_j can be performed by using a rule for numerical integration such as one of the Newton–Cotes or Gaussian rules. On the other hand, if a large number of terms is necessary, then the integration can be carried out only by a random (or pseudorandom) sampling technique and the method amounts to a Monte Carlo calculation. The number of realizations which must be sampled if the errors are to be acceptably small is often very large, unless some sort of variance reduction scheme can be employed, so that such calculations tend to be rather time consuming. It seems worthwhile to examine the question of whether the functions $\xi_j(\tau)$ can be chosen so as to minimize the number N of terms required in the sum to approximate e.

The usual approach to this type of problem involves the adoption of a least-squares criterion in which one seeks to minimize a mean-square error

$$\left\langle \int d\tau\, w(\tau) \left[e(\tau) - \sum_{j=1}^{N} \phi_j \xi_j(\tau) \right]^2 \right\rangle$$

with respect to the functions ξ_j and the random variables ϕ_j, where $w(\tau)$ is a suitable weight function. If e is Gaussian with zero mean and correlation function $R(\tau, \tau')$ then the optimum ξ_j are eigenfunctions satisfying the integral equation

$$\lambda_j \xi_j(\tau) = \int d\tau' \, R(\tau, \tau') \, w(\tau') \, \xi_j(\tau')$$

a result which is clearly not directly applicable to the white noise case. If the weight function w is chosen as unity then the series $\sum \phi_j \xi_j(\tau)$ becomes the Karhunen–Loève expansion. It should be remembered however that our aim is to find the best approximation for X and the weight function should presumably be chosen to reflect this fact. For example, if we wish to calculate $X(t)$ for a particular value t, then any information about $e(\tau)$ for τ outside the interval $(0, t)$ is irrelevant. Any effort expended in seeking to optimize the approximation for e outside $(0, t)$ would be wasted and so clearly w should be chosen to vanish outside this interval. The appropriate form to choose for w inside the interval is rather more difficult to ascertain since it depends on the way in which errors in $X(t)$ arise from errors in approximating $e(\tau)$ for τ in $(0, t)$. This in turn depends on the particular equation of motion involved. We shall see later that a variational formulation of the problem resolves this question in a natural way.

Assuming for the moment that the random force e has been represented approximately by the finite series $\sum \phi_j \xi_j(\tau)$, the next step would apparently be the numerical solution of the equation

$$\dot{X} + \beta X + \lambda X^3 = \sum_{j=1}^{N} \phi_j \xi_j(\tau)$$

followed by integration with Gaussian measure over the ϕ_j to calculate the mean values of interest. The variational method however shows that this is just the simplest of a series of approximation procedures, and numerical calculations indicate that it is the least effective.

In Section 2 the variational principle will be introduced and used, with suitable trial functions, to generate approximations for the problem under consideration. In Section 3 the results of some numerical computations based on these approximations will be presented and a comparison made with the exact data available for the Bernoulli oscillator. In Section 4 we briefly describe some further ideas based on the variational approach.

2. THE VARIATIONAL PRINCIPLE

The first problem which must be faced in seeking to provide a variational formulation of the sort of stochastic nonlinear equation con-

sidered here arises from the presence of dissipative terms. It is well known that, when such terms are present in the equation of motion, it is impossible to devise a useful functional of X alone which takes on a stationary value when X is a solution of the equation. There are several ways out of this difficulty, and some alternative ones are described briefly in the appendix, but the one which seems most convenient for our purposes is a stochastic version of that described, somewhat incompletely, by Morse and Feshbach[6] which makes use of a functional of X and a "conjugate" variable Y. It will be convenient in developing the formal theory to consider a rather more general equation of motion

$$\dot{X}(\tau) = \Lambda(X(\tau)) + e(\tau) \tag{2}$$

where Λ is a polynomial function of X.

Consider then the quantity $\mathscr{I}$ defined by

$$\mathscr{I} = \left\langle X(t) - \int_0^t d\tau\ Y(\tau)[\dot{X}(\tau) - \Lambda(X(\tau)) - e(\tau)] \right\rangle \tag{3}$$

where $X(\tau)$, $Y(\tau)$ are arbitrary random functions defined on the interval $(0, t)$ and satisfying the two end point conditions

$$\begin{aligned} X(0) &= x_0 \\ Y(t) &= 1 \end{aligned} \qquad \text{(where } x_0 \text{ is the given initial value)}$$

The variation of $\mathscr{I}$ for arbitrary variations of X and Y can be written, after integration by parts in the usual way, in the form

$$\delta\mathscr{I} = \left\langle -\int_0^t d\tau\ \delta Y(\tau)[\dot{X}(\tau) - \Lambda(X(\tau)) - e(\tau)] + \int_0^t d\tau\ \delta X(\tau)[\dot{Y}(\tau) + Y(\tau)\,\Lambda'(X(\tau))] \right\rangle$$

Thus $\delta\mathscr{I} = 0$ for variations of X and Y about the solutions of the equations

$$\begin{aligned} \dot{X}(\tau) &= \Lambda(X(\tau)) + e(\tau), & X(0) &= x_0 \\ \dot{Y}(\tau) &= -Y(\tau)\,\Lambda'(X(\tau)), & Y(t) &= 1 \end{aligned} \tag{4}$$

and conversely, if $\delta\mathscr{I} = 0$ for arbitrary variations δX, δY, then X and Y satisfy these equations for almost all realizations. Moreover the stationary value $\mathscr{I}_{\text{stat}}$ is seen to be $\langle X(t) \rangle$.

Approximations for $\langle X(t)\rangle$ are obtained by seeking a stationary value of $\mathscr{I}$ when X and Y vary over a restricted set of random functions for which $\mathscr{I}$ can be evaluated. It will be assumed that such a restricted stationary value gives a reasonable approximation to the true one if the functions X and Y are of a sufficiently general form. The variational principle can also be looked at in a slightly different way as follows: If Eqs. (4) can be solved for X and Y to an accuracy of order ε for the vast majority of realizations, then the evaluation of $\mathscr{I}$ gives an approximation for $\langle X(t)\rangle$ which is accurate to order ε^2. Although the variational principle is not an extremal one it can nevertheless provide a useful means of generating approximations. A similar situation is found in other branches of physics such as quantum scattering theory.[7]

Other moments can be dealt with in a similar way. For example, if we define the quantity

$$J=\left\langle X^n(t)-n\int_0^t d\tau\ Y(\tau)[\dot{X}(\tau)-A(X(\tau))-e(\tau)]\right\rangle$$

for random functions X and Y satisfying the conditions

$$X(0)=x_0$$
$$Y(t)=X^{n-1}(t)$$

then the stationary value of J, attained when X and Y satisfy the same equations as before but with $Y(t)=X^{n-1}(t)$, is $\langle X^n(t)\rangle$. For some trial functions it may not be possible to impose the condition $Y(t)=X^{n-1}(t)$. However it is sufficient to require that

$$\langle \delta X(t)[Y(t)-X^{n-1}(t)]\rangle=0$$

In their discussion of this variational method, in its nonstochastic form, Morse and Feshbach express the view that it is unlikely to prove to be more than a formal mathematical construction since the conjugate variable Y satisfies an equation with negative dissipation if a dissipative term appears in the equation of motion for X. However, they fail to take account of the fact that the direction of time is essentially reversed in the equation for Y since it satisfies a final rather than an initial condition. We maintain on the contrary that the quantity $Y(\tau)$ is of direct physical significance since it describes the response of the system to small changes in the applied force. This can easily be verified by writing down the equation for the variation δX arising from a variation δe,

$$\delta\dot{X}(\tau)=\delta X(\tau)\,A'(X(\tau))+\delta e(\tau)$$

If we multiply this by $Y(\tau)$ and then add to it the equation obtained by multiplying by $\delta X(\tau)$ the equation for $Y(\tau)$ we get

$$\frac{d}{d\tau}\left[Y(\tau)\,\delta X(\tau)\right] = Y(\tau)\,\delta e(\tau)$$

