Lecture 3 – The Fokker Planck equation

Last time we studied the Langevin equation

$$m\frac{dv}{dt} = -\alpha v + R(t) + F_{ext},$$

or in a slightly more general form as a stochastic differential equatio

$$dx = a(x(t);t)\,dt + b(x(t);t)\,dW.$$

Our approach was to essentially assume a realization of the random force $R(t)$, or as in homework 2, of the Wiener process dW, and then we integrated to get $v(t)$.

Note: In the first form $v(t)$ is the stochastic process that is denoted with $x(t)$ in the second form. We also looked at $x(t) = x_0 + \int_0^t v(t')\,dt'$, which is not the same as the general $x(t)$ in the second equation above. Sorry if this causes confusion.

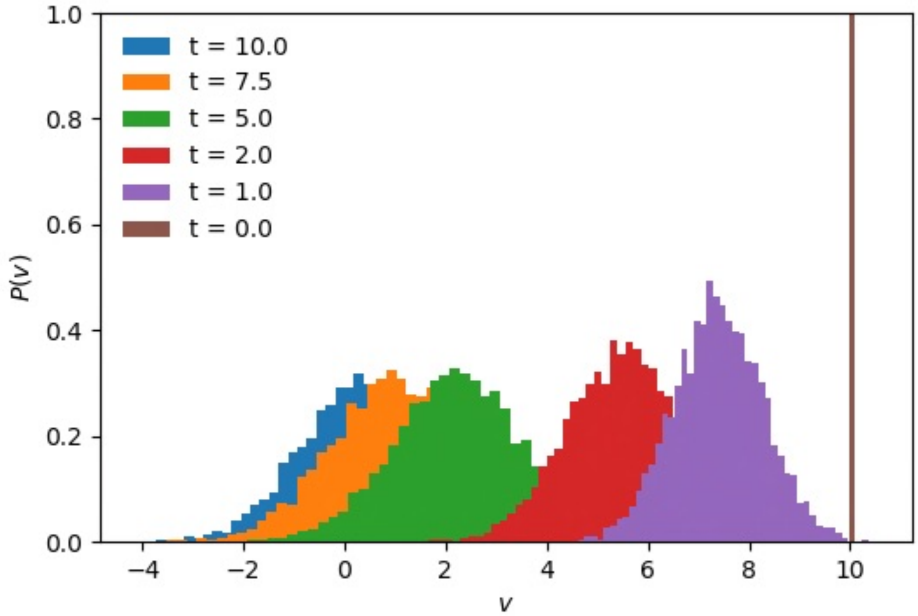
We could only write down such integrals formally, to solve them we took expectation values over the ensemble of $R(t)$ (or $W(t)$) to obtain $\langle v(t)\rangle$ (or $\langle x(t)\rangle$) etc. We can think of $v(t)$ as a stochastic variable that depends on $R(t)$ and by going through the ensemble of $R(t)$ we generated an ensemble of trajectories $v(t)$. The collection of all such trajectories $v(t)$ generates a probability distribution function $P(v,t)$ which is defined such that

$$P(v,t)\,dv = \text{probability that } v(t) \in [v, v+dv] \text{ at } t.$$

This probability density can in fact be calculated with your code for Problem 2 in problem set 1, and we can therefore plot it.

To visualize clearly the evolution, we take $v(0) = v_0 = 10$,

$(m = 1, \gamma = 0.3, T = 1.0)$. Here is how it looks like:



