Derivation of a Fokker-Planck Equation

Drift and Diffusion of a Probability Density in State Space

Andrew Forrester August 31, 2011

1 Probabilistic Properties

To derive a Fokker-Planck equation, which deals with probabilities, we should start by stating the probabilistic notation and properties we will use. The probability that event A occurs is $P[A]$. Summing over all possible mutually exclusive events must yield a total of one: $\sum_A P[A] = 1$. If you include non-exlusive events, the probability sum can exceed one, so any sums will be performed over mutually exclusive events.[1] When summations are not involved, the events named can be partially or fully simultaneous, allowing illustration using Venn diagrams with overlapping regions. The probability that both A and B occur is a joint probability $P[A, B]$. The probability that A occurs given that B occurs is a conditional probability $P[A|B]$. A joint-conditional probability property follows:

$$P[A, B] \quad = \quad P[A|B]\ P[B].$$

This property can be illustrated nicely with a Venn diagram. This property generalizes in many ways, for instance, $P[A, B, C] = P[A|B, C]\ P[B, C]$. However, we will not need any of these generalizations.

Let's say the state of a Brownian object[2] is described by a stochastic variable $A(t)$. Let the notation $P[A(t), A'(t'), A''(t'')]$ denote the probability that the object occupies the state A'' at time t'', A' at time t', and A at time t.[3] Since the object must occupy a particular state at time t', we have the following relation,

$$\sum_{A'} P[A(t), A'(t'), A''(t'')] \quad = \quad P[A(t), A''(t'')],$$

given that there are a countable number of states available. Given a continuum of states, we have

$$\int \mathrm{d}X'\ p[X(t), X'(t'), X''(t'')] \quad = \quad p[X(t), X''(t'')],$$

where p is a probability distribution or probability density instead of a probability P. This general property is displayed in the Chapman-Kolmogorov equation:

$$\int \mathrm{d}X_i\ p[X_1, \ldots, X_{i-1}, X_i, X_{i+1}, \ldots, X_n] \quad = \quad p[X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n].$$

Here are some more properties that help elucidate the meaning of the probabilities and probability distributions:

$$\begin{array}{rclrcl}
\sum_A P[A] & = & 1 & \int \mathrm{d}X\ p[X] & = & 1 \\
\sum_A P[A|B] & = & 1 & \int \mathrm{d}X\ p[X|Y] & = & 1 \\
\sum_{A,B} P[A, B] & = & 1 & \int \mathrm{d}X\, \mathrm{d}Y\ p[X, Y] & = & 1 \\
\sum_A P[A, B] & = & P[B] & \int \mathrm{d}X\ p[X, Y] & = & p[Y] \\
\sum_B P[A, B] & = & P[A] & \int \mathrm{d}Y\ p[X, Y] & = & p[X]
\end{array}$$

Note that P is always dimensionless (or has dimensions of "probability") but a probability density $p[X]$ has dimensions of $[X]^{-1}$, where $[X]$ is the dimension, such as length or energy, of the stochastic variable X. A probability density $p[X, Y, Z|U, V]$ has dimensions of $[X]^{-1}[Y]^{-1}[Z]^{-1}$.

[1] For example, let's say you flip a coin with colored sides where one side is red and the other is half-red half-blue. Take the set of events to be $A = A1 =$ "the upper side has red on it" and $A = A2 =$ "the upper side has blue on it". Although this may seem like a reasonable all-inclusive set of events, both may simultaneously occur and $\sum_A P[A] = 1.5 \neq 1$.

[2] See the Stochastic Vocabulary in the Appendix for elaboration on terms such as "Brownian object".

[3] This is slightly sloppy notation since the function P must know about both A and t. We'll eventually use more proper notation, such as $P[A; t]$, in the next section.

2 A Fokker-Planck Equation

We will now derive a very general (forward) Fokker-Planck equation (FPE), which describes the time-evolution of a probability density in state space, considering drift and diffusion but not higher-order migration. [4] In this derivation we start with an integral equation expressing two probabilistic properties we introduced in the last section, we use Taylor expansions, pull the integral out of the derivatives, and absorb the integral into definitions of some quantities to change to a differential equation, and we finally reduce the number of terms by assuming higher-order-derivative terms are negligible.

We use a d-dimensional stochastic variable $\mathbf{R}$ to represent the state of a Brownian object. It may simply represent position or it may include quantum numbers as well as position, velocity, momentum, or other state variables. We assume that the quantum numbers, if there are any, are large enough that differences in them can be considered small, so $\mathbf{R}$ can be considered continuous: $\mathbf{R} \in \mathbb{R}^d$. [5] We also assume that probability distributions such as $p[\mathbf{R}]$ are continuous and differentiable so we can take derivatives.

