Simple procedure for correcting equations of evolution: Application to Markov processes

Gregory Ryskin
Department of Chemical Engineering, Northwestern University, Evanston, Illinois 60208
(Received 28 May 1996)

A general procedure is proposed for correcting evolution equations, arising in different branches of science. Its application to Markov processes shows that the coefficients of the third- and higher-order derivatives in the Kramers-Moyal expansion are, in general, not small; nevertheless, the macroscopic-time evolution of the process is completely described by a differential equation of second order. For Brownian motion, this equation is Galilean invariant, while the Fokker-Planck equation is not. Finally, a correction is derived for the master equation. [S1063-651X(97)01411-6]

PACS number(s): 05.40.+j, 02.50.Ga, 02.20.Mp

INTRODUCTION

In this paper, a general procedure is proposed for correcting evolution equations, arising in different branches of science. It is based on the fact that the time interval, which is viewed as infinitesimal in the construction of the evolution operator, normally has an inherent lower bound, below which a different, more microscopic level of description is required. The proposed procedure improves the accuracy of the evolution operator without changing the level of description. Using this procedure and the concept of τ invariance introduced below, an evolution equation is derived for Markov processes without the usual assumptions concerning the moments of fluctuations. (Some of these assumptions are shown to be false, while others are superfluous.) The derivation shows that the coefficients of the third- and higher-order derivatives in the Kramers-Moyal expansion are, in general, not small; nevertheless, the macroscopic-time evolution of the process is completely described by a differential equation of second order. This surprising development parallels closely the emergence of the Gaussian distribution in the central limit theorem and resolves the long-standing problem of truncating the Kramers-Moyal expansion. For Brownian motion, the evolution equation derived below is Galilean invariant, while the Fokker-Planck equation is not: The expression for the diffusion coefficient in the latter yields different values in different Galilean frames.

Finally, the concept of the transition rate is shown to lack self-consistency, and a correction is derived for the master equation in the form of an operator expansion. The master equation is shown to be an approximate substitute for the differential evolution equation, which amounts to a complete reversal of the conventional view.

EQUATIONS OF EVOLUTION AND SEMIGROUPS

The behavior of a physical system is often described by an evolution equation

$$\frac{\partial f(x,t)}{\partial t}=A[f(x,t)], \tag{1}$$

where A is a linear operator, which does not depend on time. The determination of A is usually the principal task of the theory. With the evolution equation (1) is associated a one-parameter semigroup [1] of linear operators W_τ such that $f(x,t)=W_t[f(x,0)]$, with the semigroup property $W_{\tau_1+\tau_2}=W_{\tau_1}W_{\tau_2}$ and the identity operator $I=W_0$. The evolution operator A is then the *infinitesimal generator* of this semigroup, so that [1]

$$A=\lim_{\tau\to 0}\frac{W_\tau-I}{\tau} \tag{2}$$

and also

$$W_\tau=\exp(\tau A). \tag{3}$$

The evolution of the system at short times is often found via some perturbation approach and then used to determine A as indicated by Eq. (2), even though the semigroup operator W_τ may not have been introduced explicitly.

PROPOSAL

Here I want to stress the following: In every theoretical model there is normally an inherent lower bound for τ, below which the model breaks down and a more microscopic level of description becomes a necessity. Thus, within the model, the limit $\tau\to 0$ is physically impossible, and therefore the evolution operator actually obtained is

$$B=\frac{W_\tau-I}{\tau}, \tag{4}$$

where τ is small in some sense, usually in the sense of being much smaller than the typical macroscopic time interval of interest in the problem. This smallness of τ is then assumed to guarantee that B is a sufficiently good approximation for A. It would certainly be useful to find some means of estimating the error involved and of improving the approximation if necessary.

The following observation delivers both: Eq. (4) can be rewritten as $W_\tau=I+\tau B$ and then Eq. (3) yields $A=\tau^{-1}\ln(I+\tau B)$, i.e., the expansion

$$A=B-\frac{\tau}{2}B^2+\frac{\tau^2}{3}B^3-\cdots, \tag{5}$$

which can be used for evaluating and improving the approximation of the evolution operator.

