

Rules of calculus in the path integral representation of white noise Langevin equations: the Onsager–Machlup approach

Leticia F Cugliandolo¹ and Vivien Lecomte^{2,3}

¹ Sorbonne Universités, Université Pierre et Marie Curie—Paris VI, Laboratoire de Physique Théorique et Hautes Énergies (LPTHE), 4 Place Jussieu, 75252 Paris Cedex 05, France

² LIPhy, Université Grenoble Alpes & CNRS, F-38000 Grenoble, France

³ Laboratoire Probabilités et Modèles Aléatoires (LPMA), CNRS UMR 7599, Université Paris Diderot, Paris Cité Sorbonne, Bâtiment Sophie Germain, Avenue de France, F-75013 Paris, France

E-mail: leticia@lpthe.jussieu.fr and vivien.lecomte@univ-grenoble-alpes.fr

Received 19 April 2017, revised 28 June 2017

Accepted for publication 5 July 2017

Published 26 July 2017



CrossMark

Abstract

The definition and manipulation of Langevin equations with multiplicative white noise require special care (one has to specify the time discretisation and a stochastic chain rule has to be used to perform changes of variables). While discretisation-scheme transformations and non-linear changes of variable can be safely performed on the Langevin equation, these same transformations lead to inconsistencies in its path-integral representation. We identify their origin and we show how to extend the well-known Ito prescription ($dB^2 = dt$) in a way that defines a modified stochastic calculus to be used inside the path-integral representation of the process, in its Onsager–Machlup form.

Keywords: Langevin equation, Stochastic processes, path-integral formalism, Stochastic chain rule, Onsager–Machlup functional

1. Introduction

Physical phenomena are often non-deterministic, presenting a stochastic behaviour induced by the action of a large number of constituents or by more intrinsic sources of noise [1–4]. A paradigmatic example is the one of Brownian motion, the study of which is at the source of stochastic calculus. From a modelisation viewpoint, the evolution of such systems can be described by a Langevin-type equation or by the path probability of its trajectories.

An important aspect of these descriptions is that the trajectories are not differentiable in general. This peculiarity implies that the definition of the evolution equation requires special care, namely, it demands the specification of a non-ambiguous time-discretisation scheme and, moreover, it induces a modification of the rules of calculus [1–4].

The important role of the time discretisation in the Langevin equation is now clearly elucidated [5] and many results have been obtained for the construction of an associated path-integral formalism, whose functional action and Jacobian correctly take into account the choice of discretisation [6–13].

An important point in the manipulation of Langevin equations is that the usual differential-calculus chain rule for changes of variables, $d_t[u(x(t))] = u'(x(t)) d_t x(t)$, has to be modified. It is replaced by the Ito formula (or ‘stochastic chain rule’), which is itself the consequence of the Ito substitution rule $dB^2 = dt$ for an infinitesimal increment $dB = B_{t+dt} - B_t$ of a Brownian motion B_t of unit variance. Although such manipulations are well understood at the Langevin equation level, the situation is less clear for the transformation of fields performed *inside* the action functional corresponding to the Langevin equation. It is known, for instance, that the use of the stochastic chain rule in the action can yield unsolved inconsistencies, both in statistical field theory [7, 14] and in quantum field theory [14–19].

In this article, we elucidate the source of this inconsistency, focusing on the case of the Onsager–Machlup action functional corresponding to a Langevin equation for one degree of freedom, with multiplicative white noise. We find that the sole Ito substitution rule $dB^2 = dt$ proves to be insufficient to correctly perform non-linear changes of variables in the action. We identify the required generalised substitution rules and we determine that their use should be performed with extreme care, since they take different forms when applied inside the exponential of the time-discrete action, or in the prefactor of its Gaussian weight factor.

In continuous time, we show that, in general, the use of the usual stochastic chain rule inside the action yields wrong results—and this even for a Stratonovich-discretised additive-noise Langevin equation. We determine a modified stochastic chain rule that allows one to manipulate the action directly, even in continuous time.

The organisation of the article is the following. In section 2, we review the non-ambiguous construction of the Langevin equation, providing three detailed examples which illustrate the role of the Ito substitution rule. In section 3, we recall inconsistencies that appear when one manipulates the action incorrectly, and we determine the valid substitution rules. We synthesise our results in section 4. Appendices gather part of the technical details.

2. Langevin equation and stochastic calculus

In this section, we briefly review the definition of multiplicative Langevin equations. For completeness, we first describe the standard construction of an unambiguous stochastic evolution equation through time discretisation, and we then provide three examples illustrating how differential calculus is generalised for stochastic variables, following this construction.

2.1. Discretisation convention of Langevin equations

Consider a time-dependent variable $x(t)$ which verifies a Langevin equation with a force $f(x)$ and a multiplicative noise $g(x)\eta$,

$$d_t x(t) = f(x(t)) + g(x(t)) \eta(t). \quad (1)$$

The function $g(x)$, that depends in general on the value of the variable, describes the amplitude of the stochastic term of this equation. The noise $\eta(t)$ is a centred Gaussian white noise of 2-point correlator equal to $\langle \eta(t)\eta(t') \rangle = 2D\delta(t' - t)$, where D plays the role of temperature. It is well known that the Langevin equation in its continuous-time writing (1) is ambiguous: one needs to specify a ‘discretisation scheme’ in order to give it a meaning (see [3, 5] for reviews).

Such a scheme is defined in discrete time, in the zero time step limit. We denote by x_t the time-discrete variable, with now $t \in \mathbb{N}$. The central feature of the definition of the Langevin equation is the following. Upon the time step $t \leadsto t + dt$, the right-hand-side (r.h.s.) of (1) is evaluated at a value of $x = \bar{x}_t$ chosen as a weighted average between x_t and x_{t+dt} as

$$\frac{x_{t+dt} - x_t}{dt} = f(\bar{x}_t) + g(\bar{x}_t)\eta_t \quad \langle \eta_t \eta_{t'} \rangle = \frac{2D}{dt} \delta_{tt'} \quad (2)$$

where the α -discretised evaluation point is

$$\bar{x}_t = \alpha x_{t+dt} + (1 - \alpha)x_t = x_t + \alpha(x_{t+dt} - x_t). \quad (3)$$

In the time-discrete evolution (2), the noise η_t is a centred Gaussian random variable (independent from those at other times, and thus independent of x_t). Its explicit distribution reads

$$\forall t, \quad P_{\text{noise}}(\eta_t) = \sqrt{\frac{dt}{4\pi D}} e^{-\frac{1}{2} \frac{dt}{2D} \eta_t^2}. \quad (4)$$

Its form implies that the stochastic term $g(\bar{x}_t)\eta_t$ in (2) is typically of order $dt^{-1/2}$, that is much larger than $f(\bar{x}_t)$, which is of order dt^0 . This difference is at the core of the ambiguity of the equation (1): as $dt \rightarrow 0$, the deterministic contribution $f(\bar{x}_t)$ to (2) is independent of the choice of α -discretisation; however, different values of α lead to different behaviours of the stochastic term $g(\bar{x}_t)\eta_t$. Indeed, making the discretisation explicit with a superscript we see, by Taylor expansion, that

$$\begin{aligned} \left[g(\bar{x}_t^{(\bar{\alpha})}) - g(\bar{x}_t^{(\alpha)}) \right] \eta_t &= \left[g(\bar{x}_t^{(\alpha)} + (\bar{\alpha} - \alpha)(x_{t+dt} - x_t)) - g(\bar{x}_t^{(\alpha)}) \right] \eta_t \\ &= (\bar{\alpha} - \alpha)(x_{t+dt} - x_t) g'(\bar{x}_t^{(\alpha)}) \eta_t + O(dt^{-\frac{1}{2}}) \end{aligned} \quad (5)$$

is typically of order dt^0 . This shows that, in general, $g(\bar{x}_t^{(\alpha)})\eta_t$ and $g(\bar{x}_t^{(\bar{\alpha})})\eta_t$ are not equivalent in (2) when $\alpha \neq \bar{\alpha}$.

Standard discretisation choices are $\alpha = \frac{1}{2}$ (known as ‘mid-point’ or Stratonovich convention) and $\alpha = 0$ (Ito convention). The Stratonovich choice is invariant by time reversal but, as other choices of $0 < \alpha \leq 1$, yields an implicit equation (2) for x_{t+dt} at each time step. The Ito convention, yielding independent increments for $x(t)$, is often chosen in mathematics, where the construction of the corresponding ‘stochastic calculus’ [4] is done by defining a stochastic integral for the integral equation corresponding to (1).

In general, we will denote the α -discretised Langevin equation (1) as

$$dx_\alpha = f(x) + g(x) \eta. \quad (6)$$

2.2. Three examples

In this section, we review three archetypal situations illustrating the role played by the choice of α -discretisation. We explain the computations in detail, so as to start off on the right footing for understanding the origin of the apparent contradictions discussed in section 3.

2.2.1. The stochastic chain rule (or Itô formula) A first consequence of the presence of a term of order $dt^{-1/2}$ in the discrete-time Langevin equation (2) is that the usual formulæ of differential calculus have to be altered. For instance, the chain rule describing the time derivative of a function of $x(t)$ is modified as [2, 3]

$$d_t[u(x)] \stackrel{\alpha}{=} u'(x)d_t x + (1 - 2\alpha)Dg(x)^2 u''(x), \quad (7)$$

where $x = x(t)$ verifies the Langevin equation (6). It is only for the Stratonovich discretisation that one recovers the chain rule of differentiable functions. For $\alpha = 0$, the relation (7) is known as the ‘Itô formula’.

The stochastic chain rule (7) is understood as follows. Coming back to the discrete-time definition of $d_t[u(x)]$, one performs a Taylor expansion in powers of $\Delta x \equiv x_{t+dt} - x_t$, keeping in mind that, as seen from (2), Δx is of order $dt^{1/2}$; this yields

$$\begin{aligned} \frac{u(x_{t+dt}) - u(x_t)}{dt} &= \frac{u(\bar{x}_t + (1 - \alpha)\Delta x) - u(\bar{x}_t - \alpha\Delta x)}{dt} \\ &= \frac{\Delta x}{dt} u'(\bar{x}_t) + \frac{1}{2}(1 - 2\alpha) \frac{\Delta x^2}{dt} u''(\bar{x}_t) + O(dt^{\frac{1}{2}}). \end{aligned} \quad (8)$$

For a differentiable function $x(t)$, the term $\propto \Delta x^2/dt$ would be negligible in the $dt \rightarrow 0$ limit but this is not the case for a stochastic $x(t)$. The next step is to understand the continuous-time limit $dt \rightarrow 0$ of (8): the so-called ‘Itô prescription’ amounts to replacing $\Delta x^2/dt$ in this expression by its quadratic variation

$$\frac{\Delta x^2}{dt} \mapsto 2Dg(\bar{x}_t)^2 \quad \text{as } dt \rightarrow 0 \quad (9)$$

(which is not equal to the expectation value of $\Delta x^2/dt$, as occasionally read in the literature, since $g(\bar{x}_t)$ depends on the value of \bar{x}_t without averaging). Note that in equation (8), one could as well replace $\Delta x^2/dt$ by the Itô-discretised $2Dg(x_t)^2$ (or any other discretisation point) instead of the α -discretised one in (9) since this would only add terms of order $dt^{1/2}$ to (8)—hence the name ‘Itô prescription’. In this article, we will rather use the name ‘substitution rule’ for two reasons: one is that we work in a generic α -discretisation scheme; another one is that we will introduce generalisations of (9) at a later stage.

We emphasise that the substitution rule (9) has to be used with care, as will be illustrated many times in this article. The validity of its use relies on the precise definition of the chain rule (7): this identity has to be understood in an ‘ L^2 -norm’ sense, i.e. it corresponds to having $\langle [\int_0^t dt \{l.h.s. - r.h.s.\}]^2 \rangle = 0$ ($\forall t$) and not to having a point-wise equality. The precise formulation and the demonstration of (7) and (9) are given in section B.1 of appendix B, along the lines of Øksendal’s reference textbook [4].

As this sort of issues is often overlooked in the theoretical physics literature, we now explain why an argument that is regularly proposed to justify (9) is in fact invalid. One could argue that the distribution of Δx^2 in (8) is sharply peaked around its most probable value $2Dg(\bar{x}_t)^2 dt$, because its variance $\langle \Delta x^4 \rangle - \langle \Delta x^2 \rangle^2$ is of order dt^2 as read from (2) and (4); this would allow one to replace $\Delta x^2/dt$ by $2Dg(\bar{x}_t)^2$ as $dt \rightarrow 0$ in (8), hence justifying (9). However, this argument is incorrect because the variance of $\Delta x^2/dt$ is of the same order dt^0 as some other terms in the time-discrete Langevin equation (8). To understand this point in detail, it is convenient to rephrase the argument as follows. First, one notes that according to (2) and (4), the quantity $\Delta x/dt$ is dominated by its most singular contribution $g(x_t)\eta_t$ in the $dt \rightarrow 0$ limit

$$\frac{\Delta x}{dt} = g(x_t)\eta_t + O(dt^0). \quad (10)$$

In this expression, we have chosen to evaluate $g(x)$ at $x = x_t$ instead of \bar{x}_t , the difference being gathered with other terms of order $O(dt^0)$ (see (5) for a proof). This allows one to use the fact that η_t is independent of x_t in order to compute the variance of $\Delta x^2/dt$ by Gaussian integration over η_t as

$$\begin{aligned} \left\langle \left[\frac{\Delta x^2}{dt} \right]^2 \right\rangle - \left\langle \frac{\Delta x^2}{dt} \right\rangle^2 &= \left\langle \left[g(x_t)\eta_t \right]^4 \right\rangle - \left\langle \left[g(x_t)\eta_t \right]^2 \right\rangle^2 dt^2 + O(dt) \\ &= 4D^2 \left(3\langle g(x_t)^4 \rangle - \langle g(x_t)^2 \rangle^2 \right) + O(dt) \end{aligned} \quad (11)$$

and one observes that it does not vanish as $dt \rightarrow 0$ (even for a constant noise amplitude $g(x) = g$). The variance of $\Delta x^2/dt$ is thus of the same order dt^0 as other terms in equation (8); this means that the properties of the distribution of $\Delta x^2/dt$ cannot be invoked to justify the substitution rule (9). This rule has to be understood in an L^2 sense that we explain in appendix B.1. As will prove to be essential, it means that the chain rule (7) is not true ‘point-wise’ but only in a weaker sense—which has to be taken care of meticulously in the path integral action, as we discuss thoroughly in section 3.2.

Finally, we note that the substitution rule (9) is equivalently written as follows⁴

$$\eta_t^2 dt \mapsto 2D \quad (12)$$

for the discrete time white noise η_t .

2.2.2. Changing discretisation while keeping the same evolution. Since the solution $x(t)$ of the Langevin equation (1) depends crucially on the choice of α -discretisation, although this choice seems to be arbitrary, one can wonder whether $x(t)$ can also be described as the solution of another Langevin equation, with a different $\bar{\alpha}$ -discretisation and a modified force. To answer this question, one comes back to the discrete-time evolution (2) and (3)

$$\frac{x_{t+dt} - x_t}{dt} = f(\bar{x}_t^{(\alpha)}) + g(\bar{x}_t^{(\alpha)})\eta_t \quad (13)$$

where we wrote explicitly the discretisation convention in superscript. Then, writing

$$\bar{x}_t^{(\alpha)} = \bar{x}_t^{(\bar{\alpha})} + (\alpha - \bar{\alpha})(x_{t+dt} - x_t) \quad (14)$$

and expanding in powers of $\Delta x = x_{t+dt} - x_t = O(dt^{1/2})$ one obtains

$$\begin{aligned} \frac{x_{t+dt} - x_t}{dt} &= f(\bar{x}_t^{(\bar{\alpha})}) + g(\bar{x}_t^{(\bar{\alpha})})\eta_t + (\alpha - \bar{\alpha})g'(\bar{x}_t^{(\bar{\alpha})})\Delta x\eta_t + O(dt^{\frac{1}{2}}) \\ &= f(\bar{x}_t^{(\bar{\alpha})}) + g(\bar{x}_t^{(\bar{\alpha})})\eta_t + (\alpha - \bar{\alpha})g(\bar{x}_t^{(\bar{\alpha})})g'(\bar{x}_t^{(\bar{\alpha})})\eta_t^2 dt + O(dt^{\frac{1}{2}}) \end{aligned} \quad (15)$$

where we used (2) for the last line.

