An even more mathematically offensive introduction to SDEs: a bag of tricks and analytical solutions

Alexander van Meegen

A guided tour through the marvelous world of SDEs. Although SDEs might bite if you touch them in the wrong spots, they are usually quite nice and fun to play with. Since the dangerous parts are frequently and adequately addressed, the focus of this text is on the fun part—solving problems. Ideally, the tour sparks your interest and causes some further expeditions on your own.

CONTENTS

I. INTRO

We consider *stochastic differential equations* (SDEs), i.e., differential equations

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

where $\xi(t)$ is a *stochastic process* or, more colloquially, *noise*. In most cases, $\xi(t)$ is a zero-mean $\langle \xi(t) \rangle = 0$, white $\langle \xi(t)\xi(t') \rangle = 2D\delta(t - t')$, Gaussian process—a *Gaussian white noise* (GWN)—but we will go beyond this case.

The SDE is called *additive* if $g(x)$ is constant and *multiplicative* otherwise. Additive processes are simpler because the strength of the noise does not depend on the state of the system.

Since $\xi(t)$ is a stochastic process, the goal is to determine the statistics of the process $x(t)$. More concretely, the *mean* and *autocorrelation function*

$$\mu_x(t) = \langle x(t) \rangle \qquad C_x(t, t') = \langle (x(t) - \mu_x(t))(x(t') - \mu_x(t')) \rangle \tag{2}$$

are usually the most relevant quantities. In stationary (time translation invariant) cases where $\mu_x(t)$ is constant and $C_x(t, t') = C_x(t' - t)$, the *power spectrum* $S_x(f) = \lim_{T\to\infty} \frac{\langle \tilde{x}(f)\tilde{x}(f)^* \rangle}{T}$ with $\tilde{x}(f) = \int_0^T e^{2\pi i f t} x(t) dt$ is another central quantity of interest. Power spectrum and autocorrelation function are their mutual Fourier transform,

$$S_x(f) = \int_{-\infty}^{\infty} e^{2\pi i f \tau} C_x(\tau) d\tau, \tag{3}$$

which is known as the *Wiener-Khinchin theorem*. Of course one can invert Eq. (3) to get $C_x(\tau)$ from $S_x(f)$.

II. WHITE NOISE

In this entire section, $\xi(t)$ is a GWN, i.e., $\langle \xi(t)\xi(t') \rangle = 2D\delta(t - t')$, which simplifies the problem quite a bit because we can solve the SDE for given initial conditions without taking the history into account—the problem is *Markovian*.

A. Preliminaries

Before diving into the calculations, we have to briefly talk about one spot were SDE do bite.

Fokker-Planck Equation

To fully determine the statistics beyond mean and correlation function, one considers the time-dependent distribution $P_t(x)$ which obeys a partial differential equation, the *Fokker-Planck equation* (FPE)

$$\partial_t P_t(x) = -\partial_x J_t(x), \qquad J_t(x) = A(x)P - \frac{1}{2}\partial_x B(x)P, \tag{4}$$

with $B(x) > 0$ [1, chapter IX.4]. Eq. (4) has the form of a continuity equation with *probability flux* $J_t(x)$ where $A(x)$ determines the drift and $B(x)$ the diffusion.

Considering the FPE is in principle advantageous because it does not leave room for ambiguities—after all, $P_t(x)$ fully determines the statistics. This forces us to talk about the *Ito-Stratonovich dilemma*: In Ito interpretation, i.e., discretizing Eq. (1) as

$$x(t+\Delta t) - x(t) = f(x(t))\Delta t + g(x(t)) \int_t^{t+\Delta t} \xi(s) ds,$$

the flux of the corresponding FPE is [1, chapter IX.4]

$$J_t(x) = f(x)P - D\partial_x g(x)^2 P; \tag{5}$$

in Stratonovich interpretation, i.e., discretizing Eq. (1) as

$$x(t+\Delta t) - x(t) = f(x(t))\Delta t + g(\tfrac{1}{2}x(t) + \tfrac{1}{2}x(t+\Delta t)) \int_t^{t+\Delta t} \xi(s) ds,$$

the flux of the corresponding FPE is [1, chapter IX.4]

$$J_t(x) = [f(x) + Dg(x)g'(x)]P - D\partial_x g(x)^2 P. \tag{6}$$

Thus, the two interpretations differ by the *spurious drift* term $Dg(x)g'(x)$ in Eq. (6). This means we can switch between the interpretations at the expense of changing the drift term or, equivalently, $f(x)$. Looking at the spurious drift, we see immediately that it vanishes for additive processes where $g(x)$ is constant—there is no Ito-Stratonovich dilemma for additive processes. For a more extensive discussion see [1, chapter IX.5].

In the following, we avoid the FPE as much as possible since partial differential equations are nastier beasts than ordinary differential equations. The canonical reference for everything related to the FPE is [2].

