Langevin equations from time series

E. Racca
Dipartimento di Idraulica, Trasporti e Infrastrutture Civili, Politecnico di Torino, Torino, Italy

A. Porporato*
Department of Civil and Environmental Engineering, Duke University, Durham, North Carolina 27708, USA
(Received 18 August 2004; published 9 February 2005)

We discuss the link between the approach to obtain the drift and diffusion of one-dimensional Langevin equations from time series, and Pope and Ching's relationship for stationary signals. The two approaches are based on different interpretations of conditional averages of the time derivatives of the time series at given levels. The analysis provides a useful indication for the correct application of Pope and Ching's relationship to obtain stochastic differential equations from time series and shows its validity, in a generalized sense, for nondifferentiable processes originating from Langevin equations.

DOI: 10.1103/PhysRevE.71.027101 PACS number(s): 02.50.Ey, 05.40.Ca, 89.75.Kd, 92.40.−t

Measured time series are often highly fluctuating, resulting from complex, high-dimensional systems whose dynamics may not even be completely known. This justifies the interest in obtaining simple models that are able to capture the essential features of the series, such as the probability density function (PDF) and the correlation structure, being at the same time parsimonious and flexible enough to adapt to possible nonlinearities in the underlying dynamics.

In many cases, if the measured series proves to be approximately Markovian, a first modeling assumption may be represented by general one-dimensional Langevin equations. For these equations, the functional forms of the drift and diffusion terms can be easily determined directly from the time series, employing the finite-difference form of their definition together with suitable interpolations of the resulting trends. Such an approach was proposed by Friedrich *et al.* [1–6] and was already partly contained in the works of Primak *et al.* [7–9].

A different approach to model stationary time series relates its PDF to the functional form of the temporal derivatives at a given level. It is based on a relationship due to Pope and Ching [10,11] that is valid for any stationary and sufficiently smooth signal, not necessarily Markovian. Recently, the Pope and Ching formula was also used to derive the one-dimensional Langevin equation from (financial) time series, although the link with such equations and the different interpretation of the corresponding conditional averages were not rigorously assessed [12–14]. It is thus interesting to discuss the link between the Pope and Ching formula and the approach of Friedrich *et al.* and show that the Pope and Ching formula also holds, in a generalized sense, for these nondifferentiable stochastic processes.

Consider the following Langevin equation, according to the Ito interpretation:

$$\dot{x} = A(x) + \sqrt{B(x)}\xi(t), \tag{1}$$

where $A(x)$ is the drift coefficient, $B(x)$ is the diffusion term, and $\xi(t)$ is a Langevin force, i.e., white Gaussian noise with zero mean. As is well known, the PDF of $x, p(x,t)$, is given by the Fokker-Plank equation,

$$\frac{\partial p(x,t)}{\partial t} = -\frac{\partial}{\partial x}[A(x)p(x,t)] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[B(x)p(x,t)], \tag{2}$$

from which the steady-state PDF of x is obtained as

$$p(x) = \frac{N}{B(x)}\exp\left[2\int_x \frac{A(u)}{B(u)}du\right], \tag{3}$$

where N is a normalization constant.

Considering a fixed temporal interval, Δt, it is possible to show [15,16] that

$$\langle \Delta x \rangle = A(x)\Delta t, \tag{4}$$

$$\langle \Delta x^2 \rangle = B(x)\Delta t, \tag{5}$$

for $\Delta t \to 0$. The fixed Δt ensures that all the increments Δx have the same weight. It is important to stress that the difference Δx must be computed in a "causal" or "forward" way, i.e.,

$$\langle \Delta x \rangle = \langle x(t+\Delta t) - x(t) \rangle|_{x(t)}. \tag{6}$$

Moreover, as noticed by Just *et al.* [17], if the probability current vanishes, as is always the case for stationary signals ([15], p. 124), it is possible to show that

$$\langle x(t+\Delta t) - x(t) \rangle|_{x(t)} = -\langle x(t) - x(t-\Delta t) \rangle|_{x(t)}. \tag{7}$$

Equations (4) and (5) have been used to estimate drift and diffusion from time series [1–6], assuming that they are generated by Langevin processes; other authors proposed corrections to reduce the errors due to finite Δt [18].