At $t=0$ all trajectories, or all particles, have $v(0) = v_0 = 10$. Therefore

$$P(v, t=0) = \delta(v - v_0)$$

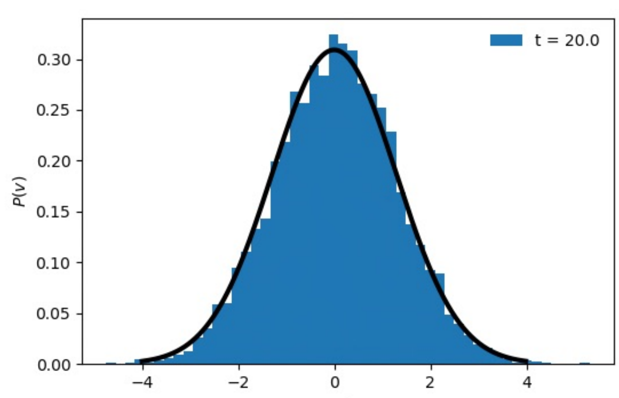
This explains the peak. As time evolves, two things happen: the distribution <u>drifts</u> toward $v=0$, and it <u>diffuses</u> and broadens, eventually reaching a width of $\sigma_v = \sqrt{\frac{T}{2m^2\gamma}} = \sqrt{\frac{5}{3}} \simeq 1.3$ At long times the

cf. p 8 in lecture 2.

distribution reaches a steady state given by the Maxwell velocity distribution which takes the form (with $m=1$ and $T = 2\gamma m k_B T$)

$$P(v) = \left(\frac{\gamma}{\pi T}\right)^{1/2} \exp\left(-\frac{v^2\gamma}{T}\right)$$

Here is a comparison:



Our goal in this lecture is to derive and understand a differential equation that gives this time evolution of $P(v,t)$

We begin with some definitions:

Stochastic processes – some definitions

See for example van Kampen, or D. Arovas

$X(t)$ a stochastic variable, discrete or continuous, that depends on the parameter t which we will take as time.

The time series $\{X(t)\}$ is called a stochastic process

The probability $P(x,t)$ that $X(t) = x$ is defined as

$$P(x,t) = \langle \delta(x - X(t)) \rangle$$

where $\langle \cdot \rangle$ is the ensemble average of the random process. The joint probability distribution

$$P(x_1,t_1;\, x_2,t_2;\, \cdots;\, x_N,t_N) = \langle \delta(x_1 - X(t_1)) \cdots \delta(x_N - X(t_N)) \rangle$$

The conditional probability of observing x_1 at t_1, x_2 at t_2, $\dots$, and x_N at t_N given that we had y_1 at τ_1, y_2 at τ_2, $\dots$, y_m at τ_m is denoted

$$P(x_1,t_1;\cdots;x_N,t_N \mid y_1,\tau_1;\,\cdots,\, y_M,\tau_M)$$

Here we assume $t_1 > t_2 > \dots > t_N > \tau_1 > \dots > \tau_M$

For example, we must have

$$P(x_1,t_1;\, x_0;\, t_0) = P(x_1,t_1 \mid x_0,t_0)\, P(x_0,t_0)$$

Markov processes

In a Markov process the conditional probability for transitioning into a state only depends on the last state, i.e.,

$$P(x_1,t_1;\dots,x_N,t_N \mid y_1,\tau_1;\dots,y_M,\tau_M)$$
$$= P(x_1,t_1;\dots,x_N,t_N \mid y_1,\tau_1)$$

A Markov process therefore lacks memory: only the most recent state determines what happens after that, the further past history does not matter.

By iteration we obtain:

$$P(x_1,t_1;\dots,x_N,t_N) = P(x_1,t_1 \mid x_2,t_2;\dots;x_N,t_N)\,P(x_2,t_2;\dots,x_N,t_N)$$
$$= P(x_1,t_1 \mid x_2,t_2)\,P(x_2,t_2 \mid x_3,t_3\dots x_N,t_N)\,P(x_3,t_3;\dots,x_N,t_N)$$
$$= \dots = P(x_1,t_1 \mid x_2,t_2)\,P(x_2,t_2 \mid x_3,t_3)\cdots P(x_{N-1},t_{N-1} \mid x_N,t_N)\,P(x_N,t_N)$$

$\Rightarrow$ All probabilities are determined by the transition probability $P(x_{n-1},t_{n-1} \mid x_n,t_n)$ and $P(x,t)$.

