From Markov Chains to the Langevin and Fokker-Planck equations

Morten Hjorth-Jensen

Spring 2026

Importance sampling, Fokker-Planck and Langevin equations

A stochastic process is simply a function of two variables, one is the time, the other is a stochastic variable X, defined by specifying

1. the set $\{x\}$ of possible values for X;
2. the probability distribution, $w_X(x)$, over this set, or briefly $w(x)$

The set of values $\{x\}$ for X may be discrete, or continuous. If the set of values is continuous, then $w_X(x)$ is a probability density so that $w_X(x)dx$ is the probability that one finds the stochastic variable X to have values in the range $[x, x + dx]$.

Importance sampling, Fokker-Planck and Langevin equations

An arbitrary number of other stochastic variables may be derived from X. For example, any Y given by a mapping of X, is also a stochastic variable. The mapping may also be time-dependent, that is, the mapping depends on an additional variable t

$$Y_X(t) = f(X, t).$$

The quantity $Y_X(t)$ is called a random function, or, since t often is time, a stochastic process. A stochastic process is a function of two variables, one is the time, the other is a stochastic variable X. Let x be one of the possible values of X then

$$y(t) = f(x, t),$$

is a function of t, called a sample function or realization of the process. In physics one considers the stochastic process to be an ensemble of such sample functions.

Random walk as a Markov chain

Consider a particle on a 1D lattice with spacing a:

$$x_n = na, \qquad n \in \mathbb{Z}.$$

Let $P_n(k)$ be the probability to be at site n after k time steps.

Assume a nearest-neighbor Markov chain:

$$\mathbb{P}(n \to n+1) = p, \qquad \mathbb{P}(n \to n-1) = q, \qquad p+q = 1.$$

Markov property: the next state depends only on the current state.

Transition matrix and Chapman–Kolmogorov

The one-step transition probabilities define a stochastic matrix T:

$$P_n(k+1) = \sum_m T_{nm}\, P_m(k),$$

with (nearest-neighbor)

$$T_{n,n-1} = p, \qquad T_{n,n+1} = q, \qquad T_{n,n} = 0.$$

The Chapman–Kolmogorov property for r steps:

$$T^{(r+s)} = T^{(r)}\, T^{(s)}.$$

This semigroup property underlies the continuum limit.

Discrete master equation

From the transition rule:

$$P_n(k+1) = p\,P_{n-1}(k) + q\,P_{n+1}(k).$$

Rewrite as an increment equation:

$$P_n(k+1) - P_n(k) = p\big(P_{n-1}(k) - P_n(k)\big) + q\big(P_{n+1}(k) - P_n(k)\big).$$

Interpretation:

- gain from neighbors $n \pm 1$,
- loss from leaving site n.

Continuous-time version (optional but standard)

Let steps occur in continuous time with rate γ. Define $P_n(t)$ and assume exponential waiting times (continuous-time Markov chain). Then

$$\frac{dP_n(t)}{dt} = \gamma\Big[pP_{n-1}(t) + qP_{n+1}(t) - (p+q)P_n(t)\Big].$$

Since $p + q = 1$:

$$\boxed{\frac{dP_n}{dt} = \gamma\Big[pP_{n-1} + qP_{n+1} - P_n\Big].}$$

This is the **master equation** for a nearest-neighbor jump process.

From lattice probabilities to a density

Define a smooth probability density $\rho(x,t)$ such that

$$P_n(t) \approx a\,\rho(x,t)\Big|_{x=na}.$$

Assume ρ varies slowly on the scale of a.

We expand

$$\rho(x \pm a, t) = \rho(x,t) \pm a\,\partial_x\rho(x,t) + \frac{a^2}{2}\partial_x^2\rho(x,t) + \mathcal{O}(a^3).$$

Taylor expansion of the master equation

Start from

$$\frac{dP_n}{dt} = \gamma\Big[pP_{n-1} + qP_{n+1} - P_n\Big].$$

Divide by a and use $P_n \approx a\rho(x,t)$:

$$\partial_t\rho(x,t) = \gamma\Big[p\rho(x-a,t) + q\rho(x+a,t) - \rho(x,t)\Big].$$

Insert Taylor expansions:

$$p\rho(x-a) + q\rho(x+a) = (p+q)\rho + (q-p)a\,\partial_x\rho + \frac{a^2}{2}(p+q)\partial_x^2\rho + \mathcal{O}(a^3)$$
$$= \rho + (q-p)a\,\partial_x\rho + \frac{a^2}{2}\partial_x^2\rho + \mathcal{O}(a^3).$$

Hence

$$\partial_t\rho = \gamma\Big[(q-p)a\,\partial_x\rho + \frac{a^2}{2}\partial_x^2\rho\Big] + \mathcal{O}(a^3).$$

Drift–diffusion equation and unbiased diffusion

Define drift velocity and diffusion coefficient

$$v \equiv \gamma(q-p)a, \qquad D \equiv \frac{\gamma a^2}{2}.$$

Then, to leading order,

$$\boxed{\partial_t \rho(x,t) = -v\,\partial_x \rho(x,t) + D\,\partial_x^2 \rho(x,t).}$$

Unbiased random walk: $p = q = \frac{1}{2} \Rightarrow v = 0$:

$$\boxed{\partial_t \rho(x,t) = D\,\partial_x^2 \rho(x,t).}$$

This is the 1D diffusion equation.

Scaling limit (diffusive scaling)

To obtain a nontrivial continuum PDE as $a \to 0$, we keep D finite:

$$D = \frac{\gamma a^2}{2} = \text{fixed}.$$

This is achieved by letting the jump rate scale as

$$\gamma \sim \frac{2D}{a^2}.$$

Interpretation:

- steps become smaller ($a \to 0$),
- jumps become more frequent ($\gamma \to \infty$),
- their combined effect yields finite diffusion.

Continuity equation form

Write diffusion as conservation of probability:

$$\partial_t \rho + \partial_x J = 0.$$

For the diffusion equation $\partial_t \rho = D\partial_x^2 \rho$, choose

$$J(x,t) = -D\,\partial_x \rho(x,t).$$

Then

$$\partial_t \rho + \partial_x(-D\partial_x \rho) = 0 \;\Rightarrow\; \partial_t \rho = D\partial_x^2 \rho.$$

This is Fick's first law: $J = -D\nabla\rho$ in 1D.

Physical interpretation and mean-square displacement

For unbiased diffusion, the second moment grows linearly:

$$\langle x^2(t)\rangle - \langle x(t)\rangle^2 = 2Dt.$$

From the Markov chain viewpoint: after k steps with step length a and time step Δt,

$$\mathrm{Var}(x_k) = ka^2, \qquad t = k\Delta t, \qquad D = \frac{a^2}{2\Delta t}.$$

This matches $2Dt$ and provides an empirical route to estimate D.