Integrating over the time interval $(0, t)$ and using the fact that $\delta X(0) = 0$, $Y(t) = 1$, we obtain

$$\delta X(t) = \int_0^t d\tau\; Y(\tau)\,\delta e(\tau)$$

so that $Y(\tau)$ determines the linear response of the system. It is interesting to note in passing that the conjugate variables introduced in this variational formulation of the problem satisfy the same equations of motion as do the conjugate operators introduced by Martin, Siggia, and Rose in their operator formalism of classical statistical dynamics.[8,9]

In order to make use of the variational method a choice of trial functions must be made and this will be done in such a way as to throw some light on the questions raised in the Introduction. In a numerical approach to the problem the random force $e(\tau)$ would be represented approximately by a finite number of quantities which might be, for example, the expansion coefficients mentioned earlier or the values assumed by $e(\tau)$ on a finite set of τ values. The quantities chosen for this purpose would normally be linear functionals of e. If the equation for X could be solved exactly, with e approximated in this way, then the solution would be a function of these same linear functionals. We are therefore led to investigate trial functions for X and Y of the form

$$\begin{aligned} X(\tau) &= F(\tau, \Phi) \\ Y(\tau) &= G(\tau, \Phi) \end{aligned} \tag{5}$$

where F and G are arbitrary functions and Φ denotes $(\Phi_1, \ldots, \Phi_N)$, a set of N linearly independent linear functionals of e on the time interval $(0, t)$ of interest. We can write

$$\Phi_j = \int_0^t d\tau\; \xi_j(\tau)\, e(\tau) \tag{6}$$

where the ξ_j may be ordinary or generalized functions. We shall often refer to these as basis functions. For the moment we shall regard the Φ_j as fixed so that, in the variational principle, only the form of the functions F and G is to be varied. The Φ_j will be referred to as reserved functionals for reasons

which will become apparent later. Since a function of the Φ_j can be rewritten as a function of N linearly independent linear combinations it is clear that, without loss of generality, we may assume the Φ_j to be mutually orthogonal. If it is assumed that, for no j do we have $\langle \Phi_j^2 \rangle = 0$, then the functionals may be normalized so that

$$\langle \Phi_j \Phi_l \rangle = \delta_{jl} \tag{7}$$

and hence the basis functions are orthonormal on $(0, t)$

$$\int_0^t d\tau\, \xi_j(\tau)\, \xi_l(\tau) = \delta_{jl} \tag{8}$$

A functional Φ_j such that $\langle \Phi_j^2 \rangle = 0$ would be zero for almost all realizations and would therefore play no useful role if included in the trial functions.

Substituting these trial functions into the expression for $\mathscr{I}$ we have

$$\mathscr{I} = \int d\phi\, \langle \delta(\phi - \Phi) \rangle \left\{ F(t, \phi) - \int_0^t d\tau\, G(\tau, \phi) \left[\dot{F}(\tau, \phi) - \Lambda(F(\tau, \phi)) - \frac{\langle e(\tau)\, \delta(\phi - \Phi) \rangle}{\langle \delta(\phi - \Phi) \rangle} \right] \right\}$$

where $\int d\phi$ denotes the N-fold integration $\int_{-\infty}^{\infty} d\phi_1 \cdots \int_{-\infty}^{\infty} d\phi_N$ and $\delta(\phi - \Phi) = \delta(\phi_1 - \Phi_1) \cdots \delta(\phi_N - \Phi_N)$.

Because of the orthonormality of the Φ_j we have

$$\langle \delta(\phi - \Phi) \rangle = \frac{1}{(2\pi)^{N/2}} e^{-(1/2)\phi^2}$$

with $\phi^2 = \phi_j \phi_j$, where a summation over repeated indices is implied here and henceforth. The ratio

$$\frac{\langle e(\tau)\, \delta(\phi - \Phi) \rangle}{\langle \delta(\phi - \Phi) \rangle} = \langle e(\tau) \rangle_\phi$$

is the conditional expectation value of $e(\tau)$ given that $\Phi_j = \phi_j$ for $j = 1, \ldots, N$. Its value may easily be found by using the Fourier representation of the δ function and the expression for the characteristic functional of a Gaussian random function. We find that

$$\langle e(\tau) \rangle_\phi = \phi_j \xi_j(\tau)$$

Thus

$$\mathscr{I} = \frac{1}{(2\pi)^{N/2}} \int d\phi\, e^{-(1/2)\phi^2} \left\{ F(t, \phi) - \int_0^t d\tau\, G(\tau, \phi)[\dot{F}(\tau, \phi) - \Lambda(F(\tau, \phi)) - \phi_j \xi_j(\tau)] \right\} \tag{9}$$

and the variation $\delta\mathscr{I}$, for variations δF, δG which are arbitrary apart from the conditions $\delta F(0, \phi) = 0$, $\delta G(t, \phi) = 0$, is given by

$$\delta\mathscr{I} = \frac{1}{(2\pi)^{N/2}} \int d\phi\, e^{-(1/2)\phi^2} \left\{ - \int_0^t d\tau\, \delta G(\tau, \phi)[\dot{F} - \Lambda(F) - \phi_j \xi_j(\tau)] + \int_0^t d\tau\, \delta F(\tau, \phi)[\dot{G} + G\Lambda'(F)] \right\}$$

The condition for a stationary point is therefore

$$\begin{aligned} \dot{F} &= \Lambda(F) + \phi_j \xi_j(\tau) \\ \dot{G} &= -G\Lambda'(F) \end{aligned} \tag{10}$$

and the stationary value, which is our approximation for $\langle X(t) \rangle$, is

$$\mathscr{I}_{\text{stat}} = \frac{1}{(2\pi)^{N/2}} \int d\phi\, e^{-(1/2)\phi^2} F(t, \phi) \tag{11}$$

Note that the quantity G need not be determined for the evaluation of $\langle X(t) \rangle$ within this approximation. This is the first of the approximation procedures to be examined in more detail later. It may easily be verified that it becomes exact in the limit as the random forcing term becomes zero.

It should be noted that, for the calculation of $\langle X(t) \rangle$ the basis functions $\xi_j(\tau)$ form an orthonormal set on the interval $(0, t)$ and so this set of functions should be chosen according to the value of t under consideration. If one avoids this complication by using the same set of basis functions for all values of t the resulting approximations are found to be slightly less accurate. This is not surprising since any information about $e(\tau)$ outside the interval $(0, t)$ is irrelevant for the evaluation of $\langle X(t) \rangle$. However there is a considerable reduction in the computation involved.

The approximation procedure derived above may be described in the following terms:

(i) The set of all realizations of the random force e is divided into subsets according to the values of the functionals Φ_j.

(ii) In each such subset X is approximated by a "constant," i.e., a quantity independent of e, this constant being taken as the solution of the equation of motion with e replaced by its conditional expectation value.

(iii) Finally $\langle X \rangle$ is calculated by performing the Gaussian averaging over the variables ϕ_j.

The functionals Φ_j are therefore seen to play a similar role to that of the reserved variables in Feynman's approximate treatment of functional integrals.[10] We shall call this approximation the piecewise mean value approximation or PMVA for short. The approximation could be obtained without making use of the variational method but the virtue of our derivation is that it leads in a natural way to improvements on the basic idea as we shall now see.

There is an obvious analogy between the treatment here of the dependence of X on the random force e and the approximate representation of an ordinary function of a single variable by means of a step function as might be used, for example, in deriving the rectangle rule for numerical quadrature. This analogy suggests two obvious ways in which the approximation might be improved:

(i) The step function could be replaced by a polygonal line function so that, in each subinterval, the function is approximated linearly.

(ii) The subdivision could be adapted to the function in question with small subintervals where the function is varying rapidly and larger ones where the variation is slower.