Some examples of Markov processes:

- Random walk where each step is independent and identically distributed.

Is Brownian motion a Markov process?

Need to be more specific. On time scales larger than τ_{micro} such that $\langle R(t) R(t') \rangle = \Gamma \delta(t-t')$, $v(t)$ as described by the Langevin equation is Markovian.

What about $x(t)$? Not in general, but yes if we observe on time scales large compared with τ_v, the velocity correlation time.

- The Wiener process, which is of course central to Brownian motion, is a Markov progress. Remember that the Wiener process is Gaussian with $\langle w(t) \rangle = w_0$ and $\langle [w(t)-w_0]^2 \rangle = t - t_0$. We can therefore write the conditional probability as

$$p(w,t \mid w_0,t_0) = \frac{1}{\sqrt{2\pi(t-t_0)}} \exp\left(- \frac{(w-w_0)^2}{2(t-t_0)} \right)$$

Note that this implies an <u>independence of increments</u>. Namely, using the Markov property

$$P(w_n,t_n; w_{n-1},t_{n-1}; \dots; w_0,t_0) = \prod_{i=0}^{n-1} P(w_{i+1},t_{i+1} \mid w_i,t_i)\, P(w_0,t_0)$$

$$= \prod_{i=0}^{n-1} \frac{1}{\sqrt{2\pi(t_{i+1}-t_i)}} \exp\left(- \frac{(w_{i+1}-w_i)^2}{2(t_{i+1}-t_i)} \right) P(w_0,t_0)$$

$$= \prod_{i=0}^{n-1} \left\{ \frac{1}{\sqrt{2\pi\Delta t_i}} \exp\left(- \frac{\Delta w_i^2}{2\Delta t_i} \right) \right\} P(w_0,t_0) = \prod_{i=0}^{n-1} P(\Delta w_i, \Delta t_i)\, P(w_0,t_0)$$

with $\Delta t_i = t_{i+1} - t_i$ and $\Delta W_i = W_{i+1} - W_i$ and $P(\Delta W_i, \Delta t_i)$ the probability of having increment ΔW_i in time interval Δt_i. This makes it convenient to simulate the Wiener process, since we can pick a fixed time step dt and then the increments ΔW_i are normally distributed with mean zero and variance dt, and all ΔW_i are independent. This fact is used to generate sample paths of the Wiener process in problem 1.2.

Chapman-Kolmogorov equation

Markov processes satisfy a relation that is central to our discussion of both the Fokker-Planck equation and the master equation. Any stochastic process, by the rules of probability theory, satisfies

$$P(x_1, t_1) = \int dx_2\, P(x_1, t_1; x_2, t_2)$$

$$= \int dx_2\, P(x_1, t_1 | x_2, t_2)\, P(x_2, t_2)$$

and similarily

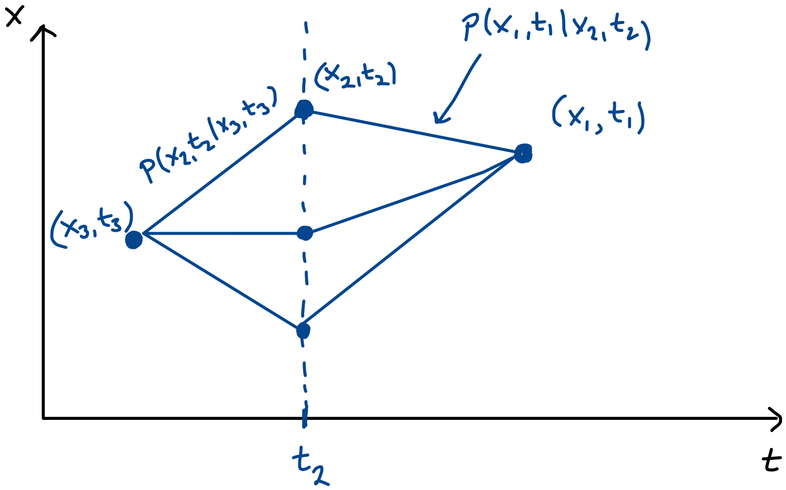
$$P(x_1, t_1 | x_3, t_3) = \int dx_2\, P(x_1, t_1 | x_2, t_2; x_3, t_3)\, P(x_2, t_2 | x_3, t_3)$$

