Diffusion in a Symmetric Bistable Potential: A Variational Approach

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

An approximation procedure for the solution of stochastic nonlinear equations, which was derived from a variational principle in a previous paper, is applied to the problem of a particle that diffuses in a symmetric bistable potential starting from the point of unstable equilibrium. The second moment $\langle \tilde{X}^2(t) \rangle$ and variance $\langle \tilde{X}^4(t) \rangle - \langle \tilde{X}^2(t) \rangle^2$ for the particle's position $\tilde{X}(t)$ are calculated as functions of the time t. Good agreement is found with results recently obtained by Baibuz *et al.* from an approximate evaluation of a path integral expression for the probability density.

1. INTRODUCTION

It is well known that a deterministic macroscopic description of the dynamics of a system evolving from an unstable state is impossible, because of the essential role played by fluctuations.[1] A simple model that illustrates this fact and has received much attention in recent years is provided by the motion of a particle diffusing in a symmetric bistable potential from the point of unstable equilibrium.[1,2] However, it appears that the formulation of an approximation procedure that is both accurate and sufficiently simple that it can be extended to systems with several degrees of freedom presents some difficulties.

Provided that the fluctuating force takes the form of white noise, two equivalent representations of such problems exist in the form of the Langevin equation and the Fokker–Planck equation, but it would appear that the former is better suited to the consideration of systems with several

---

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

degrees of freedom. For such systems the probability density is a function of several variables and is consequently difficult to deal with; moreover, it contains much more information than is normally required about the system. Other disadvantages of the Fokker–Planck equation approach are the difficulty of calculating nonsimultaneous correlations and also the difficulty of treating systems driven by random forces other than white noise.

Both the Langevin and Fokker–Planck equations have been used as the basis for various approximations for single-particle diffusion and related problems,[3⁻6] but most of these are applicable only in limiting situations when, for example, the time becomes large or the strength of the fluctuations is small. Apart from the extremely time-consuming method of direct numerical simulation, the only method that seems to be applicable over a wide range of times and parameter values is that of Baibuz *et al.*,[7] which makes use of a path integral expression for the probability density. The results obtained by these authors will be used for comparison with the calculations reported here.

Of the methods based on the Langevin equation, particular mention should be made of the scaling theory of Suzuki,[8] which will be relevant to the work described here. In its simplest form scaling theory is based on an approximate description of the particle motion in which the time interval is divided into three subintervals:

(i) An initial period $(0, \tau_0)$ when the particle oscillates about the point of unstable equilibrium. The assumption is made that, during this period, the nonlinear term in the equation of motion is negligible.

(ii) An intermediate period during which the random force is neglected, so that the particle moves deterministically toward one of the points of stable equilibrium.

(iii) A final period when the particle oscillates about a point of stable equilibrium, the nonlinear term being again neglected.

Although the sharp division into different regimes and the neglect of transitions across the instability point are oversimplifications, computer simulations[9] suggest that the basic idea is not too far from the truth if the strength of the fluctuating force is sufficiently small. The method has been further developed and generalized by de Pasquale *et al.*,[9] who have introduced a perturbation expansion about such a quasideterministic motion. However, the evaluation of the terms of this series requires the numerical calculation of stochastic integrals and is therefore rather difficult. Also, as might be anticipated, the results obtained appear to be less accurate as one approaches the critical point when the double-well potential degenerates into a single well.

Here we shall approach this problem by using an approximation procedure, which we have called the "piecewise statistical linearization approximation," or PSLA, which was derived by means of a variational principle in a previous paper[10] and was shown to give quite accurate results for the case of diffusion in a single well. The approximation provides a natural generalization of statistical linearization and may be formulated for a wide range of stochastic nonlinear systems, not necessarily driven by white noise. In Section 2 we briefly summarize the variational derivation of the PSLA and in Sections 3 and 4 we aply it to the case of a particle diffusing in a potential of the form

$$\tfrac{1}{2}\beta x^2 + \tfrac{1}{4}\lambda x^4$$

where β is negative or zero, corresponding to the double well and critical case, respectively. The quantities calculated are the second moment $\langle \tilde{X}^2(t) \rangle$, the variance $\langle \tilde{X}^4(t) \rangle - \langle \tilde{X}^2(t) \rangle^2$, and the probability density function $\langle \delta\{x - \tilde{X}(t)\} \rangle$.