Changing Variables

The Ito-Stratonovich dilemma arises for multiplicative processes. There seems to be a neat way around it if $g(x) > 0$: Dividing $\dot{x} = f(x) + g(x)\xi(t)$ by $g(x)$ and using the transformation $dy = \frac{dx}{g(x)}$, we get the additive, unambiguous SDE

$$\dot{y} = \frac{f(x(y))}{g(x(y))} + \xi(t). \tag{7}$$

Where did the dilemma go? It turns out that if one transforms variables like this, one implicitly opted for the Stratonovich interpretation [1, chapter IX.4]. If we interpret the SDE in Ito convention, we first have to subtract the spurious drift to obtain the equivalent SDE in Stratonovich interpretation, $\dot{x} = f(x) - Dg(x)g'(x) + g(x)\xi(t)$. Now we can divide by $g(x)$ and change variables to arrive at

$$\dot{y} = \frac{f(x(y))}{g(x(y))} - Dg'(x(y)) + \xi(t). \tag{8}$$

Thus, one gets a counter-intuitive additional term when transforming variables in Ito convention. The upshot is that in both interpretations one can transform a multiplicative SDE into an additive one and hence it is sufficient to consider additive SDEs.

Let's look at a simple example: geometric Brownian motion, $\dot{x} = (\mu + \xi(t))x$. We want to transform $y = \ln x$ since $dy = dx/x$. In Ito interpretation, we get $\dot{y} = \mu - D + \xi(t)$; in Stratonovich interpretation, we get $\dot{y} = \mu + \xi(t)$.

Just for completeness, for an arbitrary transformation $dy = \phi(x)dx$ or $y = \Phi(x)$, the transformed SDE in Stratonovich convention behaves as expected [1, chapter IX.4],

$$f(x) \to \phi(x)f(x), \qquad g(x) \to \phi(x)g(x), \tag{9}$$

whereas one gets an additional term in Ito convention [1, chapter IX.4],

$$f(x) \to \phi(x)f(x) + Dg(x)^2\phi'(x), \qquad g(x) \to \phi(x)g(x). \tag{10}$$

In both cases, one evaluates $x = \Phi^{-1}(y)$ after the transformation.

B. Ornstein-Uhlenbeck Process

The Ornstein-Uhlenbeck process (OU process) is the most simple SDE corresponding to the choice $f(x) = -\gamma x$ and $g(x) = 1$,

$$\dot{x} = -\gamma x + \xi(t), \tag{11}$$

where $\xi(t)$ is a zero-mean GWN with $\langle \xi(t)\xi(t')\rangle = 2D\delta(t - t')$. Note that we don't need to consider a constant term μ on the r.h.s. because we can transform $\gamma x - \mu \to \gamma x$. The most intuitive picture is an overdamped motion in a potential $U(x) = \frac{\gamma}{2}x^2$.

Greens Function Approach

Assuming fixed initial conditions x_0 and a fixed realization of $\xi(t)$, we can write down the solution of Eq. (11) since Kindergarten [1, chapter IX.1],

$$x(t) = x_0 e^{-\gamma t} + \int_0^t e^{-\gamma(t-s)}\xi(s)ds. \tag{12}$$

A formal argument for Eq. (12) is that $G(t) = \Theta(t)e^{-\gamma t}$ is the Greens function of the differential operator $\gamma + \partial_t$ and the solution follows as $x(t) = \int G(t-s)[x_0\delta(s) + \xi(s)\Theta(s)]ds$. From Eq. (12), we immediately get

$$\mu_x(t) = x_0 e^{-\gamma t} + \int_0^t e^{-\gamma(t-s)}\langle \xi(s)\rangle ds = x_0 e^{-\gamma t} \tag{13}$$

because $\xi(t)$ is zero mean. With $\Delta x(t) = x(t) - \mu_x(t) = \int_0^t e^{-\gamma(t-s)}\xi(s)ds$, we find

$$C_x(t,t') = \int_0^t \int_0^{t'} e^{-\gamma(t-s)}e^{-\gamma(t'-s')}\langle \xi(s)\xi(s')\rangle ds' ds = 2De^{-\gamma(t+t')}\int_0^{\min(t,t')} e^{2\gamma s}ds$$

where we used $\langle \xi(t)\xi(t')\rangle = 2D\delta(t-t')$ in the second equality. Evaluating the final integral, we get

$$C_x(t,t') = \frac{D}{\gamma}(e^{-\gamma(t+t'-2\min(t,t'))} - e^{-\gamma(t+t')}). \tag{14}$$

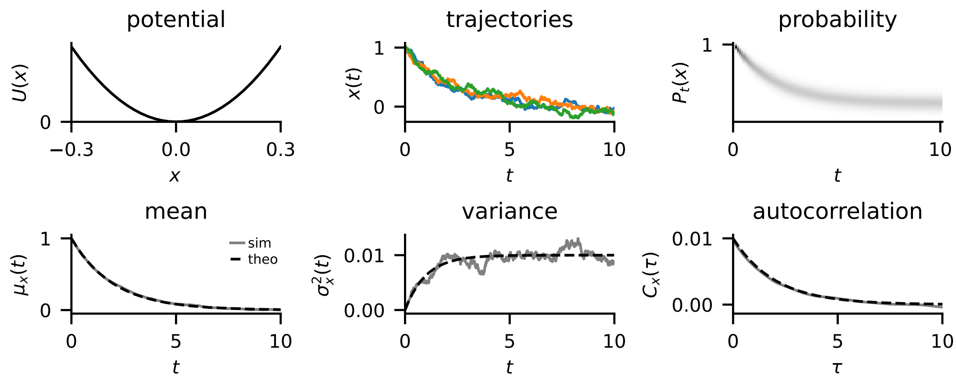

Figure 1. OU process with $\gamma = 0.5$ and $D = 0.005$ and initial condition $x_0 = 1$.