Summary

- A nearest-neighbor random walk is a Markov chain with transition probabilities p, q.
- The master equation reads

$$\frac{dP_n}{dt} = \gamma\big[pP_{n-1} + qP_{n+1} - P_n\big].$$

- With $P_n \approx a\rho(na, t)$ and Taylor expansion,

$$\partial_t \rho = -v\partial_x \rho + D\partial_x^2 \rho, \quad v = \gamma(q-p)a, \quad D = \frac{\gamma a^2}{2}.$$

- For $p = q$, one obtains the 1D diffusion equation

$$\partial_t \rho = D\,\partial_x^2 \rho,$$

equivalent to $\partial_t \rho + \partial_x J = 0$ with $J = -D\partial_x \rho$ (Fick's law).

Diffusion equation and Green's functions

For many physical systems initial distributions of a stochastic variable y tend to equilibrium distributions: $w(y,t) \to w_0(y)$ as $t \to \infty$. In equilibrium detailed balance constrains the transition rates

$$W(y \to y')w(y) = W(y' \to y)w_0(y),$$

where $W(y' \to y)$ is the probability, per unit time, that the system changes from a state $|y\rangle$, characterized by the value y for the stochastic variable Y , to a state $|y'\rangle$.
Note that for a system in equilibrium the transition rate $W(y' \to y)$ and the reverse $W(y \to y')$ may be very different.

Finding the solution to the standard diffusion equation

Consider, for instance, a simple system that has only two energy levels $\epsilon_0 = 0$ and $\epsilon_1 = \Delta E$.
For a system governed by the Boltzmann distribution we find (the partition function has been taken out)

$$W(0 \to 1) \exp\{-(\epsilon_0/kT)\} = W(1 \to 0) \exp\{-(\epsilon_1/kT)\}.$$

We get then

$$\frac{W(1 \to 0)}{W(0 \to 1)} = \exp\{-(\Delta E/kT)\},$$

which goes to zero when T tends to zero.

Initalizing

If we assume a discrete set of events, our initial probability distribution function can be given by

$$w_i(0) = \delta_{i,0},$$

and its time-development after a given time step $\Delta t = \epsilon$ is

$$w_i(t) = \sum_j W(j \to i) w_j(t = 0).$$

The continuous analog to $w_i(0)$ is

$$w(\mathbf{x}) \to \delta(\mathbf{x}),$$

where we now have generalized the one-dimensional position x to a generic-dimensional vector $\mathbf{x}$. The Kroenecker δ function is replaced by the δ distribution function $\delta(\mathbf{x})$ at $t = 0$.

Transition probabilities

The transition from a state j to a state i is now replaced by a transition to a state with position $\mathbf{y}$ from a state with position $\mathbf{x}$. The discrete sum of transition probabilities can then be replaced by an integral and we obtain the new distribution at a time $t + \Delta t$ as

$$w(\mathbf{y}, t + \Delta t) = \int W(\mathbf{y}, t + \Delta t | \mathbf{x}, t) w(\mathbf{x}, t) d\mathbf{x},$$

and after m time steps we have

$$w(\mathbf{y}, t + m\Delta t) = \int W(\mathbf{y}, t + m\Delta t | \mathbf{x}, t) w(\mathbf{x}, t) d\mathbf{x}.$$

When equilibrium is reached we have

$$w(\mathbf{y}) = \int W(\mathbf{y} | \mathbf{x}, t) w(\mathbf{x}) d\mathbf{x},$$

that is no time-dependence. Note our change of notation for W

Fourier transform

We can solve the equation for $w(\mathbf{y}, t)$ by making a Fourier transform to momentum space. The PDF $w(\mathbf{x}, t)$ is related to its Fourier transform $\tilde{w}(\mathbf{k}, t)$ through

$$w(\mathbf{x}, t) = \int_{-\infty}^{\infty} d\mathbf{k} \exp\{(i\mathbf{k}\mathbf{x})\} \tilde{w}(\mathbf{k}, t),$$

and using the definition of the δ-function

$$\delta(\mathbf{x}) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\mathbf{k} \exp\{(i\mathbf{k}\mathbf{x})\},$$

we see that

$$\tilde{w}(\mathbf{k}, 0) = 1/2\pi.$$

Fourier transformed Diffusion equation

We can then use the Fourier-transformed diffusion equation

$$\frac{\partial \tilde{w}(\mathbf{k}, t)}{\partial t} = -D\mathbf{k}^2 \tilde{w}(\mathbf{k}, t),$$

with the obvious solution

$$\tilde{w}(\mathbf{k}, t) = \tilde{w}(\mathbf{k}, 0) \exp\{\left[-(D\mathbf{k}^2 t)\right)\} = \frac{1}{2\pi} \exp\{\left[-(D\mathbf{k}^2 t)\right]\}.$$

The solution

With the Fourier transform we obtain

$$w(\mathbf{x},t) = \int_{-\infty}^{\infty} d\mathbf{k}\exp\{[i\mathbf{k}\mathbf{x}]\}\frac{1}{2\pi}\exp\{\left[-(D\mathbf{k}^2 t)\right]\} = \frac{1}{\sqrt{4\pi Dt}}\exp\{\left[-(\mathbf{x}^2/4Dt\right.$$

with the normalization condition

$$\int_{-\infty}^{\infty} w(\mathbf{x},t)d\mathbf{x} = 1.$$

Interpretation

The solution represents the probability of finding our random walker at position $\mathbf{x}$ at time t if the initial distribution was placed at $\mathbf{x} = 0$ at $t = 0$. There is another interesting feature worth observing. The discrete transition probability W itself is given by a binomial distribution. The results from the central limit theorem state that transition probability in the limit $n \to \infty$ converges to the normal distribution. It is then possible to show that

$$W(il-jl, n\epsilon) \to W(\mathbf{y}, t+\Delta t|\mathbf{x}, t) = \frac{1}{\sqrt{4\pi D\Delta t}} \exp\Big\{\Big[-((\mathbf{y}-\mathbf{x})^2/4D\Delta t)\Big]\Big\},$$

and that it satisfies the normalization condition and is itself a solution to the diffusion equation.

