Journal of Statistical Physics, Vol. 24, No. 1, 1981

Itô Versus Stratonovich

N. G. van Kampen[1]

Received January 8, 1980

A survey is given of the facts and fancies concerning the nonlinear Langevin or Itô equation. Actually, it is merely a pre-equation, which becomes an equation when an interpretation rule is added. The rules of Itô and Stratonovich differ, but both are mathematically consistent and therefore equally admissible conventions. The reason why they seem to lead to physical differences is that the Langevin approach used to arrive at the equation involves a tacit assumption. For systems with external noise this assumption can be justified, and it is then clear that the Stratonovich rule applies. Systems with internal noise, however, can only be properly described by a master equation and the Itô–Stratonovich controversy never enters. Afterward one is free to model the resulting fluctuations either with an Itô or a Stratonovich scheme, but that does not lead to any new information.

KEY WORDS: Fluctuations; stochastic differential equations; Langevin approach.

With the death of Pierre Résibois we have lost not only an eminent physicist, but also a colleague with an open mind, with whom it was possible to discuss any subject without fear that a difference of opinion would affect the mutual feelings of friendship. For this reason it may not be inappropriate to dedicate to his memory an article that is frankly argumentative, trusting that the reader will take it in the same vein that he would.

The subject is the controversy about the proper treatment of stochastic differential equations involving white noise. In spite of several expository articles,[(1,2),2] the discussion in the physical and chemical literature continues and threatens to grow to grotesque proportions. This article is another

[1] Institute for Theoretical Physics of the University at Utrecht, Utrecht, The Netherlands.

[2] Reference 2, however, is more concerned with engineering problems than with physical systems.

attempt to bring it down to the realm of reason. There is nothing new in it, everything has been said before, but I shall say it loudly and, I hope, clearly.

1. Gaussian white noise (or the "Langevin process") is the stationary stochastic process $l(t)$ having the following properties:

(i) $\langle l(t)\rangle = 0$.

(ii) $\langle l(t_1)l(t_2)\rangle = \delta(t_1 - t_2)$.

(iii) Higher moments are given by the rules of Gaussian processes. That is, odd moments vanish, and even moments are the sum of the terms one obtains by breaking them up in all possible ways into a product of pairs and applying (ii) to each pair; e.g.,

$$\begin{aligned}&\langle l(t_1)l(t_2)l(t_3)l(t_4)\rangle \\ &\quad = \delta(t_1 - t_2)\delta(t_3 - t_4) + \delta(t_1 - t_3)\delta(t_2 - t_4) + \delta(t_1 - t_4)\delta(t_2 - t_3)\end{aligned}$$

Alternatively, one may stipulate that all cumulants beyond the second vanish[3]; or that the generating functional is

$$\left\langle \exp\left[i\int_{-\infty}^{\infty} k(t)l(t)\,dt \right]\right\rangle = \exp\left\{ -\frac{1}{2}\int_{-\infty}^{\infty} [k(t)]^2\,dt \right\}$$

where $k(t)$ is an arbitrary test function.

The difficulty is that no properly defined stochastic process with these properties exists. Gaussian white noise is a singular object, just as the delta function is a singular "function." Its integral, however,

$$w(t) = \int_0^t l(t')\,dt'$$

is the perfectly respectable Wiener process, which, however, is not stationary and not differentiable.

For a physicist it is convenient to visualize $l(t)$ as a random sequence of small pulses, both positive and negative. One then has to take the limit in which these pulses have short duration and small heights, but arrive in a dense succession. (The limit exists in the sense that the integral of such a succession of pulses tends to the Wiener process; or that its characteristic functional tends to the expression given above.)

