# Fokker-Plank Equations

By Wayne Wignes

# Introduction

Random walks are commonly used to model many physical pehnomena in the physical sciences and elsewhere such as financial engineering and behavioral modelling. For an uncorrelated random walk we have a Markov process, that is to say that jumps are completely dissasociated with the past values.

One difficulty that arises when solving random walks is that (as the name implies) trajectories do not evolve deterministically, and it is found that at equilibrium this 'can introduce artificial drift velocities at equilibrium'[2], i.e. because there is a stochastic part to the Itô differential then the velocity is not deterministic. This has the effect of destroying the principle of detailed balance - a thing which is of interest in chemical processes as it allows the forward and backward rates to be compared. In contrast to the stochastic evolution of the trajectories of the particles themselves, probabilities associated with the random variables occupying a specified state *do* evolve deterministically, and this fact can be exploited so as to develop rate coefficients which respect detailed balance[1, 2].

The Fokker-plank equation describes the time evolutoin of conditional probabilities of an uncorrelated (Markov) random walk process. The *Chapman-Kolmogorov equation* (hereafter referred to as the CK equation) is an equation which was developed as a means to relate the Fokker-Plank equation to the Chemical *master equation* which utilizes the rates we hope to solve for in a way that respects detailed balance.

Section 1 of this paper will introduce the Chapman-Kolmogorov equation and how it can be used to relate the Fokker-PLank equation to the master equation. Section 2 introduces the Langevin equation as it relates the the Fokker-plank equation. The [finite volume] computational algorithm for modeling the chemical master equation is developed in section 3 along with some brief analysis' on detailed balance.

## 1 CK, Fokker-Plank, and the Master equation

For the uncorrelated random walks jumps are independent of their past, which is to say it is a Markov process; what happens next has nothing to do with what has happened in the past. We make the following definitions:

| | |
|---|---|
| $p(x_{i-1}\|x_i)$ : | Probability particle is at position $x_{i-1}$ given it was previously at $x_i$ |
| $w(x_{i-1}\|x_i)$ : | Probability *rate* of jumps from $x_{i-1}$ to $x_i$ |

As derived in the appendix, the Chapman-Komogorov equation is,

$$\boxed{\begin{aligned}\partial_t P(\mathbf{y},t|\mathbf{z},t') = &-\sum_i \frac{\partial}{\partial y_i} A_i(\mathbf{y},t)P(\mathbf{y},t|\mathbf{z},t') + \frac{1}{2}\sum_{i,j}\frac{\partial^2}{\partial y_i \partial y_j} B_{ij}(\mathbf{y},t)P(\mathbf{y},t|\mathbf{z},t') \\ &+ \int d\mathbf{x} W(\mathbf{y}|\mathbf{x},t)P(\mathbf{x},t|\mathbf{z},t') - W(\mathbf{x}|\mathbf{y},t)P(\mathbf{y},t|\mathbf{z},t')\end{aligned}} \tag{A.17}$$

Given three consecutive states $x_{i-1}, x_i, x_{i+1}$ the Chapman-Kolmogorov equation gives for this one dimensional case,

$$\begin{aligned}\frac{\partial}{\partial t}\left(P(x_i,t|x_{i+1},t')\right) = &-\frac{\partial}{\partial x_i}\left[A_i(x_i,t)P(x_i,t|x_{i+1},t')\right] + \frac{1}{2}\frac{\partial^2}{\partial x_i^2}[B_i^2(x_i,t)P(x_i,t|x_{i+1},t')] \\ &+\int dx_{i-1}W(x_i|x_{i-1},t)P(x_{i-1},t|x_{i+1},t') - W(x_{i-1}|x_i,t)P(x_i,t|x_{i+1},t')\end{aligned} \tag{1}$$

Where $A_i(x_i,t)$ represents the drift vector associated with state $i$ and $B_i(x_i,t)$ the diffusion coefficient. Notice the final integral (which is transformed into a summation for discrete states) sums over all possible starting states from which a *forward* jump might occur within our three state system[1]. Note that we have a system of three states, but we've made no such restrictions on the number of particles which might occupy those states. We will be presuming a large number (and ensemble) of identical systems from which measurements are taken and averaged.

For a process which has had time to settle into a steady state, the stochastic properties of the system become independent of time, i.e. if a process is truly random then you would expect that as time goes on how it evolves eventually has little to do with how it began. If we presume the law of large numbers applies to the averaging of measurements, then we have an ergodic process which is characterized by the condition[6],

$$\lim_{t\to\infty} p(x,t|x',0) = p(x) \tag{2}$$

An ergodic process is one for which the time average over a given interval of measurement is the same as the ensemble average. One more condition on which probability $p(x,t)$ must fulfill in order for the differential chapman-kolmogorov equation to hold is what C.W. Gardner [6] refers to as the *continuity condition*

$$\lim_{\Delta t\to\infty}\frac{1}{\Delta t}\int_{|x_i-x_{i-1}|>\varepsilon} P(x_{i-1},t+\Delta t|x_i,t)dx_{i-1} = 0 \tag{3}$$

which is a conditional form of the Lindeberg equation. A proof of the Lindeberg condition in this form is non-trivial and will not be pursued here.

The continuity condition has the implications of continuous sample paths[6] but this does not imply we have continuous *states*. The consequence of this will be that the sum on the last set of terms in equation 1 will be transformed into a discrete summation. Employing the ergodic dondition allows us to simplify the Chapman-Kolmogorov equation as follows[2],

$$\frac{\partial}{\partial t}P_i(x) = -\frac{\partial}{\partial x}\left[A_i(x)P_i(x)\right] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[B_i^2(x)P_i(x)] \quad + \quad \sum_{i-1}\Big[W(x_i|x_{i-1})P_{i-1}(x) - W(x_{i-1}|x_i)P_i(x)\Big] \tag{4}$$

Where $P_i(x)$ denotes the probability function of state $i$. We will also change notation such that $W(x_{i-1}|x_i) = r_i^{i-1}$ as this may be more intuitive than conditional probability lingo which essentially reads backward, i.e. $W(x_{i-1}|x_i)$ is the rate of going from $i$ to state $i-1$ but the two quantities are listed in reverse. $r_i^{i-1}$ however reads more naturally; the bottom index is the starting point and the upper index is the finish point for the jump. Carrying out the sum on

[1] this is likely related to the concept of 'kolmogorov forward equation'

[2] It is questionable whether it is appropriate to include the subscript $i$ on the probabilities for the differential terms. No authors I followed have this, but then again none of them attempt to show how the full chapman-kolmogorov equation transforms with the ergodic assumption

the terms involving transition rates and employing this notation, equation 4 becomes,

$$\boxed{\begin{aligned}\frac{\partial}{\partial_t}P_i(x) = &-\frac{\partial}{\partial x}\left[A_i(x)P_i(x)\right] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[B_i^2(x)P_i(x)] \\ &+ \left[r_{i-1}^i P_{i-1}(x) - r_i^{i-1}P_i(x)\right] + \left[r_{i+1}^i P_{i+1}(x) - r_i^{i+1}P_i(x)\right]\end{aligned}} \tag{5}$$

Splitting the two equations gives the well known *master equation*

$$\boxed{\partial_t P(x) = \left[r_{i-1}^i P_{i-1}(x) - r_i^{i-1}P_i(x)\right] + \left[r_{i+1}^i P_{i+1}(x) - r_i^{i+1}P_i(x)\right]} \tag{6}$$

and the [homogeneous] *Fokker-Plank equation*

$$\boxed{-\frac{\partial}{\partial x}\left[A_i(x)P_i(x)\right] + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left[B_i^2(x)P_i(x)\right] = 0} \tag{7}$$

Tyically, rather than splitting the equations, we'd say the Fokker-Plank equation gives, the probability rate when jump transitions are set to zero, and vice-versa; the Master equation gives the probability rate when the differential equation goes to zero. The latter case would seem to represent a pure jump process, while the Fokker-Plank equation - as the names of the coefficients implies - describes the drift and diffusion processes, i.e. rather than a discontinuous jump we have drift or diffusion towards another state. One begins to get an intuitive feel as to why the continuity condition is a prerequisite to teh Chapman-Kolmogorov equation.

