3.4 Path integral representation of a Langevin process

This section aims at introducing one more tool in the stochastic processes toolbox, beyond the master equation and stochastic differential equations. We have seen the importance of working with trajectories in chapter ??. This is one motivation. Another motivation comes from the fact that path integrals pervade all areas of theoretical physics (high energy, condensed matter) and whatever tool has been developed in one field can be exported to another field. We follow Janssen [51] and De Dominicis [30] who used the work of Martin, Siggia and Rose [85] to adapt path integral techniques to stochastic processes.

3.4.1 Starting from a Langevin equation

We discretize the α-discretized Langevin equation $\frac{\mathrm{d}x}{\mathrm{d}t} \overset{\alpha}{=} f(x(t)) + g(x(t))\xi(t)$ in the standard way

$$\Delta x_n = x_{n+1} - x_n = f_n \Delta t + g_n \sqrt{\Delta t}\xi_n \tag{3.69}$$

where ξ_n is a Gaussian variable with unit variance $\langle \xi_n \xi_m \rangle = \delta_{nm}$ and where $f_n = f(x_n + \alpha \Delta x_n)$, $g_n = g(x_n + \alpha \Delta x_n)$. From the discretized form we see that $\Delta x_n = O(\sqrt{\Delta t})$. The initial value x_0 is given, and the noise index runs from $i = 0$ up to $M-1$, where $M = t_{\text{obs}}/\Delta t$ is the number of time slices we have cut the time interval $[0, t_{\text{obs}}]$ into. The trajectory of interest is given by the x_j sequence, $j = 1, \dots, M$. By averaging a given observable $A(x(t_{\text{obs}}))$ one actually means that

$$\langle A(x_M) \rangle = \int \prod_{i=0}^{M-1} \frac{\mathrm{d}\xi_i}{\sqrt{2\pi}} \mathrm{e}^{-\sum_i \frac{\xi_i^2}{2}} A(x_M[\{\xi_\ell\}]) \tag{3.70}$$

where $x_M[\{\xi_\ell\}]$ is the solution of the Langevin equation for a given sequence of the random numbers ξ_i, $i = 0, \dots, M-1$. We now change variables from the ξ_i's to the x_j's. Given that

$$\xi_i = \frac{x_{i+1} - x_i - f_i \Delta t}{g_i \sqrt{\Delta t}} \tag{3.71}$$

we easily see that the Jacobian matrix $J = \left(\frac{\partial \xi_i}{x_j} \right)_{i=0,\dots,M-1, j=1,\dots,M}$ is a triangular matrix (whose diagonal and diagonal strip just below contain the only nonzero elements), whose diagonal elements are

$$\frac{\partial \xi_j}{\partial x_{j+1}} = \frac{1 - \alpha f_j' \Delta t - \alpha \frac{g_j'}{g_j}(\Delta x_j - f_j \Delta t)}{g_j \sqrt{\Delta t}} \tag{3.72}$$

For $i > j+1$, we have that $\frac{\partial \xi_i}{\partial x_{j+1}} = 0$, as the position at a given time cannot depend on the later value of the noise. Hence the determinant of the Jacobian matrix is given by the product of the diagonal entries. Hence we have that

$$\langle A(x_M) \rangle = \int \prod_{j=1}^{M} \frac{\mathrm{d}x_i}{\sqrt{2\pi}} \prod_{i=0}^{M-1} \left(\frac{1 - \alpha f_i' \Delta t - \alpha \frac{g_i'}{g_i}(\Delta x_i - f_i \Delta t)}{g_i \sqrt{\Delta t}} \right) \mathrm{e}^{-\frac{1}{2}\sum_i \left(\frac{x_{i+1} - x_i - f_i \Delta t}{g_i \sqrt{\Delta t}} \right)^2} A(x_M) \tag{3.73}$$