This looks a bit odd, it can be made a lot friendlier if we break the symmetry of the time arguments and set $t' = t + \tau$ with $\tau \geq 0$:

$$C_x(t, t+\tau) = \frac{D}{\gamma} e^{-\gamma\tau}(1 - e^{-2\gamma t}) \overset{t\to\infty}{\to} \frac{D}{\gamma} e^{-\gamma\tau}. \tag{15}$$

We of course also get the time-dependent variance from Eq. (14),

$$\sigma_x^2(t) = \frac{D}{\gamma}(1 - e^{-2\gamma t}). \tag{16}$$

Eq. (13) and Eq. (14) fully solve the problem because Eq. (11) is linear and $\xi(t)$ Gaussian and hence $x(t)$ is Gaussian as well. Hence, we can also write down the time-dependent distribution and its stationary limit

$$P_t(x) = \mathcal{N}(x \,|\, \mu_x(t), \sigma_x^2(t)) \overset{t\to\infty}{\to} \mathcal{N}(x \,|\, 0, D/\gamma) \tag{17}$$

which indeed fulfills the associated FPE.

Conveniently, we also solved the Wiener process $\dot{x} = \xi(t)$ because it follows from $\gamma \to 0$. Taking the limit in Eq. (13) and Eq. (14) leads to $\mu_x(t) = x_0$ and $C_x(t, t') = 2D\min(t, t')$. We see immediately that the Wiener process cannot become stationary; it never forgets the initial conditions and its variance $\sigma_x^2(t) = 2Dt$ diverges.

Fourier Transformation Approach

If we are only interested in the stationary statistics, there is a much faster route: the method of Rice [3]. Taking the Fourier transformation of $\dot{x} = -\gamma x + \xi$, using partial integration for the velocity term $\int_0^T e^{2\pi i f t}\dot{x}(t)dt = [e^{2\pi i f t}x(t)]_0^T - 2\pi i f \int_0^T e^{2\pi i f t}x(t)dt$, and neglecting the boundary term $[e^{2\pi i f t}x(t)]_0^T$ for large T [4], we get

$$(\gamma - 2\pi i f)\tilde{x} = \tilde{\xi}.$$

Multiplying with the complex conjugate and taking the average, we get

$$(\gamma^2 + (2\pi f)^2)\langle \tilde{x}\tilde{x}^* \rangle = \langle \tilde{\xi}\tilde{\xi}^* \rangle.$$

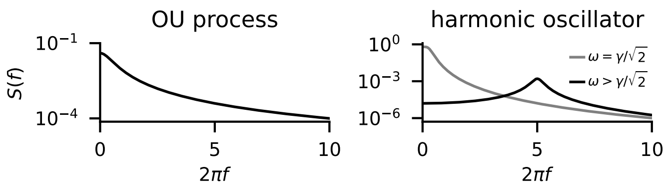

Figure 2. Power spectra of OU process and harmonic oscillator for $\gamma = 0.5$ and $D = 0.005$.

Both averages are proportional to the respective power spectrum; using the Wiener-Khinchin theorem to determine $S_\xi(f) = \int_{-\infty}^{\infty} e^{2\pi i f \tau} 2D\delta(\tau)d\tau = 2D$ we arrive at

$$S_x(f) = \frac{2D}{\gamma^2 + (2\pi f)^2}. \tag{18}$$

Indeed, this is the Fourier transformation of Eq. (15), $C_x(\tau) = \frac{D}{\gamma} e^{-\gamma|\tau|}$.

C. Other Linear Processes

Harmonic Oscillator

Both methods generalize quite well to arbitrary linear processes. First, we consider the noise-driven harmonic oscillator

$$\ddot{x} + \gamma\dot{x} + \omega^2 x = \xi(t). \tag{19}$$

This does not look like Eq. (1) but introducing $\dot{x} = v$ such that $\dot{v} = -\gamma v - \omega^2 x + \xi(t)$ immediately brings us back there, although the system is now two dimensional. We focus on the stationary statistics and use again the Fourier approach. Neglecting once again the boundary terms, we get

$$(\omega^2 - 2\pi i \gamma f - (2\pi f)^2)\tilde{x}(f) = \tilde{\xi}(f).$$

Multiplying with the complex conjugate and taking the average yields [3]

$$S_x(f) = \frac{2D}{(\omega^2 - (2\pi f)^2)^2 + (2\pi\gamma f)^2}. \tag{20}$$

Of course, we could also calculate the nonstationary statistics. Note that to this end, we need to specify x_0 and v_0 to fully determine the initial conditions.

General Linear Process

Let's generalize further an allow an arbitrary Greens function $G(t)$,

$$x(t) = \int G(t-s)\xi(s)ds, \tag{21}$$

and stick to the stationary statistics. Fourier transformation and neglecting boundary terms yields $\tilde{x}(f) = \tilde{G}(f)\tilde{\xi}(f)$ which leads to

$$S_x(f) = 2D|\tilde{G}(f)|^2. \tag{22}$$

In principle, the $G(t)$ does not have to be a causal Greens function but could be an arbitrary filter here.