These ideas can readily be incorporated into the trial functions for the problem under consideration. Taking (i) first it is clear that the appropriate trial functions are

$$X(\tau) = \tilde{k}(\tau, \Phi) + \int_0^t d\tau' \, \tilde{K}(\tau, \tau', \Phi) \, e(\tau')$$

$$Y(\tau) = \tilde{l}(\tau, \Phi) + \int_0^t d\tau' \, \tilde{L}(\tau, \tau', \Phi) \, e(\tau')$$

where $\tilde{k}$, $\tilde{K}$, $\tilde{l}$, $\tilde{L}$ are arbitrary functions. In each subset of realizations, characterized by definite values ϕ_j of the reserved functionals Φ_j, X and Y are linear functionals of e. It is more convenient to separate off the conditional mean value of e by writing

$$e(\tau) = \xi_j(\tau) \, \Phi_j + \hat{e}(\tau)$$

so that

$$\langle \hat{e}(\tau) \rangle_\phi = 0$$

$$\langle \hat{e}(\tau)\, \hat{e}(\tau') \rangle_\phi = r(\tau, \tau') = \delta(\tau - \tau') - \xi_j(\tau)\, \xi_j(\tau')$$

The functions X and Y may then be rewritten in the form

$$\begin{aligned} X(\tau) &= k(\tau, \Phi) + \int_0^t d\tau'\, K(\tau, \tau', \Phi)\, \hat{e}(\tau') \\ Y(\tau) &= l(\tau, \Phi) + \int_0^t d\tau'\, L(\tau, \tau', \Phi)\, \hat{e}(\tau') \end{aligned} \tag{12}$$

Substituting in the expression for $\mathscr{I}$ gives for the Bernoulli oscillator

$$\begin{aligned} \mathscr{I} = \frac{1}{(2\pi)^{N/2}} \int d\phi\, e^{-(1/2)\phi^2} \Bigg\{ & k(t, \phi) \\ & - \int_0^t d\tau\, l(\tau, \phi) \Bigg[\dot{k}(\tau, \phi) + \beta k(\tau, \phi) + \lambda k^3(\tau, \phi) \\ & + 3\lambda k(\tau, \phi) \int_0^t d\tau' \int_0^t d\tau''\, K(\tau, \tau', \phi)\, K(\tau, \tau'', \phi)\, r(\tau', \tau'') - \xi_j(\tau)\, \phi_j \Bigg] \\ & - \int_0^t d\tau \int_0^t d\tau'\, L(\tau, \tau', \phi) \Bigg[\int_0^t d\tau''\, \dot{K}(\tau, \tau'', \phi)\, r(\tau', \tau'') \\ & + \beta \int_0^t d\tau''\, K(\tau, \tau'', \phi)\, r(\tau', \tau'') + 3\lambda k^2(\tau, \phi) \int_0^t d\tau''\, K(\tau, \tau'', \phi)\, r(\tau', \tau'') \\ & + 3\lambda \int_0^t d\tau'' \int_0^t d\tau_2 \int_0^t d\tau_3\, K(\tau, \tau'', \phi)\, K(\tau, \tau_2, \phi)\, K(\tau, \tau_3, \phi) \\ & \times r(\tau', \tau'')\, r(\tau_2, \tau_3) - \int_0^t d\tau''\, r(\tau', \tau'')\, \delta(\tau - \tau'') \Bigg] \Bigg\} \end{aligned}$$

Equating $\delta\mathscr{I}$ to zero gives equations for k, K, l, L, but, since we do not need l and L for the calculation of the stationary value, we present only the equations for k and K. These are

$$\begin{aligned} & \dot{k}(\tau) + \beta k(\tau) + \lambda k^3(\tau) + 3\lambda k(\tau) \int_0^t d\tau_1 \int_0^t d\tau_2 \\ & \qquad \times K(\tau, \tau_1)\, K(\tau, \tau_2)\, r(\tau_1, \tau_2) = \xi_j(\tau)\, \phi_j \\ & \dot{K}(\tau, \tau') + \Bigg[\beta + 3\lambda k^2(\tau) + 3\lambda \int_0^t d\tau_1 \int_0^t d\tau_2 \\ & \qquad \times K(\tau, \tau_1)\, K(\tau, \tau_2)\, r(\tau_1, \tau_2) \Bigg] K(\tau, \tau') = \delta(\tau - \tau') \end{aligned}$$

If we define

$$p(\tau) = \int_0^t d\tau_1 \int_0^t d\tau_2 \, K(\tau, \tau_1) \, K(\tau, \tau_2) \, r(\tau_1, \tau_2)$$

$$s(\tau) = \int_0^t d\tau' \, [\beta + 3\lambda k^2(\tau') + 3\lambda p(\tau')]$$

it is easily verified that

$$K(\tau, \tau') = \theta(\tau - \tau') \, e^{-[s(\tau) - s(\tau')]}$$

where $\theta(\tau)$ is the step function

$$\theta(\tau) = \begin{cases} 1, & \tau > 0 \\ 0, & \tau < 0 \end{cases}$$

The equations can then be rewritten in the more convenient form,

$$\begin{gathered} \dot{k} + \beta k + \lambda k^3 + 3\lambda k p = \xi_j(\tau) \, \phi_j \\ \dot{s} = \beta + 3\lambda k^2 + 3\lambda p \\ p(\tau) = e^{-2s(\tau)} \left\{ \int_0^\tau d\tau' \, e^{2s(\tau')} - \sum_j \left[\int_0^\tau d\tau' \, \xi_j(\tau') \, e^{s(\tau')} \right]^2 \right\} \end{gathered} \tag{13}$$

with $k(0) = x_0$, $s(0) = 0$ and the approximation for $\langle X(t) \rangle$ is

$$\frac{1}{(2\pi)^{N/2}} \int d\phi \, e^{-(1/2)\phi^2} k(t, \phi)$$

This is the second approximation to be examined later and may be described as a piecewise statistical linearization approximation (PSLA). It may easily be verified that the approximation is exact when there is no random force or when the nonlinear term is absent from the equation of motion. For the case when $N = 0$ it is seen that one recovers the simplest nontrivial Wiener–Hermite approximation.[4] This approximation, which we shall refer to as statistical linearization, has been "rediscovered" recently and applied to the Bernoulli oscillator and other simple systems by several authors.[11–13]

It is not difficult to generate further approximations of this type by utilizing for X and Y trial functions of the form

$$k_0(\tau, \Phi) + \int_0^t d\tau_1\, k_1(\tau, \tau_1, \Phi)\, e(\tau_1)$$

$$+ \int_0^t d\tau_1 \int_0^t d\tau_2\, k_2(\tau, \tau_1, \tau_2, \Phi)\, e(\tau_1)\, e(\tau_2) + \cdots$$

$$+ \int_0^t d\tau_1 \cdots \int_0^t d\tau_n\, k_n(\tau, \tau_1, \ldots, \tau_n, \Phi)\, e(\tau_1) \cdots e(\tau_n)$$

which may be rewritten in terms of Wiener–Hermite functions.[4] However the proliferation of multivariable functions leads to a rapid increase of complexity with n. Instead of pursuing this further we now turn to the consideration of the second way of generalizing the trial functions.

It was noted earlier that the division of the set of all realizations of the random force e into subsets, which features in the approximations considered above, is determined by the choice of the reserved linear functionals Φ_j. It is apparent therefore that it would be advantageous to choose the Φ_j so as to optimize the approximation. This is readily achieved within the variational principle since it is merely necessary to allow variations in the Φ_j as well as in the form of the functions F and G in the trial functions $F(\tau, \Phi)$, $G(\tau, \Phi)$. This means that the basis functions ξ_j are subject to variation in the variational principle. It will be convenient to assume as before that the reserved functionals form an orthonormal set so that the functions ξ_j satisfy the $N(N+1)/2$ subsidiary conditions

$$\int_0^t d\tau\, \xi_j(\tau)\, \xi_l(\tau) = \delta_{jl}$$

It seems reasonable to assume that these can be imposed since a stationary point for which any of the ξ_j are null functions is unlikely to be of interest.

