A short tutorial on the path integral formalism for stochastic differential equations

Dani Martí
Group for Neural Theory,
École Normale Supérieure, Paris, France

This is a very short and incomplete introduction to the formalism of path integrals for stochastic systems. The tutorial is strongly inspired by the very nice introduction by Chow and Buice (2015), which is far more complete than this text and whose reading I highly recommend if you are interested in the subject.

Goal: Provide a general formalism to compute moments, correlations, and transition probabilities for any stochastic process, regardless of its nature. The formalism should allow for a systematic perturbative treatment.

Preliminaries: Moment-generating functionals

Single variable

Consider a single random variable X with probability density function $p(x)$, not necessarily normalized. The n-th moment of X is defined as

$$\langle X^n \rangle \equiv \frac{\int x^n p(x)\,\mathrm{d}x}{\int p(x)\,\mathrm{d}x}.$$

Instead of computing the moments one by one, we can make use of the *moment-generating function*, defined as

$$Z(\lambda) \equiv \int \mathrm{e}^{\lambda x} p(x)\,\mathrm{d}x. \tag{1}$$

With this definition, moments can be computed by taking derivatives of $Z(\lambda)/Z(0)$, where $1/Z(0)$ is just the normalization factor of the probability density:

$$\langle X^n \rangle = \frac{\mathrm{d}^n}{\mathrm{d}\lambda^n} \frac{Z(\lambda)}{Z(0)}\bigg|_{\lambda=0}. \tag{2}$$

The expansion of $Z(\lambda)$ in a power series reads then

$$\frac{Z(\lambda)}{Z(0)} = \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} \langle X^m \rangle. \tag{3}$$

• Example: moment-generating function for a Gaussian random variable X with mean μ and covariance matrix σ^2. Plugging the (non-normalized) Gaussian probability density $p(x) = \exp[-(x-\mu)^2/(2\sigma^2)]$ into Eq. (1) gives

$$Z(\lambda) = \int_{-\infty}^{\infty} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2} + \lambda x\right] \mathrm{d}x, \tag{4}$$

which can be evaluated by 'completing the square' in the exponent,

$$-\frac{1}{2\sigma^2}(x-\mu)^2 + \lambda x = -\frac{1}{2\sigma^2}(x-\mu-\lambda\sigma^2)^2 + \lambda\mu + \frac{\lambda^2\sigma^2}{2}.$$

Carrying out integral over x yields

$$\frac{Z(\lambda)}{Z(0)} = \exp\left(\lambda\mu + \frac{1}{2}\lambda^2\sigma^2\right),$$

where $Z(0) = \int_{-\infty}^{\infty} \mathrm{e}^{-x^2/(2\sigma^2)}\,\mathrm{d}x = \sqrt{2\pi}\sigma$. Once we have $Z(\lambda)$ we can easily compute the moments using Eq. (2). So the first and second moments of x are, somewhat unsurprisingly,

$$\langle X\rangle = \frac{\mathrm{d}}{\mathrm{d}\lambda}\exp\left(\lambda\mu + \frac{1}{2}\lambda^2\sigma^2\right)\bigg|_{\lambda=0} = \mu,$$
$$\langle X^2\rangle = \frac{\mathrm{d}^2}{\mathrm{d}\lambda^2}\exp\left(\lambda\mu + \frac{1}{2}\lambda^2\sigma^2\right)\bigg|_{\lambda=0} = \mu^2 + \sigma^2.$$

The *cumulants* of the random variable X, denoted by $\langle\!\langle X^n\rangle\!\rangle$, are defined by

$$\langle\!\langle X^n\rangle\!\rangle = \frac{\mathrm{d}^n}{\mathrm{d}\lambda^n}\ln\frac{Z(\lambda)}{Z(0)}\bigg|_{\lambda=0}. \tag{5}$$

