The Diffusion Approximation for Markov Processes

N. G. van Kampen*

"Warem willest du dich von uns allen
Und unsrer Meinung entfernen?"
Ich schreibe nicht, euch zu gefallen,
Ihr sollt was lernen! – *Goethe*

Markov processes in physics, chemistry and biology are often regarded as generalized diffusion and accordingly described by a Kolmogorov equation or a Langevin equation. This procedure is exact if the Lindeberg condition is obeyed. It is here argued, however, that it is at best only approximately obeyed and that therefore the mathematical theory derived from it should be replaced with an approximation procedure. Such a procedure, in the form of a power series expansion in a parameter Ω, is described. The conclusion is that the diffusion approximation is inconsistent, except for a special subclass of Markov processes. This distinction has nothing to do with the distinction between discrete and continuous variables.

Markov Processes and the Diffusion Approximation

Let $Y(t)$ be a continuous time Markov process whose range consists of the real numbers. The range may be a *discrete* countable set, or it may be *continuous*, covering a finite or infinite interval on the real line. A combination of both may also occur, but will not be considered here.

For $t > t_0$, let $P(y,t|y_0,t_0)$ be the transition probability, *i.e.*, the probability (or probability density) of $Y(t)$ conditional on $Y(t_0) = y_0$. As a consequence of the Markov property it obeys the Chapman-Kolmogorov or master equation

$$\frac{\partial P(y,t|y_0,t_0)}{\partial t} = \mathbf{W}\cdot P(y,t|y_0,t_0), \qquad (1)$$

where $\mathbf{W}$ is a linear operator acting on the y-dependence. Its matrix element $W(y|y')$ for $y \neq y'$ is the transition probability per unit time from y' to y and for simplicity is here supposed to be independent of time.

The continuous range Markov processes contain a subclass of those for which $\mathbf{W}$ has the form of a differential operator

$$\frac{\partial P}{\partial t} = -\frac{\partial}{\partial y}a(y)P + \frac{1}{2}\frac{\partial^2}{\partial y^2}b(y)P, \qquad (2)$$

with two functions $a(y)$, $b(y)$ with $b(y) > 0$. It has been proved that higher-order differential operators cannot occur (Pawula, 1967). The conditions that characterize this subclass are

$$\lim_{\Delta t\to 0}\frac{\langle \Delta Y\rangle}{\Delta t} = a\,(y)\,;\ \lim_{\Delta t\to 0}\frac{\left\langle (\Delta Y)^2\right\rangle}{\Delta t} = b\,(y)\,;$$
$$\lim_{\Delta t\to 0}\frac{\langle (\Delta Y)^{\nu}\rangle}{\Delta t} = 0$$

where y is the value of Y at any time t, $\Delta Y = Y(t+\Delta t) - Y(t)$, the averages are taken with fixed y, the limits refer to $\Delta t \to 0$ and $\nu = 3, 4, \ldots$. Under these conditions, Kolmogorov (Kolmogorov, 1931) derived Eq. 2, which is called the (second) Kolmogorov equation[1], the (nonlinear) Fokker-Planck equation, or the (generalized) diffusion equation. The third condition was later replaced by the Lindeberg condition

$$\text{Prob.}\,\{|Y\,(t+\Delta t) - Y\,(t)| > \delta\} = O(\Delta t) \qquad (3)$$

for any $\delta > 0$ (Feller, 1971). Conversely, when Eq. 2 holds these conditions are satisfied.

Unfortunately the impeccable mathematics of Kolmogorov's proof have obscured the question of the validity of the Lindeberg condition for physical processes. Some physicists are under the assumption that Eq. 2 must hold whenever the range of Y is continuous (Horsthemke, Brenig, 1977). This misconception is abetted by the fact that the term *continuous process* has been used as the name for precisely that subclass of *continuous range processes* that obey the condition Eq. 3 (Gnedenko, 1962).

*University of Utrecht, Utrecht, Holland

Re-printed from: Thermodynamics & kinetics of biological processes. Edited by I. Lamprecht and A. I. Zotin. Walter de Gruyter & Co., New York, 1982. pp. 181-195.