We therefore seek the stationary value of

$$\mathscr{I} = \frac{1}{(2\pi)^{N/2}} \int d\phi\, e^{-(1/2)\phi^2} \left\{ F(\tau, \phi) \right.$$

$$\left. - \int_0^t d\tau\, G(\tau, \phi)[\dot{F}(\tau, \phi) - \Lambda(F(\tau, \phi)) - \phi_j \xi_j(\tau)] \right\}$$

with respect to variations of F, G and the ξ_j with these subject to the above subsidiary conditions. Introducing the symmetric matrix γ of Lagrange multipliers we consider the equation

$$\delta \left[\mathscr{I} - \frac{1}{2} \gamma_{jl} \int_0^t d\tau\, \xi_j(\tau)\, \xi_l(\tau) \right] = 0$$

Equating to zero the terms in δG and δF we recover the same equations for F and G obtained before, i.e., Eqs. (10). The term involving the $\delta\xi_j$ is

$$\int_0^t d\tau\, \delta\xi_j(\tau)\left[\frac{1}{(2\pi)^{N/2}}\int d\phi\, e^{-(1/2)\phi^2}\phi_j G(\tau,\phi)-\gamma_{jl}\xi_l(\tau)\right]$$

and so we obtain the additional equation

$$\frac{1}{(2\pi)^{N/2}}\int d\phi\, e^{-(1/2)\phi^2}\phi_j G(\tau,\phi)=\gamma_{jl}\xi_l(\tau)$$

If we define

$$b_j(\tau)=\frac{1}{(2\pi)^{N/2}}\int d\phi\, e^{-(1/2)\phi^2}\phi_j G(\tau,\phi) \tag{14}$$

and the symmetric matrix C with elements given by

$$C_{jl}=\int_0^t d\tau\, b_j(\tau)\, b_l(\tau) \tag{15}$$

it is seen that $\gamma=C^{1/2}$ so that

$$\xi_j(\tau)=(C^{-1/2})_{jl}\, b_l(\tau) \tag{16}$$

These last three equations together with equations (10) for F and G and the expression (11) for the stationary value $\mathscr{I}_{\text{stat}}$ fully describe the approximation. It will be observed that the quantity G now plays an essential role in the equations thus confirming the qualitative observation which was made in the Introduction, namely, that the way in which the random force e is approximated by means of a finite set of random variables should take into account the way in which the resulting errors give rise to errors in X itself.

It is easily verified that if F, G, b, C, ξ_j satisfy these equations then so too do the transformed quantities $\tilde{F}$, $\tilde{G}$, $\tilde{b}$, $\tilde{C}$, $\tilde{\xi}_j$ defined as follows:

$$\begin{aligned}
\tilde{F}(\tau,\phi)&=F(\tau,M^{-1}\phi)\\
\tilde{G}(\tau,\phi)&=G(\tau,M^{-1}\phi)\\
\tilde{b}_j(\tau)&=M_{jl}b_l(\tau)\\
\tilde{C}&=MCM^{-1}\\
\tilde{\xi}_j(\tau)&=M_{jl}\xi_l(\tau)
\end{aligned}$$

where M is an orthogonal matrix. This is not surprising since the actual trial functions for X and Y are unchanged by this transformation. We have, for example, $\tilde{F}(\tau, \tilde{\Phi}) = F(\tau, \Phi)$ where the transformed functionals $\tilde{\Phi}_j$ are calculated with the $\tilde{\xi}_j$. This freedom could be used to ensure that the symmetric matrix C has a diagonal form.

This in the third of the approximations to be investigated numerically in the next section and will be referred to as the optimized piecewise mean value approximation (OPMVA). An alternative formulation of this approximation, and one which is likely to prove more convenient in practice, can be obtained by noting that the stationary value $\mathscr{I}_{\text{stat}}$ of $\mathscr{I}$ with respect to variations of F, G, and the ξ_j can be calculated by first seeking the conditional stationary value $\mathscr{I}^{(c)}_{\text{stat}}[\xi]$ with respect to F and G with the ξ_j held fixed. The quantity $\mathscr{I}_{\text{stat}}$ can then be determined by finding the stationary value of $\mathscr{I}^{(c)}_{\text{stat}}[\xi]$ with respect to variations in the ξ_j. This could be achieved approximately by choosing for the ξ_j suitable trial functions of a particular form containing a number of adjustable parameters. These parameters would then be varied to obtain a stationary value. This approach would be particularly advantageous for optimizing the PSLA with respect to the basis functions since in that case the full equations are rather complicated.

Our discussion so far has concentrated on the mean value $\langle X(t) \rangle$ but it is possible to generate approximations for the higher moments by using the same sorts of trial functions in the variational expressions for $\langle X^n(t) \rangle$. When this is done it is found that, for the type of trial function which yields the PMVA, we have

$$\langle X^n(t) \rangle = \frac{1}{(2\pi)^{N/2}} \int d\phi \, e^{-(1/2)\phi^2} F^n(t, \phi)$$

where F satisfies the same equation as before, i.e., Eq. (10). Using the trial function which gives the PSLA we obtain

$$\langle X^2(t) \rangle = \frac{1}{(2\pi)^{N/2}} \int d\phi \, e^{-(1/2)\phi^2} (k^2 + p)$$

$$\langle X^3(t) \rangle = \frac{1}{(2\pi)^{N/2}} \int d\phi \, e^{-(1/2)\phi^2} (k^3 + 3kp)$$

$$\langle X^4(t) \rangle = \frac{1}{(2\pi)^{N/2}} \int d\phi \, e^{-(1/2)\phi^2} (k^4 + 6k^2 p + 3p^2)$$

where k and p are determined by solving Eqs. (13). Thus the same set of equations has to be solved for calculating all the moments. This is no

longer the case for the corresponding optimized approximations since the boundary conditions differ according to the particular moment being calculated. Thus, for example, the optimum basis functions for the calculation of $\langle X^2(t)\rangle$ will be different from those involved in the calculation of $\langle X(t)\rangle$. For this reason we shall confine ourselves to the consideration of the mean value $\langle X(t)\rangle$ for the optimized approximations. It should also be noted that the satisfaction of realizability inequalities relating different moments cannot be guaranteed.

3. CALCULATIONS FOR THE BERNOULLI OSCILLATOR

Values of the first four moments $\langle X(t)\rangle$,..., $\langle X^4(t)\rangle$ have been calculated by Eaves and Reinhardt[1] by using a high-order truncation of the coupled equations for the moments. Although this method appears intractable for systems with several degrees of freedom and cannot be applied to systems driven by a random force which is not white noise it yields accurate values for the problem under consideration. These provide a valuable means of checking other approximation procedures, such as those described here, and we shall therefore limit our considerations to those values of the parameters of the problem used by these authors: namely, $\beta = \lambda = 1$, $x_0 = 2$, with $0 \leqslant t \leqslant 4$.

For the PMVA and PSLA it is necessary to choose an orthonormal set of basis functions for each value of t considered. We shall choose these to be polynomial functions of τ. When orthonormalized on the interval $(0, t)$ they may be expressed in terms of Legendre polynomials P as follows:

$$\xi_j(\tau) = \left(\frac{2j+1}{t}\right)^{1/2} P_j\left(\frac{2\tau}{t} - 1\right), \qquad j = 0, 1, 2, \ldots, N-1$$

where, for convenience, we now label from $j = 0$. The PMVA values for the first four moments have been evaluated for $N = 1, 2, 3$ using a fourth-order Runge–Kutta method for the solution of the differential equation and Simpson's rule for the numerical integration. The percentage errors in the values obtained for $\langle X(t)\rangle$ are shown graphically in Fig. 1 and the corresponding results for the other three moments in Figs. 2–4.