Making the substitutions $2A(x)/B^2(x) = \alpha(x)$ and $f(y) = B^2(x)P(x)$ we can write the Fokker plank equation as

$$\begin{aligned}-\frac{\partial}{\partial x}\left(\alpha(x)f(x)\right) + \frac{\partial^2 f(x)}{\partial x^2} &= 0 \\ -\alpha(x)f(x) + \frac{\partial f(x)}{\partial x} &= c_2\end{aligned} \tag{8}$$

which is easily solved by[4],

$$f(x) = g(x)\left[c_1 + c_2\int^x \frac{1}{g(u)}du\right] \qquad ; \qquad g(x) = e^{\int^x \alpha(t)dt}$$

Which is the steady state solution. In terms of the original quantity of interest $P(y)$ we have,

$$\boxed{\begin{aligned}P_i(x) = \frac{1}{B(x)}g(x)\left[c_1 + c_2\int^x \frac{1}{g(u)}du\right] \qquad ; \qquad & g(x) = e^{\int^x \alpha(t)dt} \\ & \alpha(t) = 2A(t)/B^2(t)\end{aligned}} \tag{9}$$

## 2 Langevin's Equation

The langevin equation reads,

$$\frac{dx}{dt} = A(x) + B(x)\frac{dw(t)}{dt} \tag{10}$$

Where $w(t)$ is the normally distributed random variable which represents the Wiener process that plays a central role in Itô calculus. $\frac{dw(t)}{dt}$ is the derivative of the stochastic variable and is known as *white noise*. In terms of thermodynamic relations the drift and diffusion coefficients typically take the form $A(x) = -\frac{D}{kT}\frac{\partial}{\partial x}V(x)$ and $B(x) = \sqrt{2D}$ where $D$ is the diffusion coefficient as derived in the field of physics (it is curious to note $B \neq D$). It is a trivial matter to generate sample paths satisfying the langevin equation; multiplying by $dt$ and discretizing the Langevin's equation gives,

$$X(t + \Delta t) = X(t) - A(x)\Delta t + B(x)\sqrt{\Delta t}N[0, 1] \tag{11}$$

Where the factor $dw(t) = \sqrt{\Delta t}N[0, 1]$ comes from the fact that $w(t)$ is normally distributed with standard deviation $\sqrt{\Delta t}$. This is briefly discussed in the appendix (see 'Lindeberg's condition'). In terms of coding in Python a random number generator was used in place of the expression $N[0, 1]$. The results of simulating various sample paths are depicted in figure 1.

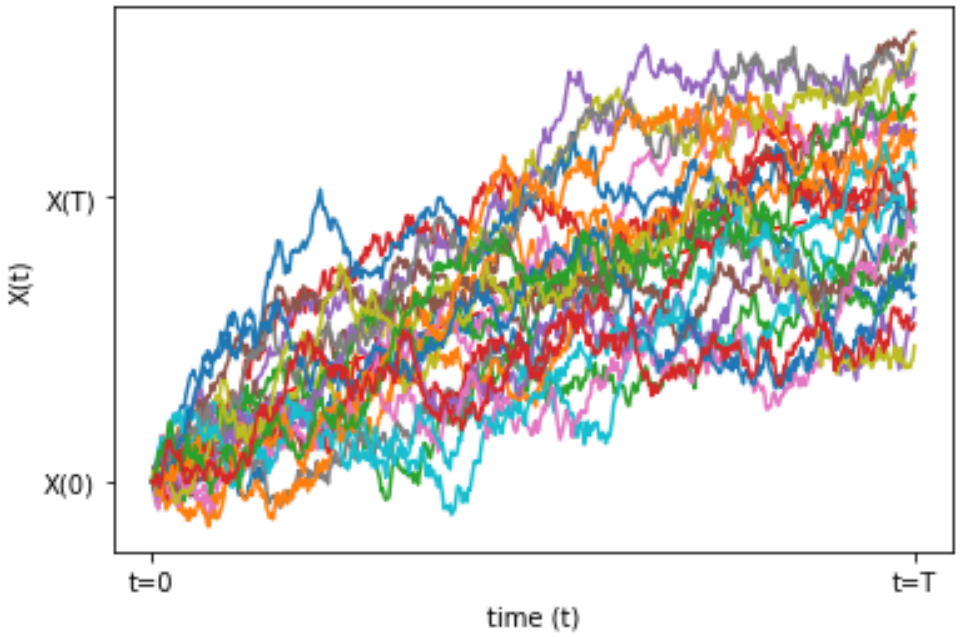


Figure 1: Figure 2: Sample paths of Langevin's process

The above algorithm for simulating sample paths, while very easy to implement, does not obey the principle of detailed balance[2]

What follows will be the development of an algorithm which does obeye the principle of detailed balance. The trade off will be that it is the probability $P(x(t))$ rather than $x(t)$ which will be solved for. To relate the probability

distribution to figure 1 one can imagine a bell shaped curve laying on its side and which peaks where sample paths in figure 1 have a tendency to cluster. To this end we insert the aforementioned values $A(x) = -\frac{D}{kT}\frac{\partial}{\partial x}V(x) = V'(x)$ and $B(x) = \sqrt{2D}$ into equation 7 to obtain the Fokker Plank equation that will be the focus of the algorithm once it is developed.

$$D\frac{\partial}{\partial y}\left[\frac{1}{kT}V'(y)P(y)\right] + \frac{D}{2}\frac{\partial^2}{\partial y^2}\left[P(y)\right] = 0 \tag{12}$$

For the moment however, we will develop the algorithm in the most general fashion possible by continuing to build off equation 7 with generalized coefficients.

# 3 Computation: Finite Volume

One may have noticed by now that the *state* $i$ is not located at the *position* $x_i$. This may seem an unnecessary convolution at first, but we will soon see the purpose of this is related to the boudary conditions we place on the probability. So for clarity, the values $r_i, P_i$ reference state $i$ while the quantities $\alpha(x_i) = \alpha_i, x_i$ reference position $x_i$.

Applying the well known relation $j = -\int \partial_t p(x) dx$ (where $j$ is the probability flux) to the Master equation gives,

$$\begin{aligned} \partial_t P(x) &= \left[r^i_{i-1}P_{i-1} - r^{i-1}_i P_i\right] - \left[r^{i+1}_i P_i - r^i_{i+1}P_{i+1}\right] \\ &= j_2 - j_1 \\ &= J \end{aligned} \tag{13}$$

Where $J = j_2 - j_1$ is the total flux, $j_1$ represents the flux between state $i$ and $i+1$, and $j_2$ represents the flux between state $i-1$ and $i$. Note that an easy rule of thumb for determining the sign of the flux terms is that flux going *into* state $i$ is positive, and flux leaving state $i$ is negative (see figure 2).