[1] but this name has also been used for the general master equation, Eq. 1

We shall argue that *for physical processes the Lindeberg condition is at best only approximately obeyed*. Hence replacing **W** with a second-order differential operator is not just a mathematical identity but an approximation. To justify it one must perform a systematic expansion of the master equation. It then turns out that one is *not* led to Eq. 2, except in a special case treated at the end of this article. In other cases Eq. 2 is inconsistent, inasmuch as it contains terms of the same order in the expansion as those that are neglected by omitting higher derivatives.

Incidentally, it may be added that Eq. 2 is mathematically equivalent with the (nonlinear) Langevin equation

$$\dot{y} = a\left(y\right) + \sqrt{b\left(y\right) l\left(t\right)}$$

where $l\left(t\right)$ is Gaussian white noise. This equation may alternatively be written in the Itô form

$$dy = a\left(y\right) dt + \sqrt{b\left(y\right)}\, dW\left(t\right)$$

where $W(t)$ is a Wiener process. As a corollary to our result we therefore also find that this equation is at best only approximately fulfilled and only in a special case. It is merely a step in the systematic approximation scheme, rather than a fundamental starting point for fluctuation theory (Graham, 1973, 1978). Thus the discussion about the correct mathematical interpretation of the Itô equation (Mortensen, 1969) is moot, because the expansion of the master equation automatically leads to well-defined equations (van Kampen, 1981).

To avoid confusion we add the obvious remark that for any given stochastic process $Y(t)$ one may write

$$\dot{y} = a\left(y\right) + f(t)$$

thereby *defining* the stochastic term $f(t)$. But not only will $f(t)$ in general not be Gaussian white noise, also *its stochastic properties depend on those of* Y (Akcasu, 1977; Bedeaux, 1977; Agarwal *et al.*, 1978; Onuki, 1978). This device is therefore of no use in finding the fluctuation of Y.

Critique of the Lindeberg Condition

For a continuous range process one must have

$$P(y,t|y_0,t_0) \to \delta(y-y_0) \quad \text{as} \quad t \to t_0.$$

The question is: How is this "limit" approached, *i.e.*, how does $P(y,t|y_0,t_0)$ behave when $t - t_0 = \Delta t$ is small? There are two fundamentally different classes.

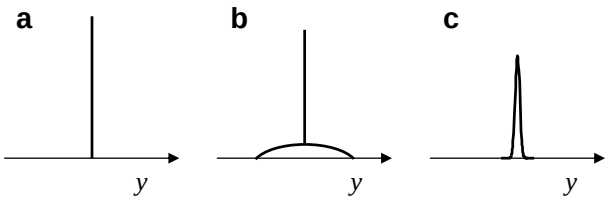

Figure 1: **Initial evolution of** $P(y, t_0 + \Delta t|y_0, t_0)$
a) Initial delta distribution at time $t = 0$.
b) Distribution at $\Delta t > 0$ for class 1 obeying Eq. 5.
c) Distribution at $\Delta t > 0$ for class 2 obeying Eq. 3.

1. Jumps in Y have some average size, independent of Δt, but the probability of a jump to occur becomes small as $\Delta t \to 0$. (The only difference with discrete range processes is that here the jump size is not restricted to discrete values.) For very small Δt the probability that a jump from y_0 to y occurs is $W(y|y_0)\Delta t$, and the probability that two or more jumps occur is of higher-order in Δt. The transition probability per unit time $W(y|y_0)$ is the probability distribution for jumps of various sizes. For any $\delta > 0$ and small Δt one has therefore

   $$\text{Prob.}\left\{|Y\left(t+\Delta t\right) - Y\left(t\right)| > \delta\right\} \propto \Delta t$$

   and therefore the Lindeberg condition, Eq. 3, does not hold.

2. Even for small Δt there are still very many, very small jumps in Y. The central limit theorem leads one to expect a Gaussian:

   $$P\left(y, t_0 + \Delta t|y_0, t_0\right) = \frac{1}{\sqrt{4\pi D \Delta t}} \exp\left[-\frac{\left(y-y_0\right)^2}{4D\Delta t}\right]$$

   with some constant D. In this case the Lindeberg condition is obeyed and of course this P does obey a differential equation

   $$\frac{\partial P}{\partial t} = D\frac{\partial^2 P}{\partial y^2}. \tag{4}$$

   However, this is the ordinary diffusion equation, which is homogeneous in y. The essential assumption implied by the Lindeberg condition is that one can choose Δt so small that: (a) many jumps occur in Δt; and (b) the width $\sigma^2 \Delta t$ covered by them is small enough that the dependence of the coefficients on y is not yet felt.