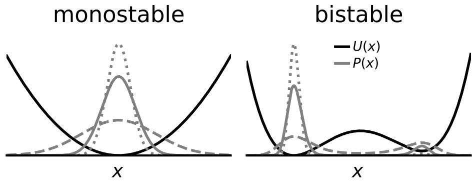

Figure 3. Stationary distribution (gray) for two potentials (black): monostable $U(x) = x^2$ and bistable $U(x) = (x-1)^2(x+1)^2 + 0.1x$. For $D = 0.1$ (solid), $D = 0.5$ (dashed), and $D = 0.05$ (dotted).

D. Nonlinear Processes

Enough linear processes, let's enter the exciting nonlinear world and consider SDEs of the form

$$\dot{x} = -\partial_x U(x) + \xi(t), \tag{23}$$

i.e., overdamped motion in a potential $U(x)$ driven by GWN. We don't consider multiplicative processes because we can transform them to additive ones and we assume that a potential $f(x) = -\partial_x U(x)$ exists because this provides a better intuition and simplifies the formulas.

Compared to the linear version, $x(t)$ is not Gaussian anymore although $\xi(t)$ is. Thus, the higher cumulants don't vanish anymore.

Stationary Distribution

Determining the stationary distribution is surprisingly easy [1, Chapter XIII.1]: In this case, $\partial_t P(x) = 0$ and the FPE reduces to $\partial_x J(x) = 0$. The flux corresponding to Eq. (23) is $J(x) = -U'(x)P - D\partial_x P$; setting $J(x) = 0$ leads to $\partial_x P(x) = -\frac{1}{D}U'(x)P(x)$ and thus

$$P(x) = Ce^{-U(x)/D} \tag{24}$$

with C determined by normalization $\int P(x)dx = 1$. This of course requires $\int e^{-U(x)/D}dx < \infty$. See Fig. 3 for two examples.

A quick consistency check with the Ornstein-Uhlenbeck process: Setting $U(x) = \frac{1}{2}\gamma x^2$, we get $P(x) = C\exp(-\gamma x^2/2D) = \mathcal{N}(x \,|\, 0, D/\gamma)$ in agreement with Eq. (17).

Despite the focus on the technical part, Eq. (24) entails an important physical interpretation that we briefly discuss. The equilibrium distribution in statistical mechanics is the Boltzmann distribution $P(\boldsymbol{p}, \boldsymbol{q}) \propto \exp(-H(\boldsymbol{p}, \boldsymbol{q})/k_B T)$. The momenta and positions decouple, hence we can integrate the momenta out. For a non-interacting system, this also holds for the positions; doing so for all but one leads to $P(q) \propto \exp(-U(q)/k_B T)$ and thus

$$D = k_B T, \tag{25}$$

the *Einstein relation* (with the constant mobility absorbed in the potential).

First-Passage Time

Next, we consider first-passage problems, i.e., the time that a process stays within a certain domain. There is a wide variety of approaches and a lot can be written about it; here we follow the arguably most simple approach, the *flux-over-population method* [5, Section IV.E]. The setup is a process constrained to the domain $(-\infty, x_b]$ and we are interested in the mean first-passage time τ from a point x_r within the domain to x_b (see Fig. 4).

For the flux-over-population method, we first alter the problem and introduce a demon which resets the process to x_r after it crossed the boundary. Now we wait until the situation becomes stationary; due to the resetting, a constant probability flux J between x_r and x_b emerges. According to the flux-over-population method, we can calculate the mean first passage time as

$$\tau = \frac{N_0}{J} \tag{26}$$

where $N_0 = \int_{-\infty}^{x_b} P(x)dx$ is the population. To obtain $P(x)$, we consider the domains $(-\infty, x_r]$ and $[x_r, x_b]$ separately, starting with the latter. For a constant flux J, the stationary FPE yields $-U'(x)P - D\partial_x P = J$. Based on the $J = 0$ solution Eq. (24), we make the ansatz $P_>(x) = C_>(x)\exp(-U(x)/D)$. This leads to $C'_>(x) = -\frac{J}{D}\exp(U(x)/D)$ which we can integrate. Taking the boundary condition $P_>(x_b) = 0$ into account, we get $C_>(x) = \frac{J}{D}\int_x^{x_b}\exp(U(y)/D)dy$ and thus

$$P_>(x) = \frac{J}{D}e^{-U(x)/D}\int_x^{x_b} e^{U(y)/D}dy.$$

In the other domain, the flux vanishes such that we get $P_<(x) = C_<\exp(-U(x)/D)$. The constant $C_<$ is determined by continuity, $P_<(x \to x_r) = P_>(x \to x_r)$, and we arrive at

$$P_<(x) = \frac{J}{D}e^{-U(x)/D}\int_{x_r}^{x_b} e^{U(y)/D}dy.$$

Now, all we have to do is to evaluate $N_0 = \int_{-\infty}^{x_r} P_<(x)dx + \int_{x_r}^{x_b} P_>(x)dx$ and, by virtue of Eq. (26), we get

$$\tau = \frac{1}{D}\int_{-\infty}^{x_r} e^{-U(x)/D}\int_{x_r}^{x_b} e^{U(y)/D}dydx + \frac{1}{D}\int_{x_r}^{x_b} e^{-U(x)/D}\int_{x}^{x_b} e^{U(y)/D}dydx.$$

This clunky expression can be simplified to

$$\tau = \frac{1}{D}\int_{x_r}^{x_b} e^{U(x)/D}\int_{-\infty}^{x} e^{-U(y)/D}dydx \tag{27}$$

by changing the order of integration and renaming $x \leftrightarrow y$. If you don't trust demons: one can derive Eq. (27) from the Kolmogorov backward equation [1, chapter XII.3].

