Small-noise approximations to the solution of the Smoluchowski equation

Moshe Gitterman* and George H. Weiss
National Institutes of Health, Bethesda, Maryland 20892
(Received 10 September 1992)

We explore the connection between the two perturbation methods that have been suggested for solving Smoluchowski-like equations with weak noise. A slight modification of the first of these methods (the size expansion of van Kampen) makes it identical to the second (the method of running coordinates). We point out that a number of modifications to both methods are possible and should be explored. We illustrate the advantages of both of these methods as applied to three solvable examples.

PACS number(s): 05.40.+j, 82.20.Fd

I. INTRODUCTION

Smoluchowski equations for the probability density $p(\mathbf{r},t)$ for the position $\mathbf{r}$ of a particle at time t subject to space-dependent diffusion and convection,

$$\frac{\partial p}{\partial t}=\sum_{i,j}\frac{\partial}{\partial x_i}\left[D_{ij}(\mathbf{r})\frac{\partial p}{\partial x_j}\right]-\sum_i\frac{\partial}{\partial x_i}[v_i(\mathbf{r})p] \tag{1}$$

arise in many applications of chemistry and physics [1,2]. There are two methods available for generating solutions to Eq. (1) in the so-called weak-noise limit. This limit is defined by the requirement that the $D_{ij}(\mathbf{r})$ are small in comparison to 1 in some suitable dimensionless set of units and $|v_i(\mathbf{r})|=O(1)$, in the same dimensionless units. The first is due to van Kampen, [2], who assumes the existence of a large parameter Ω which he uses both to expand all of the functions in Eq. (1) and as part of a transformation of coordinates. This leads, in lowest order, to an Ornstein-Uhlenbeck equation whose solution in an unbounded space is readily found. The second technique was originally suggested by Weiss and Dishon [3], and expresses the Smoluchowski equation in terms of a scaled running coordinate which serves to eliminate the velocity or the set of nonperturbative terms in Eq. (1) and forms the starting point for a systematic perturbation expansion. The method leads, in lowest order, to a diffusion equation in contrast to the Ornstein-Uhlenbeck equation associated with the van Kampen analysis. In this paper we compare the performance of these two techniques on some simple exactly solvable examples of the Smoluchowski equation, showing that neither method is intrinsically more accurate than the other. For simplicity we discuss only the one-dimensional version of Eq. (1) which will suffice to make our point, although the extension of these methods to related problems in higher dimensions does not introduce appreciably more conceptual problems.

II. DESCRIPTION OF THE APPROXIMATIONS

The Smoluchowski equation, or equivalently the Fokker-Planck equation in one dimension can be written as

$$\frac{\partial p}{\partial t}=\epsilon\frac{\partial}{\partial x}\left[D(x)\frac{\partial p}{\partial x}\right]-\frac{\partial}{\partial x}[v(x)p]\ , \tag{2}$$

where the small parameter is ϵ which multiplies the diffusion coefficient. The function $D(x)\geq 0$ can be regarded as being a spatially varying diffusion coefficient and $v(x)$ is a convective term. We implicitly assume that the coordinates in Eq. (2) are dimensionless, and the two functions that appear as coefficients in Eq. (2) are $O(1)$.

As mentioned, there are two ways to develop a perturbation expansion of the solution to Eq. (2) in the limit $\epsilon\rightarrow 0$. Both of these can be regarded as singular perturbation expansions in ϵ away from the solution of the deterministic equation

$$\dot{x}=v(x) \tag{3}$$

associated with the stochastic equation in Eq. (2). We denote the solution to this equation which corresponds to the initial condition $x(0)=x_0$ by $X(t)$.

The first technique for generating approximate solutions to Eq. (1), the so-called system size expansion, was suggested by van Kampen [2], and is equivalent to replacing the spatial variable x in Eq. (2) by a new variable y, by the substitution

$$x=X(t)+\sqrt{\epsilon}y\ . \tag{4}$$

Let the lowest-order term in the hierarchy of approximations to the true solution $p(x,t)$ in the system size expansion be denoted by $p_0(y,t)$. On substituting Eq. (4) into Eq. (2) one obtains an Ornstein-Uhlenbeck equation which takes the form

$$\frac{\partial p_0}{\partial t}=a(t)\frac{\partial^2 p_0}{\partial y^2}-b(t)\frac{\partial}{\partial y}(yp_0) \tag{5}$$

in which the coefficients $a(t)$ and $b(t)$ are expressed in terms of $D(x)$ and $v(x)$ as

$$a(t)=D(X(t))\ ,\quad b(t)=\left.\frac{dv}{dx}\right|_{x=X(t)}\ . \tag{6}$$

Systematic corrections to this lowest-order result are readily generated.