Finally, using the substitution rule (12) and sending dt to zero, one finds that the process $x(t)$, solution of the Langevin equation (6) in the α -discretisation, is also verifying another Langevin equation

$$d_t x \stackrel{\bar{\alpha}}{=} f_{\alpha \rightarrow \bar{\alpha}}(x) + g(x)\eta \quad (16)$$

⁴ Another writing is $dB_t^2 \mapsto dt$ for a Brownian motion B_t of unit variance—the relation with our discrete white noise being $\eta_t dt = (2D)^{1/2}(B_{t+dt} - B_t)$.

$$f_{\alpha \rightarrow \bar{\alpha}}(x) = f(x) + 2(\alpha - \bar{\alpha})Dg(x)g'(x) \quad (17)$$

which is understood in $\bar{\alpha}$ -discretisation and presents a modified force $f_{\alpha \rightarrow \bar{\alpha}}(x)$. One checks directly that the Fokker-Planck equations corresponding to the two Langevin equations (1), (16) and (17) are identical, illustrating the equivalence of the two corresponding processes (see for instance [5] for the special case $\alpha = 0$ and $\bar{\alpha} = 1/2$). However, we emphasise that, since we used the substitution rule (12), we have to keep in mind that the equivalence between (6), (16) and (17) is not true pointwise and this can be the source of unexpected problems, as discussed in section 3.1.

2.2.3. Infinitesimal propagator for a path integral formulation. The trajectory probability of stochastic processes described by a Langevin equation has been the focus of many studies in statistical mechanics, either from the Onsager–Machlup approach [20, 21] or from the Martin–Siggia–Rose–Janssen–DeDominicis (MSRJD) one [22–27]. The idea in the Onsager–Machlup approach (to which we restrict our present analysis) is to write the probability of a trajectory $[x(t)]_{0 \leq t \leq t_f}$ as

$$\text{Prob}[x(t)] = \mathcal{J}[x(t)] e^{-S[x(t)]}, \quad (18)$$

where $S[x(t)]$ is the ‘action’, which takes a Lagrangian form $S[x(t)] = \int_0^{t_f} dt \mathcal{L}(x, dx)$, and $\mathcal{J}[x(t)]$ is a ‘normalisation prefactor’⁵. As can be expected from the discussion at the beginning of section 2.1, the form of the action and of the normalisation prefactor will depend not only on the α -discretization of the underlying Langevin equation, but also on the discretisation convention which is used to write them. The average of a functional \mathcal{F} of the trajectory can then be written in a path integral form as

$$\langle \mathcal{F}[x(t)] \rangle = \int \mathcal{D}x \mathcal{F}[x(t)] \mathcal{J}[x(t)] \text{Prob}[x(t)] P_i(x(0)). \quad (19)$$

The path integral is understood in the Feynman sense [28]: a sum over possible trajectories which start from an initial condition sampled by a distribution $P_i(x)$. It is best depicted in a time-discrete setup in the limit of zero time step, where one integrates over the set of possible values x_t of the trajectory at discrete times $t \in dt \mathbb{N}$ separated by a time step dt , yielding

$$\prod_{t=0}^{t_f/dt-1} \{dx_t \mathbb{P}(x_{t+dt}, t+dt|x_t, t)\} \xrightarrow{dt \rightarrow 0} \mathcal{D}x \mathcal{J}[x(t)] e^{-S[x(t)]} \quad (20)$$

where $\mathbb{P}(x_{t+dt}, t+dt|x_t, t)$ is a conditional probability (or ‘infinitesimal propagator’).

In this section, we focus our attention on the infinitesimal propagator between two successive time steps, that for simplicity we take at the first time step. Our goal is to compute $\mathbb{P}(x_{dt}|x_0) \equiv \mathbb{P}(x_{dt}, dt|x_0, 0)$ and to understand how the full action and normalisation prefactor are reconstituted through (20). We note that the correct form of this propagator, taking into account the α -discretisation is well-known [6, 8–10]. Still, we derive it again by taking a pedestrian approach that illustrates the role played by the substitution rules (9) or (12)—a role that proves essential to understand in order to later find the correct rules of stochastic calculus in the action.

⁵The prefactor $\mathcal{J}[x(t)]$ can be included in the measure $\mathcal{D}x$ on trajectories, but is not exponentiated in the action in general because it does not take a Lagrangian form.

2.2.3.1. First time step: changing from the distribution of η_0 to that of x_{dt} . Let us fix the initial condition x_0 and determine the distribution of x_{dt} obtained from the discrete Langevin equation (2). This equation is an implicit equation on x_{dt} , the solution of which takes the form

$$x_{dt} = X_1(x_0, \eta_0). \quad (21)$$

Therefore, the distribution of x_{dt} reads

$$\mathbb{P}(x_{dt}|x_0) = \int d\eta_0 \delta(x_{dt} - X_1(x_0, \eta_0)) P_{\text{noise}}(\eta_0), \quad (22)$$

with the noise distribution given in equation (4). In order to integrate over η_0 , we would like to read the Dirac as a δ on the variable η_0 . Cancelling the argument of the Dirac distribution in (22) defines a function $H_0(x_0, x_{dt})$ such that

$$(x_{dt} - X_1(x_0, \eta_0)) \Big|_{\eta_0=H_0(x_0, x_{dt})} = 0. \quad (23)$$

Then, the relation (22) yields

$$\begin{aligned} \mathbb{P}(x_{dt}|x_0) &= \int d\eta_0 \frac{\delta(\eta_0 - H_0(x_0, x_{dt}))}{\left| \partial_{\eta_0} (x_{dt} - X_1(x_0, \eta_0)) \right|_{\eta_0=H_0(x_0, x_{dt})}} P_{\text{noise}}(\eta_0) \\ &= \frac{P_{\text{noise}}(H_0(x_0, x_{dt}))}{\left| \partial_{\eta_0} X_1(x_0, H_0(x_0, x_{dt})) \right|}. \end{aligned} \quad (24)$$

Note that this relation can be derived by performing a change of variables in the probability distribution P_{noise} of η_0 , to obtain the distribution of x_{dt} seen as a function of η_0 through (21). (Two ways of evaluating the denominator are recalled in appendices A.2 [10] and A.3 [8]; we follow here a different route that is better adapted for our purposes.)

2.2.3.2. Expansions in the limit $dt \rightarrow 0$. The discrete Langevin equation (2) relating (at $t = 0$) x_0 and x_{dt} is a non-linear equation for which there is no explicit solution in general. As discussed previously, in the $dt \rightarrow 0$ limit, one has $x_{dt} - x_0 = O(dt^{1/2})$ (which is true for instance for a Brownian motion when $f(x) = 0$ and $g(x) = 1$, and is checked self-consistently in general). Writing $\bar{x}_0 = x_0 + \alpha(x_{dt} - x_0)$, we then expand (2) in order to obtain $x_0 - x_{dt}$ up to order $O(dt)$ included. One deduces

$$\begin{aligned} x_{dt} - x_0 &= [f(x_0) + \eta_0 g(x_0)] dt \\ &\quad + \alpha dt (x_{dt} - x_0) \left[f'(x_0) + \eta_0 g'(x_0) + \frac{1}{2} \alpha \eta_0 (x_{dt} - x_0) g''(x_0) \right] \end{aligned} \quad (25)$$

where we used $\eta_0 = O(dt^{-1/2})$. Solving for x_{dt} , one obtains, after expansion,

$$\begin{aligned} X_1(x_0, \eta_0) &= x_0 + dt f(x_0) + dt \eta_0 g(x_0) + \alpha dt^2 [f(x_0) + \eta_0 g(x_0)] [f'(x_0) + \eta_0 g'(x_0)] \\ &\quad + \frac{1}{2} \alpha^2 dt^3 \eta_0^2 \left\{ 2f(x_0)g'(x_0)^2 + \eta_0 g(x_0)^2 g''(x_0) \right. \\ &\quad \left. + 2g(x_0) [2f'(x_0)g'(x_0) + f(x_0)g''(x_0) + \eta_0 g'(x_0)^2] \right\}, \end{aligned} \quad (26)$$

where we kept terms of high enough order in η_0 so as to ensure that the derivative w.r.t. η_0 used in (24) contains terms up to order $O(dt^2)$ included. This derivative reads

$$\partial_{\eta_0} X_1(x_0, \eta_0) = dt g(x_0) \left[1 + 2\alpha dt \eta_0 g'(x_0) + \alpha dt f'(x_0) + \alpha dt f(x_0) \frac{g'(x_0)}{g(x_0)} + \frac{3}{2} dt^2 \alpha^2 \eta_0^2 (2g'(x_0)^2 + g(x_0)g''(x_0)) \right]. \quad (27)$$

Inverting, we have

$$\frac{1}{\partial_{\eta_0} X_1(x_0, \eta_0)} = \frac{1}{|g(x_0)dt|} \left[1 - 2\alpha dt \eta_0 g'(x_0) - \alpha dt f'(x_0) - \alpha dt f(x_0) \frac{g'(x_0)}{g(x_0)} + \frac{1}{2} dt^2 \alpha^2 \eta_0^2 (2g'(x_0)^2 - 3g(x_0)g''(x_0)) \right]. \quad (28)$$

In this expression, one can now use the substitution rule (12) to derive

$$\frac{1}{\partial_{\eta_0} X_1(x_0, \eta_0)} = \frac{1}{|g(x_0)dt|} \left[1 - 2\alpha dt \eta_0 g'(x_0) - \alpha dt f'(x_0) - \alpha dt f(x_0) \frac{g'(x_0)}{g(x_0)} + D\alpha^2 dt (2g'(x_0)^2 - 3g(x_0)g''(x_0)) \right]. \quad (29)$$

For later convenience, we prefer to express the numerator of the r.h.s. in terms of \bar{x}_0 instead of x_0 . We then utilise $x_0 = \bar{x}_0 - \alpha(x_{dt} - x_0)$ and we replace $x_{dt} - x_0$ by its expression deduced from (2). All in all, the only resulting non-trivial contribution to (29) is

$$\begin{aligned} dt \eta_0 g'(x_0) &= dt \eta_0 g'(\bar{x}_0 - \alpha dt (g(\bar{x}_0)\eta_0 + f(\bar{x}_0))) \\ &= dt \eta_0 g'(\bar{x}_0) - \alpha dt^2 \eta_0^2 g(\bar{x}_0)g''(\bar{x}_0) + o(dt) \\ &= dt \eta_0 g'(\bar{x}_0) - 2\alpha D dt g(\bar{x}_0)g''(\bar{x}_0) + o(dt) \end{aligned} \quad (30)$$

and, finally, equation (29) becomes

$$\frac{1}{\partial_{\eta_0} X_1(x_0, \eta_0)} = \frac{1}{|g(x_0)dt|} \left[1 - 2\alpha dt \eta_0 g'(\bar{x}_0) - \alpha dt f'(\bar{x}_0) - \alpha dt f(\bar{x}_0) \frac{g'(\bar{x}_0)}{g(\bar{x}_0)} + D\alpha^2 dt (2g'(\bar{x}_0)^2 + g(\bar{x}_0)g''(\bar{x}_0)) \right]. \quad (31)$$

At this point one would like to exponentiate this expression, an operation that has to be performed with care since $\eta_0 dt$ is of order $dt^{1/2}$. Using the substitution rule (12) as previously, one has

$$e^{A\eta_0 dt} = 1 + A\eta_0 dt + DA^2 dt + o(dt), \quad (32)$$

$$e^{A\eta_0 dt - DA^2 dt} = 1 + A\eta_0 dt + o(dt). \quad (33)$$

These relations imply

$$\frac{1}{\partial_{\eta_0} X_1(x_0, \eta_0)} = \frac{1}{|g(x_0)dt|} \exp \left[-2\alpha dt \eta_0 g'(\bar{x}_0) - \alpha dt f'(\bar{x}_0) - \alpha dt f(\bar{x}_0) \frac{g'(\bar{x}_0)}{g(\bar{x}_0)} - D\alpha^2 dt (2g'(\bar{x}_0)^2 - g(\bar{x}_0)g''(\bar{x}_0)) \right]. \quad (34)$$

The other function that one needs to determine to compute the infinitesimal propagator (24) is $H_0(x_0, x_{dt})$, defined by equation (23), which is equivalent to equation (2) evaluated at $t = 0$. After a simple rearrangement one finds

$$\eta_0 = \frac{1}{dt} \frac{x_{dt} - x_0 - dt f(x_0)}{g(x_0) + \alpha(x_{dt} - x_0)g'(x_0)} \equiv H_0(x_0, x_{dt}). \quad (35)$$

In the denominator, one recognises an expansion around \bar{x}_0 (with the l.h.s. η_0 evaluated up to $O(dt^0)$ included, so that $P_{\text{noise}}(H_0(x_0, x_{dt}))$ contains terms up to $O(dt)$, as seen from equation (4)). This yields

$$H_0(x_0, x_{dt}) = \frac{\frac{x_{dt} - x_0}{dt} - f(\bar{x}_0)}{g(\bar{x}_0)}. \quad (36)$$

2.2.3.3. Infinitesimal propagator. Coming back to equation (24), one obtains from equations (4), (31) and (36) that

$$\begin{aligned} \mathbb{P}(x_{dt}|x_0) \stackrel{\alpha}{=} & \sqrt{\frac{dt^{-1}}{4\pi D}} \frac{1}{|g(x_0)|} \exp \left\{ -\frac{1}{2} \frac{dt}{2D} \left[\frac{\frac{x_{dt} - x_0}{dt} - f(\bar{x}_0)}{g(\bar{x}_0)} \right]^2 - \alpha dt f'(\bar{x}_0) \right. \\ & - 2\alpha dt \frac{\frac{x_{dt} - x_0}{dt} - f(\bar{x}_0)}{g(\bar{x}_0)} g'(\bar{x}_0) - \alpha dt f(\bar{x}_0) \frac{g'(\bar{x}_0)}{g(\bar{x}_0)} \\ & \left. - D\alpha^2 dt \left(2g'(\bar{x}_0)^2 - g(\bar{x}_0)g''(\bar{x}_0) \right) \right\} \end{aligned} \quad (37)$$

where the symbol $\stackrel{\alpha}{=}$ indicates that in the r.h.s. \bar{x}_0 is the α -discretised point. Recognising a double-product to complete the square, one gets

$$\begin{aligned} \mathbb{P}(x_{dt}|x_0) \stackrel{\alpha}{=} & \sqrt{\frac{dt^{-1}}{4\pi D}} \frac{1}{|g(x_0)|} \exp \left\{ -\frac{1}{2} \frac{dt}{2D} \left[\frac{\frac{x_{dt} - x_0}{dt} - f(\bar{x}_0) + 4\alpha D g(\bar{x}_0)g'(\bar{x}_0)}{g(\bar{x}_0)} \right]^2 - \alpha dt f'(\bar{x}_0) \right. \\ & \left. - \alpha dt f(\bar{x}_0) \frac{g'(\bar{x}_0)}{g(\bar{x}_0)} + D\alpha^2 dt \left(2g'(\bar{x}_0)^2 + g(\bar{x}_0)g''(\bar{x}_0) \right) \right\}. \end{aligned} \quad (38)$$