2. Our problem concerns the nonlinear Langevin equation

$$\dot{x} = f(x) + g(x)l(t) \tag{1}$$

where f and g are two given functions. For our discussion it is sufficient to

[3] This is probably what is meant by the statement in Ref. 3 that "all correlations higher than the second order vanish."

consider a single variable x, and an autonomous system, i.e., f and g do not involve t explicitly. One often writes instead of (1)

$$dx = f(x)\,dt + g(x)\,dw(t)$$

but of course that does not solve any difficulty.[4]

First, *if g does not depend on x, there is no difficulty*. Then the equation, together with a fixed initial value,

$$\dot{x} = f(x) + gl(t), \qquad x(0) = a \tag{2}$$

defines uniquely a stochastic process $x(t)$, $t \geqslant 0$. It is a Markov process, and its transition probability $P(x, t \,|\, x_0, t_0)\,dx$ (from the value x_0 at t_0 into the interval $x, x + dx$ at t) obeys the Fokker–Planck equation

$$\frac{\partial P}{\partial t} = -\frac{\partial}{\partial x} f(x)P + \frac{g^2}{2}\frac{\partial^2 P}{\partial x^2} \tag{3}$$

However, *if g does depend on x, the equation* (1) *as it stands is meaningless*: the algebraic operations indicated by the symbols cannot be carried out. The reason may be understood as follows. According to the equation, each pulse in $l(t)$ gives rise to a pulse in $\dot{x}$ and hence a jump in x. That has the effect that the value of x to be used in $g(x)$ is undetermined (and hence also the size of the jump).

3. The Itô convention[(4)] *assigns* a meaning to (1) by adding, as a matter of definition, the rule that in $g(x)$ the value of x just *before* the pulse should be taken. It is then obvious that

$$\langle g(x)l(t)\rangle = 0 \tag{4}$$

It can be shown[(5)] that, with this additional convention, (1) together with an initial value determines $x(t)$ as a Markov process, whose transition probability obeys

$$\frac{\partial P}{\partial t} = -\frac{\partial}{\partial x} f(x)P + \frac{1}{2}\frac{\partial^2}{\partial x^2}\{g(x)\}^2 P \tag{5}$$

Incidentally, readers who did not feel that (4) was obvious may multiply (5) with x and integrate so as to obtain

$$\partial_t\langle x\rangle = \langle f(x)\rangle \tag{6}$$

and compare this with the average of (1).

The Stratonovich convention[(6)] takes for x in $g(x)$ half the sum of the values before and after the jump. With this convention the expression (1) with initial value defines again a Markov process, but different from the

[4] In this form it is usually called the Itô equation, but in the present context that name might lead to confusion.

one defined by (1) with the Itô convention. The transition probability of this Markov process is determined by the Fokker–Planck equation

$$\frac{\partial P}{\partial t} = -\frac{\partial}{\partial x} f(x)P + \frac{1}{2}\frac{\partial}{\partial x} g(x)\frac{\partial}{\partial x} g(x)P \tag{7a}$$

Note that this equation may also be written

$$\frac{\partial P}{\partial t} = -\frac{\partial}{\partial x}\left\{f(x) + \frac{1}{2} g(x)g'(x)\right\}P + \frac{1}{2}\frac{\partial^2}{\partial x^2}\{g(x)\}^2 P \tag{7b}$$

which exhibits the difference from Itô's process. On multiplying (7) with x and integrating, one finds that one now has instead of (4)

$$\langle g(x)l(t)\rangle = \tfrac{1}{2}\langle g(x)\,g'(x)\rangle \tag{8}$$

which does not in general vanish.

We summarize the situation so far. *Although* (1) *looks like a differential equation, it is really a meaningless string of symbols.* It does not define a stochastic process $x(t)$.[5] Yet this "pre-equation" can be turned into an actual equation by supplementing it with an additional interpretation rule. Two different rules have been chosen by Itô and Stratonovich, turning it into two different equations, defining two different processes $x(t)$.[6] It is misleading to call (1) an equation until the interpretation rule is added. Rather, there are *two* equations: (1I) and (1S). Since these rules are merely a matter of definition, it makes no sense to argue about their being right or wrong.