For the particular case of a constant function g things simplify considerably, and the Jacobian can be rewritten as

$$\prod_j \frac{\partial \xi_j}{\partial x_{j+1}} = \frac{1-\alpha f_j' \Delta t}{g_j\sqrt{\Delta t}} = \text{Cste}^{-\alpha \sum_j f_j' \Delta t} \tag{3.74}$$

so that altogther we can write that

$$\langle A(x_M)\rangle = \int \prod_{j=1}^{M} \frac{\mathrm{d}x_i}{\sqrt{2\pi}} \prod_{i=0}^{M-1} (g^2\Delta t)^{-M/2} \mathrm{e}^{-\Delta t \sum_i \left[\frac{1}{2g^2}\left(\frac{\Delta x_i}{\Delta t}-f_i\right)^2+\alpha f_i'\right]} A(x_M) \tag{3.75}$$

Returning to a continuous time notation, it is traditional to write the $M \gg 1$ limit of Eq. (3.75) as

$$\langle A(t_{\text{obs}})\rangle = \int \mathscr{D}x A(x(t_{\text{obs}}))\mathrm{e}^{-S[x]} \tag{3.76}$$

where the dynamical action, also known as the Onsager-Machlup functional [94, 76], has the expression

$$S[x] = \int_0^{t_{\text{obs}}} \mathrm{d}t \left[\frac{1}{2g^2}(\dot{x}-f)^2 + \alpha f'\right] \tag{3.77}$$

If g were not a constant function, and even in the simplest Itō $\alpha = 0$ case where the Jacobian is unity, one would have to be careful with the continuum limit of

$$\prod_{i=0}^{M-1} \frac{1}{g} \propto \mathrm{e}^{-\sum_i \ln g(x_i+\alpha \Delta x_i)} \tag{3.78}$$

which does not seem to be easy to write, and it is often hidden inside the definition of $\mathscr{D}x$. In the particular case of g being a constant function, the integration measure over paths $\mathscr{D}x$ does not hide any $g(x)$ dependence

$$\langle A(x(t_{\text{obs}}))\rangle = \int \mathcal{D}x \mathrm{e}^{-S} A(x(t_{\text{obs}})) \tag{3.79}$$

and

$$S[x] = \int \mathrm{d}t \left[\frac{1}{2g^2}(\dot{x}-f)^2 + \alpha f'\right] \tag{3.80}$$

is known as the Onsager-Machlup [94, 76] dynamical action, while $\mathscr{D}x$ is the continuum analog of $\prod_i \mathrm{d}x_i$.

3.4.2 Dirty way for Itō

Once clean derivations of this path integral formulation have been worked out, and once it has been realized that the Itō-discretization turns the Jacobian into a mere multiplicative constant, it is possible to pretend that things could be done in three lines. Let us sketch this "derivation", which can only be solidly justified for the Itō discretization. Again we start from $\dot{x} \stackrel{0}{=} f + g\eta$. And we ask about the average of a quantity $A(t) = A(x(t))$. By definition

$$\langle A\rangle = \int \mathscr{D}\eta \mathrm{e}^{-\frac{1}{2}\int \eta^2} A(x[\eta](t)) \tag{3.81}$$

and when we write $x[\eta](t)$ we see x as the solution of the $\dot{x} \overset{0}{=} f + g\eta$ stochastic differential equation in which η is a given function. Hence $x[\eta](t)$ is a functional of η. By asserting that $\mathscr{D}\eta$ and $\mathscr{D}x$ differ only by a multiplicative constant (which is the same as asserting thet Jacobian is a constant), we immediately get

$$\langle A \rangle = \int \mathscr{D}x A(x(t_{\text{obs}}))\mathrm{e}^{-\frac{1}{2g^2}\int(\dot{x}-f)^2} \tag{3.82}$$

This is exactly the same form as that found by the more rigorous derivation based on a discretized process, on condition that $\alpha = 0$ is used throughout.

3.4.3 Defining equilibrium

Given a random process $x(t)$ extending over the time window $[0, t_{\text{obs}}]$, and given the probability $\mathscr{P}[x]$ to observe a full time realization of this process over $[0, t_{\text{obs}}]$ we define

$$\varSigma[x] = \ln \frac{\mathscr{P}[x]}{\mathscr{P}[x^{\mathrm{R}}]} \tag{3.83}$$

where the time-reversed trajectory is defined by $x^{\mathrm{R}}(t) = x(t_{\text{obs}} - t)$. The average of this quantity,

$$\langle \varSigma[x] \rangle = \int \mathscr{D}x \mathscr{P}[x] \ln \frac{\mathscr{P}[x]}{\mathscr{P}[x^{\mathrm{R}}]} \tag{3.84}$$

is called the total entropy produced over the time window $[0, t_{\text{obs}}]$. Equilibrium is achieved iff, in a stationary state, $\langle \varSigma \rangle = 0$.