A second approach that has been explored, mainly in

literature dealing with chromatographic methods, is based on a transformation to a running coordinate ξ

$$\xi=\int_{x_0}^{x}\frac{d\alpha}{v(\alpha)}-t \ . \tag{7}$$

This is obtained from the first-order differential equation which follows from Eq (2) when D is set equal to zero as motivated by the method of characteristics [see Eq. (18) below]. We will denote the solution to this equation for x in terms of ξ by $x=H(\xi+t)$ so that when there is no diffusion the position of a peak initially located at x_0 is $x=H(t)$. This permits us to identify the function $H(t)$ in terms of the solution of the dynamic equation (3) by $H(t)=X(t)$. For future analysis we also define the quantities $D(H(\beta))=\mathcal{D}(\beta)$ and $v(H(\beta))=\mathcal{V}(\beta)$ which are the diffusion coefficient and the velocity evaluated at the peak for the diffusion-free system. Since the solution to the zero-noise limit of Eq. (2) is equivalent to setting $\xi=0$ in Eq. (7) we can assert that in the weak-noise limit ξ should also be expected to be close to zero. With this in mind one can return to Eqs. (2) and (7) and introduce a stretched coordinate ρ by

$$\rho=\frac{\xi}{\sqrt{\epsilon}} \ , \tag{8}$$

further defining a reduced form of $p(\rho,t)$ by

$$\psi(\rho,t)=\mathcal{V}(\rho)p(\rho,t) \ . \tag{9}$$

The substitutions of Eqs. (7)–(9) into Eq. (2) allow us to transform that equation to the form

$$\frac{\partial\psi}{\partial t}=\frac{\partial}{\partial\rho}\left[\frac{\mathcal{D}(t+\sqrt{\epsilon}\rho)}{\mathcal{V}(t+\sqrt{\epsilon}\rho)}\frac{\partial}{\partial\rho}\left[\frac{\psi}{\mathcal{V}(t+\sqrt{\epsilon}\rho)}\right]\right] . \tag{10}$$

No approximations have been made so far. While this equation is clearly as hard to solve as is Eq. (2), it is nevertheless useful because from it one can generate a systematic series of approximations in the small parameter ϵ. The lowest-order term in this sequence, $\psi_0(\rho,t)$, satisfies an equation that is found by setting $\epsilon=0$ in Eq. (10). That is to say, ψ_0 is the solution

$$\frac{\partial\psi_0}{\partial t}=\frac{\mathcal{D}(t)}{\mathcal{V}^2(t)}\frac{\partial^2\psi_0}{\partial\rho^2} \ , \tag{11}$$

which reduces to the standard diffusion equation provided that one replaces t in this equation by the variable

$$\Delta(t)=\int_0^t\frac{\mathcal{D}(\tau)}{\mathcal{V}^2(\tau)}d\tau \ . \tag{12}$$

III. COMPARATIVE PROPERTIES OF THE TWO APPROXIMATIONS

As already mentioned, it is possible to calculate systematic higher-order corrections to both the van Kampen approximation and to that discussed by Weiss and Dishon, as well as to develop the theories for problems arising in dimensions greater than one. A natural question to ask is whether the two techniques are identical. We will show that a slightly modified version of the van Kampen approximation is equivalent, at least in lowest order, to the theory developed in [3].

Indeed, since the coordinate ξ is small whenever ϵ is small we can expand the formal relation for x in terms of ξ [or ρ as defined in Eq. (8)] to lowest order as

$$x=H(t+\xi)\sim H(t)+\dot{H}(t)\xi=X(t)+\sqrt{\epsilon}\dot{H}(t)\rho \ . \tag{13}$$

This differs from the van Kampen transformation in Eq. (4) because a time-dependent term $\dot{H}(t)$ appears as a coefficient of ρ, implying that the two approximations, as originally formulated, are not equivalent. However, the difference is apparent and not real, as may be demonstrated by slightly modifying the original van Kampen ansatz, replacing Eq. (4) by the generalization

$$x=X(t)+\sqrt{\epsilon}\varphi(X)\rho \ , \tag{14}$$

and at the same time replacing $p(x,t)$ by

$$p(x,t)=\Pi(\rho,t)/\alpha(X) \ , \tag{15}$$

where ρ is the new spatial variable and $\alpha(X)$ and $\varphi(X)$ are so far unspecified. The specific choices $\alpha(X)=\varphi(X)=v(X)$ suggested by Eq. (13) suffice to produce a diffusion equation for $\Pi_0(\rho,t)$ that is identically equal to that in Eq. (11). Hence the modified van Kampen theory can be brought into a form that is essentially equivalent to the theory presented in [3].