The approach of Pope and Ching [10,11] also relates, in a more general way, the steady-state PDF of stationary processes to its temporal increments at given levels, as

*Electronic address: amilcare@duke.edu

$$p(x) = \frac{N'}{\langle\langle \dot{x}^2 | x \rangle\rangle} \exp\left[\int_x \frac{\langle\langle \ddot{x} | u \rangle\rangle}{\langle\langle \dot{x}^2 | u \rangle\rangle} du \right], \tag{8}$$

where the dot denotes temporal derivative and the notation $\langle\langle\ \rangle\rangle$ stresses the fact that the interpretation of the conditional averages is different from those of Eqs. (4) and (5). In fact, the averages in Eq. (8) refer to an infinitesimal window between x and $x+dx$, rather than to a fixed temporal interval. Hence, each contribution in the averages of Eq. (8) needs to be weighted by the time spent in the window of width dx. This is clear if one considers that, for stationary signals, $\langle\langle \dot{x} | x \rangle\rangle$ is zero, in contrast to $\langle \Delta x \rangle / \Delta t$. In Langevin equations, the latter is equal to $A(x)$ [see Eq. (4)] and, in general, is different from zero in signals that are not symmetric in time [19]. Similar observations were made by Sokolov [20], who also reported explicit formulas for the conditional averages in Eq. (8).

Pope and Ching derived Eq. (8) under the hypothesis of twice differentiable signals. However, as shown in [21], the previous expression is also valid for signals that are only differentiable once. Along these lines, one can expect that the same equation also applies, in a generalized sense, to Langevin equations with Gaussian noise (that are nondifferentiable). The starting point to find a link between the two approaches is already partly contained in the analysis of Stolovistky and Ching [21], who derived the conditional averages for the second-order process

$$\dot{x} = v,$$

$$\dot{v} = f(x) - \gamma v + \sqrt{\gamma g(x)}\xi(t), \tag{9}$$

as $\langle\langle \dot{v} | x \rangle\rangle = \langle\langle \ddot{x} | x \rangle\rangle = f(x)$ for any γ, and $2\langle\langle v^2 | x \rangle\rangle = 2\langle\langle \dot{x}^2 | x \rangle\rangle = g(x)$ for the limiting case of $\gamma \to \infty$.

Since it is also known that, for $\gamma \to \infty$, the system (9) can be reduced to the first-order Langevin equation [15,16,21,22],

$$\dot{x} = \frac{f(x)}{\gamma} + \sqrt{\frac{g(x)}{\gamma}}\xi(t), \tag{10}$$

one also has

$$\frac{f(x)}{\gamma} = \frac{\langle\langle \ddot{x} | x \rangle\rangle}{\gamma} = \frac{\langle \Delta x \rangle}{\Delta t} = A(x) \tag{11}$$

and

$$\frac{g(x)}{\gamma} = 2\frac{\langle\langle \dot{x}^2 | x \rangle\rangle}{\gamma} = \frac{\langle \Delta x^2 \rangle}{\Delta t} = B(x). \tag{12}$$

Thus, apart from a constant and provided the conditional averages are interpreted correctly, the terms in Eqs. (3) and (8) for one-dimensional Langevin equations give the same behavior as a function of x.

The analysis of the second-order difference of x completes the link between the two approaches. For the system (9) it is possible to show that ([16], p. 215)

$$\langle \Delta x \rangle = v \Delta t \tag{13}$$

and

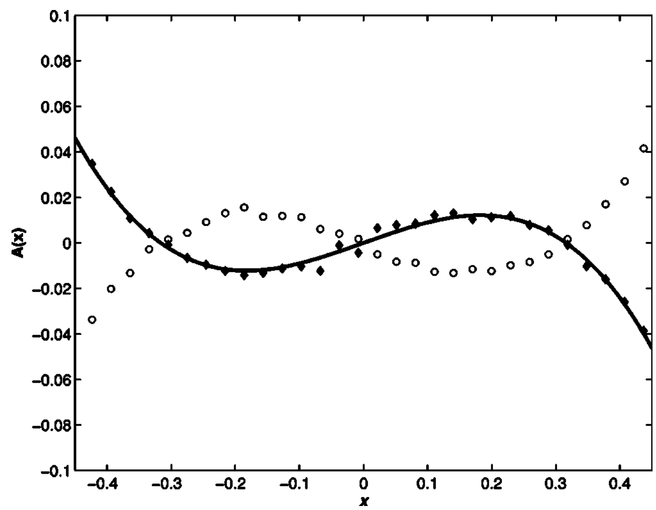

FIG. 1. Estimated values of the drift coefficient for the pitchfork bifurcation process. The solid line represents the theoretical term, the solid diamonds are the causal estimates, and the open circles are the acausal ones.

