

## The CTRW propagator

In what follows we consider 1D CTRW with zero mean and finite variance of the single jump displacement. We are interested in the long time limit of the PDF of a particle to be in position  $x$  at time  $t$ ,  $p(x, t)$ . As we already mentioned we may express the PDF as

$$p(x, t) = \sum_{n=0}^{\infty} W(n, t) p_n(x).$$

In the long time limit the main contribution is coming from trajectories involving large number of jumps. The total displacement is  $x = \sum_{i=1}^n \Delta x_i$ . Since the random variables  $\Delta x_i$ s are i.i.d. we expect that for large  $n$ ,  $p_n(x)$  will be a Gaussian. Hence, we write the displacement's PDF as:

$$p_n(x) \rightarrow \frac{1}{\sqrt{2\pi n\sigma^2}} e^{-\frac{x^2}{2n\sigma^2}} \equiv G_n(x),$$

which is the Gaussian CLT. We found (in our discussion of renewal processes) that for a diverging mean waiting time  $W(n, t)$  is given by

$$W(n, t) \sim \frac{t}{n\alpha} l_{\alpha, An, 1}(t) = \frac{t}{\alpha n^{1+1/\alpha} A^{1/\alpha}} l_{\alpha, 1, 1}\left(\frac{t}{[An]^{1/\alpha}}\right). \quad (1)$$

Hence, we expect

$$\begin{aligned} p(x, t) &\sim \int_0^{\infty} W(n, t) G_n(x) dn \\ &= \int_0^{\infty} \frac{t}{n\alpha} l_{\alpha, An, 1}(t) \frac{1}{\sqrt{2\pi n\sigma^2}} e^{-\frac{x^2}{2n\sigma^2}} dn \end{aligned} \quad (2)$$

The above derivation used heuristic arguments. In what follows we will use an alternative approach. We start with the Weiss-Montroll equation and expand the Fourier transform of the single step PDF in  $k$ :  $\tilde{f}(\mathbf{k}) = 1 - \sigma^2 k^2/2 + O(k^4)$ , and we also expand the Laplace transform of the waiting time PDF in  $s$ :  $\hat{\psi}(s) \sim 1 - As^\alpha$ . Substituting these expansions in the the Weiss-Montroll equation one finds:

$$\begin{aligned} \hat{\tilde{P}}(\mathbf{k}, s) &= \frac{1 - \hat{\psi}(s)}{s} \frac{1}{1 - \hat{\psi}(s) \tilde{f}(\mathbf{k})} \\ &\sim \frac{As^\alpha}{s} \frac{1}{1 - (1 - As^\alpha)(1 - \sigma^2 k^2/2 + O(k^4))} \\ &= \frac{As^{\alpha-1}}{\sigma^2 k^2/2 + As^\alpha} + O(s^{2\alpha-1} k^2). \end{aligned} \quad (3)$$

Note that,

$$\frac{As^{\alpha-1}}{\sigma^2 k^2/2 + As^\alpha} = \int_0^\infty As^{\alpha-1} e^{-Ans^\alpha} e^{-n\sigma^2 k^2/2} dn. \quad (4)$$

The spatial part,  $e^{-n\sigma^2 k^2/2}$ , corresponds to a Gaussian PDF with a variance equal to  $n$  times the variance of a single step. The temporal part corresponds to the one sided Levy distribution that we derived in the discussion of renewal processes. In the expansion above we assumed that the long time limit propagator of the CTRW depends only on the variance of the single step and not on the higher moments. Eq. (4) implies that this assumption is identical to the assumptions we used in deriving eq. (2).

Why aren't the higher moments of  $f$  important? Returning to eq. (3) and expanding in  $k$  we write

$$\begin{aligned} \widehat{p}(\mathbf{k}, s) &\sim_{s \rightarrow 0} \frac{As^\alpha}{s} \frac{1}{1 - (1 - As^\alpha)(1 - \sigma^2 k^2 + O(k^4))} \\ &= \frac{1}{s} \frac{1}{1 + \frac{\sigma^2 k^2}{As^\alpha} - \sigma^2 k^2 - O\left(\frac{k^4}{As^\alpha}\right) + O(k^4)} \\ &= \frac{1}{s} \left[ \frac{1}{1 + \frac{\sigma^2 k^2}{2As^\alpha} - \sigma^2 k^2 + O\left(\frac{k^4}{As^\alpha}\right)} \right] \\ &= \frac{1}{s} \left[ \sum_{n=0}^\infty \left( -\frac{\sigma^2 k^2}{2As^\alpha} + \sigma^2 k^2 + O\left(\frac{k^4}{As^\alpha}\right) \right)^n \right]. \end{aligned}$$