For a Markov process, this implies

$$\boxed{P(x_1, t_1 | x_3, t_3) = \int dx_2\, P(x_1, t_1 | x_2, t_2)\, P(x_2, t_2 | x_3, t_3)}$$

Chapman-Kolmogorov relation

Pictorially, this relation can be visualized as follows:



(Note the similarity (and differences) to a path integral).

Now we want to use the Chapman-Kolmogorov relation to derive a differential equation for $P(x,t|x_0,t_0)$. Note that in our discussion of the probability distribution for the velocities $P(v,t)$ in the Langevin equation, what we really meant was $P(v,t|v_0,0)$ since all trajectories started with velocity v_0 at $t=0$. Anyway, imagine we know $P(x,t)$ and want to find $P(x,t+\delta t)$. For this we write the Chapman-Kolmogorov relation as

$$P(x,t+\delta t|x_0,t_0) = \int dx' \, P(x,t+\delta t|x' t) \, P(x',t|x_0,t_0) \qquad (\#)$$

To make progress we need to make some assumptions about the stochastic process $x(t)$. In fact, we want to assume it is continous in the following sense. Define

$$\delta x(t) = x(t+\delta t) - x(t)$$

and assume
$$\langle \delta x(t) \rangle = A(x)\,\delta t$$
$$\langle (\delta x(t))^2 \rangle = B(x)\,\delta t$$
$$\langle (\delta x(t))^n \rangle = O((\delta t)^2) \qquad n>2$$

The Wiener process, for example, satisfies this with $A = 0$, $B = 1$.

These assumptions allow us to expand the transition probabilities $P(x, t+\delta t \mid x', t)$ in the integral as follows:

$$P(x, t+\delta t \mid x', t) = \langle \delta(x - x' - \delta x(t)) \rangle$$

$$= \langle \delta(x-x') + \delta^{(1)}(x-x')\, \delta x(t) + \tfrac{1}{2} \delta^{(2)}(x-x') (\delta x(t))^2 + \cdots$$

↑ Taylor expansion

$$= \delta(x-x') + \langle \delta x(t) \rangle\, \delta^{(1)}(x-x') + \tfrac{1}{2} \langle (\delta x(t))^2 \rangle\, \delta^{(2)}(x-x') + \cdots$$

where $\delta^{(n)}(x-x') = \frac{\partial^n}{\partial x'^n} \delta(x-x')$. Using our assumptions about the expectation values this becomes

$$= \delta(x-x') + A(x')\, \delta^{(1)}(x-x')\, \delta t + \tfrac{1}{2} B(x')\, \delta^{2}(x-x')\, \delta t + \cdots$$

Plugging this into # and using

$$\int_{-\infty}^{\infty} f(x')\, \delta^{(n)}(x-x')\, dx' = (-1)^n f^{(n)}(x)$$

which you can prove by integrating by parts, we get (suppressing the conditional x_0, t_0)

$$P(x, t+\delta t) = P(x,t) - \frac{\partial}{\partial x}\left[A(x)\, P(x,t) \right] \delta t + \frac{1}{2} \frac{\partial^2}{\partial x^2} \left[B(x)\, P(x,t) \right] \delta t + O((\delta t)^2)$$

Taking the limit $\delta t \to 0$

Fokker-Planck equation

$$\boxed{\frac{\partial P(x,t)}{\partial t} = -\frac{\partial}{\partial x}\left[A(x)\, P(x,t) \right] + \frac{1}{2} \frac{\partial^2}{\partial x^2}\left[B(x)\, P(x,t) \right]}$$

This can be put in the form of a continuity equation

$$\frac{\partial P}{\partial t} + \frac{\partial}{\partial x} j = 0$$

where $\quad j = A(x)\, P - \frac{1}{2} \frac{\partial}{\partial x} [B(x)\, P]$

From this we interpret the first term as drift and the second as diffusion.