The global prefactor $|g(x_0)|^{-1}$ in the infinitesimal propagator (38) can also be expressed in terms of $|g(\bar{x}_0)|^{-1}$. For this, one starts again from $x_0 = \bar{x}_0 - \alpha(x_{dt} - x_0)$ and replaces $x_{dt} - x_0$ by its expression deduced from (2). This leads to

$$\begin{aligned} \frac{1}{|g(x_0)|} = & \frac{1}{|g(\bar{x}_0)|} \left[1 + \alpha dt \eta_0 g'(\bar{x}_0) + \alpha dt f(\bar{x}_0) \frac{g'(\bar{x}_0)}{g(\bar{x}_0)} \right. \\ & \left. + D\alpha^2 dt \left(2g'(\bar{x}_0)^2 - g(\bar{x}_0)g''(\bar{x}_0) \right) \right]. \end{aligned} \quad (39)$$

Exponentiating in the same way as we obtained (31),

$$\begin{aligned} \frac{1}{|g(x_0)|} = & \frac{1}{|g(\bar{x}_0)|} \exp \left[1 + \alpha dt \eta_0 g'(\bar{x}_0) + \alpha dt f(\bar{x}_0) \frac{g'(\bar{x}_0)}{g(\bar{x}_0)} \right. \\ & \left. + D\alpha^2 dt \left(g'(\bar{x}_0)^2 - g(\bar{x}_0)g''(\bar{x}_0) \right) \right]. \end{aligned} \quad (40)$$

Since this relation contains a term $\propto \eta_0$, once again one has to complete the square. Coming back to equation (24), finally, many terms compensate and, instead of equation (38), one obtains a simpler expression for the infinitesimal propagator:

$$\mathbb{P}(x_{dt}|x_0) \stackrel{\alpha}{=} \sqrt{\frac{dt^{-1}}{4\pi D}} \frac{1}{|g(\bar{x}_0)|} \times \exp \left\{ -\frac{1}{2} \frac{dt}{2D} \left[\frac{\frac{x_{dt}-x_0}{dt} - f(\bar{x}_0) + 2\alpha D g(\bar{x}_0)g'(\bar{x}_0)}{g(\bar{x}_0)} \right]^2 - \alpha df'(\bar{x}_0) \right\}. \quad (41)$$

2.2.3.4. The continuous-time limit. The result (41) is well-known and can be derived in simpler ways [6, 9, 10] that are reviewed in appendix A, and that do not use (or use in a different way) the equivalent substitution rules (9) or (12). Having such different approaches leading to the same result is important in order to identify the conditions under which this substitution rule can be used; we note in particular that we used this rule in equations (29), (34) and (40) only in the prefactor of the exponential $P_{\text{noise}}(H_0)$ and not inside the exponential. As discussed throughout section 3, such restriction on the condition under which the substitution rule (9) is valid proves to be crucial.

We can read from equation (41) the continuous-time limit (20): this yields the trajectory probability in the form (18) with the so-called Onsager–Machlup action

$$S[x(t)] \stackrel{\alpha}{=} \int_0^{t_f} dt \left\{ \frac{1}{2} \frac{1}{2D} \left[\frac{dx - f(x) + 2\alpha D g(x)g'(x)}{g(x)} \right]^2 + \alpha f'(x) \right\} \quad (42)$$

where the arguments of the functions f and g are taken in α -discretisation. The associated normalisation prefactor reads

$$\mathcal{J}[x(t)] \stackrel{\alpha}{=} \prod_t \left\{ \sqrt{\frac{dt^{-1}}{4\pi D}} \frac{1}{|g(\bar{x}_t)|} \right\}. \quad (43)$$

We emphasise (and this seems to have been little stressed in the literature) that it is essential to specify the discretisation point of the normalisation prefactor $\mathcal{J}[x(t)]$, since it can yield different contributions to the action for different discretisation conventions, as should be clear from equation (40). For instance, when proving the Fluctuation-Dissipation Theorem and the Fluctuation Theorem for Langevin equations with multiplicative noise [8], one has to take into account that reversing the time changes the discretisation from α to $1 - \alpha$. This implies that when comparing the trajectory probability of a path and its time reversed, the discretisation of one of the normalisation prefactors has to be restored to α from $1 - \alpha$, which induces terms similar to those in (40) in the action without which the Fluctuation Theorem would not be verified.

We also note that Itami and Sasa have recently discussed in [10] the consequences of choosing different α -discretisations in the Langevin equation and in the action.

3. Stochastic calculus in the path integral action

In general, the different actions $S[x(t)]$ that are studied in statistical mechanics (or in quantum field theory [29]) take the form of the time integral of a ‘Lagrangian’: $S[x(t)] = \int_0^{t_f} dt \mathcal{L}(dx(t), x(t))$. This is the case, for instance, of the action (42) that we derived in the previous section and which corresponds to the α -discretised Langevin equation (6). Since the trajectories $x(t)$ that

verify the Langevin equation are not differentiable, it is natural to expect that the Lagrangian $\mathcal{L}(\mathrm{d}_t x(t), x(t))$ should be sensitive to the convention of α -discretisation for its writing, and that the differential transformations performed in the Lagrangian should incorporate terms akin to the stochastic ones $\propto (1 - 2\alpha)$ of the modified chain rule (7).

It is often assumed that the continuous-time chain rule (7) can be applied when manipulating the action (see for instance [30]) or that the formulæ (16) and (17) describing the change of discretisation in the Langevin equation can be equally used. In this section, we show

- (i) that performing a change of discretisation in the Onsager–Machlup action is possible but completely wrong if one uses the relations (16) and (17); and
- (ii) similarly, that non-linear changes of variables are allowed in the action but are also wrong if one applies the chain rule (7).

In both cases, one arrives at inconsistencies when using the incorrect transformations to manipulate the action. We describe the procedure which enables ones to elucidate the origin of such apparent contradictions of stochastic calculus. The moral of the story is that the substitution rule (9) alone is insufficient to understand the transformations of the action and that other transformations, that we derive, are needed.

3.1. From one discretisation to another

In this section, we examine the condition of validity of the transformation rules allowing one to go from an α -discretised Langevin equation (6) to an equivalent $\bar{\alpha}$ -discretised Langevin equation (16) and (17). We focus on the transformation from a generic α -discretisation to the Stratonovich one ($\alpha = 1/2$), which is often performed for the reason that the stochastic chain rule (7) takes a simple form for $\alpha = 1/2$ (yielding back the standard chain rule of differential calculus). We show that, although it seems natural to perform the same transformations in the action as at the Langevin level, such as changing discretisation through (16) and (17), the resulting action actually proves to be invalid (see figure 1 for a schematic representation of the procedure). Finally, we identify the reason why the correct rules of calculus in the action are more complex than at the Langevin level, and we determine the correct calculus to be used in the action that actually involves generalised substitution rules akin to (9).

3.1.1. Direct change of discretisation in the action. The α -discretised Langevin equation (6) $\mathrm{d}_t x \stackrel{\alpha}{=} f(x) + g(x) \eta$ is equivalent to the following Langevin equation in Stratonovich discretisation, with an α -dependent force f_α

$$\mathrm{d}_t x \stackrel{\text{Strato}}{=} f_\alpha(x) + g(x) \eta, \quad (44)$$

$$f_\alpha(x) = f(x) + 2D(\alpha - \frac{1}{2})g'(x)g(x). \quad (45)$$

This is seen, for instance, by coming back to the time-discrete definition (2) and (3) of the α -discretisation and by working with the symmetric Stratonovich discretisation point (the superscript S indicates such choice of discretisation in what follows)

$$\bar{x}_t^S = \frac{1}{2}(x_{t+dt} + x_t), \quad (46)$$

a procedure that we followed in section 2.2.2 for a generic change of discretisation: equations (16) and (17) yield the result above, i.e. equations (44) and (45) with a force $f_\alpha = f_{\alpha \rightarrow 1/2}$.

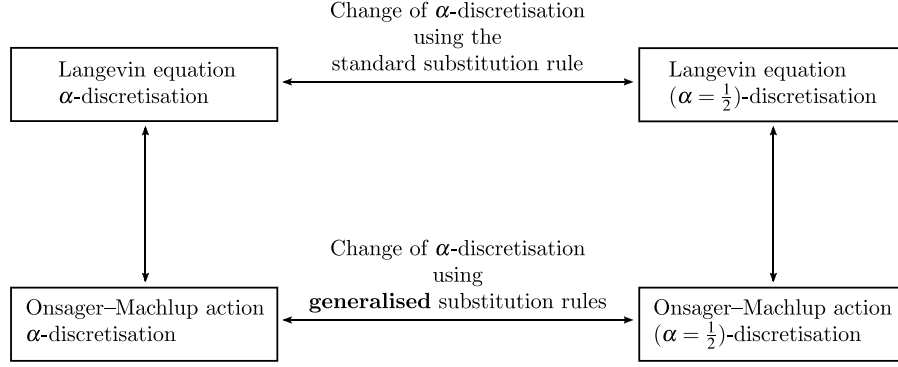


Figure 1. Schematic representation, for a change of discretisation, of the difference between the stochastic calculus performed in the Langevin equation and in the Onsager–Machlup action. The α -discretised Langevin equation (6) can be transformed by use of the substitution rule (9) into a Stratonovich-discretised one ($\alpha = 1/2$) given by equations (44) and (45). However, one cannot use such equations in the α -discretised Onsager–Machlup action (42) to get the correct Stratonovich-discretised action. Instead, to go from one action to the other, one has to use the generalised substitution rules (54)–(57) in discrete time for the infinitesimal propagator (once expanded in powers of Δx and dt), or to rely on *modified* substitution rules (74) and (75) inside the exponential of the propagator.

3.1.2. Change of discretisation in the infinitesimal propagator. Since the α -discretised Langevin equation (6) and the Stratonovich one (44) and (45) are equivalent, they must possess equivalent infinitesimal propagators. The change of discretisation in the infinitesimal propagator proves to be more involved than in the equation itself.

3.1.2.1. Expanding without throwing powers of Δx out with the bathwater. We focus, without loss of generality, on the first time step $0 \leadsto dt$. The propagator (41) in α -discretisation is

$$\mathbb{P}(x_{dt}|x_0) \stackrel{\alpha}{=} \frac{\mathcal{N}}{|g(\bar{x}_0)|} \exp \left\{ -\frac{1}{2} \frac{dt}{2D} \left[\frac{\frac{x_{dt}-x_0}{dt} - f(\bar{x}_0) + 2\alpha D g(\bar{x}_0)g'(\bar{x}_0)}{g(\bar{x}_0)} \right]^2 - \alpha dt f'(\bar{x}_0) \right\}$$

with $\mathcal{N} \equiv \sqrt{\frac{dt^{-1}}{4\pi D}}$. (47)

The aim is to determine an equivalent propagator in terms of the Stratonovich mid-point $\bar{x}_0^S = \frac{1}{2}(x_{dt} + x_0)$. We expand (47) in powers of $\Delta x \equiv x_{dt} - x_0$, using

$$\bar{x}_0 = \bar{x}_0^S + \left(\alpha - \frac{1}{2}\right)\Delta x \quad (48)$$

and keeping all terms of order dt^0 inside the exponential (note that they *define* the Gaussian weight), while putting all terms of order $O(dt^{1/2})$ and $O(dt)$ in a prefactor of this weight. In this procedure, one should remember that $\Delta x = O(dt^{1/2})$. This crucially implies that, in the exponential, the expansion of the term

$$-\frac{1}{2} \frac{1}{2D dt} \left[\frac{\Delta x}{g(\bar{x}_0^S + (\alpha - \frac{1}{2})\Delta x)} \right]^2 \quad (49)$$

generates terms of order $O(dt^{1/2})$ and $O(dt)$ which are proportional to $dt^{-1}\Delta x^3$ and $dt^{-1}\Delta x^4$. Expanding then the exponential, one gets terms up to $dt^{-2}\Delta x^6$. Explicitly, the result is

$$\begin{aligned}
& \frac{\mathbb{P}(x_{dt}|x_0)}{\frac{\mathcal{N}}{|g(\bar{x}_0^S)|} e^{-\frac{1}{2} \frac{dx}{D} \left(\frac{\Delta x}{dt} \right)^2 / g(\bar{x}_0^S)^2}} \\
& \stackrel{S}{=} 1 + \left[\frac{f(\bar{x}_0^S)}{2Dg(\bar{x}_0^S)^2} + \frac{(2-8\alpha)g'(\bar{x}_0^S)}{4g(\bar{x}_0^S)} \right] \Delta x + \left[\frac{(2\alpha-1)g'(\bar{x}_0^S)}{4Dg(\bar{x}_0^S)^3} \right] \Delta x^3 dt^{-1} \\
& + \left[-\alpha \left(\alpha Dg'(\bar{x}_0^S)^2 + f'(\bar{x}_0^S) \right) + \frac{\alpha f(\bar{x}_0^S)g'(\bar{x}_0^S)}{g(\bar{x}_0^S)} - \frac{f(\bar{x}_0^S)^2}{4Dg(\bar{x}_0^S)^2} \right] dt \\
& + \left[\frac{(-12\alpha^2 + 8\alpha - 1)g''(\bar{x}_0^S)}{8g(\bar{x}_0^S)} + \frac{(14\alpha^2 - 8\alpha + 1)Dg'(\bar{x}_0^S)^2 + (2\alpha - 1)f'(\bar{x}_0^S)}{4Dg(\bar{x}_0^S)^2} \right. \\
& \quad \left. + \frac{(3-8\alpha)f(\bar{x}_0^S)g'(\bar{x}_0^S)}{4Dg(\bar{x}_0^S)^3} + \frac{f(\bar{x}_0^S)^2}{8D^2g(\bar{x}_0^S)^4} \right] \Delta x^2 \\
& + \left[\frac{(1-2\alpha)^2g''(\bar{x}_0^S)}{16Dg(\bar{x}_0^S)^3} - \frac{(28\alpha^2 - 24\alpha + 5)g'(\bar{x}_0^S)^2}{16Dg(\bar{x}_0^S)^4} + \frac{(2\alpha-1)f(\bar{x}_0^S)g'(\bar{x}_0^S)}{8D^2g(\bar{x}_0^S)^5} \right] \Delta x^4 dt^{-1} \\
& + \frac{(1-2\alpha)^2g'(\bar{x}_0^S)^2}{32D^2g(\bar{x}_0^S)^6} \Delta x^6 dt^{-2}. \tag{50}
\end{aligned}$$

Note that we also changed the discretisation of the normalisation prefactor from $1/|g(\bar{x}_0)|$ to $1/|g(\bar{x}_0^S)|$ using a relation similar to (39). The symbol $\stackrel{S}{=}$ indicates that the r.h.s. is evaluated in the Stratonovich discretisation.

3.1.2.2. Comparison to the propagator arising from changing discretisation at the Langevin level. We would like to compare this result to that of the commutative procedure depicted in figure 1, namely,

- (i) transform the original α -discretised Langevin equation into the Stratonovich-discretised one (44) which includes an α -dependent force $f_\alpha(x)$ given by (45); and
- (ii) follow the same procedure as previously done to get the corresponding propagator, that we denote $\mathbb{P}_{f_\alpha}^S$.