In some cases, a more direct approach will be possible: Knowing W_τ for some small τ may allow one to determine A *exactly* by solving Eq. (3) instead of first constructing the approximation B and then correcting it via the expansion (5). All the results below, except Eq. (36), are obtained in this way.

It should be noted that the time-independence restriction, imposed on A in the beginning, is not crucial: It is quite obvious that variation of A or B on macroscopic time scales τ_{macro} will not alter the validity of Eqs. (3) and (5) when $\tau \ll \tau_{\text{macro}}$. Moreover, since some of the basic results of the semigroup theory remain valid for nonlinear operators [1], the expansion (5) can be expected to work even for nonlinear operators A and B. The discussion below will be confined to linear operators, however.

MARKOV PROCESSES: PRELIMINARIES

As an application of the above proposal, let us consider a Markov process. Let $w(x,t;\xi,\tau)d\xi$ denote the (transition) probability that the system (say, a Brownian particle), which was at the point (generally, in the state) x at time t, will be found between $x+\xi$ and $x+\xi+d\xi$ at the time $t+\tau$. This implies the assumption that, in order to predict (probabilistically) the position of the particle at time $t+\tau$, it will be sufficient to know *only* its position x at time t and not, say, also its velocity, or its positions at previous times, etc. (the Markov property). For this to be true, the time interval τ should be chosen according to Einstein's prescription (see [2] and [3], p. 3): τ must be ''so large that the motions performed by a particle during two consecutive time intervals τ may be considered as mutually independent events.'' Then, following Einstein, one can write for the probability density $p(x,t)$ of the stochastic variable x,

$$p(x,t+\tau)=\int d\xi\; p(x-\xi,t)w(x-\xi,t;\xi,\tau). \tag{6}$$

The limits of integration implied in Eq. (6) and throughout the paper are $-\infty$ to ∞. The transition probability density satisfies the normalization condition $\int d\xi\; w(x,t;\xi,\tau)=1$.

As pointed out by Einstein, Eq. (6) holds true only for sufficiently large τ, i.e., for $\tau \gg \tau_{\text{corr}}$, where τ_{corr} denotes the time scale of correlation between consecutive *changes* of x; otherwise the process ceases to be Markov. If x is the position of a Brownian particle, τ_{corr} is the velocity autocorrelation time (see [3], p. 45, and [4], pp. 74–76 and 206); for a large (a few micrometers) Brownian particle, Eq. (6) then holds true only for τ of order 1 sec and greater. The limit $\tau \to 0$, usually considered in order to accomplish the transition to a differential equation, should be understood to mean only that $\tau \ll \tau_{\text{macro}}$, where τ_{macro} is the macroscopic time scale that characterizes variation with time t of $p(x,t)$ and $w(x,t;\xi,\tau)$.

Let us introduce characteristic functions in the usual way

$$\psi(k,t)\equiv\int dx\; e^{ikx}p(x,t), \tag{7}$$

$$\phi(x,t;k,\tau)\equiv\int d\xi\; e^{ik\xi}w(x,t;\xi,\tau). \tag{8}$$

The characteristic function of the transition probability can be expressed via the moments of the fluctuations $\mu_m(x,t;\tau)\equiv\int d\xi\; \xi^m w(x,t;\xi,\tau)$ as

$$\phi(x,t;k,\tau)=1+\sum_{m=1}^{\infty}\frac{(ik)^m}{m!}\,\mu_m(x,t;\tau). \tag{9}$$

The alternative expression involves the cumulants of the fluctuations $\kappa_m(x,t;\tau)$,

$$\phi(x,t;k,\tau)=\exp\left[\sum_{m=1}^{\infty}\frac{(ik)^m}{m!}\,\kappa_m(x,t;\tau)\right], \tag{10}$$

where $\kappa_1=\mu_1=\langle\xi\rangle$, $\kappa_2=\mu_2-\mu_1^2=\langle\xi^2\rangle-\langle\xi\rangle^2=\sigma^2$, $\kappa_3=\mu_3-3\mu_1\mu_2+2\mu_1^3$, etc. [4,5]; $\langle\xi\rangle$ and σ^2 are the mean and the variance of the fluctuations, respectively.