Langevin equation

For a stochastic process $x(t)$ satisfying

$$dx = a(x)\, dt + b(x)\, dW$$

we have

$$\langle dx \rangle = a(x)\, dt$$

$$\langle (dx)^2 \rangle = (b(x))^2 \langle (dW)^2 \rangle = (b(x))^2\, dt$$

where we used the Ito prescription. So in this case

$A(x) = a(x)$ is the drift term

$B(x) = b^2(x)$ is the diffusion term.

This of course tells us that for the original Langevin equation

$$m \frac{dv}{dt} = -\alpha v + R(t)$$

or

$$\frac{dv}{dt} = -\gamma v + \frac{\sqrt{\Gamma}}{m} \eta(t)$$

that

$$A(v) = -\gamma v$$

$$B(v) = \Gamma / m^2$$

So the Fokker-Planck equation for the Langevin equation is

$$\boxed{\frac{\partial P}{\partial t} = \gamma \frac{\partial}{\partial v} (v P) + \frac{\Gamma}{2m^2} \frac{\partial^2 P}{\partial v^2}}$$

Alternatively, we can obtain the coefficients A and B from the Langevin equation by direct integration

$$m \int_t^{t+\delta t} \frac{dv}{dt'} dt' = -\alpha \int_t^{t+\delta t} v(t')\, dt' + \int_t^{t+\delta t} R(t')\, dt'$$

$$v(t+\delta t) - v(t) = -\gamma v(t)\, \delta t + \frac{1}{m} \int_t^{t+\delta t} R(t')\, dt'$$

Taking the ensemble average and using $\langle R \rangle = 0$

$$\langle \delta v(t) \rangle = -\gamma\, v(t)\, \delta t \quad \text{or} \quad A(v) = -\gamma v$$

Similarily we get the second moment

$$\langle (\delta v(t))^2 \rangle = \left\langle \left(-\gamma v(t) \delta t + \frac{1}{m} \int_t^{t+\delta t} R(t')\, dt' \right)^2 \right\rangle$$

$$= \frac{1}{m^2} \int_t^{t+\delta t} dt_1 \int_t^{t+\delta t} dt_2 \underbrace{\langle R(t_1) R(t_2) \rangle}_{\Gamma \delta(t_1 - t_2)} + O((\delta t)^2)$$

using $v(t)$ and $R(t)$ uncorrelated and $\langle R \rangle = 0$

$$= \frac{\Gamma}{m^2} \delta t \quad \Rightarrow \quad B(v) = \frac{\Gamma}{m^2}$$

Stationary solution

We saw in our numerical solution of the Langevin equation that at long times $P(v,t)$ evolved into the stationary solution

$$P_s(v) = P_{eq}(v) = N e^{-\frac{mv^2}{2k_B T}} = N \exp\left(-\frac{m^2 \gamma}{\Gamma} v^2\right)$$

with N a normalization constant. Let's check we reproduce this. In a steady state

$$\frac{\partial P}{\partial t} = 0$$

and

$$\gamma \frac{\partial}{\partial v}(v P_{eq}) + \frac{\Gamma}{2m^2} \frac{\partial^2 P_{eq}}{\partial v^2} = \left\{ \gamma + \frac{\Gamma}{2m^2} \frac{-2m^2\gamma}{\Gamma} \right\} \frac{\partial}{\partial v}(v P_{eq}) = 0 \quad ✓$$

$$\hookrightarrow \frac{\partial}{\partial v}\left(-\frac{2m^2\gamma}{\Gamma} v\, P_{eq} \right)$$

Why is $\langle (\delta v)^n \rangle = O((\delta t)^2)$?