The result is, of course, directly read from equation (47), where α is first replaced by $1/2$ (and hence \bar{x}_0 by \bar{x}_0^S), and then f is replaced by f_α ; this yields

$$\mathbb{P}_{f_\alpha}^S(x_{dt}|x_0) \stackrel{S}{=} \frac{\mathcal{N}}{|g(\bar{x}_0^S)|} \exp \left\{ -\frac{1}{2} \frac{dt}{2D} \left[\frac{\frac{\Delta x}{dt} - f_\alpha(\bar{x}_0^S) + Dg(\bar{x}_0^S)g'(\bar{x}_0^S)}{g(\bar{x}_0^S)} \right]^2 - \frac{1}{2} dt f'_\alpha(\bar{x}_0^S) \right\}. \tag{51}$$

By consistency, this propagator should be equal to the result (47), in the small dt limit. To check whether this is the case, we follow the same procedure as the one leading to equations (50) from (47), that is to say, we expand in powers of Δx and dt , and we replace f_α by its explicit expression in terms of f , g and α , to obtain

$$\begin{aligned}
\frac{\mathbb{P}_{f_\alpha}^S(x_{dt}|x_0)}{\frac{\mathcal{N}}{|g(\bar{x}_0^S)|} e^{-\frac{1}{2} \frac{dt}{2D} \left(\frac{\Delta x}{dt}\right)^2 / g(\bar{x}_0^S)^2}} &\stackrel{s}{=} 1 + \left[\frac{f(\bar{x}_0^S)}{2Dg(\bar{x}_0^S)^2} + \frac{(\alpha-1)g'(\bar{x}_0^S)}{g(\bar{x}_0^S)} \right] \Delta x \\
&+ \left[-\frac{f(\bar{x}_0^S)^2}{2Dg(\bar{x}_0^S)^2} - f'(\bar{x}_0^S) - \frac{2(\alpha-1)f(\bar{x}_0^S)g'(\bar{x}_0^S)}{g(\bar{x}_0^S)} \right. \\
&\quad \left. + D \left\{ (1-2\alpha)g(\bar{x}_0^S)g''(\bar{x}_0^S) + (-2(\alpha-1)\alpha-1)g'(\bar{x}_0^S)^2 \right\} \right] \frac{1}{2} dt \\
&+ \frac{(2(\alpha-1)Dg(\bar{x}_0^S)g'(\bar{x}_0^S) + f(\bar{x}_0^S))^2}{8D^2g(\bar{x}_0^S)^4} \Delta x^2.
\end{aligned} \tag{52}$$

The result is clearly different from the one in equation (50), while one expects $\mathbb{P}(x_{dt}|x_0) = \mathbb{P}_{f_\alpha}^S(x_{dt}|x_0)$ because these two propagators correspond to the same Langevin equation. In particular, the maximum power of Δx for $\mathbb{P}_{f_\alpha}^S(x_{dt}|x_0)$ in equation (52) is Δx^2 while it is Δx^6 in equation (50) for $\mathbb{P}(x_{dt}|x_0)$.

Note that if one takes for $g(x)$ a constant function g , the two propagators are still different, as checked by direct inspection (unless $\alpha = 1/2$, as it should because then there is no change of discretisation and the two computations are identical). The simple case of additive noise, thus, also requires a peculiar attention.

3.1.2.3. Appropriate substitution rules to render the two approaches compatible. As discussed in section 2.2.1, the Ito prescription amounts to using the substitution rule

$$\Delta x^2 \mapsto 2Dg(x)^2 dt \quad \text{as } dt \rightarrow 0, \tag{53}$$

where on the r.h.s., the argument x of $g(x)$ can be taken at any discretisation point, at minimal order in dt . We have seen in section 2.2.3.3 that the use of such prescription is justified as long as it is performed outside the exponential, for the determination of the infinitesimal propagator.

Therefore, in order to recover from (50) the simpler result (52) for the propagator, a natural possibility is to look for ‘generalised substitution rules’ akin to (53), but now for terms of the form $\Delta x^n dt^m$ with n, m chosen so that $\Delta x^n dt^m$ is typically of order $O(dt^{1/2})$ or $O(dt)$. One finds by direct computation that, to guarantee that (50) becomes (52), there is a *unique* prescription to replace the terms $\Delta x^n dt^m$ by standard infinitesimals of the natural form $C^{\text{st}} \times (2Dg(x)^2)^{n/2} dt$ when n is even and $C^{\text{st}} \times \Delta x (2Dg(x)^2)^{(n-1)/2}$ when n is odd. It reads

$$\Delta x^2 = 2Dg(x)^2 dt, \tag{54}$$

$$\Delta x^3 dt^{-1} = 3 \Delta x 2Dg(x)^2, \tag{55}$$

$$\Delta x^4 dt^{-1} = 3 (2Dg(x)^2)^2 dt, \tag{56}$$

$$\Delta x^6 dt^{-2} = 15 (2Dg(x)^2)^3 dt. \tag{57}$$

A justification of these generalised substitution rules, to be understood in a precise L^2 sense, is presented in appendix B. It is similar in spirit to the usual mathematical definition of the first rule (the usual Ito prescription (9)), the L^2 definition of which is also recalled in this appendix.

3.1.2.4. Discussion and comparison to a naive continuous-time computation. In section 3.1.1 we showed that the change of discretisation at the Langevin equation level requires the use of the standard substitution rule (9) (the Ito prescription). This transformation follows the upper branch in figure 1. In section 3.1.2.3 we proved that the change of discretisation at the Onsager–Machlup level (for the infinitesimal propagator) requires a full set of generalised substitution rules, given by the relations (54)–(57), that include the Ito prescription (9) but extend it with transformation rules for three other infinitesimals. This transformation follows the lower branch in figure 1. Therefore, the paths along the upper and lower branches should be followed using procedures that involve a different set of substitution rules.

The key point that explains the discrepancy between the two approaches is that when one changes the discretisation in the action, the term which is quadratic in Δx (see (49)) transforms in a non-trivial way and contributes to a higher order in powers of Δx than when one changes the discretisation in the Langevin equation (as done in section 3.1.1). Technically, the presence of a square $(\Delta x/dt)^2$ divided by the noise amplitude in the infinitesimal propagator implies that, when keeping terms of order $O(dt^{1/2})$ and $O(dt)$, higher powers of Δx are generated, as observed in (50).

An instructive observation is to draw a comparison between the Stratonovich-discretised continuous-time action corresponding to (51)

$$\mathcal{S}_{f_\alpha}^S[x(t)] \stackrel{S}{=} \frac{1}{2} \int_0^{t_f} dt \left\{ \frac{1}{2D} \left[\frac{d_t x - f_\alpha(x) + D g(x) g'(x)}{g(x)} \right]^2 - dt f'_\alpha(x) \right\} \quad (58)$$

and the result of a naive computation. First, one notes that both the right-down and the down-right branches of the commutative diagram represented in figure 1 agree with the same result (58), together with the corresponding prefactor $\prod_t |g(\bar{x}_t^S)|$; this is true provided one uses the generalised substitution rules (54)–(57). Another—naive—approach consists in attempting to arrive at this result by changing the discretisation directly in the time-continuous action, with the following procedure:

- (i) start from the continuous-time α -discretised action (42);
- (ii) use the rules (16) and (17) for the change of discretisation in the Langevin equation;
- (iii) change the discretisation of the normalisation prefactor (43) from α to Stratonovich, using a relation similar to (40)⁶.

However, as detailed in appendix D, the result of this procedure is different from (58) and is thus *incorrect*. The reason lies in the fact that the rules (16) and (17) for the change of discretisation in the Langevin equation do not involve substitution rules of high enough order in Δx : they disregard essential terms contributing to the expansion (50) that are crucial to arrive at the final correct propagator (51) (or, equivalently, to recover the correct action (58) with its associated Stratonovich-discretised normalisation prefactor). This confirms that the sole standard substitution rule (9) is not sufficient to handle successfully the path integral representation of the stochastic process, and that the generalised substitution rules (54)–(57) that we propose have to be used instead.

⁶ The relation (40) allows one to change the discretisation of the prefactor $\mathcal{J}[x(t)]$ from the Ito one ($\alpha = 0$) to the α one, but is easily adapted to change from α to Stratonovich ($\alpha = 1/2$); see equation (D.6).

3.2. Non-linear transformation

A similar apparent contradiction occurs when one attempts to use the chain rule (7) in the action, instead of restricting its use to the Langevin level. Such inconsistency was observed for non-linear field transformations for the MSRJD action in App. E of [7]⁷. In appendix C of the present article, we translate this computation to the case of the Onsager–Machlup action, and the result is the same: using the chain rule (7) in the action brings an inconsistency when changing variables. Related issues have been observed in the context of quantum field theory [14–19]. In this section, we examine the origin of this paradox. We show that it is again due to an invalid use of the Ito substitution rule (9) in the action, and we provide the correct treatment of non-linear transformations in the action, working with the infinitesimal propagator. We also propose a modified chain rule that can be used in continuous time inside the action.

3.2.1. Non-linear transformation in the Langevin equation. We start from the α -discretised Langevin equation (6) and consider an increasing C^1 function $u(x)$ which is a bijection and is used as a non-linear change of variables. The chain rule (7) implies that the function $U(t) = u(x(t))$ satisfies an α -discretised Langevin equation

$$d_t U \stackrel{\alpha}{=} u'f + (1 - 2\alpha)Dg^2u'' + u'g\eta. \quad (59)$$

This writing is a shortcut for the Langevin equation with a force F and a noise amplitude G

$$d_t U(t) \stackrel{\alpha}{=} \overbrace{u'(x_U(t))f(x_U(t)) + (1 - 2\alpha)Dg(x_U(t))^2u''(x_U(t))}^{\equiv F(U(t))} + \underbrace{u'(x_U(t))g(x_U(t))\eta(t)}_{\equiv G(U(t))} \quad (60)$$

$$x_U(t) = u^{-1}(U(t)) \quad (\text{i.e. } u(x_U(t)) = U(t)). \quad (61)$$

Our aim is to compare different procedures represented on the commutative diagram of figure 2. Concretely we take the following two paths.

- (i) The down path (on the left) that starts from the α -discretised Langevin equation (6) and arrives at the Onsager–Machlup action on $x(t)$ given by the expression in (42), which, together with its associated normalisation prefactor (43), is equivalent to the infinitesimal propagator (41).
- (ii) The right-down-left path. It starts from the Langevin equation (59), goes next to its corresponding Onsager–Machlup representation and, finally, through the application of rules that we still need to find, this path performs a non-linear transformation on the Onsager–Machlup action on $U(t)$ that should take it to the one on $x(t)$.

We first analyse these procedures at the infinitesimal propagator level.

3.2.2. Direct determination of the propagator. As in section 3.1.2, we perform the comparison by keeping only the quadratic in Δx contribution to the Gaussian weight in the exponential, and by expanding the rest in front of this weight. The propagator (41) associated to the Langevin equation (6) reads

⁷Note that this appendix is found only in the arXiv v1 preprint version.

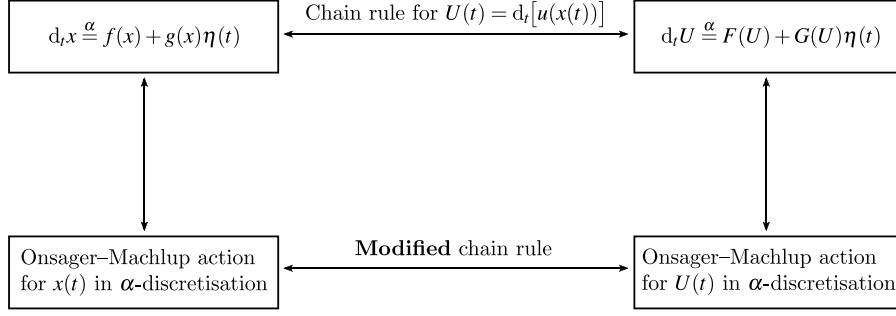


Figure 2. Schematic representation, for a non-linear (bijective) change of variables $x \mapsto u(x)$, of the difference between the stochastic calculus performed in the Langevin equation and in the Onsager–Machlup action. The α -discretised Langevin equation (6) can be transformed exploiting the chain rule (7) into a Langevin equation (59) for $U(t) = u(x(t))$. To this equation corresponds a Onsager–Machlup action, equivalent to the infinitesimal propagator (63). However, one cannot use such chain rule in the α -discretised Onsager–Machlup action (42) to derive the correct action of the process $U(t)$, as explained in appendix C. To go from one action to the other, one has to use, instead, the generalised substitution rules (54)–(57) in discrete time after expanding the action (or equivalently (74) and (75) inside the discretised action), or to rely on a *modified* chain rule for the time-continuous process as discussed in section 3.3.2. In the text, for simplicity, the lower branch of this commutative diagram is performed from right to left.

$$\mathbb{P}(x_{dt}|x_0) \stackrel{\alpha}{=} \frac{\mathcal{N}}{|g(\bar{x}_0)|} e^{-\frac{1}{2} \frac{dt}{2D} \left(\frac{\Delta x}{dt} \right)^2 / g(\bar{x}_0)^2} \times \left\{ 1 - dt \alpha f'(\bar{x}_0) + \frac{f(\bar{x}_0) - 2D\alpha g(\bar{x}_0) g'(\bar{x}_0)}{2Dg(\bar{x}_0)^2} \Delta x \right\}. \quad (62)$$

In this expansion, we have already used the standard substitution rule (9) to reexpress Δx^2 .

3.2.3. Indirect path: passing through the propagator for $U(t)$. Corresponding to the Langevin equation (60) for $U(t)$, one can write from (41) the propagator

$$\mathbb{P}_U(U_{dt}|U_0) \stackrel{\alpha}{=} \frac{\mathcal{N}}{|G(\bar{U}_0)|} \exp \left\{ -\frac{1}{2} \frac{dt}{2D} \left[\frac{\frac{U_{dt}-U_0}{dt} - F(\bar{U}_0) + 2\alpha D G(\bar{U}_0) G'(\bar{U}_0)}{G(\bar{U}_0)} \right]^2 - \alpha dt F'(\bar{U}_0) \right\}. \quad (63)$$

Since the two Langevin equations (6) and (60) are equivalent, this propagator has to be equivalent to (62). As remarked in the literature in the stochastic [7, 31] and the quantum mechanical [14–19] contexts, the application of the chain rule does *not* yield back (62) or (41). The computation describing this inconsistency for our Onsager–Machlup action of interest is recalled for completeness in appendix C