4. When a physicist encounters an expression like (1) he usually feels free to transform it to a new variable $\bar{x} = \phi(x)$ so as to obtain

$$\dot{\bar{x}} = \bar{f}(\bar{x}) + \bar{g}(\bar{x})l(t) \tag{9}$$

$$\bar{f}(\bar{x}) = f(x)\phi'(x) \tag{10a}$$

$$\bar{g}(\bar{x}) = g(x)\phi'(x) \tag{10b}$$

However, since (1) is *not* an equation, this formal transformation cannot lead to any conclusion. In particular, there is no reason why the interpretation rules applied to the transformed preequation should lead to equations that are the transforms of the original equations (1I) and (1S). It happens to be true for the Stratonovich rule, but it is not true for the Itô convention. Symbolically,

$$(\bar{1}\mathrm{S}) = \overline{(1\mathrm{S})}\,, \qquad (\bar{1}\mathrm{I}) \neq \overline{(1\mathrm{I})} \tag{11}$$

[5] In the following we always suppose the presence of a fixed initial value $x(0) = a$ when not mentioned otherwise.

[6] Other interpretation rules could be envisaged,[21] but they have not been used in physics, and are not the subject of this article.

The easiest way to see this is by looking at the corresponding Fokker–Planck equations. If one transforms (7) to the new variable $\bar{x}$, taking into account that $\bar{P}(\bar{x},t) = P(x,t)/\phi'(x)$, one finds

$$\frac{\partial \bar{P}}{\partial t} = -\frac{\partial}{\partial \bar{x}} \bar{f}(\bar{x})\bar{P} + \frac{1}{2}\frac{\partial}{\partial \bar{x}} \bar{g}(\bar{x}) \frac{\partial}{\partial \bar{x}} \bar{g}(\bar{x})\bar{P}$$

where $\bar{f}$ and $\bar{g}$ are indeed related to f and g by (10). Hence, if $x(t)$ is defined by (1S), the process $\bar{x}(t) = \phi(x(t))$ is the one defined by (9S).

However, if one similarly transforms (5) into

$$\frac{\partial \bar{P}}{\partial t} = -\frac{\partial}{\partial \bar{x}} \tilde{f}(\bar{x})\bar{P} + \frac{1}{2}\frac{\partial^2}{\partial \bar{x}^2} \{ \tilde{g}(\bar{x})\}^2 \bar{P}$$

one finds instead of (10)

$$\tilde{f}(\bar{x}) = f(x)\phi'(x) + \tfrac{1}{2}\{ g(x)\}^2 \phi''(x) \tag{12a}$$

$$\tilde{g}(\bar{x}) = g(x)\phi'(x) \tag{12b}$$

Hence, if $x(t)$ is defined by (1I), the process $\bar{x}(t)$ is the one defined by (9I) with (12) instead of (10). With these new transformation formulas one may freely transform the variables in the Itô scheme.[7] The fact that (12) differs from the transformation formulas (10), which one expects naively, is no valid reason for concluding that Itô is wrong.

5. As an example, take for x the velocity of a Brownian particle. According to Langevin,[8] the physics is described by

$$\dot{x} = -x + cl(t) \tag{13}$$

This is of the form (1), but since $g = c$ is constant, as in (2), no additional interpretation convention is needed. Both Itô and Stratonovich lead to the same Fokker–Planck equation (3), in this case

$$\frac{\partial P}{\partial t} = \frac{\partial}{\partial x} xP + \frac{c^2}{2}\frac{\partial^2 P}{\partial x^2} \tag{14}$$

Now let $\bar{x} = \frac{1}{2}x^2$ be the energy. According to Stratonovich, the transformed expression (9) is

$$\dot{\bar{x}} = -2\bar{x} + c(2\bar{x})^{1/2} l(t) \tag{15}$$

This expression requires an additional interpretation rule, for which one must of course take the Stratonovich convention, as the Stratonovich transformation rule was used. Thus (15) is equivalent to

$$\frac{\partial \bar{P}}{\partial t} = 2\frac{\partial}{\partial \bar{x}} \bar{x}\bar{P} + \frac{c^2}{2}\frac{\partial}{\partial \bar{x}} (\bar{x})^{1/2} \frac{\partial}{\partial \bar{x}} (\bar{x})^{1/2}\bar{P} \tag{16}$$

This is in fact the transformed form of (14).