2. THE PIECEWISE STATISTICAL LINEARIZATION APPROXIMATION

Let us suppose that $\tilde{X}(\tau)$ satisfies on the interval $(0, T)$ an equation of motion of the form

$$\dot{\tilde{X}}(\tau) = \Lambda(\tilde{X}(\tau)) + f(\tau) \tag{1}$$

where Λ is a polynomial function, $f(\tau)$ is a white noise function with strength parameter normalized to unity, and $\tilde{X}$ satisfies the sharp initial condition $\tilde{X}(0) = x_0$. If it is desired to calculate the expectation value $\langle F[\tilde{X}] \rangle$, where F is a functional of the solution $\tilde{X}$ of (1), then one is led to consider the quantity

$$\mathscr{I} = \left\langle F[X] - \int_0^T d\tau \; Y(\tau)\{\dot{X}(\tau) - \Lambda(X(\tau)) - f(\tau)\} \right\rangle \tag{2}$$

where X and Y are arbitrary random functions such that $X(0) = x_0$ and $Y(T) = 0$. It is then easily verified that the stationary value of $\mathscr{I}$ is $\langle F[\tilde{X}] \rangle$. Approximations may be generated by using trial functions for X and Y that are such that $\mathscr{I}$ can be evaluated.

We consider the following choice of trial functions:

$$X(\tau) = k(\tau, \Phi) + \int_0^T d\tau' \; K(\tau, \tau', \Phi) \, \hat{f}(\tau')$$

$$\tag{3}$$

$$Y(\tau) = l(\tau, \Phi) + \int_0^T d\tau' \; L(\tau, \tau', \Phi) \, \hat{f}(\tau')$$

where k, K, l, and L arbitrary, apart from the conditions needed to ensure that $X(0) = x_0$, $Y(T) = 0$; and Φ denotes a fixed set $\{\Phi_j\}$ of N linear functionals of $f(\tau)$ on the interval $(0, T)$. The functionals Φ_j may, without loss of generality, be chosen as statistically orthonormal and so may be written in the form

$$\Phi_j = \int_0^T d\tau \, \xi_j(\tau) f(\tau)$$

where the "basis" functions $\xi_j(\tau)$ are orthonormal on $(0, T)$. The Gaussian random function $\hat{f}(\tau)$ defined as

$$\hat{f}(\tau) = f(\tau) - \xi_j(\tau) \Phi_j$$

is such that its conditional expectation value, for given values ϕ_j of the functionals Φ_j, is zero:

$$\langle \hat{f}(\tau) \rangle_\phi = \frac{\langle \hat{f}(\tau) \, \delta(\phi - \Phi) \rangle}{\langle \delta(\phi - \Phi) \rangle} = 0$$

Also, it may be seen that

$$\langle \hat{f}(\tau) f(\tau') \rangle = r(\tau, \tau') = \delta(\tau - \tau') - \xi_j(\tau) \xi_j(\tau')$$

where a summation from 1 to N over repeated indices is implied.

Substituting the trial functions (3) into the expression for $\mathscr{I}$ and seeking a stationary value with respect to the arbitrary functions k, K, l, and L, one obtains equations for these functions. Those for k and K are

$$\dot{k} = \phi_j \xi_j + \Lambda(k) + \frac{p}{2!} \Lambda''(k) + \frac{3p^2}{4!} \Lambda^{(4)}(k)$$

$$+ \frac{5.3}{6!} p^3 \Lambda^{(6)}(k) + \cdots$$

$$\dot{K}(\tau, \tau', \phi) = \delta(\tau - \tau') + \left\{ \Lambda'(k) + \frac{3p}{3!} \Lambda'''(k) \right.$$

$$\left. + \frac{5.3}{5!} p^2 \Lambda^{(5)}(k) + \cdots \right\} K(\tau, \tau', \phi)$$

$$(4)$$

where p is given by

$$p(\tau, \phi) = \int_0^T d\tau' \int_0^T d\tau'' \, K(\tau, \tau', \phi) \, r(\tau', \tau'') \, K(\tau, \tau'', \phi)$$

The solution of the equation for K is

$$K(\tau, \tau', \phi) = e^{s(\tau) - s(\tau')}\theta(\tau - \tau') \qquad (5)$$

where θ is the step function defined by

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

and where

$$\dot{s}(\tau) = \Lambda'(k) + \frac{3p}{3!}\Lambda'''(k) + \frac{5.3}{5!}p^2\Lambda^{(5)} + \cdots \qquad (6)$$

with $s(0) = 0$. The quantity p may then be expressed in terms of s as follows:

$$p(\tau) = e^{2s}\left\{\int_0^\tau d\tau'\, e^{-2s(\tau')} - \sum_j \left[\int_0^\tau d\tau'\, \xi_j(\tau')e^{-s(\tau')}\right]^2\right\} \qquad (7)$$