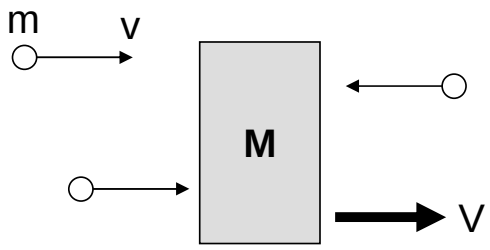

Figure 2: **The one-dimensional Rayleigh particle**

The difference between both classes is clarified in Figure 1. In class 1, after a small Δt the initial delta function is still visible since there is a chance that no jump took place. However, it has developed a foothill whose shape is $W(y|y_0)$ and accordingly the height of the delta function has decreased. In fact, to order Δt, one has

$$P\left(y, t_0 + \Delta t | y_0, t_0\right) = W\left(y|y_0\right) \Delta t$$
$$+\delta\left(y - y_0\right)\left[1 - \Delta t \int W\left(y'|y\right) dy'\right]. \qquad (5)$$

In class 2, however, the initial delta function starts to broaden immediately, and after any Δt, however small, it has changed into a narrow but smooth Gaussian-like peak. In other words, $Y(t)$ varies through infinitely many infinitely small jumps. Which picture applies to physics?

Here is an example of a physical Markov process with continuous range that belongs to class 1 and does not obey the Lindeberg condition (Figure 2). Let $Y(t)$ be the velocity V of a Brownian particle, in this connection often called the Rayleigh particle (Rayleigh, 1891; Hoare, Rahman, 1973). Let M be its mass, m the mass of the surrounding fluid molecules, T their temperature, and $v = \sqrt{kT/m}$ their thermal velocity. Obviously the number of collisions in any finite time interval will be finite (with probability one) and the jumps in V are of order

$$\frac{m}{M}\, v = \frac{1}{M}\sqrt{mkT}. \qquad (6)$$

For small intervals Δt the probability for a jump to occur is proportional to Δt, but the size of the jumps that do occur is determined by Eq. 6.

By similar considerations one can convince oneself that the Lindeberg condition is not satisfied either by the *position* of a Brownian particle, the current fluctuations in a resistor, the fluctuations of the electromagnetic field in a medium, etc. The basic reason is that fluctuations in physical processes are due to the particulate nature of matter, and the magnitude of the individual jumps is determined by the mass, charge, or spin of the particles involved and cannot therefore be infinitely small. It is obviously not satisfied by such processes as radioactive decay, emission and absorption of photons, and chemical reactions, because they have a discrete range. For the same reason problems in population dynamics and epidemiology do not satisfy Kolmogorov's requirement.

Linear and Nonlinear Problems

Write the transition probability $W(y|y')$ as a function $W(y'; \Delta y)$ of the starting point y' and the length $\Delta y = y - y'$ of the jump. As a function of Δy it is a more or less narrow peak, whose width is a measure for the size of the jumps. As to the dependence on y', we distinguish various cases.

1. *Homogeneous case*: $W(y'; \Delta y)$ for fixed Δy does not depend on y'. The only quantity with which the jump size can be compared has to be a property of the solution, *viz.*, the distance over which $P(y,t)$ varies appreciably. For homogeneous problems any solution becomes flat in the course of time, and hence Eq. 2 holds in the limit of long time. This is the reason why Einstein and Smoluchowski could use this equation for the position of a Brownian particle.

2. *Linear case*: $W(y'; \Delta y)$ depends linearly on y'. In this case the solutions tend typically to an equilibrium solution $P^{eq}(y)$ with a finite width. The validity of the diffusion approximation is now determined by the ratio of the jump size to the variation distance of $P^{eq}(y)$. This ratio is an inherent property of W; *i.e.*, it is determined by the parameters in W. In this case, therefore, a systematic approximation must be based on the expansion in powers of such a parameter. For instance, for the Rayleigh particle the equilibrium distribution is

$$P^{eq}\left(V\right) \propto \exp\left[-\frac{MV^2}{2kT}\right],$$

and has therefore a variation distance $\sqrt{kT/M}$. In order to ensure that the jump size, Eq. 6, is small, one is forced to consider values of M for

which $\sqrt{m/M} \ll 1$. This is the well-known condition for the validity of a Fokker-Planck equation for $P(V,t)$.