Using (1) the last equation from the previous section, $p[X] = \int \mathrm{d}Y\ p[X, Y]$, and (2) the joint-conditional probability property $p[X, Y] = p[X|Y]\ p[Y]$, we have

$$p[\mathbf{R}(t+\Delta t)] = \int \mathrm{d}^d R'\ p[\mathbf{R}(t+\Delta t), \mathbf{R}'(t)] = \int \mathrm{d}^d R'\ p[\mathbf{R}(t+\Delta t)|\mathbf{R}'(t)]\ p[\mathbf{R}'(t)].$$

We should switch to a more rigorous notation to aid in the impending mathematical manipulations:

$$p[\mathbf{R}, t+\Delta t] = \int \mathrm{d}^d R'\ p[\mathbf{R}|\mathbf{R}'; \Delta t, t]\ p[\mathbf{R}', t].$$

We shall approximate this exact relationship by Taylor-expanding each side, but again we should change notation. The notation above makes the joint-conditional property more apparently true, but a different notation will be more suitable for the Taylor expansions. Let's define a change-in-state holor $\boldsymbol{\xi} \equiv \Delta\mathbf{R} = \mathbf{R} - \mathbf{R}'$, denoting the change from $\mathbf{R}'$ to $\mathbf{R}$. We'll use $\boldsymbol{\xi}$ in the equations below rather than $\Delta\mathbf{R}$ because the differential $\mathrm{d}^d\xi$ looks better than $\mathrm{d}^d(\Delta R)$. We can rewrite the above equation like so:

$$p[\mathbf{R}, t+\Delta t] = \int \mathrm{d}^d\xi\ p[\mathbf{R}-\boldsymbol{\xi}, \boldsymbol{\xi}; \Delta t, t]\ p[\mathbf{R}-\boldsymbol{\xi}, t].$$

Note that $\mathrm{d}^d R' = (-1)^d \mathrm{d}^d\xi$, but if we change the direction or sense of the integrations, then we can replace $\int \mathrm{d}^d R'$ with $\int \mathrm{d}^d\xi$. [6] Now we can apply on the left side a one-dimensional Taylor expansion (See Appendix B.1) about time t in terms of Δt and apply on the right side a multi-dimensional product Taylor expansion (see Appendix B.4) about state $\mathbf{R}$ in terms of $-\boldsymbol{\xi}$, and we'll define a new quantity in the last step:

$$\begin{aligned}
\sum_{n=0}^{\infty} \frac{(\Delta t)^n}{n!} \partial_t{}^n p[\mathbf{R}, t] &= \int \mathrm{d}^d\xi \sum_{|\alpha|\geq 0} \frac{(-\xi)^\alpha}{\alpha!} \partial_R^\alpha \left\{ p[\mathbf{R}, \boldsymbol{\xi}; \Delta t, t]\ p[\mathbf{R}, t] \right\} \\
&= \sum_{|\alpha|\geq 0} (-1)^\alpha \partial_R^\alpha \left\{ \frac{1}{\alpha!} \left(\int \mathrm{d}^d\xi\ \xi^\alpha\ p[\mathbf{R}, \boldsymbol{\xi}; \Delta t, t] \right) p[\mathbf{R}, t] \right\} \\
&= \sum_{|\alpha|\geq 0} (-1)^\alpha \partial_R^\alpha \left\{ \frac{\langle \xi^\alpha \rangle_{\Delta t}(\mathbf{R}, t)}{\alpha!}\ p[\mathbf{R}, t] \right\},
\end{aligned}$$

where we've defined $\langle \xi^\alpha \rangle_{\Delta t}(\mathbf{R}, t)$ by

$$\langle \xi^\alpha \rangle_{\Delta t}(\mathbf{R}, t) \equiv \int \mathrm{d}^d\xi\ \xi^\alpha\ p[\mathbf{R}, \boldsymbol{\xi}; \Delta t, t],$$

[4] I consider this derivation an improvement over and abstraction of the derivations by Chandrasekhar[1] and Wilde and Singh[2], where the conceptual steps are made clear and the abstractions are explained in the appendices.

[5] $\mathbf{R}$ is a univalent holor of plethos d. (A holor is a mathematical entity that is made up of one or more independent quantities. A holor may be multiply-indexed, like a tensor, but its transformation properties, under rotation, say, are not necessarily specified.) $\mathbf{R}$ doesn't necessarily transform as a Euclidean vector, so I won't call it a vector to prevent confusion amongst physicists.