The stationary value of $\mathscr{I}$, which is our approximation for $\langle F[\tilde{X}]\rangle$, is given by

$$\frac{1}{(2\pi)^{N/2}}\int d\phi\, e^{-\phi^2/2}\left\langle F\left[k(\tau, \phi) + \int_0^T d\tau'\, K(\tau, \tau', \phi)\hat{f}(\tau')\right]\right\rangle_\phi$$

where the conditional expectation value in the integrand may be written in terms of k and p for many functionals F of interest. Thus, for example, for the second and fourth moments $\langle \tilde{X}^2(t)\rangle$ and $\langle \tilde{X}^4(t)\rangle$ with $t \in (0, T)$ the approximations are

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

$$\frac{1}{(2\pi)^{N/2}}\int d\phi\, e^{-\phi^2/2}[k^4(t, \phi) + 6k^2(t, \phi)\, p(t, \phi) + 3p^2(t, \phi)] \qquad (8)$$

and for the probability density function $\langle \delta\{x - X(t)\}\rangle$ we get

$$\frac{1}{(2\pi)^{(N+1)/2}}\int d\phi \exp(-\tfrac{1}{2}\phi^2)\frac{\exp\{-[x - k(t, \phi)]^2/2p(t, \phi)\}}{[p(t, \phi)]^{1/2}}$$

which is a continuous superposition of Gaussians with different means and variances.

The PSLA may be summarized in the following way:

1. The set of all realizations of the random force is divided into sub-sets according to the values ϕ_j taken by the linear functionals Φ_j.

2. In each such subset the solution of (1) is approximated by a linear functional of f as in ordinary statistical linearization.

3. Mean values of interest are calculated by a two-stage averaging process, the average over each subset being performed analytically and the further averaging by a numerical N-fold integration over the ϕ_j with Gaussian measure.

The functionals Φ_j therefore play the role of reserved linear functionals.[10]

For $N = 0$ it is clear that one recovers simple time-dependent statistical linearization, while for $N \to \infty$ we have $\xi_j(\tau) \, \xi_j(\tau') \to \delta(\tau - \tau')$ and it may be verified that $p \to 0$ while $k(\tau, \Phi)$ tends to the exact solution of (1). The hope is that useful approximations can be obtained for small N so that the numerical integration in expressions such as (8) can be performed directly without resorting to Monte Carlo methods. In a previous paper[10] this was shown to be the case for the problem of diffusion in a single-well potential.

The trial function (3) can be generalized by allowing the linear functionals Φ_j also to be subject to variation within the constraints imposed by the orthonormality conditions. If the Φ_j are otherwise unrestricted, the resulting approximations are rather complicated and involve coupled integrodifferential equations relating k, K, l, and L and the optimized basis functions. A simpler approach that we adopt here is to choose basis functions of a particular form but containing adjustable parameters that can then be varied to obtain a stationary value for the quantity of interest. This stationary value is then the required approximation. Consideration of causality shows that, if the functional F depends on $X(\tau)$ only for τ in the interval $(0, t)$, then the basis functions should be chosen to vanish outside this interval. For the single-well potential it appears that satisfactory results can be obtained with any convenient set of basis functions satisfying this requirement, but it will be seen that, for the double-well potential, this is no longer true. However, when the choice of basis functions is guided by the simple considerations of scaling theory, the PSLA is equally effective for the double well.

3. THE DOUBLE-WELL POTENTIAL

The Langevin equation now takes the form

$$\dot{\tilde{X}}(\tau) = -\beta \tilde{X}(\tau) - \lambda \tilde{X}^3(\tau) + f(\tau)$$

with $\tilde{X}(0) = 0$, where β may be taken as -1 for the double well and 0 for the critical case. The equations for the PSLA are accordingly

$$\dot{k} + \beta k + \lambda k^3 + 3\lambda kp = \phi_j \xi_j(\tau)$$

$$\dot{s} = \beta + 3\lambda k^2 + 3\lambda p \tag{9}$$

$$p(\tau) = e^{-2s(\tau)} \left\{ \int_0^\tau d\tau' \, e^{2s(\tau')} - \sum_j \left[\int_0^\tau d\tau' \, \xi_j(\tau') \, e^{s(\tau')} \right]^2 \right\}$$