Alternatively, one may transform (13) according to Itô's rules (12),

$$\dot{\bar{x}} = -2\bar{x} + \tfrac{1}{2}c^2 + c(2\bar{x})^{1/2}l(t) \tag{17}$$

This equation must then be interpreted according to Itô, which leads to

$$\frac{\partial \bar{P}}{\partial t} = \frac{\partial}{\partial \bar{x}}\left(2\bar{x} - \frac{1}{2}c^2\right)\bar{P} + \frac{c^2}{2}\frac{\partial^2}{\partial \bar{x}^2}\bar{x}\bar{P} \tag{18}$$

Clearly this is the same as (16).

Conclusion: If one starts from the well-defined equation (13), one is free to apply nonlinear transformations to the variable following either Stratonovich or Itô. Of course one must expect incorrect results if one mixes both, for instance, by transforming (13) according to Stratonovich into (15) and then interpreting (15) according to Itô. Such incorrect results must not be construed as evidence that Itô's interpretation is physically wrong.[9] In fact, in this example the physics only entered through (13), which is not subject to the controversy; the example merely demonstrates that the mathematics is consistent.

6. When the physical basis itself is not well defined the situation is less clear-cut. Before broaching that topic, the following two remarks must be made, which are relevant for the physical discussion, but are themselves merely mathematics.

First suppose one has an expression (1) with coefficient functions f_1, g_1, and one turns it into an equation by adding the Itô rule:

$$\dot{x} = f_1(x) + g_1(x)l(t) \qquad \text{(I)} \tag{19}$$

On the other hand, take another expression (1), with coefficients f_2, g_2, and the Stratonovich rule

$$\dot{x} = f_2(x) + g_2(x)l(t) \qquad \text{(S)} \tag{20}$$

Then it is possible to relate f_2, g_2 to f_1, g_1 in such a way that both equations are equivalent, i.e., with the same initial condition they determine the same process $x(t)$.

The proof again uses the equivalent Fokker–Planck equations. Equation (19) defines a Markov process whose transition probability obeys (5) with subscripts 1. Similarly, (20) defines a Markov process whose transition probability obeys (7) with subscript 2. On comparing (5) with (7b), one sees that both equations are the same if one chooses

$$g_2(x) = g_1(x), \qquad f_2(x) + \tfrac{1}{2}\,g_2(x)\,g_2'(x) = f_1(x) \tag{21}$$

Thus *the difference between the Itô and Stratonovich rules amounts to a difference in the choice of f*: it can always be compensated by an appropriate modification of the coefficient f.

7. The second remark concerns an altogether different kind of stochastic differential equation, viz.

$$x = f(x) + g(x)L(t) \tag{22}$$

where $L(t)$ is a given stationary stochastic process with zero mean and short, *but not infinitely short*, autocorrelation time τ_c. The precise form of $L(t)$ does not matter, and it is easy to construct a perfectly respectable, not singular, stochastic process having these properties. Hence (22) is a well-defined differential equation, which neither requires nor admits additional interpretation rules. One may now study the limit $\tau_c \to 0$ and ask whether one gets (1I) or (1S)—or perhaps something else.

There is one hitch, however, because the solutions of (22) are obviously not Markovian. Hence it is not possible to select a solution by fixing its initial value at $t = 0$. The question should therefore be formulated more precisely: does the *set* of solutions of (22) tend to the *set* of solutions of (1I) or (1S)?