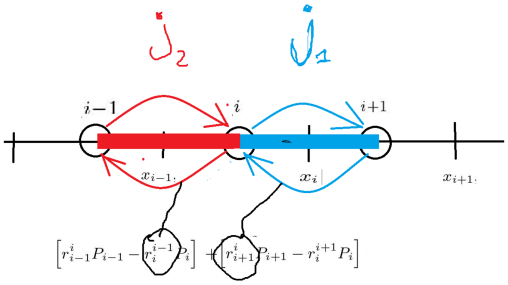


Figure 2: Flux $j_1, j_2$ for state $i$ and their relation to the master equation

When we want to use the Fokker-Plank equation to calculate the flux in a given region we integrate the homogeneous Fokker-Plank equation (equation 8) which is equivalent to the steady state case ($\partial_t p(x) = 0$) with the jump rates set to zero ($r^{i+1}_i = r^i_{i+1} = 0$). Note that the constant $-c_2$ in this equation is the flux. Why this is so is due to the well known relations for probability flux $\nabla \cdot j = -\frac{\partial P}{\partial t}$. When the probability rate is zero we have a steady state, but when the flux iteslf is zero we have equilibrium[12]. In one dimension this is to say,

$$\begin{aligned} \nabla j = -\frac{\partial P}{\partial t} \quad &\longrightarrow \quad \text{steady state} \\ j = \int \frac{\partial P}{\partial t} dx \quad &\longrightarrow \quad \text{equilibrium} \end{aligned} \tag{14}$$

Equation 8 is thus related to the flux as follows,

$$\begin{aligned} -j = -\int (\nabla j) dx &= \int \frac{\partial P_i(x)}{\partial t} dx \\ &= -\alpha(x) f(x) + \frac{\partial f(x)}{\partial x} \\ &= c_2 \end{aligned} \tag{15}$$

and this gives $j = -c_2$ in both regions.

**region 1:** $x_{i-1} < x < x_{i+1}$ **:**

.

We will first consider the interval $[x_{i-1}, x_{i+1}]$ which encompasses states $i$ and $i+1$. Equation 9 - the steady state solution - then becomes,

$$P_i(x) = \left[ c_1 + c_2 \int_{x_{i-1}}^{x} \frac{1}{g(u)} du \right] g_{i-1}(x) \qquad ; \qquad g(x) = e^{\int_{x_{i-1}}^{x} \alpha(t) dt} \tag{16}$$

Now for the sake of convenience we make the following definitions,

$$\begin{aligned} Q^{(i)} &= \int_{x_{i-1}}^{i} (x) \int_{x_{i-1}}^{x} \frac{1}{g(u)} du, & Q^{(i+1)} &= \int_{x_i}^{i+1} g(x) \int_{x_{i-1}}^{x} \frac{1}{(u)} du, \\ I^{(i)} &= \int_{x_{i-1}}^{x_i} (x) dx, & I^{(i)} &= \int_{x_{i+1}}^{x_i} (x) dx \end{aligned} \tag{17}$$

Where the superscript denotes the upper bound of the integrals (and one can infer the lower bound by subtracting one). We have two unknowns $c_1$ and $c_2$ and this warrants two boundary conditions. Integrating the probability density over the cells gives the probability of occupying those cells, respectively;

$$\int_{x_{i-1}}^{x_i} P(x) \quad = \quad c_1 I^{(i)} + c_2 Q^{(i)} \quad = \quad P_i \tag{18}$$

$$\int_{x_i}^{x_i+1} P(x) \quad = \quad c_1 I^{(i+1)} + c_2 Q^{(i+1)} \quad = \quad P_{i+1} \tag{19}$$

Solving this set of equations for the coefficient $c_2$ yeilds,

$$c_2 = -\frac{I^{(i+1)} P_i - I^{(i)} P_{i+1}}{I^{(i)} Q^{(i+1)} - I^{(i+1)} Q^{(i)}} \tag{20}$$

If within the region of interest $[x_{i-1}, x_{i+1}]$ we have $\alpha_{i-1}(t) \approx constant$ then we can make the following estimates,

$$g(x) = e^{\int_{x_{i-1}}^{x} \alpha(t) dt} \approx e^{\alpha_{i-1}(x - x_{i-1})} \tag{21}$$

Where $\alpha_{i-1} = \alpha(x_{i-1})$, i.e. we will adopt the convention that we estimate our integrals by taking the height of rectangles at the beginning of the intervals(s)[3] The quantities in equation 17 then become,

[3]One might object that the value of $\alpha$ should be different on the interval $[x_i, x_{i+1}]$ than it is on the intervale $[x_{i-1}, x_i]$, and indeed this may be so. This is an attempt to reduce the results presented in reference [1] (who does not seem to impose equilibrium conditions in order to calculate the rates) to those presented in [2] (who does impose equilibrium, i.e. $P^{eq} \propto e^{-\frac{V(x)}{kT}}$, and this simplifies the analysis). In hindsight, perhaps the two were not meant to be related, but reference [1] does not seem to explain how he goes about estimating the integrals in equation 20.

Note that $\alpha \propto V^{'}(x)$ for the Langevin equation, so evaluating the integral in the expontential may be unneccessary; we could just say $\int^{x} \alpha(t) dt \propto V(x)$. This however, would not give us the factor of $\alpha_{i-1}$ which the end result is multiplied by in reference [2] (see the rates given in equation A.7 of reference [2] or equation 26 of this paper). So we must have

$$I^{(i+1)} = \int_{x_i}^{x_{i+1}} e^{\alpha_{i-1}(x-x_{i-1})}dx = e^{\alpha_{i-1}(x_i-x_{i-1})}\Delta x = e^{\alpha_{i-1}\Delta x}\Delta x$$
$$I^{(i)} = \int_{x_{i-1}}^{x_i} e^{\alpha_{i-1}(x-x_{i-1})}dx = e^{\alpha_{i-1}\Delta x}\Delta x = \Delta x \tag{22}$$

$$\begin{aligned}
Q^{i+1} &= \int_{x_i}^{x_{i+1}} e^{\alpha_{i-1}(x-x_{i-1})} \int_{x_{i-1}}^{x} e^{-\alpha_{i-1}(t-x_{i-1})}dt \\
&= \int_{x_i}^{x_{i+1}} \int_{x_{i-1}}^{x} e^{\alpha_{i-1}(x-t)}dt \\
&= -\frac{1}{\alpha_{i-1}} \int_{x_i}^{x_{i+1}} \big(1 - e^{\alpha_{i-1}(x-x_{i-1})}\big)du \\
&= -\frac{\Delta x}{\alpha_{i-1}} + \frac{1}{\alpha_{i-1}}\Big(e^{2\alpha_i\Delta x} - e^{\alpha_{i-1}\Delta x}\Big) \\
&\approx -\frac{\Delta x}{\alpha_{i-1}}
\end{aligned} \tag{23}$$

Plugging these values into equation 19 gives,

$$\begin{aligned}
c_2 &= -\frac{I^{(i+1)}P_i - I^{(i)}P_{i+1}}{I^{(i-1)}Q^{(i+1)} - I^{(i+1)}Q^{(i)}} \\
&= -\frac{\big(e^{\alpha_{i-1}\Delta x}\Delta x\big)P_i - \big(\Delta x\big)P_{i+1}}{\big(-\Delta x\big)\big(\frac{\Delta x}{\alpha_{i-1}}\big) - \big(e^{\alpha_{i-1}\Delta x}\Delta x\big)\Big(-\frac{\Delta x}{\alpha_{i-1}}\Big)} \\
&= -\frac{\alpha_{i-1}}{\Delta x}\frac{e^{\alpha_{i-1}\Delta x}P_i - P_{i+1}}{e^{\alpha_{i-1}\Delta x} - 1} \\
&= -\frac{\alpha_{i-1}}{\Delta x}\frac{e^{\alpha_{i-1}\Delta x}P_i - P_{i+1}}{e^{\alpha_{i-1}\Delta x} - 1}
\end{aligned} \tag{24}$$

Thus we've found the flux on the interval $x_{i-1}, x_{i+1}$

$$j_2 = \frac{\partial P(x)}{\partial t} = -c_2 = \frac{\alpha_{i-1}}{\Delta x}\frac{e^{\alpha_{i-1}\Delta x}P_i - P_{i+1}}{e^{\alpha_i\Delta x} - 1} \tag{25}$$

The rates can be read off from this by comparing to ...