Equivalently, the expansion of $\ln(Z(\lambda)/Z(0))$ in power series is

$$\ln\frac{Z(\lambda)}{Z(0)} = \sum_{n=1}^{\infty}\frac{\lambda^n}{n!}\langle\!\langle X^n\rangle\!\rangle. \tag{6}$$

The relation between cumulants and moments follows from equating the expansion (6) to the logarithm of the expansion (3),

$$\sum_{n=1}^{\infty}\frac{\lambda^n}{n!}\langle\!\langle X^n\rangle\!\rangle = \ln\sum_{n=0}^{\infty}\frac{\lambda^n}{n!}\langle X^n\rangle = \ln\left(1 + \sum_{n=1}^{\infty}\frac{\lambda^n}{n!}\langle X^n\rangle\right), \tag{7}$$

expanding the right-hand side of Eq. (7), and equating coefficients of equal powers of λ. The first four cumulants are

$$\langle\!\langle X\rangle\!\rangle = \langle X\rangle$$
$$\langle\!\langle X^2\rangle\!\rangle = \langle X^2\rangle - \langle X\rangle^2 = \left\langle\big(X - \langle X\rangle\big)^2\right\rangle$$
$$\langle\!\langle X^3\rangle\!\rangle = \langle X^3\rangle - 3\langle X^2\rangle\langle X\rangle + 2\langle X\rangle^3 = \left\langle\big(X - \langle X\rangle\big)^3\right\rangle$$
$$\langle\!\langle X^4\rangle\!\rangle = \langle X^4\rangle - 4\langle X^3\rangle\langle X\rangle - 3\langle X^2\rangle^2 + 12\langle X^2\rangle\langle X\rangle^2 - 6\langle X\rangle^4 = \left\langle\big(X - \langle X\rangle\big)^4\right\rangle - 3\langle\!\langle X^2\rangle\!\rangle^2.$$

• Example: cumulant-generating function for a Gaussian random variable X with mean μ and variance σ^2. We just compute the logarithm of the moment-generating function found in the previous example,

$$\ln\frac{Z(\lambda)}{Z(0)} = \lambda\mu + \frac{1}{2}\lambda^2\sigma^2$$

and identify the cumulants through relation (6). We have $\langle\!\langle X\rangle\!\rangle = \mu$, $\langle\!\langle X^2\rangle\!\rangle = \sigma^2$, and $\langle\!\langle X^n\rangle\!\rangle = 0$ for $n > 2$.

Multidimensional variables

Let $\mathbf{X}$ be a random variable having r components $X_1, X_2, \ldots, X_r$, and with joint probability density $p(\mathbf{x}) \equiv p(x_1, x_2, \ldots, x_r)$. The moments of $\mathbf{X}$ are

$$\langle X_1^{m_1} X_2^{m_2} \cdots X_r^{m_r} \rangle = \frac{\int x_1^{m_1} x_2^{m_2} \cdots x_r^{m_r} p(\mathbf{x})\,\mathrm{d}\mathbf{x}}{\int p(\mathbf{x})\,\mathrm{d}\mathbf{x}},$$

where $\mathrm{d}\mathbf{x} \equiv \mathrm{d}x_1 \mathrm{d}x_2 \cdots \mathrm{d}x_r$. The moment-generating function of $\mathbf{X}$ is a straightforward generalization of the one-dimensional version

$$Z(\boldsymbol{\lambda}) \equiv \int \exp(\boldsymbol{\lambda}^T \mathbf{x})\, p(\mathbf{x})\,\mathrm{d}\mathbf{x}, \tag{8}$$

where here $\boldsymbol{\lambda}$ is an r-dimensional vector and $\boldsymbol{\lambda}^T \mathbf{x}$ is just the dot product between $\boldsymbol{\lambda}$ and $\mathbf{x}$. Moments are generated with

$$\langle X_1^{m_1} X_2^{m_2} \cdots X_r^{m_r} \rangle = \frac{\partial^{m_1}}{\partial x_1^{m_1}} \frac{\partial^{m_2}}{\partial x_2^{m_2}} \cdots \frac{\partial^{m_r}}{\partial x_r^{m_r}} \frac{Z(\boldsymbol{\lambda})}{Z(\mathbf{0})}\bigg|_{\boldsymbol{\lambda}=\mathbf{0}} \tag{9}$$

