Langevin dynamics and invariant measures of stochatic equations: Cheat sheet

Yann Ollivier

We consider diffusions with drift in $\mathbb{R}^n$ (stochastic differential equations), and sum up the relationship between drift, diffusion matrix, and invariant distributions, using the "right" variables. We find a decomposition into "potential" and "geometric" parts, the latter also expressing non-reversibility.

Physical Langevin-type processes on the pair (position, speed), with noise on the speed only, are very naturally interpreted in this framework.

Diffusion with drift on x. Consider a stochastic process with drift vector field F and noise covariance $\Sigma = (\Sigma^{ij})$,

$$x_{t+\mathrm{d}t} = x_t + F(x_t)\,\mathrm{d}t + \sqrt{2\,\mathrm{d}t}\,\mathcal{N}(0,\Sigma) \tag{1}$$

where F and Σ may depend on x. (We consider this process in the Itô sense, namely, the limit of the direct simulation of (1) with stepsize $\mathrm{d}t \to 0$.) Assume it has an invariant distribution, $\mu(x) = e^{-H(x)}$ (not necessarily normalized). Let us work out the relationships between H, F and Σ.

Proposition 1. *The distribution $\mu = e^{-H(x)}$ is invariant iff (using Einstein notation)*

$$F^i = -\Sigma^{ij}\partial_j H + \partial_j \Sigma^{ij} + R^i \tag{2}$$

where R^i is any vector field satisfying $\partial_i R^i = R^i \partial_i H$.

There are three notable cases:

- $\Sigma = \mathrm{Id}$, $R = 0$ and $F = -\partial H$: the "potential" case, the easiest way to build a process with invariant measure e^{-H}. (It is actually reversible, see Prop. 2.)
- The deterministic case $\Sigma = 0$, $F = R$. The deterministic flow defined by R preserves the measure $\mu = e^{-H}$ if and only if the divergence of R is equal to $R\cdot\partial H$. A typical example is a rotation in a potential $H = \|x\|^2$: then R both preserves volumes and is orthogonal to ∂H (preserves the level sets of H), so that *both* $\partial \cdot R$ and $R \cdot \partial H$ vanish. If R mixes the level sets of H, then it must stretch volumes according to the variations of H. Thus, intuitively R represents the "geometry-preserving" part of the drift F.

- The pure noise case, $F = 0$, produces an invariant measure e^{-H} that satisfies some PDE depending on the spatial variations of Σ. [1] In dimension 1 with $F = 0$ and $\Sigma = \sigma^2$, one solution[2] is $H = 2 \ln \sigma$ with $R = 0$ and $F = 0$, with invariant measure $1/\sigma^2$.

The effects on the invariant measure of the deterministic dynamics R, and of the stochastic process are thus completely decoupled.

In particular, if F and Σ and produce a diffusion with invariant distribution e^{-H}, then a diffusion with invariant distribution $e^{-H-H'}$ is obtained via $F - \Sigma\partial H' - R$ and Σ. This is analogous to the Metropolis–Hastings construction.

Proposition 2. *$\mu = e^{-H(x)}$ is reversible iff it is invariant and $R = 0$.*

Proof.
The generator of the process is the operator L such that

$$\mathbb{E}f(x_{t+\mathrm{d}t}) = f(x_t) + (Lf)\,\mathrm{d}t + o(\mathrm{d}t) \tag{3}$$

By a Taylor expansion of f one finds

$$Lf = F \cdot \nabla f + \mathrm{Tr}(\Sigma\nabla\nabla f) = F^i\partial_i f + \Sigma^{ij}\partial_i\partial_j f \tag{4}$$

Then $\mathbb{E}f(X_t) = e^{tL}f(X_0)$.

The adjoint of L acts on measures via $\int(Lf)\mu = \int f(L^*\mu)$, or

$$\mathrm{law}(x_{t+\mathrm{d}t}) = \mathrm{law}(x_t) + (L^*\mathrm{law}(x_t))\,\mathrm{d}t + o(\mathrm{d}t) \tag{5}$$

and is

$$L^*\mu = -\partial_i(F^i\mu) + \partial_i\partial_j(\Sigma^{ij}\mu) \tag{6}$$

and μ is invariant iff $L^*\mu = 0$.

Reversibility is $\int(Lf)g\mu = \int f(Lg)\mu$, or $L^*(g\mu) = L(g)\mu$.

To prove the statement about invariance, let us compute $L^*\mu$ with $\mu = e^{-H}$. Given F and Σ, define R from the formula. Then $L^*e^{-H} = \partial_i(-F^i\mu+\partial_j(\Sigma^{ij}\mu)) = \partial_i(-F^i\mu+(\partial_j\Sigma^{ij})\mu-\Sigma^{ij}(\partial_j H)\mu) = -\partial_i(R^i\mu)$ namely

$$L^*\mu = \left(-\partial_i R^i + R^i\partial_i H\right)\mu \tag{7}$$

hence the claim in Proposition 1.

[1] Namely $\Sigma^{ij}\partial_i\partial_j H - \Sigma^{ij}\partial_i H\partial_j H + 2\partial_i\Sigma^{ij}\partial_j H = \partial_i\partial_j\Sigma^{ij}$.