Considering once again the Ornstein-Uhlenbeck process with $U(x) = \frac{1}{2}\gamma x^2$, we can solve the inner Gaussian integral in Eq. (27) and arrive after a substitution at

$$\tau = \frac{\sqrt{\pi}}{\gamma}\int_{\sqrt{\frac{\gamma}{2D}}x_r}^{\sqrt{\frac{\gamma}{2D}}x_b} e^{x^2}\big(1 + \operatorname{erf}(x)\big)dx, \tag{28}$$

the infamous Siegert.

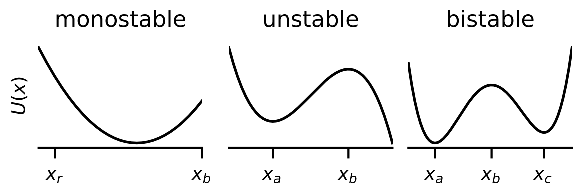

Figure 4. First passage problems for differen systems: monostable $U(x) = x^2$, unstable $U(x) = -(x-1)x(x+1)$, and bistable $U(x) = (x-1)^2(x+1)^2 + 0.1x$.

Unstable and Bistable Systems

Now let's consider specific potentials. Before diving into the math, we take a step back and think about what we would like to achieve. Let's focus on two examples: 1) An unstable system with a metastable state and an abyss. Due to the noise, the system will eventually leave the metastable state and end up in the abyss. 2) A bistable system with two minima. Due to the noise, the system will transition between these minima. In both case, want the the barrier to be high enough to separate the states; the relevant question is how long it takes until the system leaves one state and either falls into the abyss or transits to the other state (see Fig. 4).

First, let's consider the unstable potential with a minimum at $U(x_a)$, a maximum at $U(x_b)$ with $x_b > x_a$, and an abyss to the right, e.g., a cubic potential. Once the maximum is passed, the system almost has no chance to escape the abyss. Thus, the time spent in the metastable state can be approximated by the first passage time from a point before the minimum to a point slightly behind the maximum. With the noise being small, $1/D \gg 1$, we can approximate the integrals in Eq. (27) using the Laplace approximation to get [1, chapter XIII.2]

$$\tau = \frac{2\pi}{\sqrt{U''(x_a)|U''(x_b)|}} \exp\Big([U(x_b) - U(x_a)]/D\Big). \tag{29}$$

The leading exponential term is the *Arrhenius factor*: the first-passage time (or inverse rate) depends exponentially on the height of the potential barrier $\Delta U = U(x_b) - U(x_a)$ relative to the noise strength D.

Next, the bistable case with two minima at $U(x_a)$ and $U(x_c)$ separated by a barrier with maximum at $U(x_b)$. The main obstacle is to climb the maximum; thus, the transition time is approximately equal to the first passage time to from a minimum to slightly behind the maximum. Again using a Laplace approximation for the first passage time from the minimum at x_a, τ_a, and from the minimum at x_c, τ_c, the ratio of the transition times is [1, chapter XIII.1]

$$\frac{\tau_c}{\tau_a} = \sqrt{\frac{U''(x_a)}{U''(x_c)}} \exp\Big([U(x_a) - U(x_c)]/D\Big). \tag{30}$$

This time, the difference in the depth of the minima $\Delta U = U(x_a) - U(x_c)$ relative to D is the decisive factor.

III. COLORED NOISE

Colored noise problems are entirely different beasts—the innocuous change to a finite autocorrelation time of $\xi(t)$ has quite a bit of impact [1, chapter IX.7]. One immediate consequence is that we need to specify carefully how the system was prepared. In the white noise case, we can simply fix the initial condition x_0 and the problem is fully specified because the driving noise does not remember its past. However, as soon as the noise has a memory, we need to specify the shared history of noise and system; the most simple history is to assume that x and ξ are independent. In other cases, the history leads to correlations which need to be taken into account. Of course this becomes irrelevant if one considers the stationary situation in which all details of the preparation are forgotten.

A. Stationary Linear Processes

Since we can forget about the preparation in the stationary case, let us first focus on this case. Furthermore, we also consider linear processes $x(t) = \int G(t-s)\xi(s)ds$ to make our lives even more easy. In contrast to Eq. (21), $\xi(t)$ is now a zero-mean Gaussian process with stationary correlation function $C_\xi(\tau)$ and corresponding power spectrum $S_\xi(f)$. Conveniently, the Fourier transformation approach immediately generalizes: From $\tilde{x}(f) = \tilde{G}(f)\tilde{\xi}(f)$, we get [6, appendix A.19]

$$S_x(f) = |\tilde{G}(f)|^2 S_\xi(f). \tag{31}$$

Inserting the white noise spectrum $S_\xi(f) = 2D$ leads us back to Eq. (22).