• Example: moment-generating function for a random Gaussian vector $\mathbf{X}$ with zero mean and variance $\boldsymbol{\Sigma} = \mathbf{K}$, i.e., $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{K})$. If we drop, as usual, the normalization factor of the probability density we have

$$Z(\boldsymbol{\lambda}) = \int \exp\left(-\frac{1}{2}\mathbf{x}^T \mathbf{K}^{-1} \mathbf{x} + \boldsymbol{\lambda}^T \mathbf{x}\right) \mathrm{d}\mathbf{x}. \tag{10}$$

The matrix $\mathbf{K}$ is symmetric and therefore diagonalizable (if $\mathbf{K}$ happens to be non-symmetric, the contribution from its antisymmetric part to the quadratic form $\mathbf{x}^T\mathbf{K}^{-1}\mathbf{x}$ vanishes). Let us denote by $\{\mathbf{u}_i\}$ the basis of unit eigenvectors of $\mathbf{K}$, and by $\mathbf{U}$ the matrix that results from arranging the set $\{\mathbf{u}_i\}$ in columns. Because $\{\mathbf{u}\}$ is an orthonormal set, $\mathbf{U}^T\mathbf{U} = \mathbf{U}\mathbf{U}^T = \mathbf{1}$. The diagonalization condition can be summarized by $\mathbf{K}\mathbf{U} = \mathbf{U}\mathbf{D}$, where $\mathbf{D}$ is the diagonal matrix with the eigenvalue σ_i^2 at its i-th diagonal entry. In terms of the new basis, $\mathbf{K}^{-1} = \mathbf{U}\mathbf{D}^{-1}\mathbf{U}^T$, $\mathbf{x} = \sum_i c_i \mathbf{u}_i = \mathbf{U}\mathbf{c}$, and $\boldsymbol{\lambda} = \sum_i d_i \mathbf{u}_i = \mathbf{U}\mathbf{d}$, where $\mathbf{c}$ and $\mathbf{d}$ are coefficient vectors. We have

$$\exp\left(-\frac{1}{2}\mathbf{x}^T\mathbf{K}^{-1}\mathbf{x} + \boldsymbol{\lambda}^T\mathbf{x}\right) = \exp\left(-\frac{1}{2}\mathbf{c}^T\mathbf{D}^{-1}\mathbf{c} + \mathbf{d}^T\mathbf{c}\right) = \prod_i \exp\left(-\frac{c_i^2}{2\sigma_i^2} + c_i d_i\right),$$

and therefore

$$\begin{aligned} Z(\boldsymbol{\lambda}) &= \int \prod_i \exp\left(-\frac{c_i^2}{2\sigma_i^2} + c_i d_i\right) |\mathbf{U}|\,\mathrm{d}\mathbf{c}, = \prod_i \int \exp\left(-\frac{c_i^2}{2\sigma_i^2} + c_i d_i\right) \mathrm{d}c_i \\ &= \prod_i \left(2\pi\sigma_i^2\right)^{1/2} \exp\left(\frac{d_i^2\sigma_i^2}{2}\right) = |2\pi\mathbf{D}|^{1/2} \exp\big(\mathbf{d}^T\mathbf{D}\mathbf{d}\big) = |2\pi\mathbf{K}|^{1/2} \exp\big(\boldsymbol{\lambda}^T\mathbf{K}\boldsymbol{\lambda}\big) \\ &= Z(\mathbf{0}) \exp\big(\boldsymbol{\lambda}^T\mathbf{K}\boldsymbol{\lambda}\big). \end{aligned} \tag{11}$$

where in the first line we used that the Jacobian $|\mathbf{U}|$ is 1 for orthogonal transformations, and in the second line we completed the square for all the decoupled factors.