Along with the van Kampen approximation, that of Weiss and Dishon also can be modified. In fact, the choice of the running coordinate in Eq. (7) is by no means a unique one, and slightly different versions of Eq. (11) can be found by adopting different choices in place of ξ as given in Eq. (7). One can sometimes take advantage of this flexibility by putting the equation for the lowest-order term in a more convenient form than the one provided by the definition in Eq.(7). This point will be illustrated in the analysis of the example considered in the following section.

IV. ANALYSIS OF LINEAR DRIFT IN THE PRESENCE OF ADDITIVE AND MULTIPLICATIVE NOISE

As a first simple case in which one can compare how each of the two techniques works we consider that in which $D(x)=D=\text{const}$ and $v(x)=-xT$, where T is also a constant. The initial condition is taken to be $p(x,0)=\delta(x-x_0)$. The Smoluchowski equation with these specific features is

$$\frac{\partial p}{\partial t}=D\frac{\partial^2 p}{\partial x^2}+\frac{1}{T}\frac{\partial}{\partial x}(xp) \ , \tag{16}$$

which is already in the form of an Ornstein-Uhlenbeck equation. This implies that the lowest-order term in the van Kampen approximation will automatically produce the exact solution for $p(x,t)$, independent of the order of magnitude of D.

Let us next consider the approach based on a transformation to moving coordinates. The application of Eq. (7) yields the definition

$$\xi=T\ln\left[\frac{x_0}{x}\right]-t \ . \tag{17}$$

The solution of the Ornstein-Uhlenbeck equation in free space is known to be a Gaussian in x. The solution for the lowest-order approximation given in Eq. (11) will be a Gaussian in ρ, which is equivalent to the statement that it is Gaussian in $\ln(x_0/x)$ rather than in x. Hence the lowest-order approximation obtained in this way cannot be the exact solution. However, one can choose a different form for the running coordinate ξ. The particular form in Eq. (17) follows from the appropriate characteristic equation which, of this example, has the form

$$\frac{dt}{1}=-\frac{T\,dx}{x}\ . \tag{18}$$

The choice of the running coordinate in Eq. (17) is equivalent to choosing a specific form for the constant of integration in Eq. (18). However, this choice is not unique, and one can, for example, also choose the constant of integration, ξ, to be

$$\xi=xe^{t/T}\ , \tag{19}$$

i.e., at $t=0$ one has $\xi=x_0$. In this case a short calculation shows that one again finds the exact solution of the original equation in the lowest-order approximation obtained from the solution of Eq. (11).

A second instructive example is one in which the underlying dynamics consist of linear drift and a simple form of multiplicative noise. The Langevin equation describing this model is

$$\dot{x}=-kx+xn(t)\ , \tag{20}$$

where $n(t)$ is white noise with $\langle n^2(t)\rangle=1/T$, the parameter T having the dimensions of time. Equation (20) is to be interpreted in the Stratonovich sense. The corresponding Fokker-Planck equation can be solved exactly, leading to an expression for the probability density $p(x,t)$ in the form

$$p(x,t)=\frac{1}{x\sqrt{4\pi(t/T)}}\exp\left[-\frac{\{\ln(x/x_0)+kt\}^2}{4t/T}\right]\ , \tag{21}$$

from which the moments are readily calculated as

$$\langle x^n(t)\rangle=x_0^n\exp[n(n-kT)(t/T)]\ , \tag{22}$$

which increases or decreases monotonically as a function of time depending on the order of the moment. Comprehensive analyses of Eqs. (20)–(22), including a treatment of such properties as the effect of the boundary at $x=0$ and critical slowing-down behavior at some values of k, can be found in Refs. [6–8]. In this example the introduction of a running coordinate suggested in Eq. (7) leads to an exact solution while a solution based on the original van Kampen formulation has a second moment that is always monotonic in the same direction regardless of the parameters.