3. *Nonlinear case*: $W(y';\Delta y)$ is a nonlinear function of y'. This gives rise to the additional requirement that the jump size be small compared to the variation distance of W itself as a function of y'. Unless the jump size is much smaller one cannot hope to construct a general approximation method. Hence one has to find an expansion parameter by which one can reduce the jump size with respect to the scale inherent in the nonlinearity. Evidently that will lead in lowest approximation to a description of the fluctuations as if the system were linear. That is the reason for the success of the linear noise approximation in many nonlinear electronic devices. However, in order to include the effect of the nonlinearity on the fluctuation one has to go to the next approximation. Inevitably this also brings corrections due to the fact that the jumps are not infinitely small. As a result it turns out to be inconsistent to use Eq. 2 with nonlinear functions $a(y), b(y)$ without at the same time adding higher-order derivatives. These inconsistencies were responsible for the confusion in the early attempts to treat noise in nonlinear resistors (van Kampen, 1965).

   For example, the velocity fluctuations of a Rayleigh particle in equilibrium are described in the linear noise approximation by the familiar linear Fokker-Planck equation

$$\frac{\partial P\left(V,t|V_0,t_0\right)}{\partial t} = \gamma\left\{\frac{\partial}{\partial V}VP + \frac{kT}{M}\frac{\partial^2 P}{\partial V^2}\right\}. \quad (7)$$

   This corresponds to the linear damping law $\dot{V} = -\gamma V$. In the case of an ideal gas one can compute from kinetic theory

$$\gamma = \frac{4\rho A}{M}\left[\frac{2mkT}{\pi}\right]^{\frac{1}{2}}$$

   where A is the surface area of the particle and ρ the number density of the gas molecules. One can also find the first nonlinear correction

$$\dot{V} = -\gamma V\left[1 + \frac{m}{6kT}V^2 + \cdots\right]. \quad (8)$$

   This correction is of the relative order m/M which is the same parameter that measures the size of the jumps. In fact, we had to suppose $m/M \ll 1$ in order to approximate the master equation by a second order differential equation. *It is therefore inconsistent to modify Eq. 7 so as to take the nonlinear correction Eq. 8 into account without adding at the same time higher derivatives to account for the fact that the jumps are no longer infinitely small.* In fact, m/M is the only parameter in the problem, and one should therefore start from the master equation, expand systematically in powers of m/M, and expect to obtain successive nonlinear correlations and simultaneous higher derivative terms (van Kampen, 1961).

The Expansion of the Master Equation

This section describes the expansion of the master equation in general terms; for details the reader is referred to the original literature (van Kampen, 1961, 1976; McNeil, 1972; Kubo *et al.*, 1973; Kurtz, 1976; Fox, Kac, 1977). Our starting point is the master equation, Eq. 1. It is not necessary to differentiate between discrete and continuous range processes. For most actual cases in physics, chemistry, and biology it turns out that W involves a parameter Ω in such a way that for $\Omega \to \infty$ the magnitude of the individual jumps goes to zero relative to the distance in the Y-scale over which the nonlinearity becomes appreciable. In many cases, for instance the kinetics of chemical reactions, Ω is simply the size of the system, but in the example of the Rayleigh particle $\Omega \equiv M$. In electric circuits Ω is often the capacity of a condenser (van Kampen, 1960, 1961).

The precise condition on how W must depend on Ω is specified in the literature. For chemical reactions it roughly says that the kernel $W(y|y') \equiv W(y';\Delta y)$ must depend on the jump size Δy through an extensive variable, and on the starting point y' through an intensive variable. When no such parameter can be found the following expansion method cannot be used. Examples where no such parameter is available: the random velocity of one of the molecules of an ideal gas; the random walk of a diatomic molecule along its vibrational levels due to collisions (Montroll, Shuler, 1958).

When a parameter with the required properties is available it is possible to give a systematic expansion of the master equation in powers of $\Omega^{-\frac{1}{2}}$. In order to describe the result, we define the jump moments for $\nu = 1, 2, \ldots$ in the usual way

$$\begin{aligned}\alpha_\nu\left(y\right) &= \int \left(\Delta y\right)^\nu W\left(y;\Delta y\right) d\left(\Delta y\right) \\ &= \alpha_{\nu,0}\left(y\right) + \Omega^{-1}\alpha_{\nu,1}\left(y\right) + \ldots\end{aligned}$$

The expansion then leads to the following consequences.