The idea to examine the origin of this inconsistency, as done previously for the change of discretisation, is to treat the ‘dangerous’ term of the propagator $\left[\frac{U_{dt}-U_0}{dt G(\bar{U}_0)} \right]^2$ in a safe way, by expanding the propagator and putting all terms in prefactor, apart from the quadratic part defining the Gaussian weight itself. To set up the expansion, one uses that

$$\bar{U}_0 = (1 - \alpha)u(x_0) + \alpha u(x_{dt}), \quad (64)$$

$$x_0 = \bar{x}_0 - \alpha \Delta x, \quad (65)$$

$$x_{dt} = \bar{x}_0 + (1 - \alpha)\Delta x, \quad (66)$$

and one expands in powers of Δx , keeping in mind that this quantity is $O(dt^{1/2})$. The change of variables in the (conditional) probability

$$\mathbb{P}_U(U_{dt}|U_0) u'(x_{dt}) = \mathbb{P}(x_{dt}|x_0) \quad (67)$$

is also needed, where in $u'(x_{dt})$ one uses (66). After a tedious computation (where the substitution rule (53) for Δx^2 is employed though only in the prefactor), the result is that the propagator $\mathbb{P}(x_{dt}|x_0)$ obtained from (67), with \mathbb{P}_U read from (63), is

$$\begin{aligned} & \frac{\mathbb{P}(x_{dt}|x_0)}{|g(\bar{x}_0)|} e^{-\frac{1}{2} \frac{dt}{D} \left(\frac{\Delta x}{dt} \right)^2 / g(\bar{x}_0)^2} \\ & \quad \doteq 1 + \left[\frac{f(\bar{x}_0)}{2Dg(\bar{x}_0)^2} - \frac{\alpha g'(\bar{x}_0)}{g(\bar{x}_0)} - \frac{3(-1 + 2\alpha)u''(\bar{x}_0)}{2u'(\bar{x}_0)} \right] \Delta x \\ & \quad + \left[\frac{3D\alpha(-2 + 3\alpha)g(\bar{x}_0)g'(\bar{x}_0)u''(\bar{x}_0)}{u'(\bar{x}_0)} - \frac{(2\alpha f'(\bar{x}_0)u'(\bar{x}_0) + 3(-1 + 2\alpha)f(\bar{x}_0)u''(\bar{x}_0))}{2u'(\bar{x}_0)} \right. \\ & \quad \left. + \frac{Dg(\bar{x}_0)^2(3(1 - 6\alpha + 6\alpha^2)u''(\bar{x}_0)^2 + 2(1 - 3\alpha + 3\alpha^2)u'(\bar{x}_0)u^{(3)}(\bar{x}_0))}{2u'(\bar{x}_0)^2} \right] dt \\ & \quad + \frac{(-1 + 2\alpha)u''(\bar{x}_0)}{4dtDg(\bar{x}_0)^2u'(\bar{x}_0)} \Delta x^3 \\ & \quad + \left[\frac{(-1 + 2\alpha)f(\bar{x}_0)u''(\bar{x}_0)}{8dtD^2g(\bar{x}_0)^4u'(\bar{x}_0)} - \frac{\alpha(-2 + 3\alpha)g'(\bar{x}_0)u''(\bar{x}_0)}{4dtDg(\bar{x}_0)^3u'(\bar{x}_0)} \right. \\ & \quad \left. - \frac{3(7 - 32\alpha + 32\alpha^2)u''(\bar{x}_0)^2 + 4(1 - 3\alpha + 3\alpha^2)u'(\bar{x}_0)u^{(3)}(\bar{x}_0)}{48dtDg(\bar{x}_0)^2u'(\bar{x}_0)^2} \right] \Delta x^4 \\ & \quad + \frac{(1 - 2\alpha)^2u''(\bar{x}_0)^2}{32dt^2D^2g(\bar{x}_0)^4u'(\bar{x}_0)^2} \Delta x^6. \end{aligned} \quad (68)$$

This form seems to be different from (62) because it still involves the function $u(x)$ that should not be present in the microscopic propagator for x (unless of course the transformation is the identity $u(x) = x$ in which case (68) is equal to (62)). However, as checked with a direct computation, using the generalised substitution rules (55)–(57) allows one to remove all dependencies of (68) in the function $u(x)$. Strikingly, the result is the correct propagator (62).

This computation shows that one can follow without inconsistencies the different branches of figure 2 for non-linear transformations, provided that the correct expansion is done when performing the change of variables in the action (yielding (68)) and that the generalised substitution rules (55)–(57) are applied to the prefactor of the Gaussian weight $|g(\bar{x}_t)|^{-1} \exp \{ -\Delta x^2 / [4D dt g(\bar{x}_t)]^2 \}$, after the expansion of the infinitesimal propagator.

3.3. Discussion

In this section, we gather the previous results on the change of discretisation and the change of variables in a common description, aiming at understanding which are the valid rules of stochastic calculus that apply in the action. We first describe the origin of the observed issues in the infinitesimal propagator, setting down modified substitution rules than can be applied ‘inside’ the exponential of the propagator (instead of ‘outside’, on the prefactor of the Gaussian weight as done so far). We then formulate a modified chain rule in continuous time that one should apply in the path integral formalism.

3.3.1. (Generalised) substitution rules and exponentials of infinitesimals. In sections 3.1 and 3.2, we noted that the expansion of the infinitesimal propagator involves a separation between

- (i) a purely Gaussian weight (which defines the probability distribution of the increment $\Delta x = x_{dt} - x_0$) and
- (ii) a prefactor gathering all other terms, of the form $1 + O(dt^{1/2}) + O(dt)$.

We now first show explicitly that the generalised substitution rules (54)–(57) cannot be applied in the exponential and we elucidate which are the ‘modified substitution rules’ to use in the exponentiated expression. Recalling the notation, $\mathcal{N} = 1/(4\pi D dt)^{1/2}$ we denote by

$$\mathbb{P}_t^G \stackrel{\alpha}{=} \frac{\mathcal{N}}{|g(\bar{x}_t)|} e^{-\frac{1}{2} \frac{dt}{D} \left(\frac{\Delta x}{dt} \right)^2 / g(\bar{x}_t)^2} \quad (69)$$

the part of the infinitesimal propagator (taken in a given α -discretisation) which corresponds to the Gaussian distribution of Δx . Then, either for the change of discretisation (50) or for the non-linear change of variables (68), the microscopic propagator is decomposed as

$$\mathbb{P}(x_{dt}|x_0) \stackrel{\alpha}{=} \mathbb{P}_0^G \times \exp \left\{ A_0 dt + A_1 \Delta x + A_2 \Delta x^2 + A_3 \frac{\Delta x^3}{dt} + A_4 \frac{\Delta x^4}{dt} \right\} \quad (70)$$

where A_0, \dots, A_4 are functions of \bar{x}_0 taken in α -discretisation. The number of terms inside the exponential is finite, because higher-order powers of Δx and dt do not contribute at the orders $O(dt^{1/2})$ and $O(dt)$ we are interested in⁸. Note that (50) and (68) are written in an expanded form, which goes up to order Δx^6 as

$$\begin{aligned} & \exp \left\{ A_0 dt + A_1 \Delta x + A_2 \Delta x^2 + A_3 \frac{\Delta x^3}{dt} + A_4 \frac{\Delta x^4}{dt} \right\} \\ &= 1 + A_0 dt + A_1 \Delta x + \left(\frac{1}{2} A_1^2 + A_2 \right) \Delta x^2 \\ & \quad + A_3 \frac{\Delta x^3}{dt} + (A_1 A_3 + A_4) \frac{\Delta x^4}{dt} + \frac{A_3^2}{2} \frac{\Delta x^6}{dt^2}. \end{aligned} \quad (71)$$

In this form, one can then apply the generalised substitution rules (54)–(57) in a valid manner and reexponentiate the result, taking into account the orders in dt correctly. (This is similar to what we have done in (33) when treating the exponential of functions of the noise η_t only that now we deal with a function of Δx .) Denoting by $\sigma = 2Dg^2(x)$ the noise amplitude, one finds that the form (70) of the propagator becomes

⁸Of course other computations than the change of discretisation and the change of variables that we considered in sections 3.1 and 3.2 could generate larger powers of Δx , such as $\Delta x^5/dt^2$ or $\Delta x^6/dt^2$, which are respectively of order $O(dt^{1/2})$ and $O(dt)$. The modified rule that we present in the present section are easily adapted to such terms.

$$\mathbb{P}(x_{dt}|x_0) \stackrel{\alpha}{=} \mathbb{P}_0^G \times \exp \left\{ \left[A_1 + 3A_3\sigma \right] \Delta x + \left[A_0 + A_2\sigma + 3(A_3)^2\sigma^2 + 3A_4\sigma^2 \right] dt \right\} \quad (72)$$

with terms in the exponential that are order Δx (or $dt^{1/2}$) and dt only, as they should.

One observes by direct inspection that the generalised substitution rules (54)–(57) *cannot* be used directly inside the exponential of (70) in order to get the correct result (72). Indeed, the term $A_3 \frac{\Delta x^3}{dt}$ in (70) generates a quadratic contribution $\propto (A_3)^2$ in (72). The valid ‘modified substitution rule’ to use in the exponential (70) are thus

$$\left. \begin{aligned} A_2(x) \Delta x^2 &\mapsto A_2(x) 2Dg(x)^2 dt & (73) \\ A_3(x) \Delta x^3 dt^{-1} &\mapsto 3A_3(x) 2Dg(x)^2 \Delta x + 3[A_3(x) 2Dg(x)^2 \Delta x]^2 dt & (74) \\ A_4(x) \Delta x^4 dt^{-1} &\mapsto 3A_4(x) (2Dg(x)^2)^2 dt & (75) \end{aligned} \right\} \begin{array}{l} \text{valid} \\ \text{only in} \\ \text{the exp.} \end{array}$$

One observes that while the first and third line coincide with the corresponding ones in (54) and (56), the second line is different: in (55) $\Delta x^3 dt^{-1}$ is substituted by an expression which is independent of its possible prefactor, while in the exponential (70) we need to use (74) that *effectively* replaces $\Delta x^3 dt^{-1}$ by an expression which explicitly depends on its prefactor A_3 (in other words, the second term depends on $[A_3(x)]^2$).

In the formulation leading from (70) to (72) it is rather evident that the generalised substitution rules (54)–(57) cannot be applied inside the exponential: indeed one can see $e^{A_3 \Delta x^3 / dt}$ as equivalent to a moment-generating function of parameter A_3 , and the exponent in (72) as the corresponding cumulant-generating function, cut after $O(dt)$; thus, forgetting the quadratic term $\propto (A_3)^2$ in (74), which is of order dt , amounts to forgetting the term of degree 2 in the expansion of a cumulant-generating function⁹.

We note that Gervais and Jevicki [15] have also determined in a quantum-field theory context that the correct procedure to change variables (in their case, to perform a canonical transformation) requires an expansion of the exponent up to terms of order $\Delta x^4 dt^{-1}$, akin to (70). However, to our understanding, their treatment of these terms is unrelated to ours and remains perturbative in D , in contrast to our treatment which is non-perturbative.

3.3.2. Modified chain rule. The chain rule (7) allows one to deduce an α -discretised Langevin equation on a variable $U(t) = u(x(t))$ from the corresponding Langevin equation on the variable $x(t)$ after a non-linear transformation, as discussed in section 3.2. This same chain rule does not directly allow one to perform such non-linear change of variables at the level of the action (see appendix C). To understand this issue on a general footing, let us start from a Langevin equation of the form (60)

$$d_t U(t) \stackrel{\alpha}{=} F(U(t)) + G(U(t)) \eta(t) \quad (76)$$

The corresponding Onsager–Machlup weight reads

$$\prod_t \frac{\mathcal{N}}{|G(\bar{U})|} \times \exp \left\{ - \int_0^t dt \left[\frac{1}{4D} \left(\frac{d_t U - F(\bar{U}) + 2\alpha D G(\bar{U}) G'(\bar{U})}{G(\bar{U})} \right)^2 + \alpha dt F'(\bar{U}) \right] \right\} \quad (77)$$

⁹For Δx^2 and Δx^4 , the higher-order term of the cumulant expansion do not contribute because they are $o(dt)$. However, if a term in $A_5 \Delta x^5 dt^{-2}$ had been present in (70), its modified substitution rule in the exponential would present a quadratic contribution $\propto (A_5)^2$ as in (74) for $A_3 \Delta x^3 dt^{-1}$.

with each $\bar{U} = U(t)$ taken in α -discretisation. The naive approach consists in substituting $U(t)$ by $u(x(t))$ and then using the chain rule to determine the Onsager–Machlup weight for the trajectory $x(t)$. In the next paragraphs, in order to understand why this procedure fails, we come back to the microscopic propagator (63) corresponding to (77), in which we expand the square and we study separately the terms affine in $d_t U$ and the term quadratic in $d_t U$. As we now show, the result is that the standard chain rule allows one to transform the terms affine in $d_t U$, while to correctly transform the quadratic term $\propto (d_t U)^2$, one has to use a ‘modified chain rule’.

3.3.2.1. Terms affine in $d_t U$ For the infinitesimal propagator, these terms take the form

$$\mathcal{B}_1 = dt B_0(\bar{U}) + dt B_1(\bar{U}) d_t U \quad (78)$$

where the first and second terms are of orders $O(dt)$ and $O(dt^{1/2})$, respectively. In order to reexpress $e^{\mathcal{B}_1}$ in terms of the original variable $x(t)$ one can follow either of the two following approaches.

- (i) In discrete time, one takes the same path as in the previous section. Firstly, one discretises time explicitly; secondly, one expands $U_{dt} - U_0 = U(x_{dt}) - U(x_0)$ in powers of $\Delta x = x_{dt} - x_0$ around \bar{x}_0 using (65)–(66). With the usual substitution rule (9) $\Delta x^2 = 2Dg(x)^2 dt$, after reexponentiation one obtains that

$$e^{\mathcal{B}_1} \stackrel{\alpha}{=} \exp \left\{ dt B_0(u(x(t))) + dt B_1(u(x(t))) \times \left[u'(x(t)) d_t x(t) + (1 - 2\alpha) D g(x(t))^2 u''(x(t)) \right] \right\} \quad (79)$$

in the $dt \rightarrow 0$ limit. In the light of section 3.1.1, the computation involves no term in $\Delta x^3 dt^{-1}$ (nor higher order in powers of $\Delta x^n dt^m$), implying that the standard substitution rules could have also been applied *inside* the exponential.

- (ii) In continuous time, one can use the chain rule (7) inside the exponential, for $U(t) = u(x(t))$ to get the result (79). It is valid here as shown by the discrete-time computation described in the previous point.

3.3.2.2. Term proportional to $(d_t U)^2$ This term takes the form

$$\mathcal{B}_2 = -\frac{1}{2} \frac{dt}{2D} B_2(\bar{U}) (d_t U)^2 \quad \text{with } B_2 = \frac{1}{G^2}. \quad (80)$$

It is of order dt^0 and if one naively uses the chain rule to compute $d_t U = d_t[u(x(t))]$, one misses a number of terms; such computation would yield

$$e^{\mathcal{B}_2} \stackrel{\text{wrong!}}{=} \exp \left\{ -\frac{dt}{4D} B_2(u(x(t))) \left[u'(x(t)) d_t x(t) + (1 - 2\alpha) D g(x(t))^2 u''(x(t)) \right]^2 \right\} \quad (81)$$

where $g(x)^2 = 1/B_2(u(x))$. Instead, one should discretise in time, using

$$\mathcal{B}_2 = -\frac{1}{4D} dt^{-1} B_2(\bar{U}_0) (U_{dt} - U_0)^2 \quad \text{as } dt \rightarrow 0 \quad (82)$$

where $U_0 = u(x_0)$, $U_{dt} = u(x_{dt})$ and \bar{U}_0 is defined in (64). We also define the function $b_2(x) = B_2(u(x)) = 1/g(x)^2$ for lighter notations.

Expansion of \mathcal{B}_2 . Using the relations (64)–(66), one then expands (82) in powers of Δx up to order dt to find

$$\begin{aligned} \mathcal{B}_2 = & -\frac{1}{4Dg(\bar{x}_0)^2} \frac{\Delta x^2}{dt} + \frac{(-1+2\alpha)u''(\bar{x}_0)}{4Dg(\bar{x}_0)^2u'(\bar{x}_0)} \frac{\Delta x^3}{dt} \\ & + \left[-\frac{(-1+\alpha)\alpha g'(\bar{x}_0)u''(\bar{x}_0)}{4Dg(\bar{x}_0)^3u'(\bar{x}_0)} - \frac{(1+8(-1+\alpha)\alpha)u''(\bar{x}_0)^2}{16Dg(\bar{x}_0)^2u'(\bar{x}_0)^2} \right. \\ & \left. + \frac{(-1-3(-1+\alpha)\alpha)u^{(3)}(\bar{x}_0)}{12Dg(\bar{x}_0)^2u'(\bar{x}_0)} \right] \frac{\Delta x^4}{dt}, \quad (83) \end{aligned}$$

which is not obviously related to (81). We note that this expression contains a crucial term proportional to $\Delta x^3 dt^{-1}$ which, as we have discussed in section 3.3.1, has to be treated with great care. The modified substitution rule (74) has to be used here (and not the rule (55)), in order to handle correctly $\Delta x^3 dt^{-1}$ inside the exponential. We also remark that the term in $\Delta x^3 dt^{-1}$ is non-zero for an additive noise (i.e. when $g(x)$ is constant), indicating that non-linear changes of variables also have to be handled with care in this case.