$$\langle \Delta^2 x \rangle = \langle \Delta v \rangle = [f(x) - \gamma v]\Delta t, \tag{14}$$

with $\Delta t \to 0$. From the previous expressions it is clear that, for $\gamma \to \infty, \langle \Delta^2 x \rangle = \gamma[A(x) - A(x)]\Delta t = 0$. Thus, similarly to Eqs. (4) and (5), the application of Eq. (14) also corresponds to a forward (or causal) estimate of the second-order difference. In fact, writing explicitly the expression

$$\langle \Delta^2 x \rangle = \langle x(t+2\Delta t) - x(t+\Delta t) \rangle|_{x(t+\Delta t)} - \langle x(t+\Delta t) - x(t) \rangle|_{x(t)}, \tag{15}$$

we see that it tends to zero for $\Delta t \to 0$ as both averages tend to $A(x)$. Note that the same thing is obtained with a backward estimate. However, a totally different result is obtained when using a centered estimate for the second-order difference. This approach was pursued by Tang [14], who actually employed it to estimate $\langle\langle \ddot{x} | x \rangle\rangle$, instead of the correct interpretation. Interestingly, however, in the case of Langevin equations, we have, using Eq. (7),

$$\langle x(t+\Delta t) - x(t) \rangle|_{x(t)} - \langle x(t) - x(t-\Delta t) \rangle|_{x(t)} = 2\langle \Delta x \rangle|_{x(t)} = 2A(x)\Delta t, \tag{16}$$

that, apart from a multiplicative constant, agrees with $\langle\langle \ddot{x} | x \rangle\rangle$ [see Eq. (11)].

We tested the previous results using numerical simulations of the stochastic pitchfork bifurcation process, $\dot{x}(t) = \epsilon x(t) - x^3(t) + g\xi(t)$, with $\epsilon = 0.1, g = 0.05$, and integration time step $\Delta t = 0.5$. Figure 1 shows the estimate of the drift term using Eq. (4) along with the difference between the causal and acausal estimates of Eq. (7). The estimates of the second-order difference computed using Eqs. (15) and (16) are shown in Fig. 2. As expected, the forward estimate is practically zero, while the centered estimate follows very well its theoretical behavior that is proportional to the drift coefficient [see Eq. (16)]. Finally, we show a comparison between the ratios $\langle \Delta x \rangle / \langle \Delta x^2 \rangle$ and $\langle\langle \ddot{x} | x \rangle\rangle / \langle\langle \dot{x}^2 | x \rangle\rangle$. While

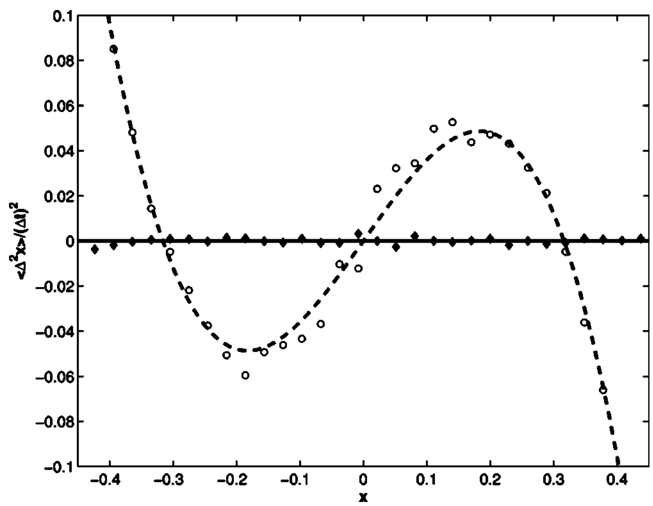

FIG. 2. Estimated values of $\langle \Delta^2 x \rangle / \Delta t^2$. The forward estimates, represented by the solid diamonds, tend to zero according to Eq. (15), whereas the centered estimates follow the theoretical behavior, $2A(x)/\Delta t$ [dashed line, Eq. (16)].