V. EXPONENTIAL DECAY OF DIFFUSION AND MOBILITY

A third slightly less trivial example based on an exactly solvable equation for diffusion-dependent peak broadening in electrophoresis [7] also allows us to compare the two approximation techniques. The diffusion equation derived in that context takes the form

$$\frac{\partial p}{\partial t}=\epsilon\frac{\partial}{\partial x}\left[e^{-x}\frac{\partial p}{\partial x}\right]-\frac{\partial}{\partial x}(e^{-x}p)\ . \tag{23}$$

When the initial condition is $p(x,0)=\delta(x)$ the exact solution has been shown in [7] to be

$$p(x,t)=\frac{1}{\epsilon t}\exp\left[\frac{x}{2\epsilon}(1+\epsilon)-\frac{1+e^x}{\epsilon t}\right]I_{1/\epsilon-1}\left[\frac{2e^{x/2}}{\epsilon t}\right]\ , \tag{24}$$

where $I_n(x)$ is a Bessel function. Realistic values for ϵ based on electrophoretic systems are the order of $\epsilon=10^{-3}$. In this regime we may make a further approximation to the last expression by using the asymptotic relation [8]

$$I_\lambda(a\lambda)\sim\frac{1}{\sqrt{2\pi\lambda}}\frac{e^{\lambda\eta(a)}}{(1+a)^{2)^{1/4}}}\left\{1+O\left[\frac{1}{\lambda}\right]\right\}\ , \tag{25}$$

which is valid for large λ. The function $\eta(a)$ is

$$\eta(a)=(1+a^2)^{1/2}+\ln\left[\frac{a}{1+(1+a^2)^{1/2}}\right]\ . \tag{26}$$

On making use of the approximation for the Bessel function in Eq. (25), collecting terms, setting $\beta=(2/t)\exp(x/2)$, and inserting the result into Eq. (24) we find

$$p(x,t)\sim\frac{1}{t\sqrt{2\pi\epsilon}}\frac{\beta}{((\beta^2+1)^{1/4}}\frac{1}{(\beta^2+1)^{1/2}-1}\exp\left[\frac{1}{\epsilon}\left\{(\beta^2+1)^{1/2}+\ln\left[\frac{(\beta^2+1)^{1/2}-1}{\beta}\right]+\frac{x}{2}\left[\frac{1+\epsilon}{\epsilon}\right]-\frac{e^x+1}{t}\right\}\right]\ . \tag{27}$$

The corresponding lowest-order approximation given by van Kampen has the form

$$p_0(x,t)\sim\frac{1}{\sigma_1(t)\sqrt{2\pi\epsilon}}\exp\left\{-\frac{[x-\ln(1+t)]^2}{2\epsilon\sigma_1^2(t)}\right\} \tag{28}$$

in which

$$\sigma_1^2(t)=1-\frac{1}{(t+1)^2} \tag{29}$$

and the lowest-order approximation furnished by the development in [3] is

$$p_0(x,t)\sim\frac{e^x}{[2\pi\epsilon(2t+t^2)]^{1/2}}\exp\left\{-\frac{(e^x-1-t)^2}{2\epsilon(2t+t^2)}\right\}\ . \tag{30}$$

A comparison of the relative errors incurred through the use of Eqs. (28) and (30) is shown in Fig. 1. From the two curves in the figure it is evident that the van Kampen ap-

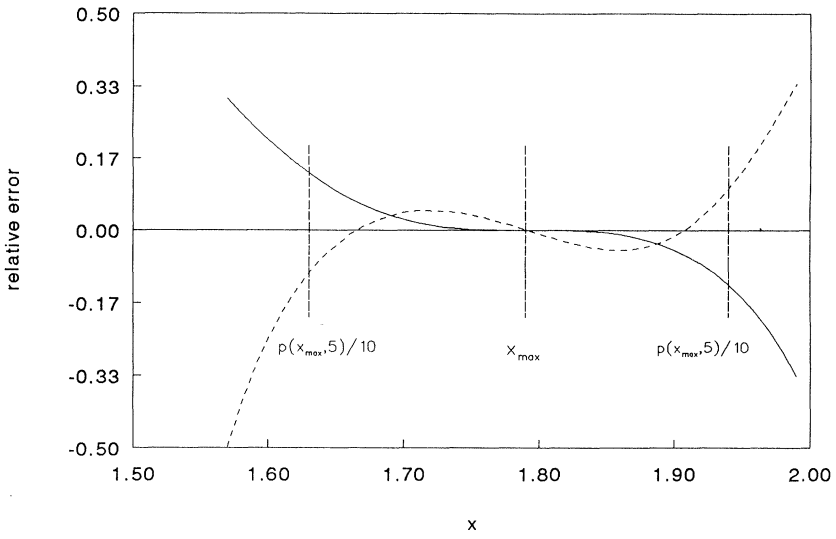

FIG. 1. A comparison of the relative error of the lowest-order approximations generated by the two perturbation techniques applied to the example in Eq. (23). The solid line denotes that produced by the van Kampen method while the dashed line corresponds to results produced by the method of running coordinates. The central of the three vertical lines indicates the position of the peak maximum, $x_{\max}$ and the two surrounding lines indicate the abscissas of the points at which $p(x,t)=p(x_{\max},t)/10$.