For any two probability distributions over some events indexed by i, say p_i and q_i, the quantity $D(p \parallel q) = \sum_i p_i \ln \frac{p_i}{q_i}$ verifies

$$D(p \parallel q) \geq 0 \text{ with equality iff } \forall i,\ p_i = q_i \tag{3.85}$$

Indeed, if the t_i's are in $[0, 1]$ such that $\sum_j t_j = 1$, then the cnvexity of the logarithm tells us that

$$\ln(\sum_j t_j x_j) \geq \sum_j t_j \ln x_j \tag{3.86}$$

and choosing $t_j = p_j$, $x_i = \frac{q_j}{p_j}$, tells us that

$$\ln(\sum_j p_j \frac{q_j}{p_j}) \geq \sum_j p_j \ln \frac{q_j}{p_j} \tag{3.87}$$

and the left hand side vanishes, hence $D(p \parallel q) \geq 0$. The quantity $D(p \parallel q)$ is the Kullback–Leibler divergence (or the relative entropy). It somehow tells us how similar the two distributions p_i and q_i are. It is tempting to see it as the distance from q to p, though $D(p \parallel q)$ is obviously not a distance in any mathematical sense (it's not even symmetric in p and q).

Hence the entropy production $\langle \Sigma[x] \rangle$ measures how the time-reversed process differs from the time-forward one. It's a measure of the length of the arrow of time.

In practice, for a Langevin process (Markov, overdamped) of the form $\dot{\mathbf{r}} = \mathbf{F} + \sqrt{2T}\boldsymbol{\eta}$, where $\mathbf{F} = -\partial_{\mathbf{r}} V + \mathbf{f}$ is the sum of a conservative force $-\partial_{\mathbf{r}} V$ and of a nonconservative one $\mathbf{f}$, we see that

$$\mathscr{P}[\mathbf{r}] = P_{\text{init}}(\mathbf{r}(0))\mathrm{e}^{-\frac{1}{4T}\int_0^{t_{\text{obs}}} \mathrm{d}t(\dot{\mathbf{r}}-\mathbf{F})^2 - \frac{1}{2}\int_0^{t_{\text{obs}}} \mathrm{d}t \partial_{\mathbf{r}}\cdot\mathbf{F}} \tag{3.88}$$

and using $\mathbf{r}^{\mathrm{R}}(t) = \mathbf{r}(t_{\text{obs}} - t)$, we also have

$$\mathscr{P}[\mathbf{r}^{\mathrm{R}}] = P_{\text{fin}}(\mathbf{r}(t_{\text{obs}}))\mathrm{e}^{-\frac{1}{4T}\int_0^{t_{\text{obs}}} \mathrm{d}t(-\dot{\mathbf{r}}-\mathbf{F})^2 - \frac{1}{2}\int_0^{t_{\text{obs}}} \mathrm{d}t \partial_{\mathbf{r}}\cdot\mathbf{F}} \tag{3.89}$$

so that

$$\Sigma[\mathbf{r}] = \frac{1}{T}\int_0^{t_{\text{obs}}} \mathrm{d}t \mathbf{F}\cdot\dot{\mathbf{r}} + \ln\frac{P_{\text{init}}(\mathbf{r}(0))}{P_{\text{fin}}(\mathbf{r}(t_{\text{obs}}))} \tag{3.90}$$