Let us first consider the case of a spatially and temporally homogeneous process $w=w(\xi,\tau)$. (This case will lead eventually to an evolution equation with constant coefficients.) Equation (6) takes the form

$$p(x,t+\tau)=\int d\xi\; p(x-\xi,t)w(\xi,\tau)\equiv W_\tau[p(x,t)],$$

which is a convolution; hence we can write

$$\psi(k,t+\tau)=\phi(k,\tau)\psi(k,t). \tag{11}$$

We also have the Chapman-Kolmogorov equation [3–5]

$$w(\xi,\tau_1+\tau_2)=\int d\xi' w(\xi-\xi',\tau_2)w(\xi',\tau_1), \tag{12}$$

which means that

$$W_{\tau_1+\tau_2}=W_{\tau_1}W_{\tau_2}, \tag{13}$$

$$\phi(k,\tau_1+\tau_2)=\phi(k,\tau_1)\phi(k,\tau_2). \tag{14}$$

Once the transition probability $w(\xi,\tau)$, or the operator W_τ, for an arbitrarily chosen (within the appropriate limits) τ is known, the evolution of the Markov process is completely determined. The converse is not true: To a single Markov process there corresponds an infinite family of probabilities $w(\xi,\tau)$ or operators W_τ.

τ INVARIANCE

Choosing a particular value for τ is akin to choosing a particular coordinate system in space-time or choosing a particular value for the phase of a wave function. In all such cases, making the choice is necessary in order to specify the physical situation precisely, but the choice is arbitrary because the physics itself does not depend on it. Normally, this means that the mathematical description of the phenomenon must be invariant under arbitrary variation of that choice, a requirement of the most fundamental importance. I now ask the following question: Can one find a form of description of a Markov process that is *manifestly invariant* under arbitrary

change of τ (a τ-*invariant* description)?

Once asked, this question is answered easily. Equations (10) and (14) imply

$$\kappa_m(\tau_1+\tau_2)=\kappa_m(\tau_1)+\kappa_m(\tau_2), \tag{15}$$

which means that, for all $m=1,2,3,\ldots,\infty$,

$$\kappa_m(\tau)=\alpha_m\tau, \tag{16}$$

where the quantities α_m are independent of τ. The complete set $\{\alpha_m\}$ thus provides a complete τ-invariant description of a given Markov process; i.e., there is one-to-one correspondence between a given process and the associated set $\{\alpha_m\}$.

Once a particular τ is chosen, the usual forms of description are easily recovered since

$$\phi(k,\tau)=\exp\left[\tau\sum_{m=1}^{\infty}\frac{(ik)^m}{m!}\alpha_m\right]. \tag{17}$$

The quantities α_m have dimensions $[x]^m/[t]$ and can be called *rate invariants* of the fluctuations.

EVOLUTION EQUATION

Let us now turn to the derivation of the differential evolution equation for $p(x,t)$, associated with the semigroup (13) and (14). Comparing Eqs. (3) and (17), we see that the Fourier image of the evolution operator is

$$\sum_{m=1}^{\infty}\frac{(ik)^m}{m!}\alpha_m, \tag{18}$$

which leads immediately to the evolution equation for $p(x,t)$, viz.,

$$\frac{\partial p}{\partial t}=\sum_{m=1}^{\infty}(-1)^m\frac{\partial^m}{\partial x^m}\left(\frac{\alpha_m}{m!}p\right). \tag{19}$$

[The coefficients $\alpha_m/m!$ could be put in front of the differentiation signs; the above form is preferable because it remains valid for a general process where $\alpha_m=\alpha_m(x,t)$, as we shall see below.]