Expansion of $e^{\mathcal{B}_2}$. The correct procedure to follow in order to first use the (simple) substitution rules (54)–(57) for Δx^n is to first expand the terms of (83) which are not in $\Delta x^2 dt^{-1}$, and to use then the substitution rules (54)–(57). Alternatively, one can use the modified ones (73)–(75) which are valid inside an exponential. Recalling the notation $g^2 = 1/b_2$, and writing

$$e^{(\dots)\Delta x^2} = \exp \left[-\frac{1}{2} \frac{b_2(\bar{x}_0) dt}{2D} \left(\frac{\Delta x}{dt} \right)^2 (u'(\bar{x}_0))^2 \right] \quad (84)$$

one obtains

$$\begin{aligned} \frac{e^{\mathcal{B}_2}}{e^{(\dots)\Delta x^2}} = & \exp \left\{ \frac{3}{2} (-1+2\alpha)u'(\bar{x}_0)u''(\bar{x}_0)\Delta x \right. \\ & + dt \left[-\frac{3D(1-2\alpha)^2u''(\bar{x}_0)^2}{4b_2(\bar{x}_0)} + \frac{3D(1-2\alpha)^2u'(\bar{x}_0)^2u''(\bar{x}_0)^2}{2b_2(\bar{x}_0)} \right. \\ & + u'(\bar{x}_0) \left(\frac{3D(-1+\alpha)\alpha b_2'(\bar{x}_0)u''(\bar{x}_0)}{2b_2(\bar{x}_0)^2} \right. \\ & \left. \left. - \frac{D(1+3(-1+\alpha)\alpha)u^{(3)}(\bar{x}_0)}{b_2(\bar{x}_0)} \right) \right] \left. \right\}. \quad (85) \end{aligned}$$

This result is completely different from the naive result (81), obtained from the use of the chain rule (7) in the exponential, that can be recast as

$$\frac{e^{\mathcal{B}_2}}{e^{(\dots)\Delta x^2}} \stackrel{\text{wrong!}}{=} \exp \left\{ \frac{1}{2} (-1+2\alpha)u'(\bar{x}_0)u''(\bar{x}_0)\Delta x - \frac{D(1-2\alpha)^2u''(\bar{x}_0)^2}{4b_2(\bar{x}_0)} dt \right\}. \quad (86)$$

(They coincide for linear transformations such that $u'' = 0$.)

The result (85) allows one to identify the correct (but complicated) form of the chain rule to be used in the exponential for terms of the form (80). Instead of the chain rule (7) that would lead to (86), one has that

$$\begin{aligned}
& -\frac{1}{2} \frac{dt}{2D} B_2(U) (d_t U)^2 \\
& \stackrel{\text{in exp.}}{\mapsto} -\frac{1}{2} \frac{dt}{2D} b_2(x) (d_t x)^2 (u'(x))^2 \\
& + \frac{3}{2} (-1 + 2\alpha) u'(x) u''(x) dt d_t x \\
& + dt \left[-\frac{3D(1-2\alpha)^2 u''(x)^2}{4b_2(x)} + \frac{3D(1-2\alpha)^2 u'(x)^2 u''(x)^2}{2b_2(x)} \right. \\
& \quad \left. + u'(x) \left(\frac{3D(-1+\alpha)\alpha b'_2(x) u''(x)}{2b_2(x)^2} - \frac{D(1+3(-1+\alpha)\alpha) u^{(3)}(x)}{b_2(x)} \right) \right]. \tag{87}
\end{aligned}$$

(We took the $dt \rightarrow 0$ limit, with the r.h.s. being α -discretised.) In order to apprehend better the difference with the naive application of the chain rule (7), one can rewrite this result as

$$\begin{aligned}
& -\frac{1}{2} \frac{dt}{2D} B_2(U) (d_t U)^2 \\
& \stackrel{\text{in exp.}}{\mapsto} -\frac{1}{2} \frac{dt}{2D} b_2(x) \left[\overbrace{u'(x) d_t x + (1-2\alpha) D g(x)^2 u''(x)}^{\text{chain rule (7)}} \right]^2 \\
& + (-1 + 2\alpha) u'(x) u''(x) dt d_t x \\
& + dt \left[\frac{3D(-1+\alpha)\alpha u'(x) b'_2(x) u''(x)}{2b_2(x)^2} \right. \\
& \quad \left. + \frac{D((1-2\alpha)^2(-1+3u'(x)^2)u''(x)^2 - 2(1+3(-1+\alpha)\alpha)u'(x)u^{(3)}(x))}{2b_2(x)} \right], \tag{88}
\end{aligned}$$

the three last lines being the terms one misses if one merely applies (7).

Special cases. One notes that this modified chain rule remains non-trivial in the three following simplified cases:

- Stratonovich discretisation ($\alpha = 1/2$):

$$\begin{aligned}
& -\frac{1}{2} \frac{dt}{2D} B_2(U) (d_t U)^2 \\
& \stackrel{\text{in exp.}}{\mapsto} -\frac{1}{2} \frac{dt}{2D} b_2(x) [u'(x) d_t x]^2 - D dt \left[\frac{3u'(x) b'_2(x) u''(x)}{8b_2(x)^2} + \frac{u'(x) u^{(3)}(x)}{4b_2(x)} \right]. \tag{89}
\end{aligned}$$

- Additive noise ($B_2(U) = B_2 = b_2(x) = b_2 = 1/g^2(x) = 1/g^2$):

$$\begin{aligned}
& -\frac{1}{2} \frac{dt}{2D} b_2 (d_t U)^2 \\
& \stackrel{\text{in exp.}}{\mapsto} -\frac{1}{2} \frac{dt}{2D} b_2 \left[u'(x) d_t x + (1-2\alpha) D g^2 u''(x) \right]^2 \\
& + (-1 + 2\alpha) u'(x) u''(x) dt d_t x \\
& + \frac{D dt}{2b_2} \left[(1-2\alpha)^2 (-1+3u'(x)^2) u''(x)^2 - 2(1+3(-1+\alpha)\alpha) u'(x) u^{(3)}(x) \right]. \tag{90}
\end{aligned}$$

- Additive noise and Stratonovich discretisation:

$$-\frac{1}{2} \frac{dt}{2D} b_2 (d_t U)^2 \xrightarrow{\text{in exp.}} -\frac{1}{2} \frac{dt}{2D} b_2 [u'(x) d_t x]^2 - D dt \frac{u'(x) u^{(3)}(x)}{4b_2}. \quad (91)$$

This last case is peculiarly striking, because one could have expected the standard chain rule of differentiable calculus to be valid in the dynamical action of an additive-noise Stratonovich-discretised Langevin equation (as it is valid at the Langevin equation level). Surprisingly, this is not the case as soon as $u^{(3)}(x) \neq 0$.

4. Outlook

The trajectory probability of Langevin processes is well described by a path-integral weight, through either the MSRJD [24, 25] or the Onsager–Machlup [20, 21] formulations. In this article we studied the behaviour of the Langevin equation and its corresponding Onsager–Machlup action under two generic transformations: a change of α -discretisation and a non-linear change of variables. The correct rules to perform these transformations at the level of the Langevin equations are well-known, they have been recalled in this article, and we verified, once again, that they are reversible.

Consistency requires that the trajectory probability constructed from the Langevin equation of a variable $u(t) = u(x(t))$ in a discretisation scheme $\bar{\alpha}$ be the same as the trajectory probability of the Langevin process of the variable $x(t)$ in another discretisation scheme α , after applying to the latter the corresponding discretisation and non-linear transformations. Figures 1 and 2 provide sketches of this statement for the discretisation scheme transformation and the non-linear transformation, respectively. However, it was observed in the literature that their use in the action could yield inconsistencies, both in the stochastic field-theory context [7, 31] and in the quantum-mechanical one [14–19]. The aim of the present article was to identify the generalisation of the Ito rule and the correct rules of calculus that ensure the reversibility of the construction.

By carefully analysing the discrete-time behaviour of the propagator corresponding to the infinitesimal evolution during a time step $dt \rightarrow 0$, we identified the source of inconsistencies and we provided procedures that allow one to perform the transformations in the action in a correct manner.

To summarise them, we now list the possible sources of issues. At the infinitesimal level, we denote the trajectory increment by $\Delta x = x_{t+dt} - x_t$ which is typically of order $dt^{1/2}$. The main source of problems is that terms of the form $\Delta x^3 dt^{-1}$, $\Delta x^4 dt^{-1}$ and $\Delta x^6 dt^{-2}$ are generated in the infinitesimal propagator upon the mentioned transformations, while they are not generated at the Langevin level. First, they have to be correctly identified, and second, one has to understand their behaviour in the $dt \rightarrow 0$ limit. We have provided generalised substitution rules (54)–(57) that allow one to do so (they generalise the usual Ito prescription $dB_t^2 = dt$ for the Brownian motion). An important point is that these relatively simple rules have to be used in the prefactor of the Gaussian weight of the infinitesimal propagator (after a $dt \rightarrow 0$ expansion), and not inside the exponential of this propagator. We have provided a simple explanation of this condition in section 3.3.1. If one insisted upon applying the transformations in the exponential, the modified substitution rules become significantly more complicated and are given in equations (73)–(75).

In the continuous-time path integral, an important consequence of the previous observations is that one cannot use the stochastic chain rule (7) to perform changes of variables. One

has, instead, to rely on a time-discrete expansion or on a modified chain rule, described in section 3.3.2. We emphasise that the application of the invalid chain rule (7) in the action yields wrong results *even for an additive-noise Stratonovich-discretised Langevin equation*. The reason for this is that under a non-linear transformation of variables the equation becomes one with multiplicative noise.

For future perspectives, we can list a number of interesting questions to address:

- (i) It would be helpful to identify similar rules that would solve inconsistencies observed when manipulating the MSRJD action [7], because many field theories (including quantum ones) are better written in this formalism or in similar ones that also involve a response field.
- (ii) The generalisation to more than one degree of freedom could be tricky [6] but should be very interesting and useful.
- (iii) Langevin equations with inertia (a second time derivative) and/or coloured noise approach in the overdamped and/or white noise limit the equation that we studied here in the Stratonovich scheme (see, e.g. [2, 32]). It would be interesting to understand how the issues discussed in the present article arise and are solved in these regularised and better behaved cases since, as we showed, even the action in the Stratonovich discretisation scheme has to be treated attentively.
- (iv) The results we have presented also encourage one to revisit the validity of some non-linear transformation used in quantum field theory [14–19], where the Lagrangians defining the action take forms that are similar to that of statistical mechanics.

Acknowledgments

We are deeply indebted to Maxence Ernout, in collaboration with whom we initiated this research work. LFC gratefully thanks Camille Aron and Gustavo Lozano for early discussions on this problem. She is a member of Institut Universitaire de France. VL gratefully thanks Eric Bertin for fruitful discussions on stochastic calculus, and acknowledges support by the ERC Starting Grant 680275 MALIG, by the ANR-15-CE40-0020-03 Grant LSD and by the UGA IRS PHEMIN project.

Appendix A. Determination of the infinitesimal propagator: other approaches

In this appendix, in order to shed a different light on the use of the Ito prescription in the determination of the infinitesimal propagator, we review other less pedestrian approaches than the one presented in section 2.2.3.

A.1. À la Lau–Lubensky

To compute $\delta(x_{dt} - X_1(x_0, \eta_0))$ in (22), it proves simpler [6] to start from the following identity, where the argument of the first delta is the equation of motion at $t = 0$:

$$\delta\left(\overbrace{\eta_0 - \frac{x_{dt} - x_0}{dt} - f(\bar{x}_0)}^{\equiv F(\eta_0, x_0, x_{dt})}\right) \stackrel{(21)}{=} \frac{1}{|\partial_{x_{dt}} F(\eta_0, x_0, x_{dt})|} \delta(x_{dt} - X_1(x_0, \eta_0)), \quad (\text{A.1})$$

where one recognises $F(\eta_0, x_0, x_{dt}) = \eta_0 - H_0(x_0, x_{dt})$ from equation (35). One thus has

$$|\partial_{x_{dt}} H_0(x_0, x_{dt})| \delta(\eta_0 - H_0(x_0, x_{dt})) \stackrel{(35)}{=} \delta(x_{dt} - X_1(x_0, \eta_0)) , \quad (\text{A.2})$$

so that finally

$$\mathbb{P}(x_{dt}|x_0) \stackrel{(22)}{=} \int d\eta_0 |\partial_{x_{dt}} H_0(x_0, x_{dt})| \delta(\eta_0 - H_0(x_0, x_{dt})) P_{\text{noise}}(\eta_0) . \quad (\text{A.3})$$

By direct computation, one obtains

$$\partial_{x_{dt}} H_0(x_0, x_{dt}) = \frac{1}{dt} \frac{1}{g(\bar{x}_0)} \left[1 - \alpha dt f'(\bar{x}_0) - (x_{dt} - x_0 - f(\bar{x}_0) dt) \alpha \frac{g'(\bar{x}_0)}{g(\bar{x}_0)} \right] \quad (\text{A.4})$$

that, using the Dirac delta in (A.3) to re-express $x_{dt} - x_0$ as a function of η_0 , implies

$$\begin{aligned} & |\partial_{x_{dt}} H_0(x_0, x_{dt})| \delta(\eta_0 - H_0(x_0, x_{dt})) \\ & \stackrel{(36)}{=} \frac{1}{dt} \frac{1}{|g(\bar{x}_0)|} \left[1 - \alpha dt f'(\bar{x}_0) - \eta_0 g'(\bar{x}_0) \alpha dt \right] \end{aligned} \quad (\text{A.5})$$

$$\stackrel{(33)}{=} \frac{1}{dt} \frac{1}{|g(\bar{x}_0)|} e^{-\alpha dt f'(\bar{x}_0) - \eta_0 g'(\bar{x}_0) \alpha dt - D[g'(\bar{x}_0)]^2 \alpha^2 dt} . \quad (\text{A.6})$$

Inserting this expression in equation (A.3), one finds exactly the same propagator given in equation (41). This provides a justification for the use of the Ito rule (12) in equations (29), (34) and (40), used in the derivation of the propagator presented in section 2.2.3.

Last, we mention that Lau and Lubensky [6] actually follow a slightly different route, which involves a Fourier transformation, but in the end their treatment is equivalent to the one we presented in this section.