It will be seen that, in general, the percentage errors become rather large as t increases especially for the odd moments. It should be remembered however that the odd moments tend to zero for large t so that the absolute errors are not necessarily large. Also it is not apparent that the approximations necessarily improve as the number N of reserved variables is increased. It seems intuitively obvious however that the approximation would become exact in the limit as $N \to \infty$. Hence it must be concluded

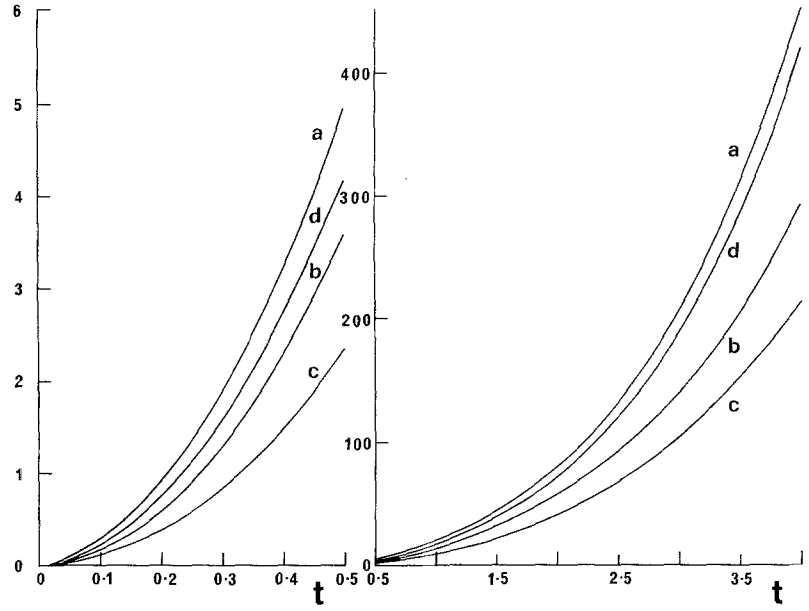

Fig. 1. The percentage errors in the values of $\langle X(t) \rangle$ given by the PMVA and OPMVA. The curves (a), (b), (c) correspond to the PMVA for $N = 1, 2, 3$, respectively, and the curve (d) to the OPMVA.

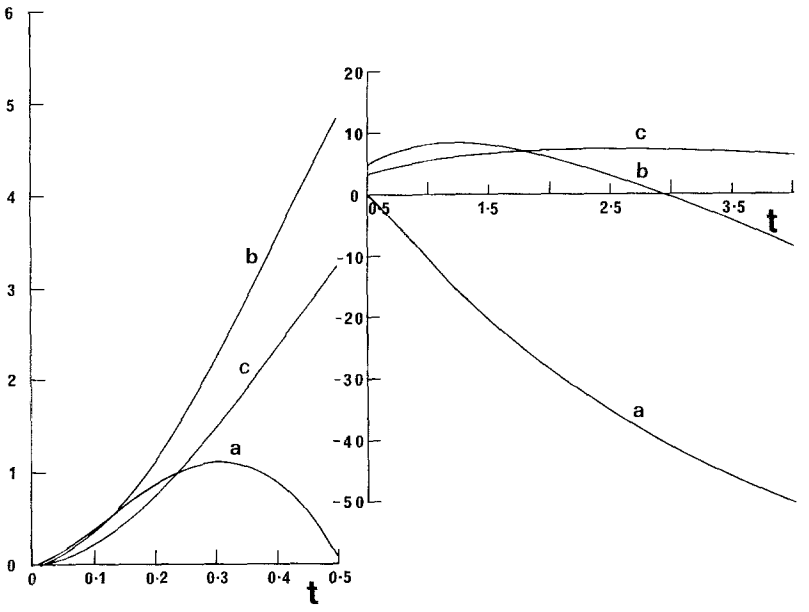

Fig. 2. The percentage errors in the values of $\langle X^2(t) \rangle$ given by the PMVA. The curves (a), (b), (c) correspond to $N = 1, 2, 3$.

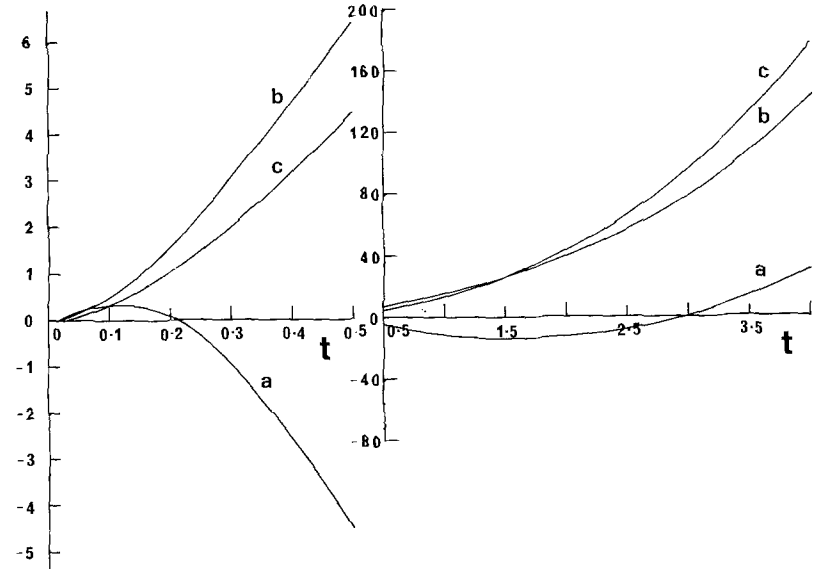

Fig. 3. The percentage errors in the values of $\langle X^3(t)\rangle$ given by the PMVA. The curves (a), (b), (c) correspond to $N = 1, 2, 3$.

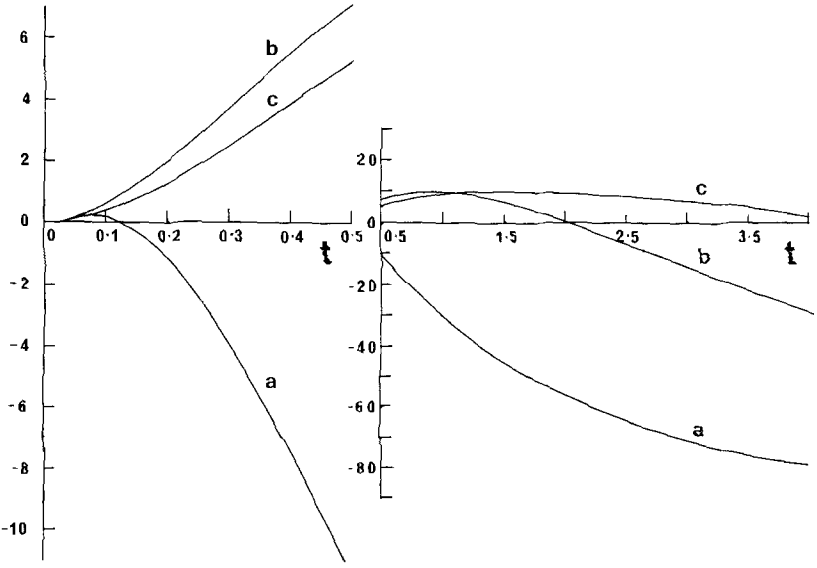

Fig. 4. The percentage errors in the values of $\langle X^4(t)\rangle$ given by the PMVA. The curves (a), (b), (c) correspond to $N = 1, 2, 3$.

that, within this scheme, it is necessary to choose N large and then to evaluate the integrals by the random sampling method. The calculation is then essentially equivalent to a Monte Carlo simulation, different variants of this method being obtained according to the choice of basis functions. Also shown in Fig. 1 are the results obtained by means of the OPMVA. The modest improvement obtained with this approximation suggests strongly that the conclusion reached above is unlikely to be altered no matter how the reserved functionals are chosen. The calculations for the OPMVA were carried out by solving the coupled equations (10), (11), (14), (15), and (16) by an iteration procedure with respect to $\xi(\tau)$. Four iterations were found to be sufficient for the required accuracy when the initial guess for $\xi(\tau)$ was taken as a constant. Other initial guesses gave a similar rapid convergence. The form of the optimum basis function which emerges from these calculations is shown in Fig. 9, which corresponds to $t = 0.6$. A curve of similar shape is obtained for other values of t.