$e^{\int_{x_{i-1}}^{x}\alpha(t)dt} \approx= e^{\alpha(x_{i-1})(x-x_{i-1})} = e^{\alpha_{i-1}(x-x_{i-1})}$ which when integrated over yeilds a factor of $\frac{1}{\alpha_{i-1}}$ in the denominator of the flux. This is neccessary to arrive at equation 26.

Again, attempting to reduce the results of reference [1] to those of reference [2] may not be appropriate. After a lot of frustrated messing around with this, I had to just do what worked. The end result (equation 26) may be compared to the rates given in equation A.7 of reference [2]

$$\boxed{\begin{aligned} r_{i+1}^{i} &= \frac{\alpha_{i-1}}{\Delta x}\frac{1}{e^{\alpha_{i-1}\Delta x}-1} \\ r_{i}^{i+1} &= \frac{\alpha_{i-1}}{\Delta x}\frac{e^{\alpha_{i-1}\Delta x}}{e^{\alpha_{i-1}\Delta x}-1} \end{aligned}} \tag{26}$$

## Region 2: $x_{i-2} < x < xi$ :

Going through the same process on the interval $x_{i-2}, x_i$ but this time with the definition of $g(x)$ discretized as,

$$g_i = e^{\int_{x_{i-2}}^{x} a(t)dt} \approx e^{\alpha_{i-2}(x-x_{i-2})} \tag{27}$$

Where $\alpha_{i-2} = \alpha(x_{i-2})$. The conditions we use to find the two constant this time are,

$$\int_{x_{i-2}}^{x_{i-1}} P(x) \quad = \quad c_1 I^{(i-1)} + Q^{(i-1)} \quad = \quad P_{i-1} \tag{28}$$

$$\int_{x_{i-1}}^{x_i} P(x) \quad = \quad c_1 I^{(i)} + Q^{(i)} \quad = \quad P_i \tag{29}$$

Solving for the the coefficient $c_2$ gives the same result as last time only the indexes are shifted,

$$c_2 = -\frac{I^{(i-1)}P_i - I^{(i)}P_{i-1}}{I^{(i)}Q^{(i-1)} - I^{(i-1)}Q^{(i)}} \tag{30}$$

Once again adopting the convention of estimating integrals by the value at the lower bound, we obtain the following values for the various parameters,

$$\begin{aligned} I^{(i+-1)} &= \int_{x_{i-2}}^{x_{i-1}} e^{\alpha_{i-2}(x-x_{i-2})} = \Delta x \\ I^{(i)} &= \int_{x_{i-1}}^{x_i} e^{\alpha_{i-2}(x-x_{i-2})} = e^{\alpha_{i-1}\Delta x}\Delta x \\ Q^{(i-1)} \quad &= \quad \int_{x_{i-2}}^{x_{i-1}}\int_{x_{i-2}}^{x} e^{\alpha_{i-2}(x-t)} \quad = \quad -\frac{1}{\alpha_{i-2}}\int_{x_{i-2}}^{x_{i-1}} \left(1 - e^{\alpha_{i-2}(x-x_{i-2})}\right) \quad = \quad 0 \\ Q^{(i)} \quad &= \quad \int_{x_{i-1}}^{x_i}\int_{x_{i-2}}^{x} e^{\alpha_{i-2}(x-t)} \quad = \quad -\frac{1}{\alpha_{i-2}}\int_{x_{i-1}}^{x_i} \left(1 - e^{\alpha_{i-2}(x-x_{i-2})}\right) \\ &= \quad -\frac{\Delta x}{\alpha_{i-2}} + e^{-\alpha_{i-2}\Delta x}\frac{\Delta x}{\alpha_{i-2}^2} \approx -\frac{\Delta x}{\alpha_{i-2}} \end{aligned} \tag{31}$$

Inserting these into equation 30 gives,

$$
\begin{aligned}
c_2 &= -\frac{I^{(i-1)}P_i - I^{(i)}P_{i-1}}{I^{(i)}Q^{(i-1)} - I^{(i-1)}Q^{(i)}} \\
&= -\frac{\big(\Delta x\big)P_i - \big(e^{\alpha_{i-2}\Delta x}\Delta x\big)P_{i-1}}{\big(e^{\alpha_{i-2}\Delta x}\Delta x\big)\big(-\frac{\Delta x}{\alpha_{i-2}}\big) - \big(\Delta x\big)\big(-\frac{\Delta x}{\alpha_{i-2}}\big)} \\
&= -\frac{\alpha_{i-2}}{\Delta x}\frac{e^{\alpha_{i-2}\Delta x}P_{i-1} - P_i}{e^{\alpha_{i-2}\Delta x} - 1}
\end{aligned} \tag{32}
$$

Which is the same result as on the interval $x_{i-1}, x_i$. So the definition of $j_1$ and $j_2$ are the same except for the the subscript on $\alpha$.

$$
\boxed{
\begin{aligned}
r_{i-1}^{i} &= \frac{\alpha_{i-2}}{\Delta x}\frac{e^{\alpha_{i-1}\Delta x}}{e^{\alpha_{i-1}\Delta x} - 1} \\
r_{i}^{i-1} &= \frac{\alpha_{i-2}}{\Delta x}\frac{1}{\Delta x e^{\alpha_{i-1}\Delta x} - 1}
\end{aligned}
} \tag{33}
$$

## Modelling Chemical Reactions

Note that the rates developed thus far are the rates calculated with the steady state conditoin $\frac{\partial_t P(x)}{\partial t} = 0$. We now attempt to model the system as it evolves when a chemical reaction is turned on. Say the system is initially in some chemical state which is characterized by the original potential in the Langevin equation $V(x_i)$. Now we are saying that $V(x_i)$ can have multiple values at any given positional coordinate $x_i$. In quantum mechanics for example, it is common that there are multiple discrete energy levels which can be occupied, and they are independent of position. Let us index this chemical state with the index $j$, so we have $V_i^j$ denoting the energy state at position $i$ and in electronic state $j$. We need to change our nationa so as to accomodate $j$. Let us denote $r_{i\to i'}^{j\to j'}$ as the jump rate from spatial state $i$ to $i'$ and electronic state $j$ to $j'$. If no arrow is indicated then this is taken as no transitions occur. The master equation developed thus far (with no electronic transitions) can then be written as,

$$\partial_t P(x) = \left[r_{i-1\to i}^j P_{i-1}^j - r_{i\to i-1}^j P_i^j\right] - \left[r_{i\to i+1}^j P_i^j - r_{i+1\to i}^j P_{i+1}^j\right] \tag{34}$$

Now we turn on a chemical reaction which can be modelled as a sum of some terms which do not make spatial transitions but which make electronic transitions[1]. This gives,

$$\partial_t P(x) = \left[r_{i-1\to i}^j P_{i-1}^j - r_{i\to i-1}^j P_i^j\right] - \left[r_{i\to i+1}^j P_i^j - r_{i+1\to i}^j P_{i+1}^j\right] + \sum_{j\neq j'}\left[r_i^{j'\to j} P_i^{j'} - r_i^{j\to j'} P_i^j\right] \tag{35}$$

The electronic transition rates may be modeled with boltzman distribution probabilities[1] which are equilibrium value probability distributions which will be discussed in the next section. Including these electronic reaction terms will cause the probability rate to be non-zero, and this one would immediately expect would negate our calculations thus far for the spatial jump rates which were dependent on a steady state calculation. However, to quote the authors of reference [1] the steady state calculation of the spatial jump rates "*.. implicitly assumes that the presence of a chemical reaction does not significantly affect the local steady-state solution. This assumption is valid when the cell size is small enough such that the diffusion within the cell is faster than the reaction...*". The authors justify this with an argument involving the difference of time scales involved in advection, diffusion, and reaction. A proof of this statemetn will not be pursued here.