1. The largest terms (lowest order in $\Omega^{-\frac{1}{2}}$) yield

$$\frac{\partial P(y,t)}{\partial t} = -\frac{\partial}{\partial y}\,\alpha_{1,0}(y)\,P. \tag{9}$$

This is the Liouville equation corresponding to the deterministic equation

$$\partial_t y = \alpha_{1,0}(y). \tag{10}$$

This equation contains no fluctuations because it is obtained in the thermodynamic limit $\Omega \to \infty$. It is therefore the macroscopic equation, which in the case of a chemical reaction is the familiar rate equation. Thus the deterministic, macroscopic, or phenomenological equation is a consequence of the master equation. Evidently it may well be nonlinear even though the master equation itself is always linear in its unknown variable P.

2. The next order in $\Omega^{-\frac{1}{2}}$ describes to lowest order the fluctuations of Y about the macroscopic value determined by Eq. 10. More precisely, let $\phi(t)$ be a solution of the macroscopic equation, Eq. 10. If one sets

$$y = \phi(t) + \Omega^{-\frac{1}{2}}\xi \tag{11}$$

the Ω-expansion gives the probability distribution $\Pi(\xi,t)$ of ξ

$$\frac{\partial \Pi(\xi,t)}{\partial t} = -\alpha'_{1,0}\left(\phi(t)\right)\frac{\partial}{\partial \xi}\xi\Pi + \frac{1}{2}\alpha_{2,0}\left(\phi(t)\right)\frac{\partial^2 \Pi}{\partial \xi^2}. \tag{12}$$

(The prime denotes the derivative of $\alpha_{1,0}$ with respect to its argument ϕ.) This is the Fokker-Planck equation whose arguments are linear in ξ but depend on time. The solution is a Gaussian and describes the fluctuations in the linear noise approximation. In the special case that ϕ is a stationary solution of Eq. 10, that is, $\alpha_{1,0}(\phi^{st}) = 0$, the coefficients of Eq. 12 are independent of time and it reduces to the familiar equation for equilibrium fluctuations, of which Eq. 7 is an example.

3. Higher-orders in $\Omega^{-\frac{1}{2}}$ add higher order terms to Eq. 12. For example, to order Ω^{-1}

$$\begin{aligned}\frac{\partial \Pi(\xi,t)}{\partial t} = \\ -\frac{\partial}{\partial \xi}\left[\alpha'_{1,0}\xi + \frac{\Omega^{-\frac{1}{2}}}{2}\alpha''_{1,0}\xi^2 + \frac{\Omega^{-1}}{3!}\alpha'''_{1,0}\xi^3\right]\Pi \\ +\frac{1}{2}\frac{\partial^2}{\partial \xi^2}\left[\alpha_{2,0} + \Omega^{-\frac{1}{2}}\alpha'_{2,0}\xi + \frac{\Omega^{-1}}{2}\alpha''_{2,0}\xi^2\right]\Pi \\ -\frac{1}{3!}\frac{\partial^3}{\partial \xi^3}\left[\Omega^{-\frac{1}{2}}\alpha_{3,0} + \Omega^{-1}\alpha'_{3,0}\xi\right]\Pi + \frac{\Omega^{-1}}{4!}\frac{\partial^4}{\partial \xi^4}\alpha_{4,0}\Pi \\ -\Omega^{-1}\left[\alpha'_{1,1}\frac{\partial}{\partial \xi}\xi\Pi + \frac{1}{2}\alpha_{2,1}\frac{\partial^2 \Pi}{\partial \xi^2}\right].\end{aligned} \tag{13}$$

All the $\alpha_{i,j}$'s are functions of $\phi(t)$. This equation leads to the following comments.