with $k(0) = 0$, $s(0) = 0$. Considering first the simple statistical linearization approximation given by $N = 0$, it is seen that the solution of (9) then gives $k = 0$ while p increases monotonically to the asymptotic value $[-\beta + (\beta^2 + 6\lambda)^{1/2}]/6\lambda$. The probability density corresponding to this choice of trial function for X is

$$[2\pi p(\tau)]^{-1/2} \exp[-x^2/2p(\tau)]$$

which is a Gaussian, permanently centered at the origin, which broadens with increasing τ. This is quite different from the double-peak density that actually develops, so it is not surprising to find that the values obtained for the moments are inaccurate except for small τ.

It is known, however, that statistical linearization works quite well if the particle starts from a point sufficiently far from the origin that there is an overwhelming probability of a particular one of the points of stable equilibrium being attained.[11] This suggests that, if the set of all realizations of the particle motion could be divided into two subsets according to which stable state is eventually attained, then a good approximation might be obtained by applying statistical linearization separately in these subsets. According to scaling theory, however, this division is determined by the solution of the linearized equation

$$\dot{X}(\tau) = X(\tau) + f(\tau)$$

in the initial period $(0, \tau_0)$, where we have now put $\beta = -1$ corresponding to the double-well potential. One is therefore led to the idea that a suitable reserved linear functional for incorporation into the PSLA might be of the form

$$\int_0^{\tau_0} d\tau \exp(-\tau) f(\tau)$$

Instead of adopting a sharp cutoff at τ_0, we shall instead extend the integration over the full range of τ values consistent with causality and include an adjustable parameter b in the exponential. Thus, the basis

function we shall employ, for calculation of the mean value of a functional depending only on $f(\tau)$ for $\tau \in (0, t)$, is given by

$$\xi(t) = \left(\frac{2b}{1 - \exp(-2bt)}\right)^{1/2} \exp(-b\tau)\, \theta(t - \tau) \tag{10}$$

To calculate $\langle \tilde{X}^2(t) \rangle$, for example, using the PSLA with $N = 1$, we therefore solve Eqs. (9), with the summation over j limited to the single term corresponding to the basis function (10), and then evaluate

$$(2\pi)^{-1} \int d\phi\, e^{-\phi^2/2} [k^2(t) + p(t)]$$

Finally, the parameter b is adjusted to give a stationary value of this quantity and this is the required approximation for $\langle \tilde{X}^2(t) \rangle$.

To calculate higher moments and the probability density, one can go through the same variational procedure. However, it will be seen that good accuracy can be obtained if the same basis function, with the same value of b, is used as in the calculation of $\langle \tilde{X}^2(t) \rangle$. As well as reducing the amount of computation, this has the further advantage that relationships such as

$$\langle \tilde{X}^2(t) \rangle = \int dx\, x^2 P(x, t)$$

and realizability inequalities such as $\langle \tilde{X}^4 \rangle \geqslant \langle \tilde{X}^2 \rangle^2$ are satisfied by the approximate quantities. This would not necessarily be so if different basis functions were used for calculating different quantities. The fact that the trial function gives quite accurate results for the mean values of several different functionals indicates that this trial function is perhaps a good approximation to the exact $\tilde{X}(\tau)$ for the vast majority of realizations.

This last observation suggests that it may be possible to calculate corrections to the approximation described above by some sort of perturbation theory. We shall not pursue this idea further here, but instead we calculate corrections by applying the PSLA with $N = 2$. The simple procedure will be adopted of taking the first basis function $\xi_1(\tau)$ as that used for the $N = 1$ calculations, while $\xi_2(\tau)$ will be chosen as a convenient function orthogonal to $\xi_1(\tau)$, such as $(\alpha_0 + \alpha_1 \tau)\, \xi_1(\tau)$ with α_0 and α_1 determined by the orthonormality conditions.