A.2. À la Itami–Sasa

In order to calculate the Jacobian $\frac{1}{|\partial_{\eta_0} X_1(x_0, H_0)|}$ arising in (24), one can proceed as follows [10]: we write the first time step $0 \leadsto dt$ of the equation of motion as

$$\begin{aligned} X_1(x_0, \eta_0) &= x_0 + f[\overbrace{\alpha X_1(x_0, \eta_0) + (1 - \alpha)x_0}^{\bar{x}_0}] dt \\ &\quad + g[\alpha X_1(x_0, \eta_0) + (1 - \alpha)x_0] \eta_0 dt . \end{aligned} \quad (\text{A.7})$$

Differentiating with respect to the noise, one obtains

$$\partial_{\eta_0} X_1 = \alpha \partial_{\eta_0} X_1 f'(\bar{x}_0) dt + \alpha \partial_{\eta_0} X_1 g'(\bar{x}_0) \eta_0 dt + g(\bar{x}_0) dt \quad (\text{A.8})$$

that implies

$$\frac{1}{|\partial_{\eta_0} X_1|} = \frac{1}{|g(\bar{x}_0)| dt} (1 - \alpha f'(\bar{x}_0) dt - \alpha g'(\bar{x}_0) \eta_0 dt) . \quad (\text{A.9})$$

Note that so far, no expansion nor approximation has been done: this result is exact. In order to exponentiate the numerator of this expression, one uses (33):

$$\frac{1}{|\partial_{\eta_0} X_1|} = \frac{1}{|g(\bar{x}_0)| dt} \exp[-\alpha f'(\bar{x}_0) dt - \alpha g'(\bar{x}_0) \eta_0 dt - D\alpha^2 g'(\bar{x}_0)^2 dt] .$$

This is the same expression as the one in equation (A.6) obtained following the Lau–Lubensky approach, and the one that we obtained in section 2.2.3.

A.3. A continuous-time derivation of the Jacobian

In the quantum-mechanical context a continuous-time formalism is used and the subtleties linked to the discretisation scheme are usually encoded in the choice of the value of the Heaviside theta function at zero, $\Theta(0) = \alpha$ [29]. In this field, the Jacobian $|\partial_{\eta_0} X_1(x_0, H_0(x_0, x_{dt}))|$ is computed with the help of the identity $\det(1 + C_{\eta_0}) = \exp \text{Tr} \ln(1 + C_{\eta_0})$ where C_{η_0} is the part of the Jacobian that depends on the noise. The expression $\ln(1 + C_{\eta_0})$ is further expanded in powers of C_{η_0} to quadratic order (so as to keep terms that are quadratic in the noise and contribute to the trace involving a time integral when the noise is delta correlated) [33]. The explicit calculation of the Jacobian along these lines was explained in appendix D in [8] and constitutes another way of arriving at the expression in equation (A.6). It is less useful for our purposes in this article since it works in continuous time and does not allow to make immediate contact with the (generalised) substitution rules in discrete time.

Appendix B. Justifying the generalised substitution rules

B.1. The usual $\Delta x^2 = 2Dg(x)^2 dt$ substitution

Stochastic calculus tells us that, when expanding infinitesimals, for a standard Brownian motion B_t , one has:

$$dB_t^2 = dt. \quad (\text{B.1})$$

For our time-discrete noise, $\eta_t^2 = 2D/dt$. For a more complex variable such as x , the solution of the Langevin equation (6), the substitution rule (9) implies

$$\Delta x^2 = 2Dg(x)^2 dt \quad (+O(dt^{3/2}) \text{ as } dt \rightarrow 0) \quad (\text{B.2})$$

where on the r.h.s., the argument x of $g(x)$ can be taken at any discretisation point, at minimal order in dt . As discussed in section 2.2.1, there is no direct argument on the distribution of Δx which allows one to use (B.2) point-wise. The meaning of this relation is to be found in an integral way. Following Øksendal [4], one uses the following ingredients:

- Two functions A_1 and A_2 of the process x are equivalent if the L^2 norm of the temporal integral of their difference is zero:

$$\begin{aligned} A_1[d_t x(t), x(t)] &= A_2[d_t x(t), x(t)] \\ \Leftrightarrow \left\langle \left(\int_0^{t_f} dt \{A_1[d_t x(t), x(t)] - A_2[d_t x(t), x(t)]\} \right)^2 \right\rangle &= 0 \end{aligned} \quad (\text{B.3})$$

$$\Leftrightarrow \left\langle \left(\sum_t dt \{A_1[\frac{\Delta x}{dt}, x_t] - A_2[\frac{\Delta x}{dt}, x_t]\} \right)^2 \right\rangle \xrightarrow{dt \rightarrow 0} 0. \quad (\text{B.4})$$

- Two Brownian increments $B_{t+dt} - B_t$ and $B_{t'+dt} - B_{t'}$ at different times $t \neq t'$ are independent:

$$\langle (B_{t+dt} - B_t)(B_{t'+dt} - B_{t'}) \rangle = \langle (B_{t+dt} - B_t) \rangle \langle (B_{t'+dt} - B_{t'}) \rangle \quad \text{if } t \neq t'. \quad (\text{B.5})$$

- The following averages are computed (e.g. à la Wick) using the Gaussian nature of B_t :

$$\langle (B_{t+dt} - B_t)^2 \rangle = dt, \quad (\text{B.6})$$

$$\langle (B_{t+dt} - B_t)^4 \rangle = 3 dt^2. \quad (\text{B.7})$$

Let us thus show that in the sense of (B.3) and (B.4), one has $\Delta x^2 = 2Dg(x)^2 dt$. For this, one computes

$$\begin{aligned} & \left\langle \left(\sum_t \{ \Delta x^2 - 2Dg(x_t)^2 dt \} \right)^2 \right\rangle \\ & \stackrel{(\text{B.6})}{=} \left\langle \left(\sum_t \{ (df(x_t) + g(x_t)\eta_t dt)^2 - 2Dg(x_t)^2 dt \} \right)^2 \right\rangle \\ & = \left\langle \left(\sum_t \{ (g(x_t)\eta_t dt)^2 - 2Dg(x_t)^2 dt \} \right)^2 \right\rangle + O(dt) \end{aligned} \quad (\text{B.8})$$

$$= \left\langle \left(\sum_t [(B_{t+dt} - B_t)^2 - dt] 2Dg(x_t)^2 \right)^2 \right\rangle + O(dt) \quad (\text{B.9})$$

$$\begin{aligned} & \stackrel{(\text{B.5})}{=} \sum_t \left\langle \left([(B_{t+dt} - B_t)^2 - dt] 2Dg(x_t)^2 \right)^2 \right\rangle + O(dt) \\ & \quad + \sum_{t \neq t'} \left\langle [(B_{t+dt} - B_t)^2 - dt] 2Dg(x_t)^2 \right\rangle \\ & \quad \times \left\langle [(B_{t'+dt} - B_{t'})^2 - dt] 2Dg(x_{t'})^2 \right\rangle \end{aligned} \quad (\text{B.10})$$

$$\stackrel{(\text{B.6})}{=} \sum_t \underbrace{\left\langle [(B_{t+dt} - B_t)^2 - dt]^2 \right\rangle}_{\stackrel{(\text{B.6}) \text{ and } (\text{B.7})}{=} 3dt^2 - 2dt^2 + dt^2} \left\langle (2Dg(x_t)^2)^2 \right\rangle + O(dt) \quad (\text{B.11})$$

$$= dt \sum_t 2dt \left\langle (2Dg(x_t)^2)^2 \right\rangle + O(dt) \quad (\text{B.12})$$

$$= O(dt) \quad (\text{B.13})$$

which goes to zero as $dt \rightarrow 0$, hence finishing the proof of (B.2).

Note that when going from (B.10) to (B.11), one cancels the sum over different time indices $t \neq t'$ using that x_t is independent of $B_{t+dt} - B_t$:

$$\begin{aligned} \left\langle [(B_{t+dt} - B_t)^2 - dt] 2Dg(x_t)^2 \right\rangle &= \left\langle (B_{t+dt} - B_t)^2 - dt \right\rangle \left\langle 2Dg(x_t)^2 \right\rangle \\ &\stackrel{(\text{B.6})}{=} 0. \end{aligned} \quad (\text{B.14})$$

In particular, the factor 2 in $\Delta x^2 = 2Dg(x)^2 dt$ is essential, because it allows one to factorise by $2Dg(x)^2$ between (B.8) and (B.9), and to obtain *in fine* the cancellation in (B.14) which makes that (B.13) is of order dt .

B.2. The generalised substitution rule $\Delta x^4 dt^{-1} = 3 (2Dg(x)^2)^2 dt$

One follows the same path, using $\langle (B_{t+dt} - B_t)^8 \rangle = 105 dt^4$, one computes

$$\begin{aligned} & \left\langle \left(\sum_t \left\{ \frac{\Delta x^4}{dt} - 3 (2Dg(x_t)^2)^2 dt \right\} \right)^2 \right\rangle \\ & \stackrel{(6)}{=} \left\langle \left(\sum_t \left\{ \frac{(df(x_t) + g(x_t)\eta_t dt)^4}{dt} - 3 (2Dg(x_t)^2)^2 dt \right\} \right)^2 \right\rangle \\ & = \left\langle \left(\sum_t \left\{ \frac{(g(x_t)\eta_t dt)^4}{dt} - 3 (2Dg(x_t)^2)^2 dt \right\} \right)^2 \right\rangle + O(dt) \end{aligned} \quad (B.15)$$

$$= \left\langle \left(\sum_t [(B_{t+dt} - B_t)^4 - 3dt^2] (2Dg(x_t)^2)^2 dt^{-1} \right)^2 \right\rangle + O(dt) \quad (B.16)$$

$$\begin{aligned} & = \dots \text{ as in (B.10), using (B.5) and the average (B.7)} \\ & = \sum_t \underbrace{\left\langle [(B_{t+dt} - B_t)^4 - 3dt^2]^2 \right\rangle}_{=105 dt^4 - 2 \times 3 \times 3 dt^4 + 9 dt^4} \left\langle (2Dg(x_t)^2)^4 dt^{-2} \right\rangle + O(dt) \end{aligned} \quad (B.17)$$

$$\begin{aligned} & = dt \sum_t 96 dt \left\langle (2Dg(x_t)^2)^4 \right\rangle + O(dt) \\ & = O(dt) \end{aligned} \quad (B.18)$$

which goes to zero as $dt \rightarrow 0$, hence finishing the proof of (56). The derivations of (55) and (57) follow in the same way.

Note that in passing from (B.16) to (B.17) we have used (i) the same independence as in (B.14) and (ii) the fact that in the double sum term

$$\begin{aligned} & \sum_{t \neq t'} \left\langle [(B_{t+dt} - B_t)^4 - 3dt^2] (2Dg(x_t)^2)^2 dt^{-1} \right\rangle \\ & \quad \times \left\langle [(B_{t'+dt} - B_{t'})^4 - 3dt^2] (2Dg(x_{t'})^2)^2 dt^{-1} \right\rangle, \end{aligned} \quad (B.19)$$

which is similar to (B.10), one again has the important cancellation

$$\begin{aligned} & \sum_t \left\langle [(B_{t+dt} - B_t)^4 - 3dt^2] (2Dg(x_t)^2)^2 dt^{-1} \right\rangle \\ & = \left\langle \sum_t [(B_{t+dt} - B_t)^4 - 3dt^2] \right\rangle \left\langle (2Dg(x_t)^2)^2 dt^{-1} \right\rangle \end{aligned} \quad (B.20)$$

$$\stackrel{(B.7)}{=} 0. \quad (B.21)$$

In particular, the factor 3 in the substitution rule $\Delta x^4 dt^{-1} = 3 (2Dg(x)^2)^2 dt$ one wants to show is essential, because it allows one to factorise by $2Dg(x)^2$ between (B.15) and (B.16), and to obtain *in fine* the cancellation in (B.21) which makes that (B.18) is of order dt . The factor 3 in $\Delta x^4 dt^{-1} = 3 (2Dg(x)^2)^2 dt$ is thus exactly the same as the one, obtained e.g. à la Wick in (B.7).

Appendix C. An inconsistency arising when applying the standard chain rule inside the dynamical action

We detail in this appendix how an invalid use of the standard stochastic chain rule (7) can lead to an inconsistency when changing variables in the dynamical action corresponding to the Langevin equation (6). This appendix is the translation to the Onsager–Machlup action of the appendix E of [7] (version v1 of the arXiv preprint) where the same inconsistency was observed in the Martin–Siggia–Rose–Janssen–DeDominicis formulation of the dynamical action.

We compare the direct path (downwards, on the left) of the commutative diagram represented on figure 2, and the indirect path where one first (top arrow) changes variables from $x(t)$ to $U(t) = u(x(t))$ in the Langevin equation, then (right arrow downwards) constructs the action, and finally (down arrow leftwards) tries to come back to the Onsager–Machlup action by applying the standard stochastic chain rule (7). On the way, one should not forget to handle correctly the change of variables in the normalisation prefactor of the action.

The direct path leads to the expression (42) of the dynamical action, together with its associated normalisation prefactor (43). The indirect path starts by obtaining the Langevin equation (60) on $U(t) = u(x(t))$ and continues by writing the corresponding the Onsager–Machlup weight (77). The last step consists in attempting to come back to the Onsager–Machlup weight for the process $x(t)$ by a change of variables in the action and in the Jacobian.

C.1. The normalisation prefactor

One can focus on the first time step $0 \leadsto dt$ without loss of generality. The change of variables in the normalisation prefactor involves two stages: (i) taking into account the factor $u'(x_{dt})$ of (67) that comes from the change of measure and (ii) actually passing from the variable \bar{U}_0 (given by equation (64)) to the variable \bar{x}_0 in the prefactor $\mathcal{N}/|G(\bar{U}_0)|$ of the Onsager–Machlup weight (77). Denoting by $J_U(\bar{x}_0)$ the elementary normalisation prefactor coming from this procedure, one has

$$J_U(\bar{x}_0) = \frac{\mathcal{N}}{|G(\bar{U}_0)|} u'(x_{dt}) \quad (\text{C.1})$$

$$= \frac{\mathcal{N}}{|g(\bar{x}_0)|} \left(u'(x_{dt}) \frac{|g(\bar{x}_0)|}{|G(\bar{U}_0)|} \right), \quad (\text{C.2})$$

where on the second line we have put in prefactor the expected contribution $\mathcal{N}/|g(\bar{x}_0)|$ of the first time step in the total path-integral normalisation prefactor (43) on $x(t)$. The other factor $u'(x_{dt}) |g(\bar{x}_0)|/|G(\bar{U}_0)|$ gives a contribution that has to be incorporated into the exponential part of the infinitesimal propagator (i.e. into the action of the path integral in the $dt \rightarrow 0$ limit). To do so, one expresses \bar{U}_0 and x_{dt} in terms of \bar{x}_0 and Δx by means of equations (64)–(66) and one expands the result up to order dt , keeping in mind that $\Delta x = O(dt^{1/2})$. One obtains from (C.2) that

$$J_U(\bar{x}_0) = \frac{\mathcal{N}}{|g(\bar{x}_0)|} \left[1 - \frac{(-1 + \alpha)u''(\bar{x}_0)}{u'(\bar{x}_0)} \Delta x + (-1 + \alpha) \left(\frac{\alpha g'(\bar{x}_0)u''(\bar{x}_0)}{2g(\bar{x}_0)u'(\bar{x}_0)} + \frac{\alpha u''(\bar{x}_0)^2}{2u'(\bar{x}_0)^2} + \frac{(-1 + \alpha)u^{(3)}(\bar{x}_0)}{2u'(\bar{x}_0)} \right) \Delta x^2 \right]. \quad (\text{C.3})$$

Then using the substitution rule (9) for Δx^2 and reexponentiating the result through (33) one gets

$$J_U(\bar{x}_0) = \frac{\mathcal{N}}{|g(\bar{x}_0)|} \exp \left\{ - \frac{(-1 + \alpha)u''(\bar{x}_0)}{u'(\bar{x}_0)} \Delta x + \left[\frac{D(-1 + \alpha)g(\bar{x}_0)^2 u''(\bar{x}_0)^2}{u'(\bar{x}_0)^2} + \frac{D(-1 + \alpha)g(\bar{x}_0)[\alpha g'(\bar{x}_0)u''(\bar{x}_0) + (-1 + \alpha)g(\bar{x}_0)u^{(3)}(\bar{x}_0)]}{u'(\bar{x}_0)} \right] \Delta t \right\}. \quad (\text{C.4})$$

Taking finally the continuous-time limit, we see that the change of variables from $U(t)$ to $x(t)$ brings a contribution $\Delta S_U[x(t)]$ to the action equal to

$$\Delta S_U[x(t)] \stackrel{\alpha}{=} \int_0^{t_f} dt \left\{ \frac{(-1 + \alpha)u''(x)}{u'(x)} dx - \frac{D(-1 + \alpha)g(x)^2 u''(x)^2}{u'(x)^2} - \frac{D(-1 + \alpha)g(x)[\alpha g'(x)u''(x) + (-1 + \alpha)g(x)u^{(3)}(x)]}{u'(x)} \right\}. \quad (\text{C.5})$$

We note that it vanishes for a linear transformation such that $u'' = 0$, or for $\alpha = 1$ and any function u . This last case is understood from (C.2), where for $\alpha = 1$ one has $\bar{x}_0 = x_{dt}$ and $\bar{U}_0 = U_{dt} = u(x_{dt})$ which implies, using (60) for the expression of G , that $G(\bar{U}_0) = u'(x_{dt})g(u(x_{dt}))$ from which the factor in parenthesis in (C.2) is equal to 1.