In Figs. 5–8 are shown the percentage errors in the values of the moments calculated by means of the PSLA for $N = 0, 1, 2, 3$ with the same choice of basis functions. It will be observed that the errors are reduced to a more acceptable level for all the moments over the whole range of t values. Moreover, in most cases, the approximations improve as N

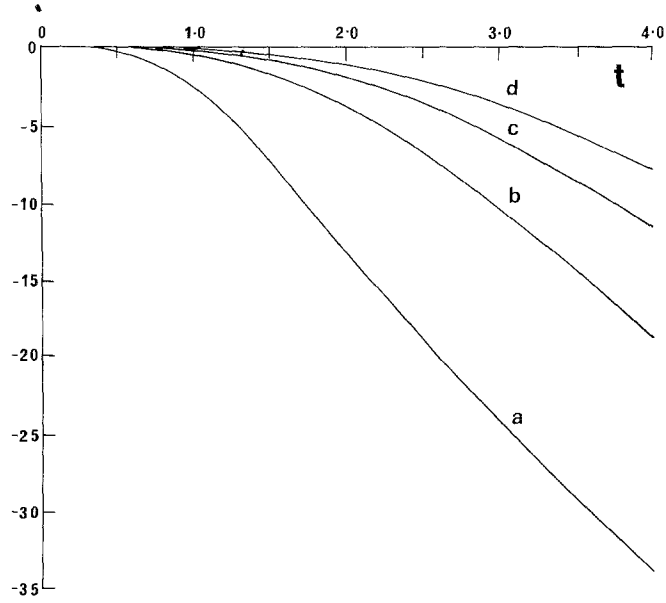

Fig. 5. The percentage errors in the values of $\langle X(t) \rangle$ given by the PSLA. The curves (a), (b), (c), (d) correspond to $N = 0, 1, 2, 3$.

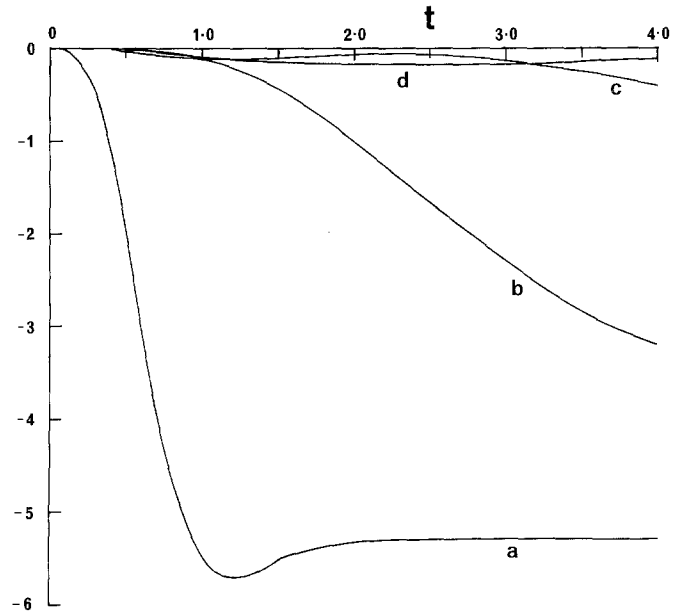

Fig. 6. The percentage errors in the values of $\langle X^2(t) \rangle$ given by the PSLA. The curves (a), (b), (c), (d) correspond to $N = 0, 1, 2, 3$.

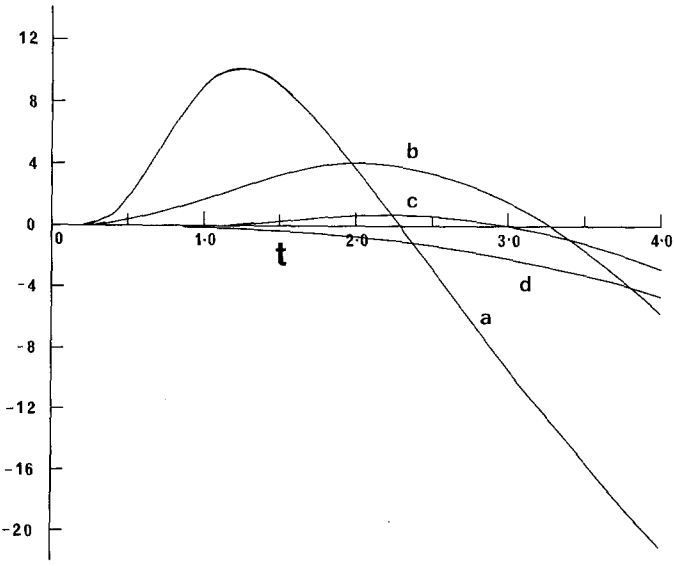

Fig. 7. The percentage errors in the values of $\langle X^3(t) \rangle$ given by the PSLA. The curves (a), (b), (c), (d) correspond to $N = 0, 1, 2, 3$.

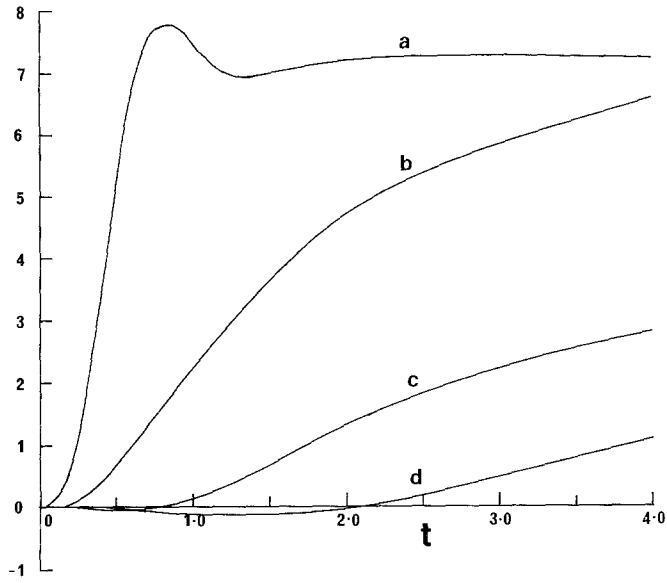

Fig. 8. The percentage errors in the values of $\langle X^4(t) \rangle$ given by the PSLA. The curves (a), (b), (c), (d) correspond to $N = 0, 1, 2, 3$.

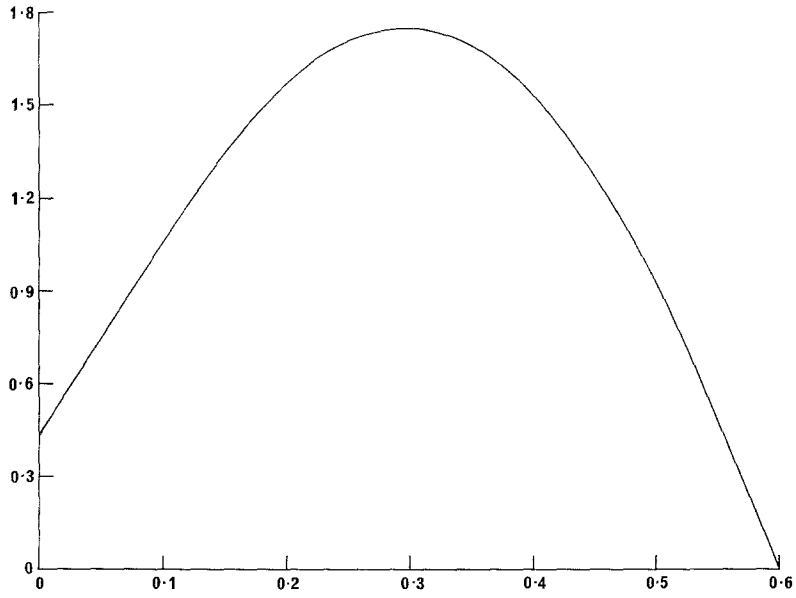

Fig. 9. Optimized basis function for PMVA with $N = 1$ for $t = 0.6$.

increases. The dependence of the results on the choice of basis functions appears not to be very great. Even when these are taken to be independent of t the errors are not increased appreciably and the computation time is considerably reduced. We conclude that, within the PSLA scheme, it is possible to obtain useful results with small values of N so avoiding the difficulties of the random sampling approach.