We see that there are two contributions to the second moment coming from the  $k^2$  terms; however, in the long time limit (equivalent to the small  $s$  limit) one of them is larger and this is the one we kept. Similarly for  $k^4$  we see that there are two contributions and the one coming from  $\sigma^4$  is much larger in the limit of small  $s$ . It is easy to see that the same argument holds for all higher moments, and therefore, the higher moments of  $f$  contribute only to the non-leading-order terms and are negligible at the long time limit that we discussed.

## The Green Function

In order to find the Green function we start with the inverse Fourier transform of eq. (3).

$$\widehat{p}(x, s) = \frac{1}{2\pi s} \int_{-\infty}^\infty e^{-ikx} \frac{1}{1 + \frac{\sigma^2 k^2}{2As^\alpha}} dk \quad (5)$$

$$= \frac{\sqrt{As^{\alpha/2-1}}}{\sigma\sqrt{2}} e^{-\frac{\sqrt{2A}}{\sigma}|x|s^{\alpha/2}} \quad (6)$$

$$= \left( -\frac{1}{\alpha|x|} \right) \frac{d}{ds} \left( e^{-\frac{\sqrt{2A}}{\sigma}|x|s^{\alpha/2}} \right). \quad (7)$$

Remembering that the exponent in the parenthesis above is the Laplace conjugate of the one sided Levy PDF we find,

$$p(x, t) = \frac{t}{\alpha |x|} l_{\alpha/2, \frac{\sqrt{2A}}{\sigma} |x|, 1}(t) = \frac{t\sigma^{2/\alpha}}{\alpha (2A)^{1/\alpha} |x|^{1+2/\alpha}} l_{\alpha/2, 1, 1}\left(\frac{t\sigma^{2/\alpha}}{|x|^{2/\alpha} (2A)^{1/\alpha}}\right). \quad (8)$$

(The last equality was derived using the relation

$$l_{\alpha, a, 1}(t) = \frac{1}{a^{1/\alpha}} l_{\alpha, 1, 1}\left(\frac{t}{a^{1/\alpha}}\right). \quad (9)$$

$$l_{\alpha, a, 1}(t) = \frac{1}{2\pi} \int_{c-i\infty}^{c+i\infty} e^{-as^\alpha} e^{st} ds \quad (10)$$

$$u^\alpha = as^\alpha \rightarrow u = sa^{1/\alpha} \rightarrow du = dsa^{1/\alpha}$$

$$\frac{1}{2\pi} \int_{c-i\infty}^{c+i\infty} e^{-as^\alpha} e^{st} ds = \frac{1}{2\pi} \int_{\bar{c}-i\infty}^{\bar{c}+i\infty} e^{-u^\alpha} e^{u(t/a^{1/\alpha})} \frac{du}{a^{1/\alpha}}$$

).

It is useful to rewrite equation (8) using a scaling function,

$$\begin{aligned} p(x, t) &= \frac{\sigma^{2/\alpha}}{\alpha (2A)^{1/\alpha} t^{\alpha/2}} \left(\frac{t}{|x|^{2/\alpha}}\right)^{1+\frac{\alpha}{2}} l_{\alpha/2, 1, 1}\left(\frac{t}{|x|^{2/\alpha}} \frac{\sigma^{2/\alpha}}{(2A)^{1/\alpha}}\right) \quad (11) \\ &= \frac{1}{t^{\alpha/2}} f\left(\frac{t}{|x|^{2/\alpha}}\right). \end{aligned}$$

## The Fokker-Planck Equation (1st derivation)

The Fokker-Planck equation can be thought of as an equation describing the evolution of the conditional probability of a stationary process. The conditional probability,

$$P_n(y_n, t_n | y_{n-1}, t_{n-1}; y_{n-2}, t_{n-2}; \dots y_1, t_1), \quad (12)$$

is the probability (in the ensemble sense) that the variable  $y(t)$  lies between  $y_n$  and  $y_n + dy_n$  at time  $t_n$ , (where  $(t_1 < t_2 \dots t_{n-1} < t_n)$ ) given that it took the values  $y_1$  at time  $t_1$ ,  $y_2$  at time  $t_2$  ... and  $y_{n-1}$  at time  $t_{n-1}$ . The probability  $p_n$  is related to the conditional probability through

$$\begin{aligned} & p_n(y_n, t_n; y_{n-1}, t_{n-1}; y_{n-2}, t_{n-2}; \dots y_1, t_1) \quad (13) \\ &= p_{n-1}(y_{n-1}, t_{n-1}; y_{n-2}, t_{n-2}; \dots y_1, t_1) \times P_n(y_n, t_n | y_{n-1}, t_{n-1}; y_{n-2}, t_{n-2}; \dots y_1, t_1). \end{aligned}$$