Although the Gaussian example is just a particular case, it is an important one because it can be computed in a closed form and because it is the basis for many perturbative schemes.

The moment-generating function $Z(\boldsymbol{\lambda})$ is an exponential of a quadratic form in λ, which implies that only moments of even order will survive (see Eq. (9)). In particular we have

$$\langle x_a x_b \rangle = \frac{\partial}{\partial x_a}\frac{\partial}{\partial x_b}\frac{Z(\boldsymbol{\lambda})}{Z(\mathbf{0})}\bigg|_{\boldsymbol{\lambda}=\mathbf{0}} = K_{ab},$$

where a, b are any two indices of the components of $\mathbf{x}$.

Moreover, it is not difficult to convince oneself that for Gaussian vectors any moment of even-order can be obtained as a sum of all the possible pairings between variables, for example:

$$\langle x_a x_b x_c x_d \rangle = \langle x_a x_b \rangle \langle x_c x_d \rangle + \langle x_a x_c \rangle \langle x_b x_d \rangle + \langle x_a x_d \rangle \langle x_b x_c \rangle.$$

This property is usually referred to as Wick's theorem, or Isserlis's theorem.

Fields

We can extend from finite-dimensional systems to functions defined on some domain $[0, T]$. The basic idea is to discretize the domain interval into n segments of length h such that $T = nh$, and then take the limits $n \to \infty$ and $h \to 0$ while keeping $T = nh$ constant. Under this limit the originally discrete set $\{t_0, t_1, \dots, t_n = T\}$ becomes dense on the interval $[0, T]$. Any t can then be approximated as $t = jh$, for some $j = 0, 1, \dots, n$ and we can identify

$$x_j \longrightarrow x(t), \qquad \lambda_j \longrightarrow \lambda(t), \qquad K_{ij} \longrightarrow K(s,t).$$

Now instead of a moment-generating function we'll have a moment-generating *functional*, an object that maps functions into real numbers, and which is formally defined as

$$Z[\lambda(t)] = \int \mathcal{D}x(s)\, \exp\left(\int \lambda(u)x(u)\,\mathrm{d}u\right) p[x(s)]$$

where we use square brackets to denote a functional and we defined the measure for integration over functions as

$$\mathcal{D}x(t) \equiv \lim_{n\to\infty} \prod_{i=0}^{n} \mathrm{d}x_i.$$

The integral in Eq. (12) is an example of *path integral*, or *functional integral* (Cartier and DeWitt-Morette, 2010). For the Gaussian case, we can write down explicitly the probability $p[x(s)]$ taking the functional versions of Eqs. (10) and (11):

$$\begin{aligned} Z[\lambda(t)] &= \int \mathcal{D}x(s) \exp\left(-\frac{1}{2}\iint x(u)K^{-1}(u,v)x(v)\,\mathrm{d}u\,\mathrm{d}v + \int \lambda(u)x(u)\,\mathrm{d}u\right) \\ &= Z[0]\exp\left(\frac{1}{2}\iint \lambda(u)K(u,v)\lambda(v)\,\mathrm{d}u\,\mathrm{d}v\right). \end{aligned} \tag{12}$$

Here $Z[0]$ is the value of the generating functional with the external source $\lambda(t)$ set to 0. The presentation here is obviously very handwavy, and there are several ill-defined quantities lying around like, for instance, the prefactor $Z[0] = \lim_{n\to\infty} |2\pi K|^{1/2}$, which is formally infinite. Fortunately the moments we derive from the moment-generating functional are well-defined.