Einstein-Smoluchenski-Kolmogorov-Chapman equation I

Let us now assume that we have three PDFs for times $t_0 < t' < t$, that is $w(\mathbf{x}_0, t_0)$, $w(\mathbf{x}', t')$ and $w(\mathbf{x}, t)$. We have then

$$w(\mathbf{x}, t) = \int_{-\infty}^{\infty} W(\mathbf{x}.t|\mathbf{x}'.t')w(\mathbf{x}', t')d\mathbf{x}',$$

and

$$w(\mathbf{x}, t) = \int_{-\infty}^{\infty} W(\mathbf{x}.t|\mathbf{x}_0.t_0)w(\mathbf{x}_0, t_0)d\mathbf{x}_0,$$

and

$$w(\mathbf{x}', t') = \int_{-\infty}^{\infty} W(\mathbf{x}'.t'|\mathbf{x}_0, t_0)w(\mathbf{x}_0, t_0)d\mathbf{x}_0.$$

Einstein-Smoluchenski-Kolmogorov-Chapman II

We can combine these equations and arrive at the famous Einstein-Smoluchenski-Kolmogorov-Chapman (ESKC) relation

$$W(\mathbf{x}t|\mathbf{x}_0 t_0) = \int_{-\infty}^{\infty} W(\mathbf{x}, t|\mathbf{x}', t')W(\mathbf{x}', t'|\mathbf{x}_0, t_0)d\mathbf{x}'.$$

We can replace the spatial dependence with a dependence upon say the velocity (or momentum), that is we have

$$W(\mathbf{v}, t|\mathbf{v}_0, t_0) = \int_{-\infty}^{\infty} W(\mathbf{v}, t|\mathbf{v}', t')W(\mathbf{v}', t'|\mathbf{v}_0, t_0)d\mathbf{x}'.$$

Forward Fokker–Planck Equation

Let $p(\boldsymbol{x}, t)$ be the probability density of a diffusion process $\boldsymbol{x} \in \mathbb{R}^d$.

The forward Fokker–Planck equation reads

$$\partial_t p(\boldsymbol{x}, t) = -\sum_{i=1}^{d} \partial_{x_i} \left[A_i(\boldsymbol{x}, t)\, p(\boldsymbol{x}, t) \right] + \sum_{i,j=1}^{d} \partial_{x_i} \partial_{x_j} \left[D_{ij}(\boldsymbol{x}, t)\, p(\boldsymbol{x}, t) \right],$$

where

- $\boldsymbol{A}(\boldsymbol{x}, t)$ is the drift vector,
- $D_{ij}(\boldsymbol{x}, t)$ is the diffusion tensor.

Operator form

Define the Fokker–Planck operator

$$L^\dagger = -\sum_i \partial_{x_i} A_i(\boldsymbol{x}, t) + \sum_{i,j} \partial_{x_i}\partial_{x_j} D_{ij}(\boldsymbol{x}, t).$$

Then the equation becomes

$$\partial_t p(\boldsymbol{x}, t) = L^\dagger p(\boldsymbol{x}, t).$$

This is a linear parabolic partial differential equation.

Definition of the Green's function

The Green's function (fundamental solution) $G(\boldsymbol{x}, t \mid \boldsymbol{x}_0, t_0)$ is defined by

$$\partial_t G = L_{\boldsymbol{x}}^{\dagger} G, \qquad G(\boldsymbol{x}, t_0 \mid \boldsymbol{x}_0, t_0) = \delta(\boldsymbol{x} - \boldsymbol{x}_0).$$

Physical interpretation:

- G is the transition probability density,
- probability to go from $\boldsymbol{x}_0$ at t_0 to $\boldsymbol{x}$ at t.

Superposition principle

Because the Fokker–Planck equation is linear, the solution for a general initial condition $p(\boldsymbol{x}, t_0)$ is

$$p(\boldsymbol{x}, t) = \int d^d \boldsymbol{x}_0 \; G(\boldsymbol{x}, t \mid \boldsymbol{x}_0, t_0)\, p(\boldsymbol{x}_0, t_0).$$

Thus the Green's function completely characterizes the time evolution.

Chapman–Kolmogorov equation

Markov processes satisfy

$$G(\boldsymbol{x}, t+\Delta t \mid \boldsymbol{x}_0, t_0) = \int d^d\boldsymbol{y}\; G(\boldsymbol{x}, t+\Delta t \mid \boldsymbol{y}, t)\, G(\boldsymbol{y}, t \mid \boldsymbol{x}_0, t_0).$$

This identity encodes the semigroup property of the propagator.

Short-time expansion

For small Δt,

$$G(\boldsymbol{x}, t+\Delta t \mid \boldsymbol{y}, t) = \delta(\boldsymbol{x}-\boldsymbol{y}) + \Delta t\, L_{\boldsymbol{y}}^{\dagger}\delta(\boldsymbol{x}-\boldsymbol{y}) + \mathcal{O}(\Delta t^2).$$

Insert into the Chapman–Kolmogorov equation and expand to first order in Δt.

Emergence of the Fokker–Planck equation

After integration by parts and taking $\Delta t \to 0$, one finds

$$\partial_t G(\boldsymbol{x}, t \mid \boldsymbol{x}_0, t_0) = L_{\boldsymbol{x}}^{\dagger}\, G(\boldsymbol{x}, t \mid \boldsymbol{x}_0, t_0),$$

which is precisely the forward Fokker–Planck equation with a delta-function initial condition.

This establishes G as the fundamental solution.

Formal operator solution

Formally, the Green's function can be written as

$$G(\boldsymbol{x}, t \mid \boldsymbol{x}_0, t_0) = \exp\left[(t - t_0)L^\dagger\right] \delta(\boldsymbol{x} - \boldsymbol{x}_0).$$

For constant drift and diffusion, this reduces to a Gaussian kernel; for general coefficients, it defines a semigroup.

Path-integral / stochastic interpretation

Using the underlying stochastic differential equation,

$$d\boldsymbol{X}_t = \boldsymbol{A}(\boldsymbol{X}_t, t)\, dt + \boldsymbol{\sigma}(\boldsymbol{X}_t, t)\, d\boldsymbol{W}_t, \quad D = \tfrac{1}{2}\boldsymbol{\sigma}\boldsymbol{\sigma}^T,$$

the Green's function satisfies

$$G(\boldsymbol{x}, t \mid \boldsymbol{x}_0, t_0) = \left\langle \delta\big(\boldsymbol{x} - \boldsymbol{X}_t\big) \right\rangle_{\boldsymbol{X}_{t_0} = \boldsymbol{x}_0}.$$

This links the Fokker–Planck equation to stochastic paths.