4. CALCULATIONS

Values of the second moment $\langle \tilde{X}^2(t) \rangle$ and the variance $\langle \tilde{X}^4(t) \rangle - \langle \tilde{X}^2(t) \rangle^2$ for the double-well potential with $\lambda = 0.01$, 0.1, and 1/3 on the

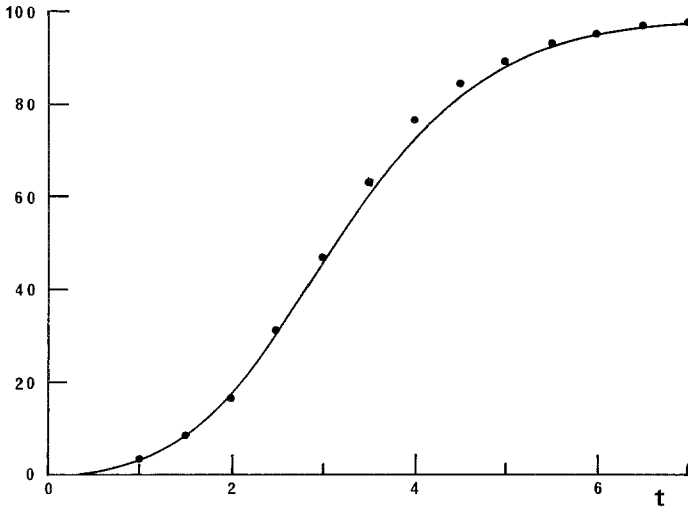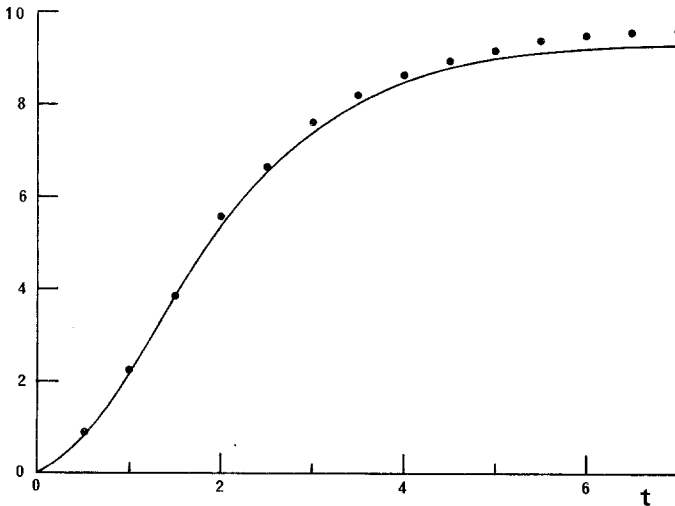Fig. 1. Values of $\langle \tilde{X}^2(t) \rangle$ for $\lambda = 0.01$ calculated with the PSLA. The $N = 1$ and $N = 2$ approximations are indistinguishable on this scale. (●) The results of Baibuz *et al.*

time interval $0 \leqslant t \leqslant 7$ have been calculated for comparison with the results of Baibuz *et al.* In addition, the second moment and variance have been evaluated for the critical case $\beta = 0$, $\lambda = 2$ for $0 \leqslant t \leqslant 1$ for comparison with the values obtained by de Pasquale *et al.*

Figures 1–3 show the values of $\langle \tilde{X}^2(t) \rangle$ for the double-well potential

Fig. 2. Values of $\langle \tilde{X}^2(t) \rangle$ for $\lambda = 0.1$ calculated with the PSLA. The $N = 1$ and $N = 2$ approximations are indistinguishable on this scale. (●) The results of Baibuz *et al.*
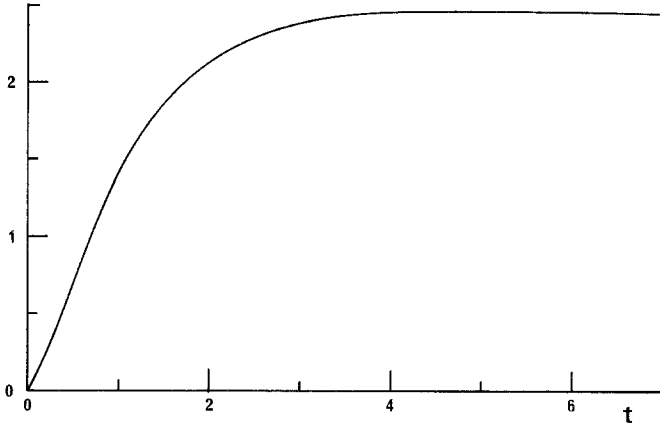
Fig. 3. Values of $\langle \tilde{X}^2(t) \rangle$ for $\lambda = 1/3$ calculated with the PSLA. The $N=1$ and $N=2$ approximations are indistinguishable.