[6] Dear reader, please let me know if something is wrong with that last step; it seems fishy to me.

which gives the elements of a mean-transition-increment holor to a state $\mathbf{R}$ at time t during a time Δt. Actually, this defines one holor for every set of multi-indices α with the same magnitude. While we're at it, let's define a set of "migration holors", which are also sometimes called kinetic terms (pg 55 of [3]) or generalized diffusion tensors[4],

$$M_{|\alpha|}^{\{\alpha\}}(\mathbf{R},t) \equiv \lim_{\Delta t\to 0} \frac{\langle \xi^\alpha\rangle_{\Delta t}(\mathbf{R},t)}{\alpha!\Delta t},$$

which are mean-transition-rate holors, giving mean rates of various transitions in state space. For more elaboration on the meaning of this notation, see Appendix C. After we divide both sides of the equation by Δt, we can write the resulting equation in terms of these migration holors. In applications, the limit in the definition above may actually have Δt go to some small non-zero value, since there may be wild fluctuations on small time-scales but a smoother behavior when averaged over a small non-zero duration.

Now, after cancelling the first term[7] in each series, dividing by Δt, and letting Δt be small enough[8], we get

$$\begin{aligned}\sum_{n=1}^{\infty} \frac{(\Delta t)^{n-1}}{n!}\partial_t{}^n p[\mathbf{R},t] &= \sum_{|\alpha|\geq 1} (-1)^\alpha \partial_R^\alpha \left\{ \frac{\langle \xi^\alpha\rangle_{\Delta t}(\mathbf{R},t)}{\alpha!\Delta t}\; p[\mathbf{R},t]\right\} \\ &= \sum_{|\alpha|\geq 1} (-1)^\alpha \partial_R^\alpha \left\{ M_{|\alpha|}^{\{\alpha\}}(\mathbf{R},t)\; p[\mathbf{R},t]\right\}.\end{aligned}$$

If we expand these series out, we can switch the right side into Einstein summation notation:

$$\begin{aligned}&\partial_t\, p[\mathbf{R},t] + \frac{\Delta t}{2}\partial_t{}^2 p[\mathbf{R},t] + \frac{(\Delta t)^2}{6}\partial_t{}^3 p[\mathbf{R},t] + \cdots \\ &\quad = -\partial_i \left\{ M_1^i(\mathbf{R},t)\; p[\mathbf{R},t]\right\} + \partial_j\partial_k \left\{ M_2^{jk}(\mathbf{R},t)\; p[\mathbf{R},t]\right\} - \partial_\ell\partial_m\partial_n \left\{ M_3^{\ell mn}(\mathbf{R},t)\; p[\mathbf{R},t]\right\} + \cdots\end{aligned}$$

Now, to get the Fokker-Planck equation we seek, we just assume all terms except the first term on the left side and the first two terms on the right side are negligible:

$$\boxed{\partial_t\, p[\mathbf{R},t] = -\partial_i \left\{ M_1^i(\mathbf{R},t)\; p[\mathbf{R},t]\right\} + \partial_j\partial_k \left\{ M_2^{jk}(\mathbf{R},t)\; p[\mathbf{R},t]\right\}}$$

As implied in the general definition above, these two migration holors are defined by

$$\begin{aligned}M_1^i(\mathbf{R},t) &\equiv \lim_{\Delta t\to 0} \frac{\langle \Delta R_i\rangle_{\Delta t}(\mathbf{R},t)}{\Delta t}, \\ M_2^{ij}(\mathbf{R},t) &\equiv \lim_{\Delta t\to 0} \frac{\langle \Delta R_i\, \Delta R_j\rangle_{\Delta t}(\mathbf{R},t)}{2\Delta t}.\end{aligned}$$

M_1^i can be called the drift holor since it describes the mean drift through state space. M_1^{ij} can be called the diffusion holor since it describes the mean diffusion through state space, like a diffusion tensor describes real-space diffusion.

3 A Note

It is important to note that the Fokker-Planck equation is a continuity equation: it is a local conservation of probability equation. Perhaps you are familiar with the charge continuity equation in electrodynamics,

$$\partial_t\, \rho = -\boldsymbol{\nabla}\cdot\mathbf{J} = -\partial_i J^i,$$

[7] Since $\left\langle \xi^{\langle 0,0,\ldots,0\rangle}\right\rangle_{\Delta t}(\mathbf{R},t) = \int \mathrm{d}^d\xi\ \xi^{\langle 0,0,\ldots,0\rangle}\ p[\mathbf{R},\boldsymbol{\xi};\Delta t,t] = \int \mathrm{d}^d\xi\ p[\mathbf{R},\boldsymbol{\xi};\Delta t,t] = 1$, the first term in the series on the right is $p[\mathbf{R},t]$, as is the first term in the series on the left.