For macroscopic times, one crucial simplification is possible. It follows from Eq. (11) that

$$\psi(k,t+n\tau)=[\phi(k,\tau)]^n\psi(k,t). \tag{20}$$

Since $\tau_{\rm corr}\ll\tau\ll\tau_{\rm macro}$, the evolution of the system over macroscopic time intervals is determined by Eq. (20) with $n\gg1$. But as $n\to\infty$, $[\phi(k,\tau)]^n\to[\gamma(k,\tau)]^n$, where $\gamma(k,\tau)\equiv\exp(ik\langle\xi\rangle)\exp(-k^2\sigma^2/2)$ is the characteristic function of a Gaussian with the same values of the mean and the variance as the actual transition probability. The proof of this statement is identical to the proof of the central limit theorem (see, e.g., [6], p. 383, and [7], p. 68).

This means that for macroscopic time intervals only the first two invariants α_1 and α_2 are important; *the values of all the rest, $\alpha_3,\alpha_4,\alpha_5,\ldots$, are immaterial*. (Note that no assumption has been made about the smallness of α_m for $m\geqslant3$; these quantities are determined by the particular Markov process and can have any values whatsoever.) Thus, for macroscopic time intervals, we are free to *replace* the actual Markov process with a different one, provided only it has the same α_1 and α_2. The most convenient replacement has $\alpha_m=0$ for $m\geqslant3$, i.e., its transition probability is Gaussian. The evolution equation for $p(x,t)$ then becomes

$$\frac{\partial p}{\partial t}=-\frac{\partial}{\partial x}(\alpha_1 p)+\frac{\partial^2}{\partial x^2}\left(\frac{\alpha_2}{2}p\right). \tag{21}$$

The emergence of the second-order differential equation (21) as a description of the macroscopic-time evolution of a Markov process is thus entirely analogous to the emergence of the Gaussian distribution in the central limit theorem.

The evolution operator in Eq. (21) is τ invariant. Provided one can evaluate the mean $\langle\xi\rangle$ and the variance σ^2 of the fluctuations taking place during an arbitrarily chosen (within the appropriate limits) interval τ, its coefficients can be calculated as $\alpha_1=\langle\xi\rangle/\tau$ and $\alpha_2/2=\sigma^2/2\tau$. Repeating this calculation for different values of τ should yield the same result; this provides a self-consistency check.

COMPARISON WITH PREVIOUS WORK

We have followed the most direct route, indicated by Eq. (3). All the literature on the subject follows instead the route described by Eq. (2). That is, one obtains for the Fourier image of A, using Eq. (9),

$$\sum_{m=1}^{\infty}\frac{(ik)^m}{m!}\lim_{\tau\to0}\frac{\mu_m}{\tau}, \tag{22}$$

which leads to the Kramers-Moyal expansion [4,5,8]

$$\frac{\partial p}{\partial t}=\sum_{m=1}^{\infty}(-1)^m\frac{\partial^m}{\partial x^m}\left[\frac{1}{m!}\left(\lim_{\tau\to0}\frac{\mu_m}{\tau}\right)p\right]. \tag{23}$$

[Formally, Eq. (22) is equivalent to Eq. (18) because Eq. (16), in conjunction with the standard formulas expressing the moments via the cumulants [5], shows that, as $\tau\to0$, $\mu_m=\kappa_m+O(\tau^2)$.] Truncated after the second-derivative term, expansion (23) produces the Fokker-Planck, or forward Kolmogorov, equation.

Lacking the insight into the true nature of the expansion coefficients, the existing derivations justify the truncation by *assuming* that for $m\geqslant3$ these coefficients are zero or negligibly small [9,7,5]. As discussed before, this assumption is false; this serious inconsistency has been known for some time (see [4], p. 199, and [8], p. 233), but remained unresolved. The other commonly made assumptions, viz., that $\lim_{\tau\to0}\mu_1/\tau$ and $\lim_{\tau\to0}\mu_2/\tau$ exist, are seen to be superfluous, as is the limit $\tau\to0$ itself.