B. Markovian Embedding

To go beyond the stationary regime, let us consider the SDE $\dot{x} = f(x) + \eta(t)$ where $\eta(t)$ is an OU process, i.e., a process with a purely exponential autocorrelation function. We already know that the OU process obeys Eq. (11), hence we can rewrite the process as

$$\dot{x} = f(x) + \eta, \qquad \tau_c\dot{\eta} = -\eta + \sqrt{2\sigma_\eta^2\tau_c}\xi(t), \tag{32}$$

where $\xi(t)$ is a zero-mean, unit-strength GWN and we chose the parameters such that the autocorrelation time of ξ is τ_c and its stationary variance σ_ξ^2. Although we started with a non-Markovian problem, we managed to convert it to a Markovian one by extending the state space—this is the idea of a *Markovian embedding* [7]. We already saw something very similar for the harmonic oscillator where we pulled the usual trick of rewriting a second-order differential equation as two coupled first-order differential equations. Since we are back at the Markovian level, we can fully specify the preparation of the system by fixing x_0 and η_0.

Of course, we are not restricted to OU process. We could for example assume that $\eta(t)$ is a noise-driven harmonic oscillator and end up with a three-dimensional state space. For more a more intricate example, see [8].

C. Kubo Oscillator

Let's finish with one of the few analytically solvable colored noise problems [7]: The Kubo oscillator

$$\dot{x} = i[\omega + \xi(t)]x. \tag{33}$$

Why is it called an oscillator? If we separate $x = \Re(x) + i\Im(x)$ and disregard the noise, we recover the familiar harmonic oscillator $\Re(\ddot{x}) = -\omega^2\Re(x)$. Note that Eq. (33) is not equivalent to $\Re(\ddot{x}) = -(\omega + \xi(t))^2\Re(x)$!

The easiest way forward is to transform $x = e^{i\theta}$ with $dx/x = id\theta$; this leads to $\dot{\theta} = \omega + \xi(t)$ which is straightforward to integrate to $\theta(t) = \theta_0 + \omega t + \int_0^t \xi(s)ds$. Here, we assumed a correlation-free preparation. Inserting into $x = e^{i\theta}$ and averaging leads to

$$\mu_x(t) = x_0 e^{i\omega t}\langle e^{i\int_0^t \xi(s)ds}\rangle = x_0 e^{i\omega t - \frac{1}{2}\int_0^t\int_0^t C_\xi(s,s')ds'ds} \tag{34}$$

where we used the characteristic functional of $\xi(t)$ in the second step and defined $x_0 = e^{i\theta_0}$. With the same trick, we can also compute the second moment:

$$\langle x(t+\tau)x^*(t)\rangle = e^{i\omega\tau}\langle e^{i\int_t^{t+\tau}\xi(s)ds}\rangle = e^{i\omega\tau - \frac{1}{2}\int_t^{t+\tau}\int_t^{t+\tau} C_\xi(s,s')ds'ds}. \tag{35}$$

Let's take a moment and appreciate the two closed-form analytical solutions valid for arbitrary correlation functions C_ξ.

For concreteness, let us assume an exponential autocorrelation, $C_\xi(\tau) = \sigma_\xi^2 e^{-\tau/\tau_c}$. Calculating the double integral in the exponent, we get $\int_0^t\int_0^t C_\xi(s-s')ds'ds = 2\sigma_\xi^2\tau_c[t - \tau_c(1-e^{-t/\tau_c})]$. Inserting this into Eq. (34), we arrive at

$$|\mu_x(t)| = e^{-\sigma_\xi^2\tau_c[t-\tau_c(1-e^{-t/\tau_c})]} \sim \begin{cases} e^{-\sigma_\xi^2\tau_c t} & \tau_c \to 0 \\ e^{-\frac{1}{2}\sigma_\xi^2 t^2} & \tau_c \to \infty \end{cases}$$

where we took the absolute value to discard the boring contributions. We see a weird effect: The mean decreases with time and it does so very rapidly for long correlation times.

The Kubo oscillator can apparently be related to the line shape of resonances [1, chapter XVI.1]. It is also very useful for dynamical mean-field theory of oscillatory units [9].

IV. MISCELLANEOUS

A. Path Integrals

Having reached the miscellanea, we can now discuss an arcane topic: path integrals. Conceptually they are very nice because they consider the entire process with its corresponding statistics. However, compared to the methods above, their practical value is limited.