in the three cases considered, calculated by means of the PSLA with $N=1$ and $N=2$ in the way described above. It will be seen that the differences between the values obtained with the two approximations are small in all cases. Figures 4–6 show the corresponding values for the variance $\langle \tilde{X}^4(t) \rangle - \langle \tilde{X}^2(t) \rangle^2$. The values obtained for the adjustable parameter b
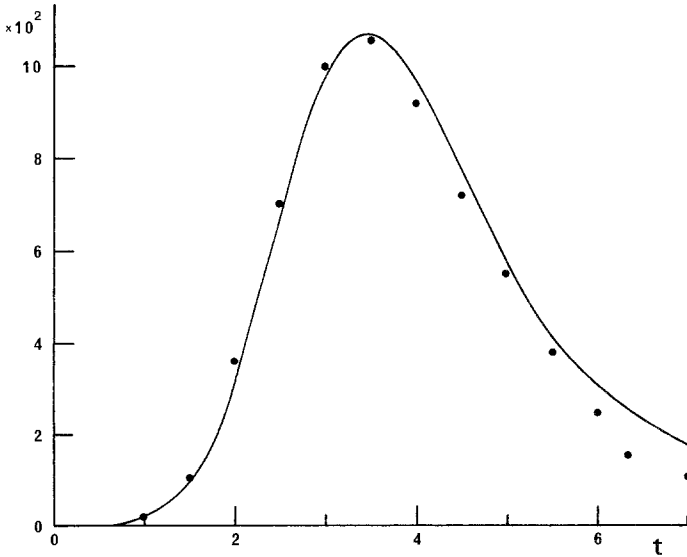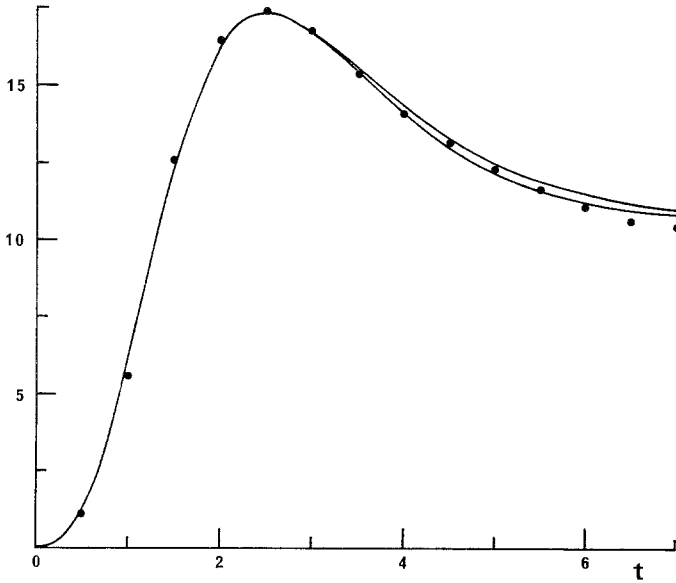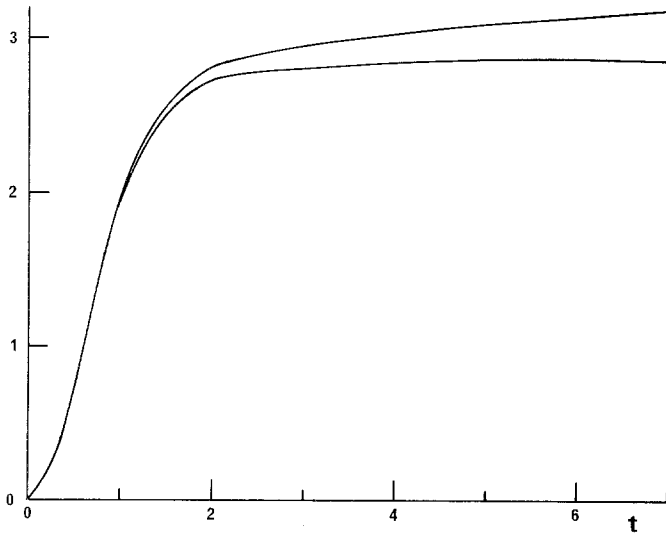

Fig. 4. Values of the variance $\langle \tilde{X}^4(t) \rangle - \langle \tilde{X}^2(t) \rangle^2$ for $\lambda = 0.01$ calculated with the PSLA. The $N=1$ and $N=2$ approximations are indistinguishable. ($\bullet$) The results of Baibuz *et al.*

Fig. 5. Values of the variance $\langle \tilde{X}^4(t) \rangle - \langle \tilde{X}^2(t) \rangle^2$ for $\lambda = 0.1$ calculated with the PSLA. The upper curve is the $N = 2$ approximation and the lower curve the $N = 1$ approximation. ($\bullet$) The results of Baibuz *et al.*