In a stationary state we have $\langle \ln\frac{P_{\text{init}}(\mathbf{r}(0))}{P_{\text{fin}}(\mathbf{r}(t_{\text{obs}}))} \rangle = 0$ and thus

$$\langle \Sigma[\mathbf{r}] \rangle = \frac{1}{T}\int_0^{t_{\text{obs}}} \mathrm{d}t \langle \mathbf{F}\cdot\dot{\mathbf{r}} \rangle = t_{\text{obs}}\langle \mathbf{F}\cdot\dot{\mathbf{r}} \rangle \tag{3.91}$$

has a very transparent interpretation: this is the power injected into the system by $\mathbf{F}$. Of course, if $\mathbf{F} = -p_{\mathbf{r}} V$ is conservative, then the entropy production vanishes owing to $\int \mathrm{d}\dot{t}\cdot\partial_{\mathbf{r}} V = V(t_{\text{obs}}) - V(0)$ which vanishes on average in a steady-state. The corresponding steady-state, without any entropy production, is an equilibrium state. When the non conservative force $\mathbf{f}$ is nonzero, then the entropy production rate is $\langle \dot{\mathbf{r}}\cdot\mathbf{f} \rangle / T$, which is the power of the dissipative forces divided by temperature (exactly what we would expect on the basis of thermodynamics, except that not only we proved the second principle, $\langle \Sigma \rangle \geq 0$, but also we see that it applies much beyond the usual framework of extensive systems).

The definition of equilibrium we have used matches others that may have been encountered, such as the detailed balance condition that states

$$\text{Prob} x(t) = x \to x(t+\Delta t) = x' = \text{Prob} x(t) = x' \to x(t+\Delta t) = x \tag{3.92}$$

3.4.4 An application of the path integral formulation to barrier crossing

A traditional and interesting application of the path integral formulation of the dynamics is barrier crossing. Think again of an overdamped particle $\dot{x} = -V' + \sqrt{2T}\eta$ evolving in a potential landscape with a minimum at its starting point x_0 and a barrier at $x_M > x_0$, such that $\Delta V = V(x_M) - V(x_0) \gg k_{\mathrm{B}} T$. The typical rate τ_K^{-1} of crossing over the barrier is proportional to the fraction of trajectories $x(t)$ sarting from x_0 and reach x_M, namely,

$$\tau_K^{-1} \simeq \int \mathscr{D}x \mathrm{e}^{-\frac{1}{4T}\int(\dot{x}+V')^2 - \frac{1}{2}\int \mathrm{d}t V''} \tag{3.93}$$

ut since we are interested in the $T \ll V$ limit, the second contribution in the exponential can safely be neglected. We now rewrite $(\dot{x} + V')^2 = \dot{x}^2 + V'^2 + 2\dot{x}V'$ and we use that $\int \dot{x}V' = V(x_M) - V(x_0) = \Delta V$. Next, since,

$$\tau_K^{-1} \simeq e^{-\frac{\beta}{2}\Delta V} \int \mathscr{D}x e^{-\frac{1}{4T}\int(\dot{x}^2 + V'^2)^2} \tag{3.94}$$

we search for the path x_c that minimizes S. This is because the $1/T$ prefactor inside the exponential tells us that out of all paths, the one that minimizes $\int(\dot{x}^2 + V'^2)$ contributes the most and that, eventually,

$$\tau_K^{-1} \simeq e^{-\frac{\beta}{2}\Delta V} \int \mathscr{D}x e^{-\frac{1}{4T}\int(\dot{x}_c^2 + V'^2(x_c))} \tag{3.95}$$

We search for the path x_c:

$$\frac{\delta}{\delta x} \int dt(\dot{x}^2 + V'^2) = 2\ddot{x}\dot{x} + 2V'V''\dot{x} \tag{3.96}$$

or $\ddot{x} = -\frac{d}{dx}(-V'^2)$. We recognize the Hamiltonian equation of motion of a particle in a a potential $-V'^2$ and we know $\dot{x}^2 + V'^2$ is a constant of motion. This number isis zero at the initial time, so that throughout time $\dot{x}^2 + V'^2 = 0$, which tells us that either $\dot{x} = -V'$ or $\dot{x} = V'$. The former solution is the gradient descent one, while the latter is the gradient ascent of interest, so that $\dot{x}_c = V'(x_c)$. We don't need to solve for x_c as a function of time, because our interest goes to $\int(\dot{x}_c^2 + V'^2(x_c)) = 2\int \dot{x}_c V'(x_c) = 2\Delta V$, and the conclusion is

$$\tau_K^{-1} \simeq e^{-\beta\Delta V}, \ \tau_K \sim e^{\beta\Delta V} \tag{3.97}$$

This is the celebrated Arrhenius formula derived by Kramers.