In order for our general derivation of the Fokker-Planck equation to hold, we needed $\langle (\delta v)^n \rangle$ for $n>2$ to be $O((\delta t)^2)$. Why is this true? Because we assumed Gaussian white noise

Gaussian white noise

It is generally not enough to specify only the mean $\langle \eta(t) \rangle$ and the variance $\langle \eta(t)\eta(t') \rangle$ to fully specify a probability distribution function. For Gaussian noise, however, it is, since all higher order correlation function are determined by these two. For example

$$\langle \eta_1, \dots, \eta_{2N} \rangle = \sum_{\text{all contractions}} \prod \langle \eta_i \eta_j \rangle$$ (Wick's theorem)

For example:

$$\langle \eta_1 \eta_2 \eta_3 \eta_4 \rangle = \langle \eta_1 \eta_2 \rangle \langle \eta_3 \eta_4 \rangle + \langle \eta_1 \eta_3 \rangle \langle \eta_2 \eta_4 \rangle + \langle \eta_1 \eta_4 \rangle \langle \eta_2 \eta_3 \rangle \quad \text{etc}$$

This is true for all Gaussian noise.

Gaussian <u>white</u> noise is further specified by having

$$\langle \eta(t)\eta(t') \rangle \propto \delta(t-t')$$

and therefore a constant power spectrum

For the Langevin equation, this means in particular that:

$$\langle (\delta v)^{2n} \rangle \sim (\delta t)^n \quad \text{and} \quad \langle (\delta v)^{2n+1} \rangle \sim (\delta t)^{n+1}$$

Spectral properties

A stationary process satisfies time translation invariance:

$$P(x_1,t_1;x_2,t_2;\cdots;x_N,t_N) = P(x_1,t_1+\tau;x_2,t_2+\tau;\cdots;x_N,t_N+\tau)$$

for all τ and N.

In particular

$$P(x,t) = P(x)$$

$$P(x_1,t_1;x_2,t_2) = P(x_1,t_1-t_2;x_2,0)$$

Suppose we are interested in knowing the power spectrum of a stationary process.

Define
$$\hat{x}_T(\omega) = \int_{-T/2}^{T/2} dt\, x(t)\, e^{i\omega t}$$

The spectral function

$$S_T(\omega) = \frac{1}{T}\langle|\hat{x}_T(\omega)|^2\rangle$$

Does the limit $S(\omega) \equiv S_T(\omega)$ exist?

$$\langle|\hat{x}_T(\omega)|^2\rangle = \int_{-T/2}^{T/2} dt \int_{-T/2}^{T/2} dt'\, \underbrace{\langle x(t)\,x(t')\rangle}_{C_x(t-t')\ \text{(stationary)}}\, e^{-i\omega(t-t')}$$

$$\left.\begin{array}{l}\tau = t-t' \\ T = \frac{1}{2}(t+t')\end{array}\right\} = \int_{-T}^{T} d\tau\, e^{-i\omega\tau}\, C(\tau)\,(T-|\tau|)$$

and therefore

$$S(\omega) = \lim_{T\to\infty} \int_{-\infty}^{\infty} d\tau\, e^{-i\omega\tau} C(\tau) \left(1 - \frac{|\tau|}{T}\right) \theta(T - |\tau|)$$

$$= \int_{-\infty}^{\infty} d\tau\, e^{-i\omega\tau} C(\tau)$$

Proving convergence nontrivial part of derivation, see Arovas p.50 or van Kampen

This is the Wiener-Khinchin theorem

$$\boxed{S(\omega) = \int_{-\infty}^{\infty} d\tau\, e^{-i\omega\tau} \langle x(t+\tau)\, x(t) \rangle}$$