Fig. 6. Values of the variance $\langle \tilde{X}^4(t) \rangle - \langle X^2(t) \rangle^2$ for $\lambda = 1/3$ calculated with the PSLA. The upper curve is the $N = 1$ approximation and the lower curve the $N = 2$ approximation.
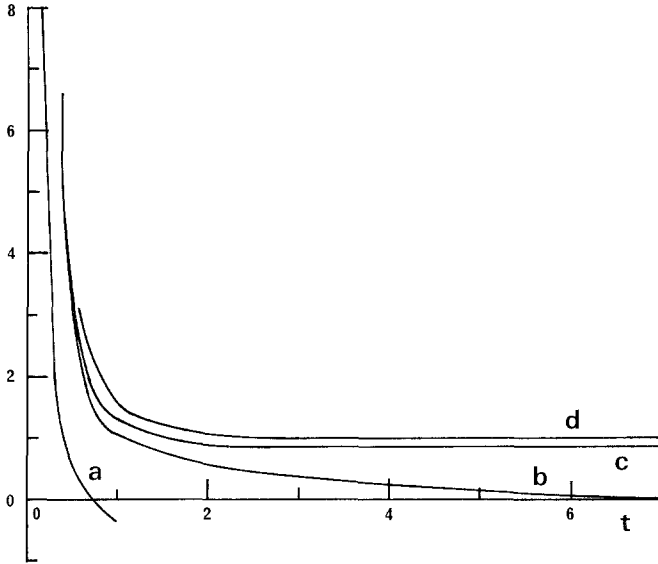
Fig. 7. The values of the variational parameter b as a function of t for various values of λ:
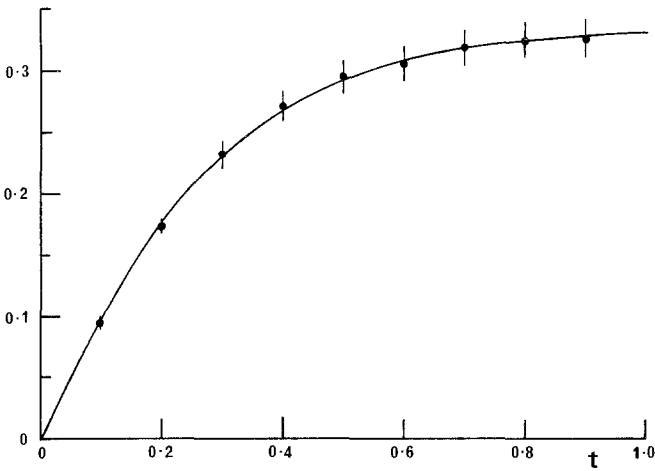(a) $\lambda = 2$, (b) $\lambda = 1/3$, (c) $\lambda = 0.1$, (d) $\lambda = 0.01$.

Fig. 8. Values of $\langle \tilde{X}^2(t) \rangle$ for $\lambda = 2$ calculated with the PSLA. The $N = 1$ and $N = 2$ approximations are indistinguishable on this scale. The points with error bars show the results obtained by de Pasquale *et al.* by direct numerical simulation.

are shown in Fig. 7 and it will be seen that b is close to 1 only when λ is small, but t is not too small. This might be expected, since small λ corresponds to the case in which the troughs of the potential are deep, so that the argument based on scaling theory is more likely to be valid. The importance of an appropriate choice of basis function, especially for small λ, has been checked by repeating these calculations using polynomial basis functions with no adjustable parameter as in previous work on the single-well potential.[10] It is found that the values obtained for the second moment and variance are then much less accurate and appear to converge only slowly to the correct values as N increases.

The results of the calculations for the critical case are shown in Figs. 8 and 9 together with values obtained by de Pasquale *et al.* using direct numerical simulation with an average over 9000 realizations being taken. Of course, in this case, the arguments that led to the particular choice of basis function do not aply and it is not surprising to find that b can become negative.