Key properties of the Green's function

- **Normalization:**
$$\int d^d\boldsymbol{x}\, G(\boldsymbol{x}, t \mid \boldsymbol{x}_0, t_0) = 1.$$
- **Positivity:** $G \geq 0$.
- **Semigroup property:**
$$\int d^d\boldsymbol{y}\; G(\boldsymbol{x}, t \mid \boldsymbol{y}, s)\, G(\boldsymbol{y}, s \mid \boldsymbol{x}_0, t_0) = G(\boldsymbol{x}, t \mid \boldsymbol{x}_0, t_0).$$

Summary

- The forward Fokker–Planck equation governs probability evolution.
- The Green's function is the fundamental solution with delta initial data.
- It is derived via the Chapman–Kolmogorov equation and short-time expansion.
- General solutions follow by convolution with the Green's function.
- The propagator has a clear stochastic and physical interpretation.

Forward Fokker–Planck equation (recap)

For $\boldsymbol{x} \in \mathbb{R}^d$, drift $\boldsymbol{A}(\boldsymbol{x}, t)$ and diffusion tensor $D_{ij}(\boldsymbol{x}, t)$:

$$\partial_t p(\boldsymbol{x}, t) = -\sum_{i=1}^{d} \partial_{x_i} \left[A_i(\boldsymbol{x}, t)\, p(\boldsymbol{x}, t) \right] + \sum_{i,j=1}^{d} \partial_{x_i} \partial_{x_j} \left[D_{ij}(\boldsymbol{x}, t)\, p(\boldsymbol{x}, t) \right].$$

Green's function (propagator) $G(\boldsymbol{x}, t \mid \boldsymbol{x}_0, t_0)$ satisfies

$$\partial_t G = L_{\boldsymbol{x}}^{\dagger} G, \qquad G(\boldsymbol{x}, t_0 \mid \boldsymbol{x}_0, t_0) = \delta(\boldsymbol{x} - \boldsymbol{x}_0),$$

and general solutions follow from

$$p(\boldsymbol{x}, t) = \int d^d \boldsymbol{x}_0 \; G(\boldsymbol{x}, t \mid \boldsymbol{x}_0, t_0)\, p(\boldsymbol{x}_0, t_0).$$

Constant-coefficient FPE (Ornstein-free drift)

Take constant drift $\boldsymbol{A}(\boldsymbol{x},t)=\boldsymbol{a}$ and constant diffusion $D_{ij}(\boldsymbol{x},t)=D_{ij}$ (symmetric, positive definite). Then

$$\partial_t p(\boldsymbol{x},t) = -\boldsymbol{\nabla}\cdot(\boldsymbol{a}\,p) + \sum_{i,j}\partial_{x_i}\partial_{x_j}\big(D_{ij}p\big) = -\boldsymbol{a}\cdot\boldsymbol{\nabla}p + \boldsymbol{\nabla}\cdot\big(D\,\boldsymbol{\nabla}p\big),$$

where D is the diffusion matrix.

Goal: find the Green's function $G(\boldsymbol{x},t\mid\boldsymbol{x}_0,t_0)$.

Shift to a comoving frame

Define the comoving coordinate

$$\boldsymbol{y} = \boldsymbol{x} - \boldsymbol{a}(t - t_0), \qquad \tilde{G}(\boldsymbol{y}, t) \equiv G(\boldsymbol{x}, t \mid \boldsymbol{x}_0, t_0).$$

Then using $\partial_t \boldsymbol{y} = -\boldsymbol{a}$ and $\boldsymbol{\nabla}_{\boldsymbol{x}} = \boldsymbol{\nabla}_{\boldsymbol{y}}$,

$$\partial_t G = \partial_t \tilde{G} - \boldsymbol{a} \cdot \boldsymbol{\nabla}_{\boldsymbol{y}} \tilde{G}.$$

Plugging into the constant-coefficient FPE gives cancellation of the drift term:

$$\partial_t \tilde{G} = \boldsymbol{\nabla}_{\boldsymbol{y}} \cdot \left(D \, \boldsymbol{\nabla}_{\boldsymbol{y}} \tilde{G} \right).$$

So drift converts the problem into pure diffusion in $\boldsymbol{y}$.

Fourier-space solution (anisotropic diffusion)

Solve

$$\partial_t \tilde{G}(\boldsymbol{y}, t) = \boldsymbol{\nabla} \cdot (D \boldsymbol{\nabla} \tilde{G}), \qquad \tilde{G}(\boldsymbol{y}, t_0) = \delta(\boldsymbol{y} - \boldsymbol{x}_0).$$

Fourier transform in $\boldsymbol{y}$:

$$\hat{G}(\boldsymbol{k}, t) = \int d^d \boldsymbol{y}\; e^{-i\boldsymbol{k}\cdot\boldsymbol{y}}\; \tilde{G}(\boldsymbol{y}, t).$$

Then $\boldsymbol{\nabla} \mapsto i\boldsymbol{k}$ implies

$$\partial_t \hat{G}(\boldsymbol{k}, t) = -(\boldsymbol{k}^T D \boldsymbol{k})\, \hat{G}(\boldsymbol{k}, t), \qquad \hat{G}(\boldsymbol{k}, t_0) = e^{-i\boldsymbol{k}\cdot\boldsymbol{x}_0}.$$

Hence

$$\hat{G}(\boldsymbol{k}, t) = \exp\Big(- i\boldsymbol{k}\cdot\boldsymbol{x}_0 - (t - t_0)\, \boldsymbol{k}^T D \boldsymbol{k} \Big).$$

Inverse Fourier transform: Gaussian propagator

Invert the transform (Gaussian integral):

$$\tilde{G}(\boldsymbol{y},t)=\frac{1}{(4\pi(t-t_0))^{d/2}\sqrt{\det D}}\exp\left(-\frac{1}{4(t-t_0)}(\boldsymbol{y}-\boldsymbol{x}_0)^T D^{-1}(\boldsymbol{y}-\boldsymbol{x}_0)\right).$$

Transform back to $\boldsymbol{x}$ using $\boldsymbol{y}=\boldsymbol{x}-\boldsymbol{a}(t-t_0)$:

$$G(\boldsymbol{x},t\mid\boldsymbol{x}_0,t_0)=\frac{1}{(4\pi\Delta t)^{d/2}\sqrt{\det D}}\exp\left(-\frac{1}{4\Delta t}\left(\boldsymbol{x}-\boldsymbol{x}_0-\boldsymbol{a}\Delta t\right)^T D^{-1}\left(\boldsymbol{x}-\right.\right.$$

with $\Delta t=t-t_0>0$.

Mean: $\mathbb{E}[\boldsymbol{X}_t]=\boldsymbol{x}_0+\boldsymbol{a}\Delta t$, Covariance: $\mathrm{Cov}(\boldsymbol{X}_t)=2D\,\Delta t$.