The answer is of course: (1S). For that was the way in which Stratonovich arrived at his rule in the first place. This answer has repeatedly been rederived using either some form of perturbation theory or the theory of stochastic differential equations.[10] It can also be obtained by the following simple argument.

Since $L(t)$ is a mathematically well-defined stochastic function, one may freely transform x by the ordinary rules of calculus, which are the same rules (10) that apply to the Stratonovich calculus. Choosing $\phi'(x) = 1/g(x)$, one obtains

$$\dot{\bar{x}} = \bar{f}(\bar{x}) + L(t)$$

If one now varies $L(t)$ in such a way as to approach $l(t)$, one obtains in the limit an equation of type (2), which is equivalent with (3). Transforming back to x, one obtains (1S), since the Stratonovich calculus is invariant for (10).

The conclusion is that (22), in the limit $L(t) \to l(t)$, reduces to (1S). It must be emphasized that *this is true only when the functions f and g are fixed, i.e., they are not altered while the limit is approached.* It would also be possible to construct a different limiting scenario, in which $L(t) \to l(t)$, while f varies in such a way that in the limit one obtains (1I). That is the reason why the limiting property derived in this remark cannot be considered as a universal argument in favor of Stratonovich.

8. Why do these plain mathematical facts create so much confusion among physicists and chemists? In particular, how is it possible that two equivalent mathematical formulations appear to lead to physically different conclusions? The reason is that the customary physical argument by which one arrives at (1) involves a tacit assumption, which fixes the function f and thereby breaks the equivalence between both formulations. I shall outline this argument—to be called "the Langevin approach."

One has a system whose macroscopic, deterministic equation of motion is known to be

$$\dot{x} = F(x) \tag{23}$$

One then realizes that, for some reason, fluctuations occur, so that (23) is not exact; rather, x fluctuates around the values given by it. To take these fluctuations into account, one supplements (23) with a fluctuating term, which one assumes to have the same form as in (1)

$$\dot{x} = F(x) + g(x)l(t) \tag{24}$$

The coefficient $g(x)$ determines the magnitude of the fluctuations and must be found from physical considerations. These often indicate that g must indeed depend on x. For instance, the fluctuations in the number of electrons arriving on an anode will be roughly proportional to the square root of that number.

Having arrived at (24), one realizes that its meaning is not well defined, and that an interpretation rule has to be added. The resulting process $x(t)$ depends on the choice of that rule, owing to the fact that *one has assumed the first term to be identical with the macroscopic law*. At this point the controversy starts, and various justifications are adduced to justify the choice of either Itô or Stratonovich.

Clearly something went wrong in this Langevin approach. No amount of physical intuition and acumen suffices to justify the meaningless string of symbols (24). If the physical argument were sound, it ought to lead to an *equation*, that is, to (24), *including* the required interpretation rule. According to the remark in Section 6, such an argument must consider at the same time the form of the function f. *There is no good reason why f should be identical with the macroscopic F*: the only requirement is that the resulting equation describes the behavior of x and its fluctuations as they actually occur in the system.

For a sound physical starting point it is indispensable to examine the mechanism from which the fluctuations arise in more detail than is done in the Langevin approach. In fact, this could have been expected a priori, because the fluctuations contain more information about the system than the macroscopic equation alone. It so happens that in linear cases, such as (13), the only additional information is the constant c, and its value can be obtained by confronting the resulting $\langle x^2 \rangle$ with the known equilibrium distribution (fluctuation–dissipation theorem). This fact has led to the view that for nonlinear cases a similar bare minimum of additional information should suffice, but that is not true. The difficulties and inconsistencies that result from this erroneous idea were first revealed by the work of MacDonald[(11),7] but are still haunting the literature.

9. As a *first category* consider those systems in which the fluctuations are due to an *external* source. The word "external" is used to indicate that (i) the noise source is not influenced by the system itself; (ii) there is a parameter which permits one in principle to turn off the noise.