- The higher order terms do not only add nonlinear corrections to the coefficients of Eq. 12, but *at the same time higher derivatives appear*. At no stage of the expansion does the nonlinear Fokker-Planck equation, Eq. 2, emerge: it is inconsistent to take the nonlinearity of $a(y)$ and $b(y)$ into account while neglecting higher derivatives.
- The fluctuations are no longer Gaussian, and the equilibrium auto-correlation function is no longer a simple exponential. Rather, each order of Ω^{-1} adds one more exponential, and therefore one more Debye relaxation term to the spectral density of the fluctuations.
- To order Ω^{-1} it is no longer true that $\langle\xi\rangle = 0$, and hence $\langle Y(t)\rangle$ no longer coincides with the solution $\phi(t)$ of the macroscopic equation. To put it differently, $\langle Y(t)\rangle$ is not a solution of the macroscopic equation, Eq. 10. Nor can one find some modified differential equation for $\langle Y(t)\rangle$ that includes the corrections of order Ω^{-1}. Instead one has to solve two coupled equations for the average and the variance of $Y(t)$ (Kubo *et al.*, 1973).
- As mentioned before, the fourth order equation, Eq. 13, cannot serve as an exact master equation of any Markov process. It is therefore not the exact equation for a Markov process that in some way approximates the original process; rather it is an approximate equation for the exact P. Only those solutions of Eq. 13 are meaningful for which the higher derivative terms, not included in Eq. 13, are actually small. Taken at

face value, the equation has also other solutions, which may become negative or infinite, but they are not approximations to the actual solution of Eq. 1. The safest way to avoid them is to treat the higher order terms in Eq. 13 by means of perturbation theory (Siegel, 1960).

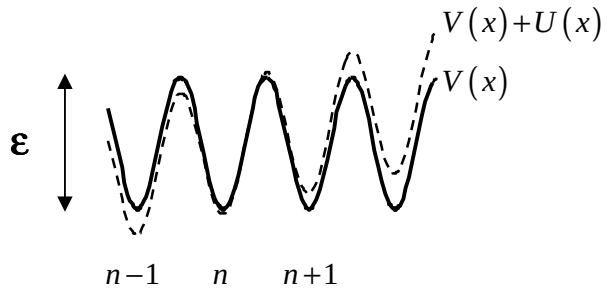

Figure 3: **The potential for a particle diffusing through a crystal.**

The Diffusion Type

The Ansatz, Eq. 11, is based on the anticipation that the fluctuations will be of relative order $\Omega^{-\frac{1}{2}}$. Its *a posteriori* justification is provided by the fact that according to Eq. 12 the vales of ξ are in fact of order unity. Nonetheless it can happen that they grow in time and become so large that the formal expansion in powers of $\Omega^{-\frac{1}{2}}$ becomes illusory after a certain initial period. It is easily seen from Eq. 12 that this will not be so provided that $\alpha'_{1,0}(\phi(t)) < 0$. This inequality is also the condition that the solution $\phi(t)$ of Eq. 10 is stable (asymptotically stable in the sense of Lyapunov). When the macroscopic solution is stable the fluctuations do not grow and the Ω-expansion can be used for all times. The physical reason for this is clear. For the precise statement of this criterion one has to distinguish between local and global stability, but these complications are not relevant here. Nor do we want to discuss the various cases in which the criterion is violated, such as critical points (Kubo *et al.*, 1973) and bistable systems (Suzuki, 1977, 1978; van Kampen, 1978), but only one special case.

If it happens that $\alpha_{1,0}(y) \equiv 0$, the macroscopic solutions are: $\phi(t) =$ constant. They are not asymptotically stable, not even locally. According to the linear noise approximation, Eq. 12, the fluctuations grow proportionally with $\sqrt{t}$, just as in Brownian motion. The above expansion can therefore be used only for a limited time, roughly for $t < \Omega$. For master equations of this type, however, the Ω-expansion takes a different form.

The point is that it is no longer true that Eq. 10 is the leading term in the expansion, because this term vanishes. Consequently, it is no longer possible to split off a macroscopic part as in Eq. 11, and this affects the orders of the remaining terms. Rearranging them accordingly one now obtains as the leading term in the Ω-expansion of the master equation (van Kampen, 1977)

$$\frac{\partial P}{\partial t} = \Omega^{-1}\left[-\frac{\partial}{\partial y}\alpha_{1,1}(y)\,P + \frac{1}{2}\frac{\partial^2}{\partial y^2}\alpha_{2,0}(y)\,P\right], \tag{14}$$

which is precisely the generalized diffusion equation, Eq 2. The variation occurs on a slower time scale than in Eq. 9, because it is now purely a net effect of fluctuations, rather than due to a concerted drift.