Once the transition rates have been calculated, and with a reaction turned on, we now have a time evolving probability function suitable for computation and to which if one desires is suitable for the Crank-Nicholson method. Putting all of the rate constants into some matrix $\mathbf{M}$ we have the following matrix equation;

$$\boxed{\frac{\partial_t \mathbf{P}(x)}{\partial t} = \mathbf{MP}} \tag{36}$$

This method is second order accurate and stable so long as the flux is finite. A detailed proof this will not be pursued here - see the appendix of references [1] and [2]. Regrettably, neither will examples for which we might apply the above algorithm be pursued. For some interesting and suprisingly useful algorithms involving chemical reactions see reference[2].

## Detailed balance

In chemical reactions it is of interest to relate forward and backward rates when possible. As previously discussed the Langevin equation finds many applications in biological and chemical systems, and we now employ it. Recall that the Langevin equation has Fokker-Plank coefficients $A(x) = -\frac{-DV^{'}(x)}{kT}$ and $B(x) = \sqrt{2D}$. Using Backward difference for the derivative this gives,

$$\alpha_{i-1}\Delta x = \frac{2A(x_{i-1})}{B(x_{i-1})}\Delta x = -\frac{V^{'}(x_{i-1})}{kT}\Delta x = -\frac{V_i - V_{i-1}}{\Delta x}\frac{\Delta x}{kT} = -\frac{V_i - V_{i-1}}{kT} \tag{37}$$

Dividing the rates given in 26 and using the above relation,

$$\frac{r_i^{i+1}}{r_{i+1}^i} = e^{-\frac{V_{i+1}-V_i}{kT}\Delta x} \qquad \longrightarrow \qquad r_i^{i+1}e^{-\frac{V_i}{kT}\Delta x} = r_{i+1}^i e^{-\frac{V_{i+1}}{kT}\Delta x} \tag{38}$$

This is known as the principle of detailed balance[11, 2, 12]. Moving everything to one side agives a result which can be compared to equation 13,

$$\frac{\partial P_i(x)}{\partial t} = \left(r_i^{i+1}e^{-\frac{V_i}{kT}\Delta x} - r_{i+1}^i e^{-\frac{V_{i+1}}{kT}\Delta x}\right) = 0 \tag{39}$$

We can compare this to equation 13 and infer a solution of the general form $P_i \propto e^{-\frac{V_i}{kT}}$ where the 'proportional to' symbol accounts for the fact that the probabilities need to be be normalized. It is readily verified from the steady state solution in equation 9 that with $\alpha(t) \propto V^{'}(t)$ and with $V(x) \approx V(x_i) = V_i = constant$ on a given interval, then our solution reduces to the following,

$$P_i^{eq} = c_1 e^{-\frac{V_i}{kT}} + c_2 \tag{40}$$

Now recall that $\nabla j = 0$ implies steady state (in one dimension) while $j = 0$ implies equilibrium. But we also found that $j = -c_2$, hence setting $c_2 = 0$ gives the equilibrium solution, which as we suspected is$P_i^{eq} \propto e^{-\frac{V_i}{kT}}$.

Any time one has a probability of the form $P \propto e^{-\frac{V_i}{kT}}$ equilibrium can be inferred. This is because this probability has the form of the *Boltzman distribution* which characterizes the canonical ensemble in statistical physics. This probability arises from minimizing entropy (in accordance with the second law of thermodynamics) in a fairly rigorous way by using the method of Lagrange multipliers. The canonical ensemble primarily implies two things; a constant temperature (or large heat bath which regulates the system of interest) and a large number of 'mental copies' of the system. This is not necessarily to say you need have a very long chain of identical systems, instead it implies we are averaging over a large number of measurements on what can be taken to be identical systems (it is the average value which is in reality measured in many molecular scale experiments).

But in order for this 'mental copy' averaging to work, we must require that the time average be the same as the ensemble average, i.e. we have an ergodic system.

# 4 Two Dimensional Fokker-Plank

Given a potential of the form $V(x,y) = V_1(x) + V_2(y)$, in two dimensions the steady state Fokker plank equation is separable,

$$-\frac{\partial}{\partial y}\left\{\frac{1}{kT}X(x) + \frac{\partial B(x)}{\partial x}\right\} - \frac{\partial}{\partial y}\left\{\frac{1}{kT}Y(y) + \frac{\partial Y(y)}{\partial y}\right\} = 0 \tag{41}$$

Where the above form arose from proposing a solution of the form $P(x,y) = X(x)Y(y)$. The solution for both $X(x)$ and $Y(y)$ are then found to be the same form as equation 9 - the steady state solution to the one dimensional case.

$$Y(y) = \left[b_1 + b_2\int^y \frac{1}{g_y(u)}du\right] g_y(y) \qquad ; \qquad g_y(y) = e^{-\frac{V_2(y)}{kT}} \tag{42}$$

$$X(x) = \left[c_1 + c_2\int^x \frac{1}{g_x(u)}du\right] g_x(x) \qquad ; \qquad g_x(x) = e^{-\frac{V_1(x)}{kT}} \tag{43}$$

Similar to probability conditions we proposed in equation 18 and 19 we have,

$$\int_{y_{j-1}}^{y_j} Y(y)dy \int_{x_{i-1}}^{x_i} X(x)dx = P_{i,j} \tag{44}$$

$$\int_{y_{j-1}}^{y_j} Y(y)dy \int_{x_i}^{x_i+1} X(x)dx = P_{i+1,j} \tag{45}$$

We see from 44 and 45 that there is a problem; we have two equations and four unknown constants to solve for. Xing, Wang, and Oster [1] whose work was mostly followed in the derivation for the 1-d case claim to use 'Mean field theory' to resolve this ... somehow. Having come up short on grasping Mean Field Theory, I attempted what made sense to me; I added two equations by integrating other squares.

Using the same notation as was utilized in equations 22, 23, and 31 to abreviate the integrations, integrating over four cells instead of only two yields the following set of equations,

$$\begin{aligned}
\left(b_1 I_y^{j-1} + b_2 Q_y^{j-1}\right)\left(c_1 I_x^{i-1} + c_2 Q_x^{i-1}\right) &= P_{i,j-1} \\
\left(b_1 I_y^{j} + b_2 Q_y^{j}\right)\left(c_1 I_x^{i-1} + c_2 Q_x^{i-1}\right) &= P_{i-1,j} \\
\left(b_1 I_y^{j} + b_2 Q_y^{j}\right)\left(c_1 I_x^{i} + c_2 Q_x^{i}\right) &= P_{i,j} \\
\left(b_1 I_y^{j} + b_2 Q_y^{j}\right)\left(c_1 I_x^{i+1} + c_2 Q_x^{j+1}\right) &= P_{i,j+1}
\end{aligned} \tag{46}$$

...Attempts to solve this system resulted in cancellations rather than isolating variables.

# 5 Conclusion

The mathematics involved in random walk processes are non-trivial. Here it has been shown that in just one dimension a number of idealized assumptions needed to be made in order to make use the of the Chapman-Kolmogorov equation. These assumptions were that the steady state solution is a good approximation, the system is ergodic, and that the cell size is small enough to make it such that turning on chemical reactions does not significantly affect the steady state solution. As shown in the derivation offered in the appendix, the Chapman Kolmogorov equation arose from invoking the Markov property which is to say that future jumps are not correlated to past jumps, so even using this equation implies idealized assumptions.