Continuity equation and probability current

Write the FPE as a continuity equation

$$\partial_t p + \boldsymbol{\nabla} \cdot \boldsymbol{J} = 0,$$

with probability current

$$J_i(\boldsymbol{x}, t) = A_i(\boldsymbol{x}, t)\, p(\boldsymbol{x}, t) - \sum_{j=1}^{d} \partial_{x_j}\big(D_{ij}(\boldsymbol{x}, t)\, p(\boldsymbol{x}, t)\big).$$

Integrating over a domain Ω gives

$$\frac{d}{dt}\int_\Omega p\, d^d\boldsymbol{x} = -\int_{\partial\Omega} \boldsymbol{J}\cdot\boldsymbol{n}\, dS,$$

so boundary conditions control probability loss/gain through $\partial\Omega$.

Reflecting vs absorbing boundaries

Reflecting (no-flux) boundary:

$$\boldsymbol{J} \cdot \boldsymbol{n}\big|_{\partial\Omega} = 0.$$

Interpretation: probability is conserved inside Ω; trajectories reflect at the boundary.

Absorbing boundary (killing):

$$p(\boldsymbol{x}, t)\big|_{\partial\Omega} = 0.$$

Interpretation: probability reaching the boundary is removed; total probability in Ω decays.

In both cases, G must satisfy the same boundary condition in $\boldsymbol{x}$.

Absorbing states and survival probability

If $\partial\Omega$ is absorbing, the total probability in Ω is the *survival probability*

$$S(t \mid \boldsymbol{x}_0, t_0) = \int_\Omega d^d\boldsymbol{x}\; G(\boldsymbol{x}, t \mid \boldsymbol{x}_0, t_0).$$

Its decay rate is the boundary flux:

$$\frac{dS}{dt} = -\int_{\partial\Omega} \boldsymbol{J}\cdot\boldsymbol{n}\, dS.$$

The corresponding **first-passage-time density** is

$$f(t \mid \boldsymbol{x}_0, t_0) = -\frac{dS}{dt}.$$

Method of images: 1D absorbing wall (worked example)

To illustrate absorbing boundaries explicitly, consider 1D drift–diffusion on $x > 0$:

$$\partial_t p = -a\,\partial_x p + D\,\partial_x^2 p, \qquad p(0,t) = 0, \qquad p(x,t_0) = \delta(x - x_0),\ x_0 > 0.$$

Free-space Green's function:

$$G_0(x,t \mid x_0,t_0) = \frac{1}{\sqrt{4\pi D\Delta t}} \exp\left[-\frac{(x - x_0 - a\Delta t)^2}{4D\Delta t}\right].$$

Absorbing boundary enforced by images:

$$G(x,t \mid x_0,t_0) = G_0(x,t \mid x_0,t_0) - G_0(x,t \mid -x_0,t_0).$$

Then $G(0,t \mid x_0,t_0) = 0$ holds identically.

Remarks on higher-dimensional absorbing boundaries

In $d > 1$, explicit closed forms are generally available only for special geometries:

- half-space (planar boundary) $\rightarrow$ image methods (with modifications under drift/aniso diffusion),
- sphere/ball $\rightarrow$ eigenfunction expansions,
- general domains $\rightarrow$ spectral methods / numerical PDE / Monte Carlo.

Conceptually, the Green's function in a domain Ω is the **killed transition density**:

$$G_\Omega(\boldsymbol{x}, t \mid \boldsymbol{x}_0, t_0) = \mathbb{E}\left[\delta(\boldsymbol{x} - \boldsymbol{X}_t)\, \mathbf{1}_{\{\tau_{\partial\Omega} > t\}} \;\middle|\; \boldsymbol{X}_{t_0} = \boldsymbol{x}_0\right],$$

where $\tau_{\partial\Omega}$ is the first hitting time of the absorbing boundary.

Summary

- Constant drift and diffusion yield an explicit Gaussian propagator:

$$G(\boldsymbol{x}, t \mid \boldsymbol{x}_0, t_0) \propto \exp\left(-\frac{1}{4\Delta t}(\boldsymbol{x} - \boldsymbol{x}_0 - \boldsymbol{a}\Delta t)^T D^{-1}(\boldsymbol{x} - \boldsymbol{x}_0 - \boldsymbol{a}\Delta t)\right).$$

- Boundary conditions are imposed through the probability current $\boldsymbol{J}$:
  - reflecting: $\boldsymbol{J} \cdot \boldsymbol{n} = 0$,
  - absorbing: $p = 0$ on $\partial\Omega$.
- Absorbing boundaries define survival and first-passage-time distributions via boundary flux.

Langevin equation: physical picture

A Langevin equation models the dynamics of a coarse-grained variable $\boldsymbol{X}_t$:

systematic forces $+$ random kicks from the environment.

Example (overdamped Brownian motion in a medium):

$$\gamma\dot{\boldsymbol{X}}_t = \boldsymbol{F}(\boldsymbol{X}_t, t)+\boldsymbol{\xi}(t), \qquad \langle\xi_i(t)\rangle = 0, \qquad \langle\xi_i(t)\xi_j(t')\rangle = 2\gamma^2 D_{ij}\,\delta(t-t').$$

After dividing by γ, this becomes a drift $+$ noise equation.

Ito SDE in d dimensions

We use the Ito SDE

$$d\boldsymbol{X}_t = \boldsymbol{A}(\boldsymbol{X}_t, t)\, dt + \boldsymbol{\sigma}(\boldsymbol{X}_t, t)\, d\boldsymbol{W}_t,$$

where

- $\boldsymbol{A}(\boldsymbol{x}, t)$ is the drift vector,
- $\boldsymbol{W}_t$ is m-dimensional Wiener process,
- $\boldsymbol{\sigma}(\boldsymbol{x}, t) \in \mathbb{R}^{d\times m}$ sets noise amplitudes.

The diffusion tensor is

$$D(\boldsymbol{x}, t) \equiv \frac{1}{2}\boldsymbol{\sigma}(\boldsymbol{x}, t)\boldsymbol{\sigma}(\boldsymbol{x}, t)^T, \qquad D_{ij} = \frac{1}{2}\sum_{k=1}^{m}\sigma_{ik}\sigma_{jk}.$$

Transition density and what we want

Define the transition probability density (propagator)

$$G(\boldsymbol{x}, t \mid \boldsymbol{x}_0, t_0) = \mathbb{P}\big(\boldsymbol{X}_t \in d^d\boldsymbol{x} \,\big|\, \boldsymbol{X}_{t_0} = \boldsymbol{x}_0\big)/d^d\boldsymbol{x}.$$

For an initial density $p(\boldsymbol{x}, t_0)$, the density at later times is

$$p(\boldsymbol{x}, t) = \int d^d\boldsymbol{x}_0 \; G(\boldsymbol{x}, t \mid \boldsymbol{x}_0, t_0)\, p(\boldsymbol{x}_0, t_0).$$

Goal: derive a PDE for $p(\boldsymbol{x}, t)$ (the Fokker–Planck equation).