A Markov process is one for which future probabilities are determined by the most recently known value, and do not depend on the previous history.

$$P_n(y_n, t_n | y_{n-1}, t_{n-1}; y_{n-2}, t_{n-2}; \dots y_1, t_1) = P_2(y_n, t_n | y_{n-1}, t_{n-1}). \quad (14)$$

Stationary Markov processes are, therefore, completely characterized by  $p_1(y)$  and  $P_2(y_2, t | y_1) \equiv P_2(y_2, t' + t | y_1, t') = p_2(y_2, t; y_1, 0) / p_1(y_1, 0)$  (note that the stationary nature of the process implies the time translation invariance).

The importance of Markov processes is not that all physical processes are Markovian, but that the analysis of Markov processes is considerably simpler. For example, in the description of a Brownian motion in terms of sharp molecular kicks, the  $x$  - *velocity* of the particle is Markovian (the probability of  $u(t + \delta t)$  depends only on  $u(t)$  and the molecular collisions during the time interval  $\delta t$ ); on the other hand, the position  $x(t)$  is not Markovian, because  $x(t + \delta t)$  depends on  $x(t)$  and  $u(t)\delta t \approx x(t) - x(t - \delta t)$ , as well as the molecular collisions. It is of course easy to formulate the problem in terms of  $u(t)$  and then derive properties of  $x(t)$  using the relation between the two variables. The question of whether a random process is Markovian or not might depend on the level of the description. For example, in the coarse-grained description of Brownian motion where we discuss  $x(t)$  as a random walk (i.e., on a time scale larger than the relaxation time of the velocity), the random process  $x(t)$  becomes a Markov process.

The conditional probabilities satisfy the Chapman-Kolmogorov equation:

$$P_2(y_3, t_3 | y_1, t_1) = \int_{-\infty}^{\infty} dy_2 P_2(y_2, t_2 | y_1, t_1) P_3(y_3, t_3 | y_2, t_2; y_1, t_1), \quad (15)$$

where  $t_1 < t_2 < t_3$ . This relation corresponds to integrating over all the possible values of the variable,  $y(t)$ , at the intermediate time  $t_2$ . For a Markov process the  $P_3$  is given by  $P_2$  and the equation above takes the form of the Smoluchowski equation

$$P_2(y_3, t_3 | y_1, t_1) = \int_{-\infty}^{\infty} dy_2 P_2(y_2, t_2 | y_1, t_1) P_2(y_3, t_3 | y_2, t_2). \quad (16)$$

This equation describes the temporal evolution of the conditional probability. If in a small time interval, only small changes in  $y$  can occur, this time evolution of a Markov random process can be rewritten as a differential equation known as the Fokker-Planck equation. An example where this is the case is the Brownian motion of a heavy particle: in a small time interval the velocity of the heavy particle is only changed by a small amount by the small number of molecular collisions. On the other hand, for the velocity distribution of molecules in a gas, each binary collision can change the velocity by a large amount, and the Fokker-Planck equation does not apply. Thus, the Fokker-Planck equation is appropriate for the fluctuations of macroscopic degrees of freedom.

In order to write the Smoluchowski equation as a differential equation we start by rewriting it using the following change of variables  $y_1 \rightarrow y_0$ ,  $y_3 \rightarrow y$ ,  $y_2 \rightarrow y - \xi$ , and  $t_2 - t_1 \rightarrow t$ ,  $t_3 - t_2 \rightarrow \tau$

$$P_2(y, t + t_1 + \tau | y_0, t_1) = \int_{-\infty}^{\infty} d\xi P_2(y - \xi, t + t_1 | y_0, t_1) P_2(y, t_2 + \tau | y - \xi, t_2), \quad (17)$$

taking into account the stationary nature of the process, we write it as

$$P_2(y, t + \tau | y_0, 0) = \int_{-\infty}^{\infty} d\xi P_2(y - \xi, t | y_0, 0) P_2(y, \tau | y - \xi, 0). \quad (18)$$