4. FURTHER CONSIDERATIONS

The trial functions which we have used here were chosen to throw some light on the questions posed in the Introduction and by no means exhaust all the possibilities. An obvious generalization would be obtained by allowing the reserved linear functionals to be time dependent so that the trial functions for X and Y, to generate an analog of the PMVA, would be of the form

$$X(\tau) = F(\tau, \Phi(\tau))$$
$$Y(\tau) = G(\tau, \Phi(\tau))$$

where $\Phi(\tau)$ denotes a set $(\Phi_1(\tau),..., \Phi_N(\tau))$ of time-dependent linear functionals of e which can be written in the form

$$\Phi_j(\tau) = \int d\tau' \, \xi_j(\tau, \tau') \, e(\tau')$$

An attractive feature of such a trial function is that it can be made to reflect the causal relationship between X and e if we choose the $\xi_j(\tau, \tau')$ to vanish for $\tau < \tau'$. Note however that this imposes an incorrect causal relationship between Y and e. Considering for simplicity the case in which the ξ_j are fixed we obtain from the variational principle the equation for F

$$\dot{F} - \Lambda(F) + \frac{\partial F}{\partial \phi_j} A_{jl}(\tau) \, \phi_l = \phi_j \psi_j(\tau)$$

where

$$\psi_j(\tau) = \xi_j(\tau, \tau)$$

and

$$A_{jl}(\tau) = -A_{lj}(\tau) = \int d\tau' \, \dot{\xi}_j(\tau, \tau') \, \xi_l(\tau, \tau')$$

It will be seen that the equation involves partial derivatives if $N > 1$. However these can be removed by replacing the reserved functionals $\Phi_j(\tau)$

by an equivalent set $\tilde{\Phi}_j(\tau)$ obtained by means of a time-dependent orthogonal transformation $M(\tau)$. If M is now chosen to satisfy the equation

$$\dot{M} = -MA$$

which is compatible with the orthogonality of M since A is antisymmetric, then the transformed A vanishes identically. Thus the equation for F can always be reduced to an ordinary differential equation. A preliminary investigation of the resulting approximations indicates some improvement over those obtained with time-independent reserved functionals. Similar considerations can be applied to the PSLA but no calculations have yet been carried out.

Mention has already been made of the possibility of going beyond statistical linearization and considering higher-order piecewise Wiener–Hermite approximations. We could obtain a more restricted, but mathematically simpler, approximation by including in the trial functions terms such as

$$\int_0^t d\tau' \int_0^t d\tau'' \, H(\tau, \tau', \tau'', \Phi)\, e(\tau')\, e(\tau'')$$

where H is a suitably chosen function containing adjustable parameters which would be varied to obtain a stationary value of $\mathscr{I}$. This would remove the complication of having to deal with unknown functions of several variables.

An unfortunate but apparently inescapable feature of systematic approximation procedures for stochastic nonlinear problems is the rapid increase in the amount of computation involved in evaluating the approximations as one progresses from the simplest in the sequence to more and more elaborate ones. Sooner or later the approximations cease to be competitive with direct numerical simulation of the problem. The procedures considered here are no exception but the approximations with one or two reserved functionals are likely to be attractive. To achieve the same order of accuracy as one gets with the PSLA with $N = 1$ by means of direct numerical simulation of the problem one needs to sample many thousands of realizations and the computation time required is hundreds of times greater.

The question then arises of whether these approximations can usefully be applied in other situations. Preliminary calculations have been carried out for diffusion in a double-well potential in which the particle starts from the point of unstable equilibrium. This is a much more stringent test of approximations than the single-well potential problem so that, for example,

ordinary statistical linearization fails completely. However the PSLA with $N=1$ is found to give quite accurate results. The dependence on the choice of basis function is more sensitive so that some optimization with respect to this function is needed. Another problem for which calculations have been carried out, using the PSLA, with promising results is that concerned with the motion of a marked fluid particle in a Gaussian random flow. This work will be described in a future paper.

The higher approximations with $N>2$ can of course be evaluated approximately by using the random sampling technique for the calculation of the N-fold integrals. Indeed, if N is large, such a calculation is essentially a Monte Carlo simulation of the problem. It would be interesting to compare such a simulation based on the PSLA with a more conventional one. However the simulation could be made much more efficient by using one of the simple approximations as the basis of a variance reduction scheme. The method of correlation sampling[14] for the evaluation of the integral $\int_a^b f(x)\,dx$ is based on the simple observation that

$$\int_a^b f(x)\,dx = \int_a^b g(x)\,dx + \int_a^b [f(x)-g(x)]\,dx$$

If g is a function which approximates f in the interval in question, and which is such that the integral $\int_a^b g(x)\,dx$ can be found accurately, then the calculation of the second integral on the right-hand side by the random sampling technique provides a more effective method for approximating $\int_a^b f(x)\,dx$ than the direct evaluation of this integral by random sampling. The reason of course is that the variance of $(f-g)$ is expected to be much less than that of f. To apply this idea to the problem considered here let k_j denote the solution of the PSLA with j reserved functionals. Then we can write

$$\frac{1}{(2\pi)^{N/2}} \int d\phi_1 \cdots \int d\phi_N\, e^{-(1/2)(\phi_1^2 + \cdots + \phi_N^2)} k_N(t, \phi_1, \ldots, \phi_N)$$

$$= \frac{1}{(2\pi)^{N/2}} \int d\phi_1\, e^{-(1/2)\phi_1^2} k_1(t, \phi_1)$$

$$+ \frac{1}{(2\pi)^{N/2}} \int d\phi_1 \cdots \int d\phi_N\, e^{-(1/2)(\phi_1^2 + \cdots + \phi_N^2)} [k_N(t, \phi_1, \ldots, \phi_N) - k_1(t, \phi_1)]$$

The first integral on the right-hand side would be evaluated by Simpson's rule, for example, and the second integral by sampling, using Gaussian pseudorandom numbers for the ϕ_j, to give the approximation for $\langle X(t)\rangle$. In situations where ordinary statistical linearization is a fair approximation

one could replace k_1 by k_0 in the above. A useful indication of the error involved in the simple approximation would also be obtained by estimating the variance of the second integral in the usual way.

5. CONCLUSION

Although more remains to be done in assessing the relative merits of the various approximations derived here we feel that it has been clearly demonstrated that the variational approach to stochastic nonlinear problems is worthy of further investigation. In particular it would appear that piecewise statistical linearization with a small number of reserved variables, which are subject to some degree of optimization, might profitably be employed for more complicated systems. Perhaps the greatest virtue of the variational method is its comparative mathematical "honesty" since it involves no dubious decoupling assumptions, no formal manipulations of divergent series, and no inbuilt preconceived physical assumptions. Finally it should be mentioned that, although the whole discussion here has been confined to a particular system with a single degree of freedom, the theory may be readily generalized in a number of ways:

(i) Systems with several degrees of freedom may be considered.

(ii) The random forcing term may not take the form of white noise.

(iii) The random force may appear in the equations of motion in a stochastically nonlinear manner.

(iv) The possibility also exists of treating non-Gaussian forcing provided that the characteristic functional of the force is known in closed form or the force is given as a finite Wiener–Hermite series in terms of a white noise function.