Chapman–Kolmogorov (Markov property)

Markovity implies (for $t_0 < s < t$)

$$G(\boldsymbol{x}, t \mid \boldsymbol{x}_0, t_0) = \int d^d\boldsymbol{y}\; G(\boldsymbol{x}, t \mid \boldsymbol{y}, s)\, G(\boldsymbol{y}, s \mid \boldsymbol{x}_0, t_0).$$

This semigroup property is the starting point for a short-time expansion.

Infinitesimal increment from the SDE

Over a short interval Δt,

$$\Delta \boldsymbol{X} \equiv \boldsymbol{X}_{t+\Delta t} - \boldsymbol{X}_t = \boldsymbol{A}(\boldsymbol{X}_t, t)\,\Delta t + \boldsymbol{\sigma}(\boldsymbol{X}_t, t)\,\Delta \boldsymbol{W},$$

with $\Delta \boldsymbol{W} \sim \mathcal{N}(\boldsymbol{0}, \Delta t\, I)$.
Conditional moments (Ito):

$$\mathbb{E}[\Delta X_i \mid \boldsymbol{X}_t = \boldsymbol{x}] = A_i(\boldsymbol{x}, t)\,\Delta t + o(\Delta t),$$

$$\mathbb{E}[\Delta X_i \Delta X_j \mid \boldsymbol{X}_t = \boldsymbol{x}] = (\boldsymbol{\sigma}\boldsymbol{\sigma}^T)_{ij}(\boldsymbol{x}, t)\,\Delta t + o(\Delta t) = 2D_{ij}(\boldsymbol{x}, t)\,\Delta t + o(\Delta t).$$

Higher conditional moments are $o(\Delta t)$ for a diffusion process.

Kramers–Moyal expansion (idea)

For a Markov process with smooth transition density, expand $p(\boldsymbol{x}, t + \Delta t)$ in terms of jump moments:

$$p(\boldsymbol{x}, t + \Delta t) = \int d^d \boldsymbol{y}\; p(\boldsymbol{y}, t)\, G(\boldsymbol{x}, t + \Delta t \mid \boldsymbol{y}, t).$$

Let $\boldsymbol{\xi} = \boldsymbol{x} - \boldsymbol{y}$ and expand around $\boldsymbol{\xi} = 0$:

$$p(\boldsymbol{y}, t) = p(\boldsymbol{x} - \boldsymbol{\xi}, t) = p(\boldsymbol{x}, t) - \xi_i \partial_{x_i} p + \frac{1}{2} \xi_i \xi_j \partial_{x_i} \partial_{x_j} p + \cdots$$

and average powers of $\boldsymbol{\xi}$ using the conditional moments from the SDE.

Derivation: keep terms up to $\mathcal{O}(\Delta t)$

Using the conditional moments:

$$\langle \xi_i \rangle = A_i(\boldsymbol{x}, t)\,\Delta t, \qquad \langle \xi_i \xi_j \rangle = 2D_{ij}(\boldsymbol{x}, t)\,\Delta t,$$

and neglecting higher moments ($o(\Delta t)$), we obtain

$$p(\boldsymbol{x}, t+\Delta t) - p(\boldsymbol{x}, t) = -\partial_{x_i}\big(A_i(\boldsymbol{x}, t)p(\boldsymbol{x}, t)\big)\Delta t + \partial_{x_i}\partial_{x_j}\big(D_{ij}(\boldsymbol{x}, t)p(\boldsymbol{x}, t)\big)\Delta t + o$$

Divide by Δt and take $\Delta t \to 0$.

Forward Fokker–Planck equation (result)

We arrive at the forward Fokker–Planck equation:

$$\boxed{\partial_t p(\boldsymbol{x},t) = -\sum_{i=1}^{d} \partial_{x_i}\Big(A_i(\boldsymbol{x},t)\,p(\boldsymbol{x},t)\Big) + \sum_{i,j=1}^{d} \partial_{x_i}\partial_{x_j}\Big(D_{ij}(\boldsymbol{x},t)\,p(\boldsymbol{x},t)\Big)}$$

Identification:

- Drift A_i comes from the mean increment of ΔX_i.
- Diffusion D_{ij} comes from the covariance of increments.

Backward generator and Ito's formula

For a smooth test function $f(\boldsymbol{x}, t)$, Ito's formula gives

$$df(\boldsymbol{X}_t, t) = \left(\partial_t f + A_i \partial_{x_i} f + D_{ij} \partial_{x_i} \partial_{x_j} f\right) dt + \left(\partial_{x_i} f\right) \sigma_{ik} \, dW_k.$$

Define the (backward) generator

$$(Lf)(\boldsymbol{x}, t) = A_i(\boldsymbol{x}, t) \partial_{x_i} f(\boldsymbol{x}, t) + D_{ij}(\boldsymbol{x}, t) \partial_{x_i} \partial_{x_j} f(\boldsymbol{x}, t).$$

Then $\mathbb{E}[f(\boldsymbol{X}_t, t)]$ evolves according to L.

Adjoint relation: generator vs Fokker–Planck

The Fokker–Planck operator $L^\dagger$ is the adjoint of L:

$$\int d^d\boldsymbol{x}\,(Lf)\,p = \int d^d\boldsymbol{x}\,f\,(L^\dagger p) \quad \text{(up to boundary terms)}.$$

Explicitly,

$$(L^\dagger p)(\boldsymbol{x},t) = -\partial_{x_i}\big(A_i p\big) + \partial_{x_i}\partial_{x_j}\big(D_{ij} p\big),$$

so that $\partial_t p = L^\dagger p$.

This is the clean operator link between Langevin (SDE) and FPE (PDE).

Probability current and boundary conditions

Write the FPE as a continuity equation:

$$\partial_t p + \boldsymbol{\nabla} \cdot \boldsymbol{J} = 0,$$

with current

$$J_i(\boldsymbol{x}, t) = A_i(\boldsymbol{x}, t)\, p(\boldsymbol{x}, t) - \partial_{x_j}\big(D_{ij}(\boldsymbol{x}, t)\, p(\boldsymbol{x}, t)\big).$$

Boundary conditions are formulated in terms of $\boldsymbol{J}$:

- Reflecting: $\boldsymbol{J} \cdot \boldsymbol{n} = 0$ on $\partial\Omega$,
- Absorbing: $p = 0$ on $\partial\Omega$.