where  $\tau$  will be a small time increment. In this equation, we are studying the conditional probability of getting to  $y$  at time  $t + \tau$  in terms of the probability of getting to a “nearby” location,  $y - \xi$ , at time  $t$  and from there to  $y$  in the small time increment  $\tau$ . Now we expand the left hand side in a Taylor expansion in  $\tau$

$$P_2(y, t + \tau | y_0, 0) \simeq P_2(y, t | y_0, 0) + \tau \frac{\partial P_2(y, t | y_0, 0)}{\partial t}, \quad (19)$$

and rewrite equation (18) as

$$\tau \frac{\partial P_2(y, t | y_0, 0)}{\partial t} = -P_2(y, t | y_0, 0) + \int_{-\infty}^{\infty} d\xi P_2(y - \xi, t | y_0, 0) P_2(y, \tau | y - \xi, 0). \quad (20)$$

Next, we expand  $P_2(y - \xi, t | y_0, 0)$  in  $\xi$ . Note that we are not expanding  $P_2(y, \tau | y - \xi, 0)$  in small  $\xi$ —for fixed  $y$  this function decreases rapidly with  $\xi$ , and it is precisely because of this that we only need to know  $P_2(z, t | y_0, 0) P_2(z + \xi, \tau | z, 0)$  for  $z$  near  $y$ . Thus, we write

$$\begin{aligned} P_2(y - \xi, t | y_0, 0) P_2(y, \tau | y - \xi, 0) &= P_2(z, t | y_0, 0) P_2(z + \xi, \tau | z, 0) |_{z=y-\xi} \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n \xi^n}{n!} \frac{\partial^n}{\partial y^n} [P_2(y, t | y_0, 0) P_2(y + \xi, \tau | y, 0)] \end{aligned} \quad (21)$$

(In the last line we used the Taylor expansion

$$P_2(z, t | y_0, 0) P_2(z + \xi, \tau | z, 0) |_{z=y-\xi} = f(z) |_{z=y-\xi} = \sum_{n=0}^{\infty} \frac{(-1)^n \xi^n}{n!} \frac{\partial^n f(y)}{\partial y^n}$$

).

Substituting the above expression in equation (20) we get

$$\tau \frac{\partial P_2(y, t | y_0, 0)}{\partial t} = -P_2(y, t | y_0, 0) + \int_{-\infty}^{\infty} d\xi \sum_{n=0}^{\infty} \frac{(-1)^n \xi^n}{n!} \frac{\partial^n}{\partial y^n} [P_2(y, t | y_0, 0) P_2(y + \xi, \tau | y, 0)]. \quad (22)$$

The term due to  $n = 0$  cancels the other term on the RHS (

$$P_2(y, t|y_0, 0) \int_{-\infty}^{\infty} d\xi P_2(y + \xi, \tau|y, 0) = P_2(y, t|y_0, 0)$$

) and we get

$$\frac{\partial P_2(y, t|y_0, 0)}{\partial t} = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial y^n} \left( P_2(y, t|y_0, 0) \lim_{\tau \rightarrow 0} \frac{1}{\tau} \int_{-\infty}^{\infty} d\xi \xi^n P_2(y + \xi, \tau|y, 0) \right). \quad (23)$$

If the random process evolves through the effect of many small changes, only the first two moments  $n = 1, 2$  of  $\xi$  will contribute, with higher moments increasing as  $\tau^p$  with  $p > 1$  giving no contribution as  $\tau \rightarrow 0$ . Hence we obtain the Fokker-Planck equation:

$$\frac{\partial P_2(y, t|y_0, 0)}{\partial t} = -\frac{\partial}{\partial y} [P_2(y, t|y_0, 0) A(y)] + \frac{1}{2} \frac{\partial^2}{\partial y^2} [P_2(y, t|y_0, 0) B(y)]. \quad (24)$$

where,

$$A(y) = \lim_{\tau \rightarrow 0} \frac{1}{\tau} \int_{-\infty}^{\infty} d\xi \xi P_2(y + \xi, \tau|y, 0), \quad (25)$$

$$B(y) = \lim_{\tau \rightarrow 0} \frac{1}{\tau} \int_{-\infty}^{\infty} d\xi \xi^2 P_2(y + \xi, \tau|y, 0). \quad (26)$$

Note that for ballistic motion we expect  $\xi \propto \tau$  and for regular white noise  $\xi^2 \propto \tau$ .