Examples are: an experiment done in a moving train; propagation of electromagnetic waves through the turbulent atmosphere; the growth of a specie under the influence of the vicissitudes of the weather. An electric circuit with an added noise generator[(13)] is another example, provided that one is sure that the generator is not affected by the currents in the circuit. The nonlinear transformer with noisy input is also of this type.

For systems with external noise one can define an equation (23) as the equation of motion of the system when the noise is turned off. As the noise is never completely white, it should be described by a stochastic function $L(t)$. If $g(x)$ is the response of the system to the added influence, it will obey

$$\dot{x} = F(x) + g(x)L(t) \tag{25}$$

If it is true that the autocorrelation τ_c of L is much shorter than the other time scales occurring, one may replace (25) with its limiting form, which is, according to Section 7, the pre-equation (24) *together with the Stratonovich rule*.

The precise conditions for the applicability of the limiting form are first that $\tau_c \ll \tau_m$, where the "mechanical" time scale τ_m is of order $x/F(x)$. This condition, however, can be eliminated by going to the interaction representation, i.e., by taking the integral of the motion of (23) as a new variable. The second condition is essential: $g\tau_c \ll 1$, where g is a number typical for the magnitude of $g(x)$. The effect of higher orders of $g\tau_c$ can also be computed,[(14)] but that is outside the scope of the stochastic differential equation (24S).

10. The *second category* is formed by systems with *internal* noise. That is, the fluctuations are due to the fact that the system itself consists of discrete particles; they are an inherent part of the very mechanism by which the state of the system evolves and cannot be turned off by manipulating a parameter. Examples: chemical reactions, Brownian particles, lasers.[8]

[7] See Ref. 12 for a review of the early history.

[8] The distinction between internal and external noise was made in a different connection by Mori.[(15)]

As it is impossible for systems with internal noise to eliminate the fluctuations, no precise definition for a deterministic equation (23) exists. *One therefore must describe the evolution of the entire system as a stochastic process* $x(t)$.[9] In those cases in which it is reasonable to assume that $x(t)$ is a Markov process the evolution equation is a master equation having the general form

$$\frac{\partial P(x,t)}{\partial t} = \int \{ W(x \mid x')P(x',t) - W(x' \mid x)P(x,t)\} \, dx' \qquad (26)$$

Here $P(x,t) \equiv P(x,t \mid x_0, t_0)$ is the transition probability between t_0 and t, while $W(x \mid x')\Delta t$ is the same transition probability taken for a time interval Δt which is so small that P does not vary much, but large enough for the Markov assumption to hold:

$$P(x, t + \Delta t \mid x', t) = W(x \mid x')\Delta t + \delta(x - x')\left[1 - \Delta t \int W(x' \mid x)\, dx'\right] + o(\Delta t)$$

The master equation determines the entire evolution, including fluctuations. The explicit form of $W(x \mid x')$ reflects the properties of the special system considered. In many cases, such as chemical reactions, the variable x takes only integral values, but in Brownian motion, for instance, it has a continuous range.[10]

Yet we know that most systems also have a deterministic evolution equation, e.g., the rate equation for a chemical reaction. How is that fact related to the full stochastic description (26)? The answer is that the $W(x \mid x')$ for these systems involves a parameter Ω such that for large Ω the fluctuations are relatively small. In the limit $\Omega \to \infty$ they are negligible and the solution of (26) is at all times a delta peak. The position of the peak moves according to a deterministic equation of the form (23). The function F is therefore only defined in this "macroscopic limit."[11]

11. Let us call this zeroth order. The fluctuations are contained in the next order of $1/\Omega$. A systematic expansion of (26) yields in this first order in $1/\Omega$ a perfectly well-defined Fokker–Planck equation for them. The Itô–Stratonovich controversy never enters. This confirms the conclusion that it has no physical content.