In applications, the values of the coefficients in the Fokker-Planck equation are often found by requiring that the equation produce correct results for the cases where $p(x,t)$ is known, e.g., in thermodynamic equilibrium. However, there exist situations where the connection of the coefficients with fluctuations is important, and then new problems arise. Since the limit $\tau\to0$ is physically impossible, the values of μ_m/τ that are actually used correspond to some small but finite τ (see [4], pp. 195 and 196). The drift coefficient μ_1/τ happens to equal α_1 exactly because $\mu_1=\kappa_1=\langle\xi\rangle$, but the dif-

fusion coefficient is calculated incorrectly whenever a finite τ is used. Namely, one obtains $\langle \xi^2 \rangle / 2\tau$ instead of the correct value $\sigma^2/2\tau$. (The two values coincide in the absence of the drift, however.) The error is likely to be negligible when τ is very small, but can be significant in problems with large τ_{corr}, such as motion of large Brownian particles.

More importantly, the Fokker-Planck expression $\langle \xi^2 \rangle / 2\tau$ for the diffusion coefficient *destroys the Galilean invariance* of the evolution equation. Let a Brownian motion be observed from two different Galilean frames F and F', where F' moves with respect to F with constant velocity v, directed along the x axis. The Galilean transformation of the coordinates is then $x' = x - vt$, $t' = t$. The probability distribution of the particle's position is a Galilean scalar $p'(x',t') = p(x,t)$. The evolution equation satisfied by this distribution is Galilean invariant if it takes the same form in both F and F'. It is not hard to show that this will be the case if and only if the drift coefficient transforms as velocity, while the diffusion coefficient is a Galilean scalar. Since $\xi' = \xi - v\tau$, so that $\langle \xi' \rangle / \tau = \langle \xi \rangle / \tau - v$, the drift coefficient $\langle \xi \rangle / \tau$ does transform properly. But the diffusion coefficient will transform properly only if the expression $\sigma^2/2\tau$ is used for it: It is easy to check that $\sigma'^2 = \sigma^2$, while $\langle \xi'^2 \rangle \neq \langle \xi^2 \rangle$. That is, the Fokker-Planck expression for the diffusion coefficient yields different values in different Galilean frames. The difference can be arbitrarily large because it depends on the relative velocity v.

Note also that the fundamental solution of Eq. (21) is the normal distribution with the variance $\alpha_2 t = (t/\tau)\sigma^2$, i.e., the variance of the sum of t/τ independent jumps equals the sum of their variances, in perfect agreement with one of the basic theorems of probability. The corresponding Fokker-Planck result is the sum of the second moments, which is wrong.

We see that in the Einstein relation connecting fluctuations with the diffusion coefficient (and through it with the viscosity, or dissipation), the mean square of the fluctuations must be replaced by their variance. There is little doubt that the same is true for other fluctuation-dissipation relations as well.

It should be possible to correct the Fokker-Planck equation *a posteriori* by including one more term of the expansion (5), truncated after the second derivative. The resulting correction is

$$-\frac{\tau}{2}\left(-\frac{\langle \xi \rangle}{\tau}\frac{\partial}{\partial x}\right)^2 = -\frac{\langle \xi \rangle^2}{2\tau}\frac{\partial^2}{\partial x^2}; \tag{24}$$

this does lead to Eq. (21) because $\langle \xi^2 \rangle - \langle \xi \rangle^2 = \sigma^2$.

THE INHOMOGENEOUS CASE

Our discussion so far has been confined to homogeneous processes; now we relax this restriction and allow w and ϕ to vary with x and t on *macroscopic* length and time scales. [A reader who wants to skip this part should move to Eq. (32).] It will be convenient to use $y \equiv x - \xi$; then Eqs. (6) and (8) become

$$p(x,t+\tau) = \int dy\ p(y,t) w(y,t;x-y,\tau), \tag{25}$$

$$\phi(y,t;k,\tau) \equiv \int dx\ e^{ik(x-y)} w(y,t;x-y,\tau). \tag{26}$$

Multiplying both sides of Eq. (25) by e^{ikx}, integrating over x, and then denoting by x the remaining integration variable, one obtains

$$\psi(k,t+\tau) = \int dx\ e^{ikx} \phi(x,t;k,\tau) p(x,t). \tag{27}$$

This can be written as

$$\psi(k,t+\tau) = \Phi_\tau(t)[\psi(k,t)], \tag{28}$$

$$\Phi_\tau(t)[\ ] \equiv \int dx\ e^{ikx} \phi(x,t;k,\tau) \mathfrak{F}^-[\ ], \tag{29}$$

where $\mathfrak{F}^-$ stands for inverse Fourier transform. For a spatially homogeneous process, the linear operator $\Phi_\tau(t)$ reduces to multiplication by $\phi(t;k,\tau)$.