## Alternative Derivation of the Fokker-Planck equation

Let us consider the set of variables  $\{a_1, a_2, \dots\}$  denoted for convenience by the vector  $\mathbf{a}$ . The equations of motion are

$$\frac{da_j}{dt} = v_j(a_1, a_2, \dots) + F_j(t), \quad (27)$$

or using the vector notation

$$\frac{d\mathbf{a}}{dt} = \mathbf{v}(\mathbf{a}) + \mathbf{F}(t). \quad (28)$$

At the beginning, no special requirements are imposed on the noise free part of the dynamics except that it is Markovian (i.e., it is a function of  $\mathbf{a}(t)$  and not the history). However, we do require that the noise is white and Gaussian.

$$\begin{aligned}\langle \mathbf{F}(t) \rangle &= 0, \\ \langle \mathbf{F}(t) \mathbf{F}(t') \rangle &= 2\mathbf{B}\delta(t-t').\end{aligned}\tag{29}$$

Our goal is to find the probability distribution of the values of  $\mathbf{a}$  at time  $t$ ,  $f(\mathbf{a}, t)$ . Further, what we really want is the average of this probability distribution over the noise. Here, we start by recognizing that  $f(\mathbf{a}, t)$  is a conserved quantity

$$\int f(\mathbf{a}, t) d\mathbf{a} = 1 \quad \text{for all } t.\tag{30}$$

Whenever a conservation law of this kind is encountered, we expect that the time derivative of the conserved quantity is balanced by the divergence of the flux—the conserved quantity times the velocity. In our case, the spatial coordinates are  $\mathbf{a}$ , the probability density at  $\mathbf{a}$  is  $f(\mathbf{a}, t)$  and the velocity is  $d\mathbf{a}/dt$ . The conservation law is written as

$$\frac{\partial f}{\partial t} + \frac{\partial}{\partial \mathbf{a}} \cdot \left( f \frac{\partial \mathbf{a}}{\partial t} \right) = 0.\tag{31}$$

Replacing the time derivative of  $\mathbf{a}$  according to equation (28) we get

$$\frac{\partial f(\mathbf{a}, t)}{\partial t} = -\frac{\partial}{\partial \mathbf{a}} \cdot (\mathbf{v}(\mathbf{a})f(\mathbf{a}, t) + \mathbf{F}(t)f(\mathbf{a}, t)).\tag{32}$$

The above equation contains a random element and is called a stochastic differential equation. We will use this equation to derive the equation for the dynamics of the noise averaged  $f$ . We define an operator  $L$  by its action on any function  $\Phi$ ,

$$L\Phi = \frac{\partial}{\partial \mathbf{a}} \cdot (\mathbf{v}(\mathbf{a})\Phi).\tag{33}$$

This operator is used to write the noise-free equation,

$$\frac{\partial f(\mathbf{a}, t)}{\partial t} = -Lf(\mathbf{a}, t)\tag{34}$$

with the formal solution (note that  $L$  is independent of  $t$ )

$$f(\mathbf{a}, t) = e^{-tL}f(\mathbf{a}, 0).\tag{35}$$

Adding the noise term to the equation we write it as

$$\frac{\partial f(\mathbf{a}, t)}{\partial t} = -Lf(\mathbf{a}, t) - \frac{\partial}{\partial \mathbf{a}} \cdot (\mathbf{F}(t)f(\mathbf{a}, t)),\tag{36}$$

with the formal solution

$$f(\mathbf{a}, t) = e^{-tL}f(\mathbf{a}, 0) - \int_0^t e^{-(t-s)L} \frac{\partial}{\partial \mathbf{a}} \cdot (\mathbf{F}(s)f(\mathbf{a}, s)) ds.\tag{37}$$

It is important to notice that  $f(\mathbf{a}, t)$  depends only on the noise at earlier times ( $s < t$ ). Substituting the formal solution for  $f(\mathbf{a}, t)$  in equation (36) we get

$$\begin{aligned} \frac{\partial f(\mathbf{a}, t)}{\partial t} &= -L f(\mathbf{a}, t) - \frac{\partial}{\partial \mathbf{a}} \cdot (\mathbf{F}(t) e^{-tL} f(\mathbf{a}, 0)) \\ &+ \frac{\partial}{\partial \mathbf{a}} \cdot \left( \mathbf{F}(t) \int_0^t e^{-(t-s)L} \frac{\partial}{\partial \mathbf{a}} \cdot (\mathbf{F}(s) f(\mathbf{a}, s)) ds \right). \end{aligned} \quad (38)$$