Nevertheless, the algorithm developed in equation 36 is a very simple yet very useful tool for modeling a wide number of molecular phenomena such as a molecule traversing a linear polymer chain. Random walks can also be used to model any number of less tangible processes such as stock-prices or decision making processes. While lacking in computational results, the focus of this paper was to present a solid introduction to what is a rhobust and widely applicable algorithm which is capable of making use of the Chapman-Kolmogorov equations.

# 6 APPENDIX

## 6.1 Lindeberg Condition

We wish to establish the equivalence between the Continuity condition (equation 3) and something called the Lindeberg condition. Reference [6] claims they are equivalent.

Let $X_1, X_2, ..., X_n$ denote a set of random independent variables with finite variance. For the moment we are assuming the individual pdf's have different forms, i.e. they are not identical (not i.i.d.). Independence between variables implies the variables may be arranged into a triangular array. This is true in general because a matrix of full rank has the Gaussian reduced form of a diagonal matrix. Let us call such a diagonal matrix $A_{n,k}$.

$$A_{n,k} = \begin{pmatrix} X_{11} & 0 & 0 & \dots 0 & \\ X_{21} & X_{22} & 0 & \dots 0 & \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ X_{d1} & X_{d2} & X_{d3} & \dots & X_{dn} \end{pmatrix}$$

$A_{n,k}$ may be denoted as $A_{n,k} = X_{nk} I_{k<n}$ Where $I_{k<n}$ denotes the Indicator function which has value 1 when $k < n$ zero otherwise.

If it is the *average* value of these variables we are concerned with measuring, then it is of interest to know under what conditions that the average value of such a set will give a normal distribution $N(\sigma^2, \mu)$. Explicitly, we wish to know under what conditions a random variable $X$ may be considered to have a distribution of the form,

$$X \sim N[0,1]: \qquad P(X) \propto e^{-x^2/2\sigma} \tag{A.5}$$

That is, we wish to know under what conditions the central limit theorem applies. *Lindeberg's condition* gives the criteria that a set of random and independent variables must meet in order for the central limit theorem to apply.

Let $Y_{nk} = X_{nk}$ , $T_n = \sum_{n=1}^{n} Y_{nk}$, and $S_n^2 = Var(T_n) = n\sigma^2$. The Lindeberg condition states that if the following condition,

$$\lim_{n\to\infty} \frac{1}{s_n^2} \sum_{n=1}^{n} E[X_{nk}^2 I_{|X_{nk}| \geq \varepsilon\sigma\sqrt{n}}] = 0 \tag{A.6}$$

is met then this implies[7],

$$\lim_{n\to\infty} \frac{T_n}{s_n} = \frac{\sqrt{n}\sum_{n=1}^{n} X_{nk}}{\sigma} \to N[0,1] \tag{A.7}$$

Here the Indicator function has value $I_A = 1$ when $x$ belongs to set $A$ and is zero otherwise. Obviously the indicator function will go to zero upon taking the limit inside of the summation so long as $X_{nk}$ are of finite value and number. For independent random variables, the variance of a sum is the sum of the variances of $X_{nk}$, hence so long as $Var[X_{nk}] < \infty$ then $s_n$ is also of finite value and the Lindeberg condition (A.6) is satisfied, so we can presume A.7 holds.

We want to show the following continuity condition

$$\lim_{\Delta t\to\infty} \frac{1}{\Delta t} \int_{|x-z|>\varepsilon} P(x, t+\Delta t|z, t)dx = 0 \tag{A.8}$$

is equivalent to [A.6] as is claimed in [6]. The continutity condition is the starting point of the Chapman-Kolmogorov equation to come. An attempted proof of the equivalence is as follows;

$$\begin{aligned}\int_{|x-z|>\varepsilon} P(x,t+\Delta t|z,t)d\mathbf{x} &= \int_{|x-z|>\varepsilon} \frac{|\mathbf{x}-\mathbf{z}|^{2+\xi}}{|\mathbf{x}-\mathbf{z}|^{2+\xi}} P(\mathbf{x},t+\Delta t|\mathbf{z},t)d\mathbf{x} = 0 \\ &\leq \frac{1}{\varepsilon^{2+\xi}} \int |\mathbf{x}-\mathbf{z}|^{2+\xi} P(\mathbf{x},t+\Delta t|\mathbf{z},t)d\mathbf{x} = 0 \\ &\vdots \\ &= \frac{1}{\varepsilon^{2+\xi}} E\left[|\mathbf{x}-\mathbf{z}|^{2+\xi}\right] \quad \rightarrow \quad 0\end{aligned}$$

Where the omitted steps likely involve some concepts from probability theory (e.g. conditional expectation value identities). Also see equation 2.14 or reference [8] for a delta function transform involving marginal pdf's which allows for the addition of two stochastic variables. This attempted proof of the equivalence was largely an attempt to relate the results of references [7] and [9] (see top of pages 54-55). [6] [whose proof of the Chapman Kolmogorov equation I followed] begins his proof by simply stating the equivalence of equation A.6 and A.8, and he cites for this equivalence the work of Skorokhod [10] whose work was unavailable to me.

Noting that equation A.6 is to be read as, 'if for all $\varepsilon > 0$ there exists some $n > 0$ for which this condition is true...', this has the form of the the Lyapunov condition, which is a stronger condition and implies the Lindeberg condtion [7]. This is not difficult to see; given that $|\mathbf{x}-\mathbf{z}| < \varepsilon$ we have,

$$|\mathbf{x}-\mathbf{z}|^2 = \frac{|\mathbf{x}-\mathbf{z}|^{2+\delta}}{|\mathbf{x}-\mathbf{z}|^{\delta}} < \frac{|\mathbf{x}-\mathbf{z}|^{2+\delta}}{\varepsilon^{\delta}}$$

and the Lindeberg condition applied to the difference of two variables becomes,

$$\begin{aligned}&\lim_{n\to\infty} \frac{1}{s_n^2} \sum_{n=1}^{n} E\Big[|X_{nk}-Z_{nk}|^2 I_{|X_{nk}|\geq\varepsilon\sigma\sqrt{n}}\Big] \\ &< \frac{1}{\varepsilon^{\delta}} E\Big[|X_{nk}-Z_{nk}|^{2+\delta}\Big] \quad \rightarrow \quad 0\end{aligned}$$

The above result follows from the assumption that the variables are of finite value and number, hence their expected value is finite and the denominator will overcome the numerator. It would be too much of a digression from the authors paper to go further into these proofs, so we conclude by noting that the Lyapunov exponent relates the consecutive limit of two infinitesimally close variables and can be used to infer when chaos sets in. Naturally, one can begin to see the parallels between this and stochastic (random) processes.

Note that *somewhwere* in this derivation is likely the key to a very important property of stochastic variables, namely that they have variance $\sigma^2 = \Delta t$.

Finally, in lieu of having completed a satisfactory proof of the equivalence between the continuity condition and the Lindeberg condition, it will be demonstrated that the associated Gaussian (normal) probability density function does indeed vanish quicker than $\Delta t \to 0$ which is what the continuity condition implies. We want to show that,

$$\lim_{\Delta t\to 0} \frac{1}{\sqrt{2\pi\Delta t}} e^{-\frac{x^2}{2\Delta t}} = 0$$

and this is trivial using the more rigorous mathematical definition of a limit which states that, if for all $\varepsilon > 0$ there exists some $\delta > 0$ for which the following condition holds,

$$|f(x) - f(a)| < \varepsilon$$

then the limit of $f(x)$ as $x \to a = f(a)$. For the Gaussian distribution we have $f(a) \to 0$ as $x \to 0$ which gives,

$$\begin{aligned} \left| \frac{1}{\sqrt{2\pi\Delta t}} e^{-\frac{x^2}{2\Delta t}} - 0 \right| &< \varepsilon \\ \left| e^{-\frac{x^2}{2\Delta t}} \right| &< \varepsilon\sqrt{2\pi\Delta t} \\ |x| &< \sqrt{2\Delta t \ln\left(\varepsilon\sqrt{2\pi\Delta t}\right)} \, |* \end{aligned}$$

So, given any $\Delta t > 0$ and $\varepsilon > 0$ we can choose $\delta = x$ such that this condition is satisfied, and this completes the proof of the continuity condition (for normally distrubted variables).