[9] This was emphasized by Green.[(16)]

[10] A continuous range is necessary but by no means sufficient for the process $x(t)$ to be a "continuous Markov process" in the technical sense that its sample paths are (with probability 1) continuous functions of t. These are the processes for which the master equation is a Fokker–Planck equation. They do not occur in nature.[(17)]

[11] In many cases, such as chemical reactions, Ω is the size of the system and $\Omega \to \infty$ is the thermodynamic limit. However, this is not universally so, e.g., in Brownian motion $\Omega = (m/M)$. The term "system size expansion" is therefore undesirable.[(18)]

One may now ask: once the stochastic process $x(t)$ is found by means of the Ω expansion, can it be modeled *a posteriori* by Eq. (1I) or (1S)? The answer is that this can be done as far as the first order in $1/\Omega$. According to Section 6, one is free to choose either one, if one does not a priori fix the function f. The Langevin approach assumes that f is identical with the macroscopic F, but when the noise is internal there is no good reason for that. For, the function F is defined only in the limit $\Omega \to \infty$, while the fluctuations belong to the next order, so that f may well differ from it by terms of order $1/\Omega$.

The higher order fluctuations cannot be faithfully modeled by an equation of the form (1I) or (1S). Of course one can always write an equation (25) and consider it as a *definition* of $L(t)$. It then turns out, however, that the stochastic properties of $L(t)$ are not only complicated, but also that its correlations depend on the solution $x(t)$ itself.[(19)] That fact is of course fatal for the practical use of the Langevin approach. Even apart from that, why insist on forcing a stochastic process into the Langevin framework when its stochastic properties have already been computed?[12]

Finally it must be said that a system with internal noise need not have a suitable parameter Ω. And even if it has one, the expansion only works when it is stable, more precisely, when the macroscopic equation turns out to be globally asymptotically stable in the Lyapunov sense. A fully satisfactory treatment of fluctuations in unstable systems is not yet available in spite of an extensive literature.

12. *Conclusions*. The expression (1) is a pre-equation and does not by itself determine a solution $x(t)$. It can be turned into a real equation by defining an additional interpretation rule. Two different rules lead to two different equations (1I) and (1S), defining two different processes. Both are mathematically consistent, but the Itô rule requires a somewhat unexpected transformation rule for f when x is subjected to a nonlinear transformation. The difference between both interpretations can be compensated by a suitable alteration of f.

An apparent physical difference is suggested by the Langevin approach, because it assumes f to be identical with the F occurring in the deterministic equation (23). For systems with *external* noise this identification can be justified. In that case the argument that physical noise is at best only approximately white shows that the Stratonovich rule is the correct one. [Of course there is also an equivalent equation (1I) with another f, different from F.]

Systems with *internal* noise can be treated adequately only by a full

[12] The argument is often used that the Langevin equation has been so useful; by the same argument one might rebuild one's car in the image of a horse.

stochastic description (26), which takes the actual physical cause of the fluctuations into account. When the system is endowed with a suitable parameter Ω a macroscopic equation can be defined. The next order in $1/\Omega$ gives a unique description of the fluctuations (at least in stable systems), in which the Itô–Stratonovich controversy never occurs. It is possible to construct *a posteriori* an equation such as Eq. (1I) or Eq. (1S) that reproduces the same results. In higher orders of $1/\Omega$, however, that can no longer be done.[13]

The final conclusion is that a physicist cannot go wrong by regarding the Itô interpretation as one of those vagaries of the mathematical mind that are of no concern to him. It merely served to point out that (1) is not a meaningful equation, and thereby warn him against glib applications of the Langevin approach to nonlinear equations. Rather, (1) only occurs in systems with external noise as a limiting form of (25). In that case the Stratonovich rule applies, but it emerges automatically when (25) is solved according to the methods developed for general stochastic differential equations.[(14)] From a physical point of view the Itô–Stratonovich controversy is moot.