Consider Eqs. (28) and (29) and write $\phi(x,t;k,\tau)$ in the form (17), where now $\alpha_m = \alpha_m(x,t)$. If a function $g(t)$ satisfies the relation $g(t+\tau) = e^{\alpha\tau} g(t)$ and α varies on a time scale that is much greater than τ, then $d \ln g/dt = \alpha$, so that $dg/dt = \alpha g$. Similarly, Eq. (28) yields

$$\frac{\partial \psi(k,t)}{\partial t} = \int dx\ e^{ikx} \left[\sum_{m=1}^{\infty} \frac{(ik)^m}{m!}\ \alpha_m\right] \mathfrak{F}^-[\psi(k,t)]$$

$$= \sum_{m=1}^{\infty} \frac{(ik)^m}{m!} \int dx\ e^{ikx} \alpha_m(x,t) p(x,t), \tag{30}$$

and then the inverse Fourier transform leads to Eq. (19).

To proceed further, we write, instead of Eq. (20),

$$\psi(k,t+n\tau) = \Phi_\tau(t+(n-1)\tau)[\cdots \Phi_\tau(t+2\tau)$$

$$[\Phi_\tau(t+\tau)[\Phi_\tau(t)[\psi(k,t)]]]\cdots]. \tag{31}$$

For small k, Eq. (10) can be written as $\phi(x,t;k,\tau) = \gamma(x,t;k,\tau)[1 + O(k^3)]$, where $\gamma(x,t;k,\tau)$ is again the characteristic function of a Gaussian with the same values of the mean and the variance. If we now replace all ϕ's in Eq. (31) by γ's, the expression on the right-hand side will differ from $\psi(k,t+n\tau)$ by a factor that can be estimated as $[1+O(k^3)]^n$. As $n \to \infty$, this factor will tend to 1 because the range of k where the above expression has non-negligible values will be shrinking as $n^{-1/2}$ due to the second exponential in γ's. (The rigorous version of the proof sketched above should impose certain conditions on the transition probability, similar to the ones required by the general central limit theorem.) Thus we are led again to Eq. (21), but this time its coefficients are functions of x and t.

In retrospect, the emergence of Eq. (21) appears almost self-evident: Whatever might be the actual transition probabilities, after a Brownian particle experienced many microscopic jumps, the probability distribution of its position should be the same as if all the transition probabilities had been Gaussian, with the same values of the mean and the variance as the actual ones (the central limit theorem). Consequently, for macroscopic time intervals Eq. (19) can be replaced by Eq. (21).

THE MASTER EQUATION

Consider now the master equation for Markov processes, usually written in the form [4,8]

$$\frac{\partial p(x,t)}{\partial t}=\int d\xi[p(x-\xi,t)r(x-\xi,t;\xi)-p(x,t)r(x,t;-\xi)]$$
$$\equiv M[p(x,t)], \tag{32}$$

where the *transition rate* $r(x,t;\xi)$, also called *transition probability per unit time*, is defined by the equation (see [4], p. 96, and [8], p. 229)

$$r(x,t;\xi)-\left[\int d\xi\ r\right]\delta(\xi)=\lim_{\tau\to 0}\frac{w(x,t;\xi,\tau)-\delta(\xi)}{\tau} \tag{33}$$

and must satisfy the condition $r(x,t;\xi)\geqslant 0$ (see [4], pp. 96 and 98). Let us inquire whether the concept of transition *rate* is a viable one. Note that $\int d\xi\ r$ has dimensions of inverse time and define $\theta\equiv[\int d\xi\ r]^{-1}$. Consider a homogeneous process. The product $\theta r(\xi)$ satisfies the normalization condition and can be viewed as a probability distribution; its characteristic function is

$$\rho(k)=1+\theta\sum_{m=1}^{\infty}\frac{(ik)^m}{m!}\lim_{\tau\to 0}\frac{\mu_m}{\tau}=1+\sum_{m=1}^{\infty}\frac{(ik)^m}{m!}\alpha_m\theta$$
$$=1+\ln\phi(k,\theta). \tag{34}$$