To show that the occurrence of $\alpha_{1,0} \equiv 0$ is not just a freak, consider the following model for diffusion in a solid (Figure 3). A particle moves in a periodic potential $V(x)$ by jumping from each equilibrium site n to either $n+1$ or $n-1$ with probabilities per unit time a_n and b_n. Thus it performs a continuous time random walk with jump probability per unit time

$$W(n|n') = a_{n'}\delta_{n,n'-1} + b_{n'}\delta_{n,n'+1}$$
$$a_n = b_n = A\exp[-\varepsilon/kT].$$

Now let the particles be charged and subject to an external field U. The fact that U is supposed to vary on a macroscopic scale is expressed by writing it as a function of $y = n/\Omega$, where Ω is the number of sites n per cm. The field $U(y)$ modifies the height of the potential peaks, so that the new jump probabilities are

$$a_n = A\exp\left\{-\frac{1}{kT}\left[\varepsilon + U\left(\frac{n-\frac{1}{2}}{\Omega}\right) - U\left(\frac{n}{\Omega}\right)\right]\right\}$$
$$b_n = A\exp\left\{-\frac{1}{kT}\left[\varepsilon + U\left(\frac{n+\frac{1}{2}}{\Omega}\right) - U\left(\frac{n}{\Omega}\right)\right]\right\}.$$

Hence the first jump moment is

$$\alpha_1(y) = -a_n + b_n = A\exp\left[-\frac{\varepsilon}{kT}\right]\times$$
$$\times\left\{\exp\left[\frac{U'(y)}{2kT\Omega} + O\left(\Omega^{-2}\right)\right] - \exp\left[\frac{-U'(y)}{2kT\Omega} + O\left(\Omega^{-2}\right)\right]\right\}.$$

On expanding the exponential one sees that the lowest order $\alpha_{1,0}(y)$ cancels so that the diffusion approximation applies.

To find the explicit formula we need the next term

$$\Omega^{-1}\alpha_{1,1}(y) = A\exp\left[-\frac{\varepsilon}{kT}\right] \times \frac{U'(y)}{kT\Omega}.$$

Also one has

$$\alpha_2(y) = a_n + b_n = 2A\exp\left[-\frac{\varepsilon}{kT}\right] + O\left(\Omega^{-2}\right).$$

Hence the diffusion equation, Eq. 14, has the form

$$\frac{\partial P}{\partial \tau} = -\frac{\partial}{\partial y}U'(y)P + kT\frac{\partial^2 P}{\partial y^2}, \qquad (15)$$

where the constant factor $(A/\Omega kT)\exp[-\varepsilon/kT]$ has been absorbed in the time scale. The conclusion is that Eq. 15 is a consistent first approximation for diffusion in an external field $U(y)$ when the jump size is small compared to the variation distance of $U(y)$. In the limit it is of course exact (Feller, 1968).

The example indicates the basic reason for the applicability of the diffusion approximation. In contrast with *e.g.* the Rayleigh particle here the nonlinearity is due to some external agency, physically distinguished from the noise source. That makes it possible to choose Ω in such a way that for large Ω the jumps become arbitrarily small compared to the variation length of the nonlinear field U. Another example is a nonlinear pendulum with damping and fluctuations caused by the surrounding air. It has not yet been investigated whether or not the nonlinearities in hydrodynamics can be disentangled in this way from the fluctuations associated with the transport coefficients.

Summary

1. In physical situations that can be described by a continuous range Markov process the Lindeberg condition is not satisfied. One can therefore not conclude as a mathematical theorem that the diffusion equation, Eq. 2, must hold, (nor that the corresponding Langevin equation holds) by means of a systematic method. The Ω-expansion appears to be rather universally applicable.

2. For the Ω-expansion no distinction need be made between continuous and discrete ranges. The first term yields the macroscopic equation, Eq. 10, and the next one the linear Fokker-Planck equation, Eq. 12, for the fluctuations. Higher-orders add nonlinear terms to the coefficients of this equation but simultaneously also higher derivatives. The result is valid for all times provided that $\alpha'_{1,0} < 0$.