Example: overdamped Brownian motion in a potential

Overdamped Langevin equation:

$$\gamma \dot{\boldsymbol{X}}_t = -\boldsymbol{\nabla} U(\boldsymbol{X}_t) + \boldsymbol{\xi}(t), \qquad \langle \xi_i(t)\xi_j(t')\rangle = 2\gamma k_B T\, \delta_{ij}\delta(t-t').$$

Equivalently (Ito SDE),

$$d\boldsymbol{X}_t = -\frac{1}{\gamma}\boldsymbol{\nabla} U(\boldsymbol{X}_t)\, dt + \sqrt{2D}\, d\boldsymbol{W}_t, \qquad D = \frac{k_B T}{\gamma}.$$

Thus

$$\boldsymbol{A}(\boldsymbol{x}) = -\frac{1}{\gamma}\boldsymbol{\nabla} U(\boldsymbol{x}), \qquad D_{ij} = D\,\delta_{ij}.$$

Corresponding Fokker–Planck equation

Insert $\boldsymbol{A}, D$:

$$\partial_t p(\boldsymbol{x}, t) = \boldsymbol{\nabla} \cdot \left(\frac{1}{\gamma} \boldsymbol{\nabla} U(\boldsymbol{x})\, p(\boldsymbol{x}, t) \right) + D\, \nabla^2 p(\boldsymbol{x}, t).$$

In equilibrium (detailed balance), the stationary solution is the Gibbs density

$$p_{\mathrm{eq}}(\boldsymbol{x}) \propto e^{-\beta U(\boldsymbol{x})}, \qquad \beta = (k_B T)^{-1},$$

which satisfies $\boldsymbol{J} = 0$.

Summary

- Start from Ito Langevin dynamics:

$$d\boldsymbol{X}_t = \boldsymbol{A}(\boldsymbol{X}_t, t)\, dt + \boldsymbol{\sigma}(\boldsymbol{X}_t, t)\, d\boldsymbol{W}_t, \quad D = \tfrac{1}{2}\boldsymbol{\sigma}\boldsymbol{\sigma}^T.$$

- Short-time increment statistics imply:

$$\mathbb{E}[\Delta X_i] = A_i \Delta t, \qquad \mathbb{E}[\Delta X_i \Delta X_j] = 2D_{ij}\Delta t.$$

- Keeping terms up to $\mathcal{O}(\Delta t)$ yields the forward FPE:

$$\partial_t p = -\partial_i(A_i p) + \partial_i \partial_j (D_{ij} p).$$

- Operator viewpoint: L (Ito generator) and $L^\dagger$ (Fokker–Planck) are adjoints.

Importance sampling, Fokker-Planck and Langevin equations

We will now derive the Fokker-Planck equation. We start from the ESKC equation

$$W(\mathbf{x}, t|\mathbf{x}_0, t_0) = \int_{-\infty}^{\infty} W(\mathbf{x}, t|\mathbf{x}', t')W(\mathbf{x}', t'|\mathbf{x}_0, t_0)d\mathbf{x}'.$$

Define $s = t' - t_0$, $\tau = t - t'$ and $t - t_0 = s + \tau$. We have then

$$W(\mathbf{x}, s + \tau|\mathbf{x}_0) = \int_{-\infty}^{\infty} W(\mathbf{x}, \tau|\mathbf{x}')W(\mathbf{x}', s|\mathbf{x}_0)d\mathbf{x}'.$$

Importance sampling, Fokker-Planck and Langevin equations

Assume now that τ is very small so that we can make an expansion in terms of a small step xi, with $\mathbf{x}' = \mathbf{x} - \xi$, that is

$$W(\mathbf{x}, s|\mathbf{x}_0) + \frac{\partial W}{\partial s}\tau + O(\tau^2) = \int_{-\infty}^{\infty} W(\mathbf{x}, \tau|\mathbf{x} - \xi)W(\mathbf{x} - \xi, s|\mathbf{x}_0)d\mathbf{x}'.$$

We assume that $W(\mathbf{x}, \tau|\mathbf{x} - \xi)$ takes non-negligible values only when ξ is small. This is just another way of stating the Master equation!!

Importance sampling, Fokker-Planck and Langevin equations

We say thus that $\mathbf{x}$ changes only by a small amount in the time interval τ. This means that we can make a Taylor expansion in terms of ξ, that is we expand

$$W(\mathbf{x},\tau|\mathbf{x}-\xi)W(\mathbf{x}-\xi,s|\mathbf{x}_0) = \sum_{n=0}^{\infty}\frac{(-\xi)^n}{n!}\frac{\partial^n}{\partial x^n}\left[W(\mathbf{x}+\xi,\tau|\mathbf{x})W(\mathbf{x},s|\mathbf{x}_0)\right].$$

Importance sampling, Fokker-Planck and Langevin equations

We can then rewrite the ESKC equation as

$$\frac{\partial W}{\partial s}\tau = -W(\mathbf{x}, s|\mathbf{x}_0) + \sum_{n=0}^{\infty} \frac{(-\xi)^n}{n!} \frac{\partial^n}{\partial x^n} \left[W(\mathbf{x}, s|\mathbf{x}_0) \int_{-\infty}^{\infty} \xi^n W(\mathbf{x} + \xi, \tau|\mathbf{x}) d\xi \right].$$

We have neglected higher powers of τ and have used that for $n = 0$ we get simply $W(\mathbf{x}, s|\mathbf{x}_0)$ due to normalization.

Importance sampling, Fokker-Planck and Langevin equations

We say thus that $\mathbf{x}$ changes only by a small amount in the time interval τ. This means that we can make a Taylor expansion in terms of ξ, that is we expand

$$W(\mathbf{x},\tau|\mathbf{x}-\xi)W(\mathbf{x}-\xi,s|\mathbf{x}_0) = \sum_{n=0}^{\infty}\frac{(-\xi)^n}{n!}\frac{\partial^n}{\partial x^n}\left[W(\mathbf{x}+\xi,\tau|\mathbf{x})W(\mathbf{x},s|\mathbf{x}_0)\right].$$

Importance sampling, Fokker-Planck and Langevin equations

We can then rewrite the ESKC equation as

$$\frac{\partial W(\mathbf{x}, s|\mathbf{x}_0)}{\partial s}\tau = -W(\mathbf{x}, s|\mathbf{x}_0) + \sum_{n=0}^{\infty} \frac{(-\xi)^n}{n!} \frac{\partial^n}{\partial x^n} \left[W(\mathbf{x}, s|\mathbf{x}_0) \int_{-\infty}^{\infty} \xi^n W(\mathbf{x}+\xi, \tau$$

We have neglected higher powers of τ and have used that for $n = 0$ we get simply $W(\mathbf{x}, s|\mathbf{x}_0)$ due to normalization.