proximation is more accurate in the trailing edge of the cure and, more importantly, in the neighborhood of the peak, while the approximation produced by the theory of Weiss and Dishon is more accurate in the leading edge of the curve. We have not considered whether this holds true as well for initial conditions other than the δ function at $x=0$.

VI. DISCUSSION

We have described the simplest formulation of two techniques for generating approximate solutions to parabolic equations in a limit in which diffusive effects are small. Both techniques can be extended to lead to a perturbation expansion that goes in powers of $\epsilon^{1/2}$. The results of our limited investigation do not suggest that either expansion in its original formulation is uniformly superior to the other, but we have also shown that different extensions of the two expansions are possible which have not been explored to any great extent. It remains as an open question whether this flexibility can be exploited and, if so, how, knowing only the form of the Smoluchowski equation.

A number of investigators have proposed other methods for generating solutions to parabolic equations with weak diffusion effects. However, all of these methods generally reduce to one or the other of the two methods that have been described or do not allow one to develop a systematic perturbation expansion. For example, de Pasquale, Tartaglia, and Tombesi suggested what they termed the "quasideterministic approach which ... is the mapping between the original stochastic process and a new process which is associated with the initial condition" [9]. Moreover, they characterized their method as being a "non-perturbative expansion" [10]. However, their technique can be shown to be equivalent to the lowest order of the running coordinate approximation in the form of Eq. (19).

Two final comments should be made. First, both techniques are practical only for the solution of problems that involve diffusion in an unbounded space. When absorbing or reflecting boundaries are required in the underlying physical picture both techniques are generally inapplicable due to difficulties associated with boundary conditions. The difficulty in using the van Kampen method is traceable to the fact that no solution is known to the Ornstein-Uhlenbeck equation with time-dependent coefficients in the presence of boundaries. The difficulty with the technique based on moving coordinates is that any boundaries which, in the formulation of the problem in physical coordinates, are fixed in space become time dependent after the coordinate transformation indicated in Eq. (7). Our second comment is that we expect that neither of the two mentioned techniques provides an approximation that is uniform in time. This difficulty is not unique to the two formulations that we have discussed, but is rather a failing of all perturbative techniques that do not specifically incorporate a technique for extending their validity over longer periods of time. The work of Suzuki as summarized in [11] provides one (*ad hoc*) approach to overcoming that particular problem.

The problems remaining open in this general area relate to the incorporation of the effects of boundaries into the perturbation formalism and that of enlarging the range in time over which the perturbation expansion is able to furnish accurate results.

*Permanent address: Dept. of Physics, Bar-Ilan University, Ramat-Gan 52100, Israel.

[1] C. W. Gardiner, *Handbook of Stochastic Methods,* 2nd ed. (Springer-Verlag, Berlin, 1985).

[2] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amstrdam, 1981).

[3] G. H. Weiss and M. Dishon, Sep. Sci. **7**, 585 (1972); **8**, 337 (1973).

[4] G. H. Weiss and D. Rodbard, Sep. Sci. **7**, 217 (1972).
[5] *Handbook of Mathematical Functions,* edited by M. Abramowitz and I. A. Stegun (Dover, New York, 1965).
[6] A. Schenze and H. Brand, Phys. Rev. A **20**, 1628 (1979).
[7] M. Suzuki, K. Kaneko, and F. Sasagawa, Prog. Theor. Phys. **65**, 828 (1981).
[8] R. Graham and A. Schenze, Phys. Rev. A **25**, 1731 (1981).
[9] F. de Pasquale, P. Tartaglia, and P. Tombesi, Phys. Rev. A **31**, 2447 (1985).
[10] F. de Pasquale, P. Tartaglia, and P. Tombesi, Z. Phys. **43**, 353 (1981).
[11] M. Suzuki, Adv. Chem. Phys. **46**, 195 (1981).