ACKNOWLEDGMENT

The authors are grateful to Dr. R. H. Kraichnan for helpful comments on this work.

APPENDIX: ALTERNATIVE VARIATIONAL METHODS

Apparently the most straightforward variational method for the type of problem considered here is that described by Kraichnan[15–17] in which one seeks to minimize a mean-square error given by

$$E = \left\langle \int_0^t d\tau\, \rho(\tau)[\dot{X}(\tau) - A(X(\tau)) - e(\tau)]^2 \right\rangle$$

where $\rho(\tau)$ is a positive weight function. If Λ is a polynomial function, E can be expressed directly in terms of the correlation functions up to some finite order of the random functions $X(\tau)$, $e(\tau)$. The minimization can then be carried out with respect to these correlation functions provided that they can be chosen to satisfy the realizability constraints. Alternatively of course one can take suitable trial functions for the random function $X(\tau)$ itself. A difficulty however is that the time dependence of this trial function cannot be left unrestricted and must be given in some parametrized form. For example, if we were to use for $X(\tau)$ a trial function of the form $F(\tau, \Phi)$ considered earlier, then the variation δE would involve δF and $\delta \dot{F}$. The usual integration by parts procedure for the removal of the latter is not available to us since δF does not necessarily vanish at the upper limit of the τ integration. Other disadvantages of this approach are the appearance of an arbitrary weight function ρ and the fact that quantities of interest are not calculated as stationary values. The importance of formulating variational principles in such a way that the quantities of interest are expressed as stationary values of functionals of trial functions has been stressed in a recent review by Gerjuoy *et al.*[18]

A variational method, closely related to this one, but which removes the difficulty of restricted τ dependence has been familiar for some time in quantum mechanics where it is known as the Frenkel variational method.[19] It can also be formulated for classical stochastic problems.[20] Let us assume that the time derivative in the equation of motion has been approximated by a forward difference with time step h so that the equation is replaced by the difference equation

$$\frac{X_{n+1} - X_n}{h} = \Lambda(X_n) + e_n$$

The best approximation of a particular form is obtained by a step-by-step least-squares procedure. Assuming that $X_1 \cdots X_n$ have already been determined, X_{n+1} is then found by minimizing the mean-square error for the next time step to give

$$\delta \left\langle \left[\frac{X_{n+1} - X_n}{h} - \Lambda(X_n) - e_n \right]^2 \right\rangle = 0$$

Since X_n is not subject to variation we obtain

$$\left\langle \delta X_{n+1} \left[\frac{X_{n+1} - X_n}{h} - \Lambda(X_n) - e_n \right] \right\rangle = 0$$

and taking the limit as $h \to 0$ to return to the original equation of motion gives

$$\langle \delta X(\tau)[\dot{X}(\tau) - A(X(\tau)) - e(\tau)] \rangle = 0$$

It may easily be verified that, if the arbitrary functions subject to variation appear only linearly in the trial function $X(\tau)$, then this variational method yields the same approximations as does the one we have used. This is no longer the case when the arbitrary functions appear nonlinearly in X. Thus we can recover the PMVA and PSLA by this method but not the corresponding optimized approximations.

Other variational methods can be obtained by first formulating the problem in terms of functional integrals. For the simple system considered here this is very easy to do since the associated Fokker–Planck equation can be rewritten as a Bloch equation by means of a simple transformation and the probability density function can therefore be expressed as a Feynman path integral with real exponent.[21,22] The Feynman variational principle can then be applied directly. For more complex dissipative systems a functional integral representation of the probability density can still be obtained but cannot in general be expressed in the standard Feynman form. If the equation is

$$\dot{X}_\alpha = A_\alpha(X) + e_\alpha(\tau)$$

where $e_\alpha(\tau)$ is Gaussian with zero mean and correlation function

$$\langle e_\alpha(\tau)\, e_\beta(\tau') \rangle = d_{\alpha\beta}\delta(\tau - \tau')$$

then the functional integral for the conditional probability density $P(x, t \mid x', t')$ that the system is in state x at time t given that the state was x' at time t' is given by[23]

$$\int_{\substack{x(t)=x \\ x(t')=x'}} D[x] \int D[y]\, e^{-S[x,y]}$$

where

$$S[x, y] = \int_{t'}^{t} d\tau \, \{ i y_\alpha(\tau)[\dot{x}_\alpha(\tau) - A_\alpha(x(\tau))] - \tfrac{1}{2} d_{\alpha\beta}\, y_\alpha(\tau)\, y_\beta(\tau) \}$$

and the y are auxiliary functions which generate the response functions of the system. The quantity S is complex so that the inequality on which the Feynman extremal principle is based is inapplicable. However it may easily be verified that $\int e^{-S}$ is still given by the stationary value, with respect to S_0, of the quantity

$$W = e^{-\langle S - S_0 \rangle_0} \int e^{-S_0}$$

where

$$\langle S - S_0 \rangle_0 = \int (S - S_0)\, e^{-S_0} \Big/ \int e^{-S_0}$$

Approximations to the conditional probability density P can be obtained by seeking the stationary value of W as S_0 is allowed to vary over a restricted set of functionals such that W is well defined and can be evaluated.

An alternative procedure is to first carry out the integration with respect to y in the expression for P. The resulting functional integral is then of a form to which the Feynman extremal principle can be applied. Attempts to find useful approximations in this way have however proved disappointing.[24]

REFERENCES

1. J. O. Eaves and W. P. Reinhardt, *J. Stat. Phys.* **25**:127 (1981).
2. J. B. Morton and S. Corrsin, *J. Stat. Phys.* **2**:153 (1970).
3. R. H. Kraichnan, *J. Fluid Mech.* **41**:189 (1970).
4. W. C. Meecham and A. Siegel, *Phys. Fluids* **7**:1178 (1964).
5. M. Shugard, J. C. Tully, and A. Nitzan, *J. Chem. Phys.* **69**:336 (1978).
6. P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw–Hill, New York, 1953).
7. B. L. Moiseiwitsch, *Variational Principles* (Wiley-Interscience, New York, 1966).
8. P. C. Martin, E. D. Siggia, and H. A. Rose, *Phys. Rev. A* **8**:423 (1973).
9. R. Phythian, *J. Phys.* **A8**:1423 (1975).
10. R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw–Hill, New York, 1965).
11. W. D. Iwan and A. B. Mason, *Int. J. Non-Linear Mech.* **15**:71 (1980).
12. M. C. Valsakumar, K. P. N. Murthy, and G. Ananthakrishna, *J. Stat. Phys.* **30**:617 (1983).
13. B. J. West, G. Rovner, and K. Lindenberg, *J. Stat. Phys.* **30**:633 (1983).
14. J. H. Halton, *SIAM Rev.* **12**:1 (1970).
15. R. H. Kraichnan, *Phys. Rev. Lett.* **42**:1263 (1979).
16. R. H. Kraichnan, in *Nonlinear Dynamics*, H. G. Helleman, ed. (New York Academy of Sciences, New York, 1980).
17. R. H. Kraichnan, in *Theoretical Approaches to Turbulence*, D. L. Dwoyer *et al.*, ed. (Springer, New York, 1985).
18. E. Gerjuoy, A. R. P. Rau, and L. Spruch, *Rev. Mod. Phys.* **55**:725 (1983).
19. J. Frenkel, *Wave Mechanics: Advanced General Theory* (Oxford University Press, London, 1934).
20. R. Phythian, *J. Fluid Mech.* **53**:469 (1972).
21. B. Caroli, C. Caroli, and B. Roulet, *J. Stat. Phys.* **26**:83 (1981).
22. V. F. Baibuz, V. Yu. Zitserman, and A. N. Drozdov, *Physica* **127A**:173 (1984).
23. R. Phythian, *J. Phys.* **A10**::777 (1977).
24. R. Phythian and W. D. Curtis, *J. Phys.* **A13**:1575 (1980).