The probability density function $\langle \delta\{x - \tilde{X}(t)\}\rangle$ has also been calculated as a function of x by means of the PSLA with $N = 1$ for the case of $\lambda = 1/3$. Figure 10 shows the results for $t = 0.3$, 0.8, and 6.4 compared with those of Baibuz *et al.*
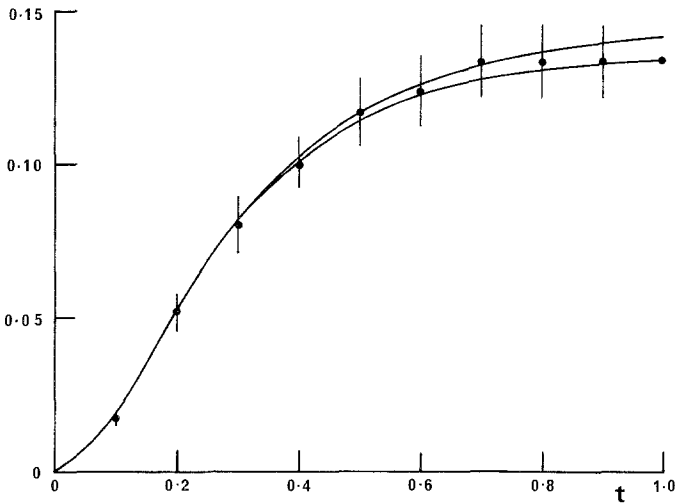

Fig. 9.   Values of the variance $\langle \tilde{X}^4(t)\rangle - \langle \tilde{X}^2(t)\rangle^2$ for $\lambda = 2$ calculated with the PSLA. The upper curve corresponds to $N = 1$ and the lower curve to $N = 2$. The points with error bars show the results obtained by de Pasquale *et al.* by direct numerical simulation.
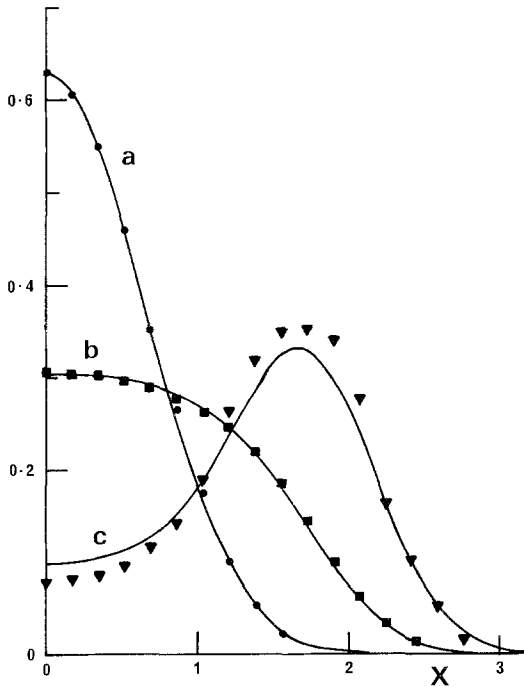
Fig. 10. The probability density function for $\lambda = 1/3$, for (a) $t = 0.3$, (b) $t = 0.8$, and (c) $t = 6.4$. ($\bullet$, $\blacksquare$, $\blacktriangledown$) The results of Baibuz *et al.* for these three cases.

5. CONCLUSION

The method described here appears to embody in a natural way the ideas of statistical linearization and scaling theory and incorporates them into a calculation scheme that seems more efficient than direct numerical simulation. It gives quite accurate results for both low-order moments and the probability density function, without involving an excessive amount of computation, for the double-well potential and the critical case as well as for the single well. In a future paper we hope to examine the application of this approach to systems with more degrees of freedom.

REFERENCES

1. N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, 1981).
2. N. G. Van Kampen, *J. Stat. Phys.* **17**:71 (1977).
3. H. D. Vollmer and H. Risken, *Physica* **110A**:106 (1982).

4. F. Haake, *Phys. Rev. Lett.* **41**:1685 (1978).
5. B. Caroli, C. Caroli, and B. Roulet, *J. Stat. Phys.* **26**:83 (1981).
6. U. Weiss and W. Haeffner, in *Functional Integration*, J. P. Antoine and E. Tirapegui, eds. (Plenum, New York, 1980).
7. V. F. Baibuz, V. Yu. Zitserman, and A. N. Drozdov, *Physica* **127A**:173 (1984).
8. M. Suzuki, *J. Stat. Phys.* **16**:477 (1977).
9. F. de Pasquale, P. Tartaglia, and P. Tombesi, *Z. Phys. B* **43**:353 (1981).
10. R. Phythian and W. D. Curtis, *J. Stat. Phys.* **42**:1019 (1986).
11. M. C. Valsakumar, K. P. N. Murthy, and G. Ananthakrishna, *J. Stat. Phys.* **30**:617 (1983).