Now we need to compute the moments, for which we need the notion of functional derivative. Although functional derivatives can be defined rigorously, here we only need to remember the following axioms

$$\frac{\delta\lambda(t)}{\delta\lambda(s)} = \delta(t-s) \quad \left(\text{equivalent to} \quad \frac{\partial\lambda_j}{\partial\lambda_i} = \delta_{ij}\right),$$

$$\frac{\delta}{\delta\lambda(t)}\int \lambda(s)x(s)\,\mathrm{d}s = x(t) \quad \left(\text{equivalent to} \quad \frac{\partial}{\partial\lambda_i}\sum_j \lambda_j x_j = x_i\right),$$

where $\delta(\cdot)$ is the Dirac delta, or point-mass functional, and δ_{ij} is the Kronecker delta: $\delta_{ij} = 1$ if $i = j$, 0 otherwise.

As usual, moments are obtained from taking functional derivatives on the moment-generating functional:

$$\left\langle \prod_i x(t_i) \right\rangle = \prod_i \frac{\delta}{\delta\lambda(t_i)} \frac{Z[\lambda]}{Z[0]}\bigg|_{\lambda(t)=0}.$$

• Example: given the moment-generating functional of a Gaussian process, Eq. (12), the two-point correlation function is

$$\langle x(t)x(s)\rangle = \frac{\delta}{\delta\lambda(t)} \frac{\delta}{\delta\lambda(s)} \frac{Z[\lambda]}{Z[0]}\bigg|_{\lambda(u)=0} = K(t,s).$$

More generally, any (even-order) moment will be of the form

$$\left\langle \prod_i x(t_i) \right\rangle = \prod_i \frac{\delta}{\delta\lambda(t_i)} \frac{Z[\lambda]}{Z[0]}\bigg|_{\lambda(t)=0} = \sum_{\substack{\text{all possible}\\ \text{pairings}}} K(t_{i_1}, t_{i_2}) \cdots K(t_{i_{2s-1}}, t_{i_{2s}}),$$

because of Wick's theorem.

Functional formalism for stochastic differential equations

Imagine we have a stochastic process defined by the Langevin equation

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(x,t) + g(x,t)\,\xi(t), \tag{13}$$

with initial condition $x(t_0) = x_0$, and with $x(t)$ defined on the domain $t \in [t_0, T]$. The symbol $\xi(t)$ denotes a source of white noise of zero mean and unit variance: $\langle \xi(t)\rangle = 0$, $\langle \xi(t)\xi(t')\rangle = \delta(t-t')$. Equation (13) is to be interpreted as the Itō stochastic differential equation

$$\mathrm{d}x = f(x,t)\,\mathrm{d}t + g(x,t)\,\mathrm{d}B_t, \tag{14}$$

where $\mathrm{d}B_t$ is a Brownian stochastic process. According to the convention for an Itō stochastic process, $g(x,t)$ is *non-anticipating* (or *adapted*), meaning that when we evaluate the integrals over time, $g(x,t)$ is independent of B_τ for $\tau > t$. To explicitly enforce our choice of initial condition we rewrite Eq. (13) as

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(x,t) + g(x,t)\,\xi(t) + x_0\delta(t-t_0). \tag{15}$$

The goal now is to use the path integral formalism to compute the moments and correlation functions of $x(t)$.

Probability density functional of a stochastic process

We need first the probability density functional for $x(t)$ given the initial condition $x(t_0) = x_0$. We denote this functional by $P[x(t)|x_0]$. This functional is also called the *transition probability* and is the equivalent of the probability density function for finite-dimensional variables, because it assigns a probability measure to each realization $x(t)$ obeying $x(0) = x_0$. To work on less shaky grounds, we start by discretizing version of Eq. (15) and work on a finite-dimensional system. For a small time step h, and according to the Itō prescription, Eq. (15) is to be interpreted as

$$x_{i+1} - x_i = f_i(x_i)h + g_i(x_i)\xi_i\sqrt{h} + x_0\delta_{i0}, \tag{16}$$

where $i \in \{0, 1, \ldots, N\}$, $T = Nh$, δ_{ij} is the Kronecker delta, and ξ_i is a discrete random variable that satisfies $\langle\xi_i\rangle = 0$ and $\langle\xi_i\xi_j\rangle = \delta_{ij}$. The subindices in f_i and g_i mean that functions $f(x,t)$ and $g(x,t)$ are evaluated at time $t = ih$. Note that the discretized stochastic variable vector x_i depends on both the discretized white noise process ξ_i and the initial condition x_0.