This is impossible, except in the trivial case of the δ-function distribution. In all other cases, $\phi(k,\theta)\to 0$ as $k\to\pm\infty$ (see [10], p. 514), and thus $\rho(k)$ cannot satisfy the condition $|\rho(k)|\leqslant 1$ required of a characteristic function (see [10], p. 499). That is, $r(\xi)$ defined by Eq. (33) cannot be non-negative. For example, if the transition probability is Gaussian, the first two moments of $\theta r(\xi)$ are equal to the first two cumulants of $w(\xi,\theta)$, while all the higher ones vanish. Then $r(\xi)$ is a linear combination of $\delta(\xi)$ and its first and second derivatives; the derivatives of $\delta(\xi)$ violate the non-negativity condition (see [11], p. 12).

Of course, the limit $\tau\to 0$ is never actually taken; instead, some small but finite value of τ is used in calculations (see [4], p. 98). If the symbol $\lim_{\tau\to 0}$ in Eq. (33) is dropped (in which case the δ functions become superfluous and $\theta=\tau$) and the result is substituted into Eq. (32), the latter becomes

$$\frac{\partial p(x,t)}{\partial t}=\frac{\int d\xi\ p(x-\xi,t)w(x-\xi,t;\xi,\tau)-p(x,t)}{\tau}, \tag{35}$$

which is a finite-difference approximation for the evolution equation, based on Eq. (6). That means M fits the pattern of Eq. (4). To improve the accuracy of the master equation, one can use the expansion (5), viz.,

$$\frac{\partial p}{\partial t}=M[p]-\frac{\tau}{2}M^2[p]+\frac{\tau^2}{3}M^3[p]-\cdots . \tag{36}$$

This must converge to Eq. (19), which can be replaced by Eq. (21) as discussed above. Thus the master equation is but an approximate substitute for the differential evolution equation (21), a complete reversal of the conventional view [4].

ACKNOWLEDGMENT

This work was supported by the National Science Foundation.

[1] J. A. Goldstein, *Semigroups of Linear Operators and Applications* (Oxford University Press, New York, 1985).
[2] A. Einstein, Ann. Phys. (Leipzig) **17**, 549 (1905) [reprinted, with English translation, in *The Collected Papers of Albert Einstein*, edited by John Stachel (Princeton University Press, Princeton, NJ, 1989), Vol. 2, Doc. 16].
[3] C. W. Gardiner, *Handbook of Stochastic Methods*, 2nd ed. (Springer, Berlin, 1985).
[4] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1992).
[5] H. Risken, *The Fokker-Planck Equation*, 2nd ed. (Springer, Berlin, 1989).
[6] J. Mathews and R. L. Walker, *Mathematical Methods of Physics*, 2nd ed. (Benjamin/Cummings, Menlo Park, CA, 1970).
[7] R. Iranpour and P. Chacon, *Basic Stochastic Processes: The Mark Kac Lectures* (Macmillan, New York, 1988).
[8] D. T. Gillespie, *Markov Processes* (Academic, Boston, 1992).
[9] A. N. Kolmogorov, Math. Ann. **104**, 415 (1931) [English translation in *Selected Works of A. N. Kolmogorov*, edited by A. N. Shiryayev (Kluwer, Dordrecht, 1992), Vol. II, pp. 62–108].
[10] W. Feller, *An Introduction to Probability Theory and Its Applications*, 2nd ed. (Wiley, New York, 1971), Vol. 2.
[11] M. J. Lighthill, *Introduction to Fourier Analysis and Generalised Functions* (Cambridge University Press, Cambridge, 1958).