[2] A diffusion may have several distinct invariant measures due to infinite-mass measures. For instance, in dimension 1, the "right-moving" process $x_{t+\mathrm{d}t} = x_t + \mathrm{d}t + \sqrt{2\,\mathrm{d}t}\,\mathcal{N}(0,1)$ has reversible measure e^x (with $H = -x$, $R = 0$) but the uniform measure is invariant as well (with $H = 0$, $R = 1$), though not reversible: in the second case, the drift is interpreted as a measure-preserving, non-reversible translation ($R = 1$) while in the first case it is interpreted as a reversible potential. For this same process on a circle, only the uniform measure is left.

To prove the statement about reversibility, observe that $L^*(g\mu) = -(\partial_i g)F^i\mu + 2\partial_i g\partial_j(\Sigma^{ij}\mu) + (\partial_i\partial_j g)\Sigma^{ij}\mu + gL^*\mu$. Consequently, for invariant μ,

$$\begin{aligned} L^*(g\mu) - L(g)\mu &= -2F^i(\partial_i g)\mu + 2\partial_i g\partial_j\Sigma^{ij}\mu - 2(\partial_i g)\Sigma^{ij}(\partial_j H)\mu && (8)\\ &= -2R^i\partial_i g && (9)\end{aligned}$$

□

Process on (x, v)**.** Let us apply the above to the variable (x, v) instead of x. This allows us to build processes that reach an invariant distribution over x, but with noise only on v.

Namely, consider a Hamiltonian $H(x, v)$ (for instance $H(x, v) = U(x) + \|v\|^2/2$) and let us build processes with invariant measure $e^{-H(x,v)}$.

Take $\Sigma = ((0\ 0)\ (0\ \ \mathrm{Id}))$ over (x, v), and denote $R = (R_x, R_v)$. Computing F, the process is

$$\begin{cases} x_{t+\mathrm{d}t} = x_t + R_x\,\mathrm{d}t \\ v_{t+\mathrm{d}t} = v_t - \partial_v H\,\mathrm{d}t + R_v\,\mathrm{d}t + \sqrt{2\,\mathrm{d}t}\,\mathcal{N}(0, \mathrm{Id}) \end{cases} \tag{10}$$

and the constraint on R is

$$\partial_x \cdot R_x + \partial_v \cdot R_v = R_x \cdot \partial_x H + R_v \cdot \partial_v H \tag{11}$$

The well-known Hamiltonian dynamics is the choice

$$R_x = \partial_v H, \qquad R_v = -\partial_x H \tag{12}$$

Then, *both* sides vanish in the constraint on R. (Indeed, the left side vanishes since Hamiltonian dynamics preserve volumes on (x, v), and the right side vanishes as Hamiltonian dynamics preserve H.) If A is any symmetric matrix not depending on (x, v), then $R_x = A\partial_v H$ and $R_v = -A\partial_x H$ will cancel both sides as well.

Another choice is to take fields R_x, R_v that would already satisfy the invariance constraints for x and v alone, namely, $\partial_x \cdot R_x = R_x \cdot \partial_x H$ and likewise for v.

Moreover, since the constraint on R is linear, one can add up these two possibilities. In particular, if R_x satisfies the invariance constraint $\partial_x \cdot R_x = R_x \cdot \partial_x H$ on x alone (possible only if $\partial_x H$ does not depend on v, such as $H = U(x) + \|v\|^2/2$), then the system

$$\begin{cases} x_{t+\mathrm{d}t} = x_t + \partial_v H\,\mathrm{d}t + R_x\,\mathrm{d}t \\ v_{t+\mathrm{d}t} = v_t - \partial_v H\,\mathrm{d}t - \partial_x H\,\mathrm{d}t + \sqrt{2\,\mathrm{d}t}\,\mathcal{N}(0, \mathrm{Id}) \end{cases} \tag{13}$$

has the desired invariant measure $e^{-H(x,v)}$. (In that case v is the deviation from the reference speed R_x, not the speed/momentum.)

Thus basically any stochastic dynamics on x can be lifted to (x, v) without noise on x.

Some degrees of freedom are just "gauges", i.e., redefinitions of the variable v (e.g., the Hamiltonian dynamics with $H = U(x) + \|v - V(x)\|^2/2$ is just a shift on the definition of v, does not actually produce speed $V(x)$ at x with the Hamiltonian choice for R). I don't know if, up to these degrees of freedom, the only possibilities are Hamiltonian + separable.

Physical units. The dynamics of a particle of mass m in a potential $U(x)$, at temperature T with an equilibration time τ, is

$$\begin{cases} x_{t+\mathrm{d}t} = x_t + v\,\mathrm{d}t \\ v_{t+\mathrm{d}t} = v_t - \frac{1}{\tau} v_t\,\mathrm{d}t - \frac{1}{m}\partial_x U\,\mathrm{d}t + \sqrt{\frac{2\,\mathrm{d}t T}{m\tau}}\mathcal{N}(0, \mathrm{Id}) \end{cases} \tag{14}$$

(choice $H = (1/T)(U + m\|v\|^2/2)$ and $\Sigma = T/m\tau$ and $R_x = (T/m)\partial_v H$ and $R_v = -(T/m)\partial_x H$ above).

An interesting choice is $T = 0$, $\tau = m \to 0$ which converges to the gradient descent of U.