# 7 Chapman Kolmogorov Equation

Chapman Attempted to bridge the gap between the differential form of komogorov's equation and the discrete form of the probability flux, or the *master equation*.

Consider the random walk in which it is given that a particle is located at point $x_1$ at time $t$. At time $t + \Delta t$ we wish to know the probability that it jumps to a neighbor $x_3$ that is two steps away (so there is some intermediary $x_2$ sitting between them). In terms of conditional probabilities we consider the reverse situation; given that a particle occupies position $x_3$ at time $t_3$, what is the probability that it came from $x_1, t_1$? We denote the conditional probability as $P(x_1, t_1|x_3, t_3)$. It is a requirement of the joint pdf that it be normalized with respect to all variables involved, and we can exploit this fact to write,

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

Combining this with the usual conditional probability identity (Baye's Rule) gives,

$$\begin{aligned} P(x_1, t_1, |x_3, t_3) &= \int dx_2 P(x_1, t_1, x_2, t_2|x_3, t_3) \\ &= \int dx_2 \frac{P(x_1, t_1, x_2, t_2, x_3, t_3)}{P(x_3, t_3)} \\ &= \int dx_2 \frac{P(x_1, t_1, x_2, t_2, x_3, t_3)}{P(x_2, t_2, x_3, t_3)} \frac{P(x_2, t_2, x_3, t_3)}{P(x_3, t_3)} \\ &= \int dx_2 P(x_1, t_1|x_2, t_2, x_3, t_3) P(x_2, t_2|x_3, t_3) \end{aligned} \tag{A.9}$$

Here is where the uncorrelated random walk and the correlated random walk diverge. For the random walk we have a *Markov process* which is to say that what happens next is in no way affected by previous jumps; the future is independent of the past, but only on the current position. In terms of conditional probabilities this implies we can write the above relation as,

$$P(x_1, t_1, |x_3, t_3) = \int dx_2 P(x_1, t_1|x_2, t_2) P(x_2, t_2|x_3, t_3) \qquad \longrightarrow \qquad uncorrelated \quad walk \tag{A.10}$$

The *conditional Lindeberg condition*[**?**] gives the criteria that a collection of random variables must meet in order for them to be of a normal distribution. An alternative form of the Lindeberg condition is that, for all $\varepsilon > 0$, if the condition,

$$\begin{aligned} \lim_{\Delta t \to \infty} \frac{1}{\Delta t} \int_{|\mathbf{x}-z|>\varepsilon} |x - z|^2 P(\mathbf{x}, t + \Delta t|z, t) dx = 0 \\ \int_{|\mathbf{x}-z|>\varepsilon} d\mathbf{x} W(\mathbf{x}|\mathbf{z}, t) = 0 \end{aligned} \tag{A.11}$$

is satisfied then the assumption of normality (equation 7) applies[4]. $W(\mathbf{x}|\mathbf{z}, t)$ in the above represents the transition rate. Essentially what this condition is saying is that the transition probability goes to zero faster than $\Delta t$ (see figure 2), and this is somewhat intuitively obvious by looking at the form of 7. Equation 10 is also referred to as the continuity condition as it implies we have continuous sample paths (but not states) for the probability. For an *attempted* formal proof of this equivalence of the two forms of the Lindeberg condition (see equations A.6 and A.8).

[4]For the interested reader, what follows is a tedious derivation in which some of calculations are not 100%. The entirety of the derivation to come follows C.W. Gardner[6]

Kolmogorov began his derivation by making the assumption that the first and second moments $E[|\mathbf{x}\text{-}\mathbf{y}|]$ and $E[|\mathbf{x}\text{-}\mathbf{y}|^2]$do not vanish. That is to say,

i)

$$\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|\mathbf{x}-\mathbf{y}|<\varepsilon} d\mathbf{x}(x_i-y_i)P(\mathbf{x},t+\Delta t|\mathbf{y},t)=A_i(\mathbf{y},t)+O(\varepsilon)$$

ii)

$$\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|\mathbf{x}-\mathbf{y}|<\varepsilon} d\mathbf{x}(x_i-y_i)(x_j-y_j)P(\mathbf{x},t+\Delta t|\mathbf{y},t)=B_{i,j}(\mathbf{y},t)+O(\varepsilon)$$

Where $A_i(\mathbf{y},t)$ represents the *drift vector* and $B_{ij}(\mathbf{y})$ the *diffusion matrix*. All higher moments will vanish. Consider the third moment,

$$\begin{aligned}
&\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|\mathbf{x}-\mathbf{y}|<\varepsilon}|\mathbf{x}-\mathbf{y}|^3 d\mathbf{x}\\
&=\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|\mathbf{x}-\mathbf{y}|<\varepsilon}|\mathbf{x}-\mathbf{y}||\mathbf{x}-\mathbf{y}|^2 d\mathbf{x}\\
&\leq\lim_{\Delta t\to 0}\varepsilon\frac{1}{\Delta t}\int|\mathbf{x}-\mathbf{y}|^2 d\mathbf{x}\\
&=\varepsilon B_{ij}(\mathbf{y},t)\\
&=0
\end{aligned}\tag{A.12}$$

The second to last line resulted from an application of condition ii. The last line results from the strict mathematical statement which was stipulated in 10 that *for all* $\varepsilon>0$ this condition should hold, and the only scalar value whose magnitude is less than $\varepsilon$ *for all* $\varepsilon>0$ is simply zero.

Conditions *i* and *ii* are just assumptions. Similar to the Schrodinger equation, there were likely good reasons why Kolmogorov made these assumptions (probably related to the Lindeberg condition), but one need not waste too much time trying to ask why these assumptions were made. Many mathematicians are today playing around with alternative schemes.

We now take the expectation value of the time derivative of the continuity condition (10) and apply the Markov property,

$$\begin{aligned}
&\partial t\int d\mathbf{x}f(\mathbf{x})P(\mathbf{x},t|\mathbf{z},t')\\
&\quad=\lim_{\Delta t\to 0}\frac{1}{\Delta t}\{\int d\mathbf{x}f(\mathbf{x})[P(\mathbf{x},t+\Delta t|\mathbf{z},t')-P(\mathbf{x},t|\mathbf{z},t')]\}\\
&\quad=\lim_{\Delta t\to 0}\frac{1}{\Delta t}\left\{\int d\mathbf{x}\int d\mathbf{y}f(\mathbf{x})P(\mathbf{x},t+\Delta t|\mathbf{y},t)P(\mathbf{y},t|\mathbf{z},t')-\int d\mathbf{y}f(\mathbf{y})P(\mathbf{y},t|\mathbf{z},t')\right\}
\end{aligned}\tag{A.13}$$

Where in the last line we used the Chapman-Kolmogorov Equation on the first term and $E[f(\mathbf{x})]=E[f(\mathbf{z})]$ on the second which follows from the fact that expectation values are taken over all space. on the second term. Taylor expanding the function $f(x)$ gives,

$$f(\mathbf{x})=f(\mathbf{y})+\sum_i\frac{\partial f(\mathbf{y})}{\partial y_i\partial y_j}(x_i-y_i)+\frac{1}{2}\sum_{i,j}\frac{\partial^2 f(\mathbf{y})}{\partial y_i^2}(x_i-y_i)(x_j-y_j)+|\mathbf{x}\text{-}\mathbf{y}|^2\frac{\partial^2 f(\mathbf{y})}{\partial y_i^2}\tag{A.14}$$