It may be that there exist systems whose fluctuations are neither internal nor external in the sense used above. Such systems would have to be discussed separately, but they are still subject to the general principle that *mathematical manipulations must be based on a physical picture of the noise source*. Projection operator techniques are purely formal and cannot therefore provide a justification for the use of stochastic methods.

REFERENCES

1. A. H. Gray and T. K. Caughey, *J. Math. Phys.* **44**:288 (1965).
2. R. E. Mortensen, *J. Stat. Phys.* **1**:271 (1969).
3. N. S. Goel and N. Richter-Dyn, *Stochastic Models in Biology* (Academic Press, New York, 1974).
4. K. Itô, *Proc. Imp. Acad. Tokyo* **20**, no. 8 (1944); **22**, no. 2 (1946); *Mem. Am. Math. Soc.* **4**:1 (1951).
5. J. L. Doob, *Stochastic Processes* (Wiley, New York, 1953), pp. 273ff.
6. R. L. Stratonovich, *SIAM J. Control* **4**:362 (1966); *Topics in the Theory of Random Noise. I* (Gordon and Breach, New York 1963), pp. 89ff.
7. K. Itô, *Nagoya Math. J.* **1**:35 (1950); **3**:55 (1951); L. Arnold, *Stochastische Differentialgleichungen* (Oldenbourg, Vienna, 1973), p. 103.

[13] The higher orders, and in fact the entire Markov process defined by (26), can be reproduced faithfully by the stochastic equation $\dot{x} = f(x) + M(x, t)$ if one takes for $M(x, t)$ a suitable "compound Poisson process" whose stochastic properties depend on x, rather than Gaussian white noise.[(20)] Actually one may even take $f = 0$, so that M merely represents the sequence of jumps of x.

8. P. Langevin, *C. R. Acad. Sci. Paris* **146**:530 (1908); G. E. Uhlenbeck and L. Ornstein, *Phys. Rev.* **36**:823 (1930).
9. B. J. West, A. R. Bulsara, K. Lindenberg, V. Seshadri, and K. E. Shuler, *Physica* **97A**:211 (1979).
10. E. Wong and M. Zakai, *Ann. Math. Stat.* **36**:1560 (1965).
11. D. K. C. MacDonald, *Phil. Mag.* **45**:63 (1954); *Phys. Rev.* **108**:541 (1957).
12. N. G. van Kampen, in *Fluctuation Phenomena in Solids*, Burgess, ed. (Academic Press, New York, 1965).
13. T. Kawakubo, S. Kabashima, and Y. Tsuchiya, *Prog. Theor. Phys. Suppl.* **64**:150 (1978).
14. N. G. van Kampen, *Phys. Reports* **24**:171 (1976).
15. H. Mori, *Prog. Theor. Phys.* **53**:1617 (1975).
16. M. S. Green, *J. Chem. Phys.* **20**:1281 (1952).
17. N. G. van Kampen, in *Thermodynamics and Kinetics of Biological Processes*, A. I. Zotin, ed. (Moscow, 1980).
18. N. G. van Kampen, *Can. J. Phys.* **39**:551 (1961); in *Advances in Chemical Physics*, Vol. 34, I. Prigogine and S. A. Rice, eds. (Wiley, New York, 1976), p. 245; R. Kubo, K. Matsuo, and K. Kitahara, *J. Stat. Phys.* **9**:51 (1973).
19. Z. A. Akcasu, *J. Stat. Phys.* **16**:33 (1977); D. Bedeaux, *Phys. Lett.* **62A**:10 (1977).
20. A. Onuki, *J. Stat. Phys.* **19**:325 (1978); P. Hanggi, *Z. Phys.* **B36**:271 (1980); see also D. L. Snyder, *Random Point Processes* (Wiley, New York, 1975), pp. 185ff.
21. E. J. McShane, *Proc. Nat. Acad. Sci. US* **63**:275, 1084 (1969); S. K. Srinivasan, *SM Archives* **3**:325 (1978).