Probability Density Functional

The first thing we would like to do is to measure the probability of a realization of a stochastic process. Let's consider the most simple case first, the Wiener process $\dot{W} = \xi(t)$. Integrating over a short period $\Delta t = t_i - t_{i-1}$, we get the Wiener increments $W_i - W_{i-1} = \int_{t_{i-1}}^{t_i}\xi(s)ds$ which are by construction independent, zero-mean Gaussian with variance $2D\Delta t$. Thus, the joint density of N such Wiener increments starting from W_0 is

$$P_W(W_1,\dots,W_N\,|\,W_0) = \prod_{i=1}^N\left(\frac{1}{\sqrt{4\pi D\Delta t}}\right)\exp\left(-\frac{1}{4D\Delta t}\sum_{i=1}^N(W_i - W_{i-1})^2\right). \tag{36}$$

Now we fix $T = N\Delta t$ and let Δt shrink and N grow proportionally. Informally taking the limit $\Delta t \to 0$, we rewrite the sum in the exponent as $\sum_{i=1}^{N} \Delta t \left(\frac{W_i - W_{i-1}}{\Delta t}\right)^2 \to \int_0^T \dot{W}^2 dt$ and arrive at our first probability density functional (PDF)

$$P_W[W] = C \exp\left(-\frac{1}{4D}\int_0^T \dot{W}^2 dt\right) \tag{37}$$

where we lumped the multiplicative prefactor into the constant C.

A SDE is nothing but a transformation of the stochastic process ξ to the stochastic process $x = x[\xi]$. Thus, we expect that their respective PDFs are related by $P_\xi[\xi]\mathcal{D}\xi = P_x[x]\mathcal{D}x$. We already found the PDF of ξ since $\dot{W} = \xi$ is linear such that the Jacobian is not relevant: $P_\xi[\xi] = C\exp(-\frac{1}{4D}\int_0^T \xi^2 dt)$. All that's left to calculate to get $P_x[x] = \frac{\mathcal{D}\xi}{\mathcal{D}x} P_\xi[\xi[x]]$ is the Jacobian. For the SDE $\dot{x} = f(x) + \xi(t)$, the Jacobian turns out to be $\exp(-\frac{1}{2}\int_0^T f'(x)dt)$ and we get [10]

$$P[x] = C \exp\left(-\frac{1}{4D}\int_0^T \left[(\dot{x} - f(x))^2 + 2Df'(x)\right] dt\right). \tag{38}$$

Curiously, the Jacobian does depend on the discretization although there is no Ito-Stratonovich dilemma for additive processes: The above result holds in Stratonovich discretization; in Ito discretization the Jacobian is one. Going back to the respective underlying discretization, it can be shown that the difference is irrelevant [10], as it should be. Again, the advantage of the Stratonovich discretization is that the simple transformation rule $P_\xi[\xi]\mathcal{D}\xi = P_x[x]\mathcal{D}x$ holds [10].

Having determined the PDF, we can now calculate things—for example the transition probability,

$$P_T(x_f) = \int_{x_0}^{x_f} \mathcal{D}x\, P[x], \tag{39}$$

by integrating over all paths that obey the initial condition $x(0) = x_0$ and end at $x(T) = x_f$.

A problem is that the constant C is infinite in the limit $\Delta t \to 0$. This can be fixed by considering ratios of PDF's: Dividing $P[x]$ by $P_W[x]$, which is equivalent to the corresponding Radon-Nicodim derivative, we get

$$\frac{dP}{dP_W}[x] = \exp\left(-\frac{1}{4D}\int_0^T \left[f(x)^2 - 2\dot{x}f(x) + 2Df'(x)\right] dt\right). \tag{40}$$

Apparently, Eq. (40) is a direct consequence of the *Ghirsanov theorem* [10], which would put all of this on solid mathematical footing.

Ornstein-Uhlenbeck Process

To drive the point of the comparably little practical value home, let's solve the OU proces using path integrals. With $f(x) = -\gamma x$ and using $2\gamma\int_0^T \dot{x}x dt = \gamma[x^2]_0^T$, we get from Eq. (38) the PDF

$$P_T(x_f) = Ce^{-\frac{\gamma}{4D}\left[x_f^2 - x_0^2\right]} \int_{x_0}^{x_f} \mathcal{D}x\, e^{-\frac{1}{4D}\int_0^T\left[\dot{x}^2 + \gamma^2 x^2\right]dt} \tag{41}$$

where we absorbed the constant contribution from the Jacobian in C. To solve the path integral, we expand the path around the most likely path x_*, i.e., $x = x_* + \delta x$ with $\delta x(0) = \delta x(T) = 0$, and make a linear shift of the integration variable $x \to \delta x$. Since the exponent is quadratic, there are only three types of terms: 1) terms quadratic in x_* and independent of δx which we can pull out of the integral, 2) mixed terms which vanish by definition of x_*, and 3) terms quadratic in δx and independent of x_*. Since δx does not depend anymore on x_f and x_0, the remaining integral with an exponent quadratic in δx can be absorbed in C and we arrive at

$$P_T(x_f) = Ce^{-\frac{\gamma}{4D}\left[x_f^2 - x_0^2\right]} e^{-\frac{1}{4D}\int_0^T \left[\dot{x}_*^2 + \gamma^2 x_*^2\right] dt}.$$

All that's left is to find the most likely path x_*. To this end, we identify the integrand in the exponent as a Lagrangian and consider the associated Euler-Lagrange equation:

$$\mathcal{L}(\dot{x}, x) = \dot{x}^2 + \gamma^2 x^2 \qquad \overset{\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}} = \frac{\partial \mathcal{L}}{\partial x}}{\Rightarrow} \qquad \ddot{x}_* = \gamma^2 x_*. \tag{42}$$