Importance sampling, Fokker-Planck and Langevin equations

We simplify the above by introducing the moments

$$M_n = \frac{1}{\tau}\int_{-\infty}^{\infty} \xi^n W(\mathbf{x}+\xi,\tau|\mathbf{x})d\xi = \frac{\langle[\Delta x(\tau)]^n\rangle}{\tau},$$

resulting in

$$\frac{\partial W(\mathbf{x},s|\mathbf{x}_0)}{\partial s} = \sum_{n=1}^{\infty}\frac{(-\xi)^n}{n!}\frac{\partial^n}{\partial x^n}\left[W(\mathbf{x},s|\mathbf{x}_0)M_n\right].$$

Importance sampling, Fokker-Planck and Langevin equations

When $\tau \to 0$ we assume that $\langle [\Delta x(\tau)]^n \rangle \to 0$ more rapidly than τ itself if $n > 2$. When τ is much larger than the standard correlation time of system then M_n for $n > 2$ can normally be neglected. This means that fluctuations become negligible at large time scales.
If we neglect such terms we can rewrite the ESKC equation as

$$\frac{\partial W(\mathbf{x}, s|\mathbf{x}_0)}{\partial s} = -\frac{\partial M_1 W(\mathbf{x}, s|\mathbf{x}_0)}{\partial x} + \frac{1}{2}\frac{\partial^2 M_2 W(\mathbf{x}, s|\mathbf{x}_0)}{\partial x^2}.$$

Importance sampling, Fokker-Planck and Langevin equations

In a more compact form we have

$$\frac{\partial W}{\partial s} = -\frac{\partial M_1 W}{\partial x} + \frac{1}{2}\frac{\partial^2 M_2 W}{\partial x^2},$$

which is the Fokker-Planck equation! It is trivial to replace position with velocity (momentum).

Importance sampling, Fokker-Planck and Langevin equations

Langevin equation

Consider a particle suspended in a liquid. On its path through the liquid it will continuously collide with the liquid molecules. Because on average the particle will collide more often on the front side than on the back side, it will experience a systematic force proportional with its velocity, and directed opposite to its velocity. Besides this systematic force the particle will experience a stochastic force $\mathbf{F}(t)$. The equations of motion are

- $\frac{d\mathbf{r}}{dt} = \mathbf{v}$ and
- $\frac{d\mathbf{v}}{dt} = -\xi\mathbf{v} + \mathbf{F}$.

Importance sampling, Fokker-Planck and Langevin equations

Langevin equation

From hydrodynamics we know that the friction constant ξ is given by

$$\xi = 6\pi\eta a/m$$

where η is the viscosity of the solvent and a is the radius of the particle . Solving the second equation in the previous slide we get

$$\mathbf{v}(t) = \mathbf{v}_0 e^{-\xi t} + \int_0^t d\tau e^{-\xi(t-\tau)}\mathbf{F}(\tau).$$

Importance sampling, Fokker-Planck and Langevin equations

Langevin equation

If we want to get some useful information out of this, we have to average over all possible realizations of $\mathbf{F}(t)$, with the initial velocity as a condition. A useful quantity for example is

$$\langle \mathbf{v}(t) \cdot \mathbf{v}(t) \rangle_{\mathbf{v}_0} = v_0^{-\xi 2t} + 2 \int_0^t d\tau e^{-\xi(2t-\tau)} \mathbf{v}_0 \cdot \langle \mathbf{F}(\tau) \rangle_{\mathbf{v}_0}$$

$$+ \int_0^t d\tau' \int_0^t d\tau e^{-\xi(2t-\tau-\tau')} \langle \mathbf{F}(\tau) \cdot \mathbf{F}(\tau') \rangle_{\mathbf{v}_0}.$$

Importance sampling, Fokker-Planck and Langevin equations

Langevin equation

In order to continue we have to make some assumptions about the conditional averages of the stochastic forces. In view of the chaotic character of the stochastic forces the following assumptions seem to be appropriate

$$\langle \mathbf{F}(t) \rangle = 0,$$

and

$$\langle \mathbf{F}(t) \cdot \mathbf{F}(t') \rangle_{\mathbf{v}_0} = C_{\mathbf{v}_0} \delta(t - t').$$

We omit the subscript $\mathbf{v}_0$, when the quantity of interest turns out to be independent of $\mathbf{v}_0$. Using the last three equations we get

$$\langle \mathbf{v}(t) \cdot \mathbf{v}(t) \rangle_{\mathbf{v}_0} = v_0^2 e^{-2\xi t} + \frac{C_{\mathbf{v}_0}}{2\xi}(1 - e^{-2\xi t}).$$

For large t this should be equal to 3kT/m, from which it follows that

Importance sampling, Fokker-Planck and Langevin equations

Langevin equation

Integrating

$$\mathbf{v}(t) = \mathbf{v}_0 e^{-\xi t} + \int_0^t d\tau e^{-\xi(t-\tau)} \mathbf{F}(\tau),$$

we get

$$\mathbf{r}(t) = \mathbf{r}_0 + \mathbf{v}_0 \frac{1}{\xi}(1 - e^{-\xi t}) + \int_0^t d\tau \int_0^\tau \tau' e^{-\xi(\tau-\tau')} \mathbf{F}(\tau'),$$

from which we calculate the mean square displacement

$$\langle (\mathbf{r}(t) - \mathbf{r}_0)^2 \rangle_{\mathbf{v}_0} = \frac{v_0^2}{\xi}(1 - e^{-\xi t})^2 + \frac{3kT}{m\xi^2}(2\xi t - 3 + 4e^{-\xi t} - e^{-2\xi t}).$$

Importance sampling, Fokker-Planck and Langevin equations

Langevin equation

For very large t this becomes

$$\langle (\mathbf{r}(t) - \mathbf{r}_0)^2 \rangle = \frac{6kT}{m\xi} t$$

from which we get the Einstein relation

$$D = \frac{kT}{m\xi}$$

where we have used $\langle (\mathbf{r}(t) - \mathbf{r}_0)^2 \rangle = 6Dt$.