C.2. The change of variables in the action

We can now combine the contribution $\Delta S_U[x(t)]$ obtained in the previous section and the change of variables from $U(t) = u(x(t))$ to $x(t)$ in the action $S[U(t)]$ of the process $U(t)$. The expression of $S[U(t)]$ is read from (77). The correct procedure to follow is discussed in section 3.2 following a discrete-time approach. One can also apply the continuous-time modified chain-rule discussed in section 3.3.2. Both approaches yield back the correct action (42) for the process $x(t)$.

If one improperly applies the chain rule (7) to determine $d_t U = d_t [u(x(t))]$ in (77), one finds a result for the action $S_U[x(t)]$, in which there are supplementary terms compared to the correct result $S[x(t)]$ given by (42); that is

$$S_U[x(t)] \stackrel{\alpha}{=} \Delta S_U[x(t)] + S[U(t)] \Big|_{U(t)=u(x(t))} \quad (\text{C.6})$$

$$\stackrel{\alpha}{=} S[x(t)] + \int_0^{t_f} dt \left\{ \frac{(-1 + 2\alpha)u''(x)}{u'(x)} dx - \frac{3D(-1 + \alpha)\alpha g(x)g'(x)u''(x)}{u'(x)} + \frac{D(1 + (-1 + \alpha)\alpha)g(x)^2 u''(x)^2}{u'(x)^2} + \frac{D(-1 - 3(-1 + \alpha)\alpha)g(x)^2 u^{(3)}(x)}{u'(x)} \right\}. \quad (\text{C.7})$$

The terms in the time integral should be absent if the procedure had been correct (At the very least, the result should not depend on the function u). These terms are equivalent to the terms in equation (E.18) of the appendix E in [7]. Their presence is due to the fact

that, when using the chain rule (7) as we did, one discards terms proportional to $\Delta x^3 dt^{-1}$, $\Delta x^4 dt^{-1}$ and $\Delta x^6 dt^{-2}$ that are present in equation (50) when following the correct procedure. The supplementary terms vanish when a linear change of variables is applied, i.e. when $u'' = 0$.

Special cases—One notes that these supplementary terms still remain present in the three following simplified cases:

(i) Stratonovich discretisation ($\alpha = 1/2$):

$$S_U[x(t)] \stackrel{S}{=} S[x(t)] + \int_0^{t_f} dt \left\{ \frac{3Dg(x)g'(x)u''(x)}{4u'(x)} + \frac{3Dg(x)^2u''(x)^2}{4u'(x)^2} - \frac{Dg(x)^2u^{(3)}(x)}{4u'(x)} \right\} \quad (C.8)$$

(ii) Additive noise ($g(x) = g$ is a constant):

$$S_U[x(t)] \stackrel{\alpha}{=} S[x(t)] + \int_0^{t_f} dt \left\{ \frac{(-1 + 2\alpha)u''(x)}{u'(x)} d_t x + \frac{Dg^2(1 + (-1 + \alpha)\alpha)u''(x)^2}{u'(x)^2} + \frac{Dg^2(-1 - 3(-1 + \alpha)\alpha)u^{(3)}(x)}{u'(x)} \right\} \quad (C.9)$$

(iii) Additive noise and Stratonovich discretisation:

$$S_U[x(t)] \stackrel{S}{=} S[x(t)] + \int_0^{t_f} dt Dg^2 \left\{ \frac{3u''(x)^2}{4u'(x)^2} - \frac{u^{(3)}(x)}{4u'(x)} \right\} \quad (C.10)$$

This last case is surprising because, as often described, the additive-noise Stratonovich-discretised Langevin equation is the better behaved in terms of the rules of differential calculus. But in spite of this fact, as we have shown, the standard chain rule of differential calculus *cannot* be used inside the corresponding Onsager–Machlup action (although this rule is valid at the Langevin equation level).

Appendix D. An inconsistency arising when applying the Langevin rule for changing discretisation inside the dynamical action

In this appendix, we study how the α -discretised Langevin equation (6) can be described by a path-integral probability written in a different $\bar{\alpha}$ -discretisation. The direct procedure to follow is to change the discretization in the Langevin equation first [this yields (16) with a modified force $f_{\alpha \rightarrow \bar{\alpha}}(x)$ given by (17)], and to write the corresponding trajectory weight. One reads its action from (42) as

$$S_{\alpha\bar{\alpha}}[x(t)] \stackrel{\bar{\alpha}}{=} \int_0^{t_f} dt \left\{ \frac{1}{2} \frac{1}{2D} \left[\frac{d_t x - f_{\alpha \rightarrow \bar{\alpha}}(x) + 2\bar{\alpha}Dg(x)g'(x)}{g(x)} \right]^2 + \bar{\alpha}f'_{\alpha \rightarrow \bar{\alpha}}(x) \right\}, \quad (D.1)$$

where the arguments of the functions f and g are taken in $\bar{\alpha}$ -discretisation. The associated normalisation prefactor reads, from (43)

$$\mathcal{J}[x(t)] \stackrel{\bar{\alpha}}{=} \prod_t \left\{ \sqrt{\frac{dt^{-1}}{4\pi D}} \frac{1}{|g(\bar{x}_t)|} \right\}. \quad (D.2)$$

The correct way of performing the change in discretisation in the original α -discretized action (42) (together with the prefactor (43)) was described in section 3.1.2, going through the infinitesimal propagator in discrete time and using the generalised substitution rules (54)–(57). The computation is done in the special case $\bar{\alpha} = 1/2$ but also works for any value of $\bar{\alpha}$.

The naive procedure discussed in section 3.1.2.4 consists in applying the two following steps.

- For the action: start from the (α -discretized) action (42) and assume that one can replace the force f by the effective force $f_{\alpha \rightarrow \bar{\alpha}}$ appearing in the Langevin equation (16). This would yield an exponential contribution to the action of the form

$$\int_0^{t_f} dt \left\{ \frac{1}{2} \frac{1}{2D} \left[\frac{d_t x - f_{\alpha \rightarrow \bar{\alpha}}(x) + 2\alpha D g(x) g'(x)}{g(x)} \right]^2 + \bar{\alpha} f'_{\alpha \rightarrow \bar{\alpha}}(x) \right\}. \quad (\text{D.3})$$

- For the normalisation prefactor: change the discretisation from α to $\bar{\alpha}$ by going to discrete time (and considering the first time step) and writing, with explicit discretisation points, that

$$\frac{1}{|g(\bar{x}_0^{(\alpha)})|} = \frac{1}{|g(\bar{x}_0^{(\bar{\alpha})})|} \frac{|g(\bar{x}_0^{(\bar{\alpha})})|}{|g(\bar{x}_0^{(\alpha)})|}. \quad (\text{D.4})$$

Then, using $\bar{x}_0^{(\alpha)} = \bar{x}_0^{(\bar{\alpha})} + (\alpha - \bar{\alpha})\Delta x$ to expand the second fraction in powers of $\Delta x = x_{dt} - x_0$ up to order Δx^2 , using the substitution rule and reexponentiating the result gives

$$\frac{1}{|g(\bar{x}_0^{(\alpha)})|} = \frac{1}{|g(\bar{x}_0^{(\bar{\alpha})})|} e^{\frac{(-\alpha + \bar{\alpha})g'(\bar{x}_0^{(\bar{\alpha})})}{g(\bar{x}_0^{(\bar{\alpha})})} \Delta x + D(\alpha - \bar{\alpha})^2 [g'(\bar{x}_0^{(\bar{\alpha})})^2 - g(\bar{x}_0^{(\bar{\alpha})})g''(\bar{x}_0^{(\bar{\alpha})})] \Delta x^2}. \quad (\text{D.5})$$

Taking the continuous-time limit, this gives the following contribution to the action (with x being $\bar{\alpha}$ -discretised):

$$\int_0^{t_f} dt \left\{ (\alpha - \bar{\alpha}) \frac{g'(x)}{g(x)} d_t x - D(\alpha - \bar{\alpha})^2 [g'(x)^2 - g(x)g''(x)] \right\}. \quad (\text{D.6})$$

Finally, adding (D.3) and (D.6) yields a candidate $\tilde{S}_{\alpha\bar{\alpha}}[x(t)]$ for the $\bar{\alpha}$ -discretised action of the α -discretised Langevin equation. If this procedure had been correct, one would have had recovered the same action $S_{\alpha\bar{\alpha}}[x(t)]$ as in (D.1), but, by direct inspection, one finds that

$$\begin{aligned} \tilde{S}_{\alpha\bar{\alpha}}[x(t)] = & S_{\alpha\bar{\alpha}}[x(t)] \\ & + \int_0^{t_f} dt \left\{ \frac{(\alpha - \bar{\alpha}) [(-f(x) + 2d_t x)g'(x) + g(x)(f'(x) + 2D\bar{\alpha}g'(x)^2)]}{g(x)} \right. \\ & \left. + 3Dg(x)(\alpha - \bar{\alpha})^2 g''(x) \right\}. \end{aligned} \quad (\text{D.7})$$

References

- [1] Stratonovich R L 1992 *Nonlinear Nonequilibrium Thermodynamics I* (Berlin: Springer)
- [2] Gardiner C W 1994 *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences (Springer Series in Synergetics vol 13)* 2nd edn (Berlin: Springer)
- [3] van Kampen N G 2007 *Stochastic Processes in Physics and Chemistry (North-Holland Personal Library)* 3rd edn (Amsterdam: Elsevier)
- [4] Øksendal B *Stochastic Differential Equations: an Introduction with Applications* (Universitext) 6th edn (Berlin: Springer) (corrected printing edition 2013 OCLC: 935584333)
- [5] van Kampen N G 1981 Itô versus stratonovich *J. Stat. Phys.* **24** 175–87
- [6] Lau A W C and Lubensky T C 2007 State-dependent diffusion: thermodynamic consistency and its path integral formulation *Phys. Rev. E* **76** 011123
- [7] Aron C, Barci D G, Cugliandolo L F, González Arenas Z and Lozano G S 2014 Dynamical symmetries of Markov processes with multiplicative white noise (arXiv:1412.7564v1 [cond-mat])
- [8] Aron C, Barci D G, Cugliandolo L F, González Arenas Z and Lozano G S 2016 Dynamical symmetries of Markov processes with multiplicative white noise *J. Stat. Mech.* **053207**
- [9] Aron C, Barci D G, Cugliandolo L F, González-Arenas Z and Lozano G S 2014 Magnetization dynamics: path-integral formalism for the stochastic landau-lifshitz-gilbert equation *J. Stat. Mech.* **P09008**
- [10] Itami M and Sasa S-I 2017 Universal form of stochastic evolution for slow variables in equilibrium systems *J. Stat. Phys.* **167** 46–63
- [11] Arenas Z G and Barci D G 2010 Functional integral approach for multiplicative stochastic processes *Phys. Rev. E* **81** 051113
- [12] Moreno M V, Arenas Z G and Barci D G 2015 Langevin dynamics for vector variables driven by multiplicative white noise: a functional formalism *Phys. Rev. E* **91** 042103
- [13] Barci D G, Arenas Z G and Moreno M V 2016 Path integral approach to nonequilibrium potentials in multiplicative Langevin dynamics *Europhys. Lett.* **113** 10009
- [14] Langouche F, Roekaerts D and Tirapegui E 1982 *Functional Integration and Semiclassical Expansions* (Dordrecht: Kluwer)
- [15] Gervais J L and Jevicki A 1976 Point canonical transformations in the path integral *Nucl. Phys. B* **110** 93–112
- [16] Salomonson P 1977 When does a non-linear point transformation generate an extra $o(\hbar^2)$ potential in the effective Lagrangian? *Nucl. Phys. B* **121** 433–44
- [17] Langouche F, Roekaerts D and Tirapegui E 1979 Functional integrals and the Fokker-Planck equation *Nuovo Cimento B* **53** 135–59
- [18] Alfaro J and Damgaard P H 1990 Field transformations, collective coordinates and BRST invariance *Ann. Phys.* **202** 398–435
- [19] Apfeldorf K M and Ord C 1996 Coordinate redefinition invariance and extra terms *Nucl. Phys. B* **479** 515–26
- [20] Onsager L and Machlup S 1953 Fluctuations and irreversible processes *Phys. Rev.* **91** 1505–12
- [21] Machlup S and Onsager L 1953 Fluctuations and irreversible process II. Systems with Kinetic energy *Phys. Rev.* **91** 1512–5
- [22] Martin P C, Siggia E D and Rose H A 1973 Statistical dynamics of classical systems *Phys. Rev. A* **8** 423–37
- [23] De Dominicis C 1976 Techniques de renormalisation de la théorie des champs et dynamique des phénomènes critiques *J. Phys. Colloques* **37** 247
- [24] Janssen H-K 1976 On a Lagrangean for classical field dynamics and renormalization group calculations of dynamical critical properties *Z. Phys. B* **23** 377–80
- [25] Bausch R, Janssen H K and Wagner H 1976 Renormalized field theory of critical dynamics *Z. Phys. B* **24** 113–27
- [26] Janssen H-K 1979 Field theoretical methods applied to critical dynamics *Lecture Notes in Physics: Dynamical Critical Phenomena and Related Topics* vol 104, ed C P Enz (Berlin: Springer) p 26
- [27] De Dominicis C and Peliti L 1978 Field-theory renormalization and critical dynamics above T_c : Helium, antiferromagnets, and liquid-gas systems *Phys. Rev. B* **18** 353–76
- [28] Feynman R P 1948 Space-time approach to non-relativistic quantum mechanics *Rev. Mod. Phys.* **20** 367

- [29] Zinn-Justin J 2002 *Quantum Field Theory and Critical Phenomena (Int. Series of Monographs on Physics Clarendon Press vol 113)* 4th edn (Oxford: Oxford University Press)
- [30] Tang Y, Yuan R and Ao P 2014 Summing over trajectories of stochastic dynamics with multiplicative noise *J. Chem. Phys.* **141** 044125
- [31] Langouche F, Roekaerts D and Tirapegui E 1981 General langevin equations and functional integration *Field Theory, Quantization and Statistical Physics: in Memory of Bernard Jouvét* (Dordrecht: Springer)
- [32] Aron C, Biroli G and Cugliandolo L F 2010 Symmetries of generating functionals of langevin processes with colored multiplicative noise *J. Stat. Mech.* **P11018**
- [33] Arnold P 2000 Symmetric path integrals for stochastic equations with multiplicative noise *Phys. Rev. E* **61** 6099–102