Formally, the joint probability density functional for the vector $\mathbf{x}$ can be written as the point mass constrained at the *particular* solution of the stochastic differential equation:

$$P(\mathbf{x}|\boldsymbol{\xi};x_0) = \prod_{i=0}^{N} \delta\big(x_{i+1} - x_i - f_i(x_i)h - g_i(x_i)\xi_i\sqrt{h} - x_0\delta_{i0}\big), \tag{17}$$

where we introduced the vectors $\mathbf{x} = (x_0, x_1, \ldots, x_N)$ and $\boldsymbol{\xi} = (\xi_0, \xi_1, \ldots, \xi_{N-1})$. By *particular solution* we mean that associated with a particular realization $\boldsymbol{\xi}$. In general we are interested in the joint probability density function $P(\mathbf{x}|x_0)$, which results from integrating out the noise

$$P(\mathbf{x}|x_0) = \int P(\mathbf{x}|\boldsymbol{\xi};x_0) \prod_{j=0}^{N} p(\xi_j)\,\mathrm{d}\xi_j \equiv \int P(\mathbf{x}|\boldsymbol{\xi};x_0)\,p(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}. \tag{18}$$

To compute Eq. (18), with $P(\mathbf{x}|\boldsymbol{\xi};x_0)$ given by Eq. (17), we introduce the Fourier representation of the Dirac delta

$$\delta(x) = \frac{1}{2\pi}\int \mathrm{e}^{-\mathrm{i}kx}\,\mathrm{d}k$$

for every factor in Eq. (17). We get

$$P(\mathbf{x}|\boldsymbol{\xi};x_0) = \int \prod_{j=0}^{N} \frac{\mathrm{d}k_j}{2\pi} \exp\Big[-\mathrm{i}\sum_j k_j\big(x_{j+1} - x_j - f_j(x_j)h - g_j(x_j)\xi_j\sqrt{h} - x_0\delta_{j0}\big)\Big].$$

To go further we need to specify the statistical properties of the noise source, determined by the probability density of ξ_i, $p(\xi_i)$. Let us assume for the moment that the noise is Gaussian and white, for which

$$p(\xi_i) = \frac{1}{\sqrt{2\pi}} \exp\left(-\xi_i^2/2\right).$$

The transition probability Eq. (18) reads then

$$\begin{aligned}\int P(\mathbf{x}|\boldsymbol{\xi};x_0)\prod_{j=0}^{N} p(\xi_j)\mathrm{d}\xi_j &= \int \prod_{j=0}^{N} \frac{\mathrm{d}k_j}{2\pi} \exp\Big[-\mathrm{i}\sum_j k_j\big(x_{j+1} - x_j - f_j(x_j)h - x_0\delta_{j0}\big)\Big] \\ &\times \int \prod_{j=0}^{N} \frac{\mathrm{d}\xi_j}{\sqrt{2\pi}} \exp\Big(\mathrm{i}k_j g_j(x_j)\xi_j\sqrt{h} - \xi_j^2/2\Big).\end{aligned}$$