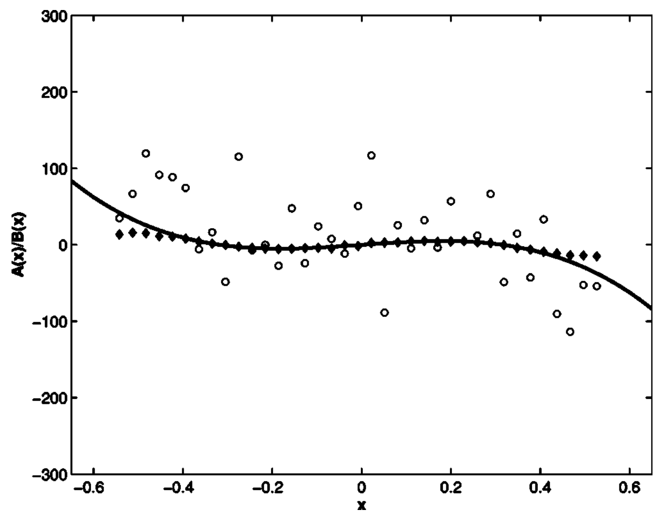

FIG. 3. Comparison among the estimated values of the ratio $\langle\langle \ddot{x} | x \rangle\rangle / \langle\langle \dot{x}^2 | x \rangle\rangle$ using Sokolov's formulas (open circles), the ratio of the estimated drift and diffusion terms, $\langle \Delta x \rangle / \langle \Delta x^2 \rangle$, using Eqs. (4) and (5) (solid diamonds), and the theoretical value, $A(x)/B(x)$ (solid line), for the pitchfork bifurcation process.

the behavior is similar in both cases and in agreement with their theoretical value, $A(x)/B(x)$, it is clear that the estimate of the conditional averages in Eq. (8) using Sokolov's formulas is statistically less efficient than the direct use of Eqs. (4) and (5) (see Fig. 3).

In summary, we showed the link between the approach to obtain drift and diffusion of Langevin equations from time series and the Pope and Ching formula for stationary processes. We stressed the importance of the correct interpretation of the estimators used and proved the validity (in a generalized sense) of the Pope and Ching formula also for nondifferentiable one-dimensional Langevin processes.

[1] R. Friendrich and J. Peinke, Phys. Rev. Lett. **78**, 863 (1997).
[2] R. Friedrich and J. Peinke, Physica D **102**, 147 (1997).
[3] S. Siegert, R. Friedrich, and J. Peinke, Phys. Lett. A **243**, 275 (1998).
[4] J. Gradisek, S. Siegert, R. Friedrich, and I. Grabec, Phys. Rev. E **62**, 3146 (2000).
[5] R. Friedrich, J. Peinke, and Ch. Renner, Phys. Rev. Lett. **84**, 5224 (2000).
[6] R. Friedrich, S. Siegert, J. Peinke, St. Lück, M. Siefert, M. Lindemann, J. Raethjen, G. Deuschl, and G. Pfister, Phys. Lett. A **271**, 217 (2003).
[7] S. Primak, Phys. Rev. E **61**, 100 (2000).
[8] V. Kontorovich and V. Lyandres, IEEE Trans. Signal Process. **43**, 2372 (1995).
[9] S. Primak and V. Lyandres, IEEE Trans. Signal Process. **46**, 1229 (1998).
[10] S. B. Pope and E. S. C. Ching, Phys. Fluids A **5**, 1529 (1993).
[11] E. S. C. Ching, Phys. Rev. Lett. **70**, 283 (1993).
[12] L.-H. Tang and Z.-F. Huang, Physica A **288**, 444 (2000).
[13] Z.-F. Huang, Physica A **287**, 405 (2000).
[14] L.-H. Tang, Physica A **324**, 272 (2003).
[15] C. W. Gardiner, *Handbook of Stochastic Methods* (Spinger-Verlag, Berlin, 1983).
[16] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (North Holland, Amsterdam, 1992).
[17] W. Just, H. Kantz, M. Ragwitz, and F. Schmüser, Europhys. Lett. **62**, 28 (2003).
[18] M. Ragwitz and H. Kantz, Phys. Rev. Lett. **87**, 254501 (2002).
[19] Consider, for example, an asymmetric sawtooth signal in which the rising part is steeper than the falling one. The signal is stationary and thus $\langle\langle \dot{x} | x \rangle\rangle$ is zero while $\langle \Delta x \rangle$ is positive and equal to the difference between the two slopes.
[20] I. M. Sokolov, Phys. Rev. E **60**, 3402 (1999).
[21] G. Stolovitzky and E. S. C. Ching, Phys. Lett. A **255**, 11 (1999).
[22] Notice that this is independent of the Ito and Stratonovitch interpretation.