<u>Examples</u>

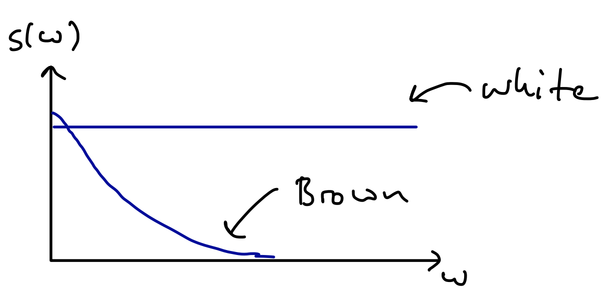
* White noise : $C(\tau) = \Gamma\delta(\tau) \quad \Rightarrow \quad S(\omega) = \Gamma$

* Brown noise : velocity correlator of Brownian motion

$$C(\tau) = A e^{-\gamma|\tau|}$$

$$\Rightarrow \quad S(\omega) = A \int_{-\infty}^{\infty} e^{-i\omega\tau - \gamma|\tau|}\, d\tau$$

$$= A \left[\frac{1}{i\omega + \gamma} + \frac{1}{-i\omega + \gamma} \right] = A\, \frac{2\gamma}{\omega^2 + \gamma^2}$$



Fokker-Planck equation and change of variables in a stochastic diff. eq.

Here is an alternative way of obtaining the Fokker-Planck equation.

Suppose $x(t)$ is a stochastic variable that satisfies the equation

$$dx = a(x)\,dt + b(x)\,dW$$

Now assume we are interested in some function of x, say $f(x)$.

How do we make a change of variables in differential form? Namely, what is df? Using the Ito description

$$\begin{aligned} df[x(t)] &= f[x(t)+dx(t)] - f[x(t)] \\ &= f'[x(t)]\,dx(t) + \tfrac{1}{2} f''[x(t)]\,[dx(t)]^2 + \ldots \\ &= f'[x(t)]\left(a(x)\,dt + b(x)\,dW\right) + \tfrac{1}{2} f''(x(t))\,(b(x))^2 (dW)^2 + O(dt^{3/2}) \end{aligned}$$

And using $(dW)^2 = dt$ we obtain Ito's formula:

$$\boxed{df[x(t)] = \left\{ a(x) f'(x) + \tfrac{1}{2}[b(x)]^2 f''(x) \right\} dt + f'(x)\,b(x)\,dW}$$

Note that this is very different from the usual change of variables in regular differential equations.

We now obtain the time evolution

$$\frac{d}{dt}\langle f[x(t)]\rangle = \frac{\langle df[x(t)]\rangle}{dt} = \left\langle a(x)\frac{\partial f}{\partial x} + \frac{1}{2}(b(x))^2 \frac{\partial^2 f}{\partial x^2} \right\rangle \qquad (\#)$$

since $\langle dW\rangle = 0$. But we can also write

$$\langle f[x(t)]\rangle = \int dx\, f(x)\, P(x,t|x_0,t_0)$$

and therefore (noting on the left $x(t)$ is a stochastic variable, on the right x is its value)

$$\frac{d\langle f[x(t)]\rangle}{dt} = \int dx\, f(x)\, \partial_t P(x,t|x_0,t_0)$$

which according to #

$$= \int dx \left(a(x) \frac{\partial f}{\partial x} + \frac{1}{2}(b(x))^2 \frac{\partial^2 f}{\partial x^2} \right) P(x,t|x_0,t_0)$$

$$= \int dx \; f(x) \left\{ -\frac{\partial}{\partial x}\left[a(x) P(x,t|x_0,t_0)\right] + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left[(b(x))^2 P(x,t|x_0,t_0)\right] \right\}$$

Since $f(x)$ is arbitrary, this means that (dropping the conditional x_0, t_0 in the notation)

$$\boxed{\frac{\partial P}{\partial t} = -\frac{\partial}{\partial x}\left(a(x) P(x,t)\right) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left[(b(x))^2 P(x,t)\right]}$$

which is again the same Fokker-Planck equation we derived earlier. Note that using the Ito prescription was essential to get this form; had we used the Stratonovich prescription the resulting Fokker-Planck equation would look (slightly) different.