3. At no stage does Eq. 2 appear: it is not a consistent approximation in this expansion. Of course Eq. 2 may lead to correct results, namely if one merely uses it in the linear approximation, such as to make use only of those features of it which coincide with the linear noise approximation. But it is incorrect to take it seriously beyond that, and for instance to conclude from it

$$P^{eq}(y) = \frac{\text{constant}}{b(y)}\exp\left[-2\int^{y}\frac{a(y')}{b(y')}dy'\right]. \quad (16)$$

4. In the special case $\alpha'_{1,0} \equiv 0$ however, the lowest term in the Ω-expansion does have the form of Eq. 2, or more precisely Eq. 14. Typically this happens when the nonlinearity stems from a different physical cause than the fluctuations, so that both can be disentangled by a suitable choise of Ω. In this case, Eq. 16 is valid apart from corrections of order Ω^{-1}.

References

[1] R. F. Pawula. Approximation of the linear Boltzmann equation by the Fokker-Planck equation. *Physical Review*, 162:186–188, 1967.

[2] A. N. Kolmogorov. Uber die analytischen Methoden in der Wahrscheinlichkeitsrechnung. *Mathematische Annalen*, 104:415–458, 1931.

[3] W. Feller. *An Introduction to Probability Theory and Its Applications, Volume 2.* John Wiley & Sons, 1971.

[4] W. Horsthemke and L. Brenig. Nonlinear Fokker-Planck equation as an asymptotic representation of the master equation. *Zeitschrift fur Physik B*, 27:341–348, 1977.

[5] B. V. Gnedenko and A. Ya. Khinchin. *An Elementary Introduction to the Theory of Probability.* Dover Publications, 1962.

[6] R. E. Mortensen. Mathematical problems of modeling stochastic nonlinear dynamic systems. *Journal of Statistical Physics*, 1:271–296, 1969.

[7] N. G. van Kampen. The validity of nonlinear Langevin equations. *Journal of Statistical Physics*, 25:431–442, 1981.

[8] Z.A. Akcasu. Fluctuations of macrovariables in nonlinear systems - Langevin equation approach. *Journal of Statistical Physics*, 16:33–58, 1977.

[9] D. Bedeaux. Equivalence of the master equation and the Langevin equation. *Physics Letters A*, 62:10–12, 1977.

[10] A. Onuki. Langevin equation with multi-Poissonian noise. *Journal of Statistical Physics*, 19:325–332, 1978.

[11] S. J. W. Rayleigh. Dynamical problems in illustration of the theory of gases. *Philosophical Magazine*, 32: 424–45, 1891.

[12] M. R. Hoare and M. Rahman. Spectral theory of Rayleigh's piston: 1. Discrete spectrum. *Journal of Physics A*, 6: 1461–1478, 1973.

[13] N. G. van Kampen. in *Fluctuation Phenomena in Solids*, R. E. Burgess ed. Academic Press, 1965.

[14] N. G. van Kampen. A power series expansion of the master equation. *Canadian Journal of Physics*, 39:551, 1961.

[15] N. G. van Kampen. The expansion of the master equation. *Advances in Chemical Physics*, 34:245, 1976.

[16] R. Kubo, K. Matsuo and K. Kitahara. Fluctuation and relaxation of macrovariables. *Journal of Statistical Physics*, 9:51–96, 1973.

[17] T. G. Kurtz. Limit theorems and diffusion approximations for density dependent Markov chains. *Math Prog Stud*, 9:51–96, 1976.

[18] N. G. van Kampen. Non-linear thermal fluctuations in a diode. *Physica*, 26:585–604, 1960.

[19] E. W. Montroll and K. E. Shuler. The application of the theory of stochastic processes to chemical reactions. *Advances in Chemical Physics*, 1:361-399, 1958.

[20] A. Siegel. Differential-operator approximations to the linear Boltzmann equation. *Journal of Mathematical Physics*, 1:378–390, 1960.

[21] M. Suzuki. Anomalous fluctuation and relaxation in unstable systems. *Journal of Statistical Physics*, 16:477–504, 1977.

[22] M. Suzuki. Theory of instability, nonlinear Brownian motion and formation of macroscopic order. *Physics Letters A*, 67:339–341, 1978.

[23] N. G. van Kampen. Escape and splitting probabilities in diffusive and nondiffusive Markov processes. *Progress of Theoretical Physics - Supplement*, 64:389–401, 1978.