Gardner [2] claims that the second order diagonal term goes to zero as $|\mathbf{x}\text{-}\mathbf{y}|\to 0$. Why this is so I am not entirely sure - it probably relates to the continuity condition, and this *almost* make sense if continuity implies continuous and

bounded derivatives (in which case $|\mathbf{x}\text{-}\mathbf{y}| \to 0$ would imply that the whole term goes to zero), but it does not - it is entirely possible to have a continuous sample paths but discontinuous (purely vertical jumps). As is in line with the Lindeberg/continuity condition, we divide the integration into two regions; $|\mathbf{x}-\mathbf{y}| < \varepsilon$ and $|\mathbf{x}-\mathbf{y}| \geq \varepsilon$. Assuming the second order diagonal term to be zero, substituting 15 into 14 gives[6],

$$\begin{aligned}
\partial t \int d\mathbf{x} f(\mathbf{x}) P(\mathbf{x},t|\mathbf{z},t') = & \\
& \iint_{|\mathbf{x}\text{-}\mathbf{y}|<\varepsilon} d\mathbf{x}d\mathbf{y} \left\{ \sum_i (x_i - y_i)\frac{\partial f}{\partial y_i} + \frac{1}{2}\sum_{i,j} \frac{\partial^2 f}{\partial y_i \partial y_j}(x_i - y_i)(x_j - y_j) \right\} \\
& \qquad \times P(\mathbf{x},t+\Delta t|\mathbf{y},t) P(\mathbf{y},t|\mathbf{z},t') \\
& + \iint_{|\mathbf{x}\text{-}\mathbf{y}|\geq\varepsilon} d\mathbf{x}d\mathbf{y} f(\mathbf{x}) P(\mathbf{x},t+\Delta t|\mathbf{y},t) P(\mathbf{y},t|\mathbf{z},t') \\
& + \iint_{|\mathbf{x}\text{-}\mathbf{y}|<\varepsilon} d\mathbf{x}d\mathbf{y} f(\mathbf{y}) P(\mathbf{x},t+\Delta t|\mathbf{y},t) P(\mathbf{y},t|\mathbf{z},t') \\
& - \iint d\mathbf{x}d\mathbf{y} f(\mathbf{y}) P(\mathbf{x},t+\Delta t|\mathbf{y},t) P(\mathbf{y},t|\mathbf{z},t')
\end{aligned} \tag{A.15}$$

Integrating by parts and noting that the probability density goes to zero at infinity, we have,

$$\begin{aligned}
\partial t \int d\mathbf{x} f(\mathbf{x}) P(\mathbf{x},t|\mathbf{z},t') = & \int d\mathbf{y} f(\mathbf{y}) \left\{ \sum_i A_i(\mathbf{y}) \frac{\partial f}{\partial y_i} + \frac{1}{2}\sum_{i,j} B_{ij}(\mathbf{y}) \frac{\partial^2 f}{\partial y_i \partial y_j} \right\} P(\mathbf{y},t|\mathbf{z},t') \\
& + \iint_{|\mathbf{y}-\mathbf{x}|\geq\varepsilon} d\mathbf{x}d\mathbf{y} f(\mathbf{y}) \left\{ W(\mathbf{y},t|\mathbf{x},t) P(\mathbf{x},t|\mathbf{z},t') - W(\mathbf{x},t|\mathbf{y},t) P(\mathbf{y},t|\mathbf{z},t') \right\} + O(\varepsilon)
\end{aligned} \tag{A.16}$$

So far we've considered $f(\mathbf{y})$ to be an arbitrary function, hence we can define it to simply be one within the region of integration and zero at the boundary and outside of it. The reason for choosing it to be zero at the boundary relates to the fact that one cannot assume taking differentials of $A_i, B_{i,j}$ is possible on the boundary[2]. . For interior points we then have the final relation,

$$\begin{aligned}
\boxed{\partial_t P(\mathbf{y},t|\mathbf{z},t') = -\sum_i \frac{\partial}{\partial y_i} A_i(\mathbf{y},t) P(\mathbf{y},t|\mathbf{z},t') + \frac{1}{2}\sum_{i,j} \frac{\partial^2}{\partial y_i \partial y_j} B_{ij}(\mathbf{y},t) P(\mathbf{y},t|\mathbf{z},t')} & \\
+ \int d\mathbf{x} W(\mathbf{y}|\mathbf{x},t) P(\mathbf{x},t|\mathbf{z},t') - W(\mathbf{x}|\mathbf{y},t) P(\mathbf{y},t|\mathbf{z},t') &
\end{aligned} \tag{A.17}$$

If the drift vector and diffusion matrix are both zero we recover the *master equation*

$$\boxed{\partial_t P(\mathbf{y},t|\mathbf{z},t') = \int d\mathbf{x} \left\{ W(\mathbf{y},t|\mathbf{x},t) P(\mathbf{x},t|\mathbf{z},t') - W(\mathbf{x},t|\mathbf{y},t) P(\mathbf{y},t|\mathbf{z},t') \right\}} \tag{A.18}$$

If $W(\mathbf{x},t|\mathbf{y},t) = W(\mathbf{y},t|\mathbf{x},t) = 0$ we get the Fokker-Plank Equation

$$\boxed{\partial_t P(\mathbf{y},t|\mathbf{z},t') = -\sum_i \frac{\partial}{\partial y_i} A_i(\mathbf{y},t) P(\mathbf{y},t|\mathbf{z},t') + \frac{1}{2}\sum_{i,j} B_{ij}(\mathbf{y},t) P(\mathbf{y},t|\mathbf{z},t')} \tag{A.19}$$

which is diffusion equation. Note that none of the equations derived thus far are stochastic but are entirely deterministic.

# References


[1] Jianhua Xing, Hongyun Wang,y and George Oster *From Continuum Fokker-Planck Models to Discrete Kinetic Models.*

[2] Wang, H., C. Peskin, and T. Elston *A robust numerical algorithm for studying biomolecular transport processes* 2003. . J. Theor. Biol. 221:491–511.

[3] HI JIN AND XIN WEN. *HAMILTONIAN-PRESERVING SCHEMES FOR THE LIOUVILLE EQUATION WITH DISCONTINUOUS POTENTIALS.*

[4] Pasquali, Sara. *The stochastic logistic equation : stationary solutions and their stability.* Rendiconti del Seminario Matematico della Università di Padova, tome 106 (2001), p. 165-18

[5] E. Zauderer. *Correlated Random Walks, Hyperbolic Systems, and Fokker-Plank Equations.* olytechnic University, Six Metrotech Center, 1998.

[6] C.W. Gardner. *Handbook of Stochastic Methods, Third Ed.*. Springer, 2003.

[7] Lehmann, Ferguson. *CLT, Part II: Independent but not identically distributed.* available @: http://personal.psu.edu/drh20/asymp/fall2002/lectures/ln04.pdf

[8] J.L. García-Palacios
`Introduction to the theory of stochastic processes and Brownian motion problems` Universidad de Zaragonza, 2004

[9] G.A. Pavliotis
`Applied Stochastic Processes` Imperial College London, 2009

[10] J.L. García-Palacios
`The Theory of Stochastic Processes II` Authors: Gikhman, Iosif I., Skorokhod, Anatoli V., 1980

[11] Masayuki Ohzeki and Akihisa Ichiki
`Mathematical understanding of detailed balance condition violation and its application to Langevin dynamics` 2015 J. Phys.: Conf. Ser. 638 012003

[12] Not listed
`Lecture 12: Detailed balance and Eigenfunction methods`. available at: http://cims.nyu.edu/ holmes/teaching/asa19/handout_ Lecture12_ 2019.pdf