Now we take the average over the noise. The initial distribution  $f(\mathbf{a}, 0)$  is not affected by the noise, therefore, after taking the average the term involving the initial distribution vanishes. The third term on the RHS of the equation contains two explicit terms of the noise ( $\mathbf{F}(t), \mathbf{F}(s)$ ) and, also, implicit terms in  $f(\mathbf{a}, s)$  (only noise terms at times earlier than  $s$ ). The noise is Gaussian and Delta correlated, that means that on averaging we can pair the explicit term  $\mathbf{F}(t)$  with  $\mathbf{F}(s)$  and also with the implicit terms in  $f(\mathbf{a}, s)$ . However, since  $f(\mathbf{a}, s)$  contains only noise terms at times earlier than  $s$ , the latter pairing doesn't yield any contribution to the average. The equation for the average is written as

$$\begin{aligned} \frac{\partial \langle f(\mathbf{a}, t) \rangle}{\partial t} &= -L \langle f(\mathbf{a}, t) \rangle \\ &+ \frac{\partial}{\partial \mathbf{a}} \cdot \left( \int_0^t e^{-(t-s)L} 2\mathbf{B} \delta(t-s) \cdot \frac{\partial}{\partial \mathbf{a}} \langle f(\mathbf{a}, s) \rangle ds \right). \\ &= -L \langle f(\mathbf{a}, t) \rangle + \frac{\partial}{\partial \mathbf{a}} \cdot \left( \mathbf{B} \cdot \frac{\partial}{\partial \mathbf{a}} \langle f(\mathbf{a}, t) \rangle \right). \end{aligned} \quad (39)$$

This is the Fokker-Planck equation. At this point  $\mathbf{B}$  can be any function of  $\mathbf{a}$ . It is convenient to remove the averaging notation  $\langle \rangle$  and write the equation as

$$\frac{\partial f(\mathbf{a}, t)}{\partial t} = -\frac{\partial}{\partial \mathbf{a}} \cdot (\mathbf{v}(\mathbf{a}) f(\mathbf{a}, t)) + \frac{\partial}{\partial \mathbf{a}} \cdot \left( \mathbf{B} \cdot \frac{\partial}{\partial \mathbf{a}} f(\mathbf{a}, t) \right). \quad (40)$$

## Brownian motion of a particle in a potential

The Langevin equation describing the dynamics of a Brownian particle moving in a potential  $U(x)$  is

$$\frac{dx}{dt} = \frac{p}{m}; \quad \frac{dp}{dt} = -U'(x) - \zeta \frac{p}{m} + F_p(t) \quad (41)$$

and the fluctuation-dissipation theorem is

$$\langle F_p(t) F_p(t') \rangle = 2\zeta k_B T \delta(t - t'). \quad (42)$$



The quantities that go into the general Fokker-Planck equation are

$$\begin{aligned} \mathbf{a} &= \begin{pmatrix} x \\ p \end{pmatrix}; & \mathbf{v}(\mathbf{a}) &= \begin{pmatrix} p/m \\ -U'(x) - \zeta \frac{p}{m} \end{pmatrix}; \\ F(t) &= \begin{pmatrix} 0 \\ F_p(t) \end{pmatrix}; & \mathbf{B} &= \begin{pmatrix} 0 & 0 \\ 0 & \zeta k_B T \end{pmatrix}; \end{aligned}$$

The Fokker-Planck equation may be written as

$$\begin{aligned} \frac{\partial f(x, p, t)}{\partial t} &= -\frac{\partial}{\partial x} \left( \frac{p}{m} f(x, p, t) \right) - \frac{\partial}{\partial p} \left( \left( -U'(x) - \zeta \frac{p}{m} \right) f(x, p, t) \right) \\ &\quad + \zeta k_B T \frac{\partial^2}{\partial p^2} f(x, p, t). \end{aligned} \tag{43}$$

It is possible to show that the equilibrium solution is

$$f_{eq}(x, p) = \frac{1}{Q} e^{-H(x, p)/k_B T} \text{ where } Q = \iint dx dp e^{-H(x, p)/k_B T},$$

and

$$H = \frac{p^2}{2m} + U(x).$$

Note that  $Q$ , is the partition function and  $f_{eq}(x, p)$  is the Boltzmann distribution.