With the ansatz $x_*(t) = Ae^{\gamma t} + Be^{-\gamma t}$ and a bit of annoying manipulations, we get the solution

$$x_*(t) = \frac{\sinh(\gamma t)}{\sinh(\gamma T)} x_f + \frac{\sinh(\gamma(T-t))}{\sinh(\gamma T)} x_0 \tag{43}$$

which obeys the boundary conditions due to $\sinh(0) = 0$. Now we need to plug x_* into the integral; fortunately we don't have to solve the integral since $\int_0^T \mathcal{L}(\dot{x}_*, x_*) dt = [\dot{x}_* x_*]_0^T$ due to the Euler-Lagrange equation. Inserting this, we arrive after further manipulations at

$$P_T(x) = Ce^{-\frac{x_f^2 - 2x_f x_0 e^{-\gamma T} + \ldots}{2\frac{D}{\gamma}(1 - e^{-2\gamma T})}}$$

which corresponds to a Gaussian with mean $x_0 e^{-\gamma T}$ and variance $\frac{D}{\gamma}(1 - e^{-2\gamma T})$ in agreement with Eq. (17).

Escape Paths

To end with an example where path integrals shine, let us consider a problem where their global view on the process is necessary: determining the most likely path the leads out of a metastable state, the escape path. We consider the SDE $\dot{x} = -U'(x) + \xi(t)$ corresponding to overdamped, noisy motion in the potential $U(x)$. The corresponding PDF follows from Eq. (38)

$$P_T(x_f) = Ce^{-\frac{1}{2D}\left[U(x_f) - U(x_0)\right]} \int_{x_0}^{x_f} \mathcal{D}x \, e^{-\frac{1}{4D}\int_0^T \left[\dot{x}^2 + U'(x)^2\right] dt + \frac{1}{2}\int_0^T U''(x) dt} \tag{44}$$

where we used $\int_0^T \dot{x} U'(x) dt = [U(x)]_0^T$ and separated the Jacobian for convenience. For small noise $D \ll 1$, $\frac{1}{4D}\int_0^T \left[\dot{x}^2 + U'(x)^2\right] dt$ dominates the exponent such that the relevant Langrangian for the most likely path and the associated Euler-Lagrange equation are [11]

$$\mathcal{L}(\dot{x}, x) = \dot{x}^2 + U'(x)^2 \qquad \overset{\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}} = \frac{\partial \mathcal{L}}{\partial x}}{\Rightarrow} \qquad \ddot{x}_* = U'(x_*) U''(x_*). \tag{45}$$

Solving the Euler-Lagrange equation with the boundary conditions $x_*(0) = x_0$ and $x_*(T) = x_f$ yields the escape path.

Whether or not we are able to determine the escape path depends on the potential and our ability to solve the corresponding boundary value problem. Conveniently, we already considered the case $U(x) = \frac{\gamma}{2}x^2$ above, leading to Eq. (43). Thus, Eq. (43) provides the most likely path of the voltage of a LIF neuron to the threshold deep in the noise-driven regime.

REFERENCES

[1] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry, Third Edition (North-Holland Personal Library)*, 3rd ed. (North Holland, 2007).

[2] H. Risken, *The Fokker-Planck Equation* (Springer Verlag Berlin Heidelberg, 1996).

[3] M. C. Wang and G. E. Uhlenbeck, On the theory of the brownian motion ii, Rev. Mod. Phys. **17**, 323 (1945), reprinted in [12].

[4] B. Lindner, A brief introduction to some simple stochastic processes, in *Stochastic Methods in Neuroscience*, edited by C. Laing and G. J. Lord (OUP Oxford, 2009).

[5] P. Hänggi, P. Talkner, and M. Borkovec, Reaction-rate theory: fifty years after kramers, Rev. Mod. Phys. **62**, 251 (1990).

[6] F. Rieke, D. Warland, R. de Ruyter van Steveninck, and W. Bialek, *Spikes: Exploring the Neural Code* (The MIT Press, Cambridge, MA, 1997).

[7] P. Hänggi and P. Jung, Colored noise in dynamical systems, Adv. Chem. Phys. **89**, 239 (1995).

[8] S. Vellmer and B. Lindner, Theory of spike-train power spectra for multidimensional integrate-and-fire neurons, Phys. Rev. Res. **1**, 023024 (2019).

[9] A. van Meegen and B. Lindner, Self-consistent correlations of randomly coupled rotators in the asynchronous state, Phys. Rev. Lett. **121**, 258302 (2018).

[10] R. L. Stratonovich, Some markov methods in the theory of stochastic processes in nonlinear dynamical systems, in *Noise in Nonlinear Dynamical Systems*, Vol. 1, edited by F. Moss and P. V. E. McClintock (Cambridge University Press, 1989) pp. 16–71.

[11] M. I. Dykman, P. V. E. McClintock, V. N. Smelyanski, N. D. Stein, and N. G. Stocks, Optimal paths and the prehistory problem for large fluctuations in noise-driven systems, Phys. Rev. Lett. **68**, 2718 (1992).

[12] N. Wax, ed., *Selected Papers on Noise and Stochastic Processes* (Dover Publications, New York, 1954).