We can integrate the last integral by completing the square, getting

$$\begin{aligned}P(\mathbf{x}|x_0) &= \int \prod_{j=0}^{N} \left\{\frac{\mathrm{d}k_j}{2\pi}\right\} \exp\left\{-\sum_j \mathrm{i}k_j \left[\frac{x_{j+1}-x_j}{h} - f_j(x_j) - x_0\frac{\delta_{j0}}{h}\right] h + \sum_j \frac{1}{2}(\mathrm{i}k_j)^2 g_j^2(x_j)h\right\}, \\ &\equiv \int \prod_{j=0}^{N} \left\{\frac{\mathrm{d}\tilde{x}_j}{2\pi\mathrm{i}}\right\} \exp\left\{-\sum_j \tilde{x}_j \left[\frac{x_{j+1}-x_j}{h} - f_j(x_j) - x_0\frac{\delta_{j0}}{h}\right] h + \sum_j \frac{1}{2}\tilde{x}_j^2 g_j^2(x_j)h\right\}.\end{aligned}$$

where for convenience we introduced a new complex variable $\tilde{x}_i = \mathrm{i}k_i$. If we now take the continuum limit $h \to 0$ and $N \to \infty$ while keeping $T = Nh$ constant, and use the prescriptions

$$\begin{array}{l} x_i \longrightarrow x(t) \\ \tilde{x}_i \longrightarrow \tilde{x}(t) \end{array}, \qquad \frac{x_{i+1} - x_i}{h} \longrightarrow \dot{x}(t), \qquad h\sum_{i=0}^{N} \longrightarrow \int_0^T \mathrm{d}t$$

$$\int \prod_{j=0}^{N} \left\{ \frac{\mathrm{d}\tilde{x}_j}{2\pi \mathrm{i}} \right\} \longrightarrow \int \mathcal{D}\tilde{x}(t), \qquad x_0 \frac{\delta_{i0}}{h} \longrightarrow x_0 \delta(t), \quad f_i(x_i) \longrightarrow f\big(x(t), t\big)\,,$$

we obtain the probability density *functional*

$$P[x(t)|x_0] = \int \mathcal{D}\tilde{x}(t) \exp\big(-S[x, \tilde{x}]\big), \tag{19}$$

where we have defined the *action*

$$S[x, \tilde{x}] \equiv \int \left\{ \tilde{x}(t) \Big[\dot{x}(t) - f(x(t), t) - x_0 \delta(t - t_0) \Big] - \frac{1}{2} \tilde{x}^2(t) g^2(x(t), t) \right\} \mathrm{d}t. \tag{20}$$

The action $S[x, \tilde{x}]$ is commonly known as the Martin-Siggia-Rose-Janssen-de Dominicis action. The new field $\tilde{x}(t)$ is called the auxiliary field, or response field. Note that because of the definition $\mathrm{i}k_i \to \tilde{x}(t)$, the integrals over $\tilde{x}(t)$ are to be evaluated along the *imaginary* axis, which is why no explicit imaginary units appear in the action.

The expectation of any observable quantity $\mathcal{O}[x(t)]$ can now be formally represented as

$$\langle \mathcal{O}[x] \rangle = \int \mathcal{D}x(t)\, \mathcal{O}[x(t)] P[x(t)|x_0] = \int \mathcal{D}x(t)\, \mathcal{D}\tilde{x}(t)\, \mathcal{O}[x(t)] \exp\big(-S[x, \tilde{x}]\big). \tag{21}$$

Moment-generating functional

We define the moment-generating functional analogously to how we defined the finite-dimensional version, Eq. (8), noting that scalar products become time integrals in function space:

$$\boldsymbol{\lambda}^T \boldsymbol{x} \longrightarrow \int \lambda(t) x(t)\, \mathrm{d}t.$$

The moment-generating functional for $x(t)$ and $\tilde{x}(t)$ is by definition the average of

$$\mathcal{O}[x(t), \tilde{x}(t)] = \exp\bigg(\int \tilde{\lambda}(t) x(t)\, \mathrm{d}t + \int \lambda(t) \tilde{x}(t)\, \mathrm{d}t \bigg).$$

Combining this equation with (21) we have

$$\begin{aligned} Z[\lambda, \tilde{\lambda}] &\equiv \bigg\langle \exp\bigg(\int \tilde{\lambda}(t) x(t)\, \mathrm{d}t + \int \lambda(t) \tilde{x}(t)\, \mathrm{d}t \bigg) \bigg\rangle \\ &= \int \mathcal{D}x(t)\, P[x(t)|x_0] \exp\bigg(\int \tilde{\lambda}(t) x(t)\, \mathrm{d}t + \int \lambda(t) \tilde{x}(t)\, \mathrm{d}t \bigg) \\ &= \int \mathcal{D}x(t)\, \mathcal{D}\tilde{x}(t)\, \exp\bigg(-S[x, \tilde{x}] + \int \tilde{\lambda}(t) x(t)\, \mathrm{d}t + \int \lambda(t) \tilde{x}(t)\, \mathrm{d}t \bigg). \end{aligned} \tag{22}$$

In Physics the fields $\lambda(t)$ and $\tilde{\lambda}(t)$ are called *source fields*, and are traditionally denoted by $J(t)$ and $\tilde{J}(t)$.

What if noise is not white or Gaussian?

The formalism can incorporate noise other than Brownian. Because the noise source is a stochastic process itself, it can be characterized by a probability functional (see (19))

$$P[\xi(t)|\xi_0, t_0] = \int \mathcal{D}\tilde{\xi}(t) \exp\bigl(-S[\xi]\bigr) = \exp\bigl(-S[\xi]\bigr),$$

where we assumed that the action of the stochastic process depends on ξ only, $S[\xi, \tilde{\xi}] = S[\xi]$. So, for example, the action for a white Gaussian process is

$$S[\xi] = \frac{1}{2} \int \xi^2(t)\, \mathrm{d}t.$$

Given the probability functional for the noise we can now write the transition probability for $x(t)$ for any general noise source:

$$\begin{aligned} P[x(t)|x_0, t_0] &= \int \mathcal{D}\xi(t)\, \delta\bigl[\dot{x}(t) - f(x,t) - \xi(t) - x_0\delta(t-t_0)\bigr] \exp\bigl(-S[\xi(t)]\bigr) \\ &= \int \mathcal{D}\xi(t)\mathcal{D}\tilde{x}(t)\, \exp\biggl(-\int \tilde{x}(t)\bigl[\dot{x}(t) - f(x,t) - x_0\delta(t-t_0)\bigr]\, \mathrm{d}t \\ &\qquad + \int \tilde{x}(t)\xi(t)\, \mathrm{d}t - S[\xi(t)]\biggr). \end{aligned}$$

We can further simplify this expression noting that

$$\int \mathcal{D}\xi(t) \exp\biggl(-S[\xi(t)] + \int \tilde{x}(t)\xi(t)\, \mathrm{d}t\biggr)$$

is nothing but the definition of the generating functional for $\xi(t)$ with source $\tilde{x}(t)$, i.e., $Z[\tilde{x}(t)] \equiv \exp W[\tilde{x}(t)]$. We can thus write

$$\begin{aligned} P[x(t)|x_0, t_0] &= \int \mathcal{D}\tilde{x}(t) \exp\biggl(-\int \tilde{x}(t)\bigl[\dot{x}(t) - f(x,t) - x_0\delta(t-t_0)\bigr]\, \mathrm{d}t + W[\tilde{x}(t)]\biggr) \\ &\equiv \int \mathcal{D}\tilde{x}(t) \exp\bigl(-S_W[x, \tilde{x}]\bigr). \end{aligned}$$

The moment-generating functional for the stochastic process would then be given by Eq. (22), with the action replaced by $S_W[x, \tilde{x}]$.

We have now defined all the quantities we need to compute the moments. Either if you want to learn how to actually compute them for several relevant examples, or if you want to grasp the power of the path-integral formalism in the general case, I avidly encourage you to read (Chow and Buice, 2015).

References

Cartier P, DeWitt-Morette C (2010) Functional integration: action and symmetries. Cambridge University Press

Chow C, Buice M (2015) Path integral methods for stochastic differential equations. The Journal of Mathematical Neuroscience 5(1):8, 10.1186/s13408-015-0018-5