[8] If Δt must go to zero, then only one term will remain on the left side.

where ρ is the charge density and $\mathbf{J}$ is the charge current density (in real space). Well, the Fokker-Planck equation can be written in this form and we can solve for the probability current density (in state space):

$$\partial_t \, p[\mathbf{R},t] \quad = \quad -\partial_i \left\{ M_1^i(\mathbf{R},t)\; p[\mathbf{R},t] - \partial_j \left[M_2^{ij}(\mathbf{R},t)\; p[\mathbf{R},t] \right] \right\} ,$$

$$J^i = M_1^i(\mathbf{R},t)\; p[\mathbf{R},t] - \partial_j \left[M_2^{ij}(\mathbf{R},t)\; p[\mathbf{R},t] \right] .$$

The total probability (usually assumed to be one) could be lost or gained at the boundaries of the state space if there is a loss or gain in the number of Brownian objects, or the Fokker-Planck equation could be reformulated in terms of number-density rather than probability density (of Brownian objects). Furthermore, one could add a source-sink term if there is injection/creation or loss/annihilation of Brownian objects within the boundaries of the state space.

4 A Warning

It seems that many times when the Fokker-Planck equation is used, it comes in a simplified form where the diffusion holor has been pulled out of one derivative:

$$\partial_t \, p[\mathbf{R},t] \quad = \quad \partial_i \left\{ -M_1^i(\mathbf{R},t)\; p[\mathbf{R},t] + M_2^{ij}(\mathbf{R},t)\; \partial_j \, p[\mathbf{R},t] \right\} .$$

In the cases where M_2 is constant or nearly constant, this is perfectly justifiable. But this (often unexplained) sleight of pen is performed in cases where M_2 is dependent on $\mathbf{R}$, and I think the reason for this move is a kind of "pragmatism": the equation may just be too hard to solve in its original form, so M_2 is pulled out of one derivative to make it solvable.

References

[1] S. Chandrasekhar: Stochastic Problems in Physics and Astronomy, *Modern Reviews of Physics*, 1943 January, Vol. 15, No. 1, pp. 1–89.

[2] Richard E. Wilde and Surjit Singh: *Statistical Mechanics: Fundamentals and Modern Applications*, John Wiley & Sons (1998)

[3] V. S. Anishchenko, V. V. Astakhov, A. B. Neiman, T. E. Vadivasova, and L. Schimansky-Geier: *Nonlinear Dynamics of Chaotic and Stochastic Systems: Tutorial and Modern Developments*, Springer (2002)

[4] C. Liu, R. Bammer, B. Acar, and M. E. Moseley: *Generalized Diffusion Tensor Imaging (GDTI) Using Higher Order Tensor (HOT) Statistics*, <http://cds.ismrm.org/ismrm-2003/0242.pdf> (accessed 2009 February 13)

5 Appendices

A Stochastics Vocabulary

- **Variable** – an algebraic symbol that can potentially take on a value from a set of values.

  When placed in an equation or inequality, the potential values are restricted to a subset of the original values, so as to satisfy the equation or inequality.
- **Process** – implies a change over time; A continuous action or series of changes over time.
  - **Continuous-time process** – a process where change occurs continuously over time.
  - **Discrete-time process** – (**Chain**) a process where change occurs in short bursts over time (as with occasional high-impulse collisions) or perhaps where the observation of the system occurs at discrete times so the change appears in discrete amounts. (think "chain of events" or "chain reaction")
- **Stochastic** – involving chance, randomness, or probability.
  - **Stochastic variable** X – a variable that takes on one of a set of possible values with a given set of probabilities. (A.k.a. random variable. I prefer "stochastic variable" over "random variable" because "random" seems to connote a uniform probability distribution, but the distribution may take any form. An even clearer name would be "probabilistic variable", but that's two more syllables.)

    Examples: the result of a die toss, the number of heads in a coin-tossing game, the energy eigenvalue of a 1D harmonic oscillator in quantum mechanics
    - **Probability distribution** $p[X]$ – (or probability density, as long as X is continuous) The function from the value space to the probability space that gives the probability of (or probability density at) any given value of the stochastic variable X. (In other words, the totality of the values of the random variable and its associated probabilities.)
  - **Stochastic process** – the change of a stochastic variable over time.

    An instance of the process can be described with $X(t)$, or the general process involving X can be described with $p[X(t)] = p[X, t]$.
    - The development of the stochastic variable $X(t)$ is not (apparently) governed by a deterministic equation.
    - The development of the associated probability distribution $p[X, t]$, however, is (or can be) governed by a deterministic equation.
- **Brownian motion** – the seemingly random movement of particles suspended in a liquid or gas or the mathematical model used to describe such random movements.
  - The randomly moving particles, called Brownian particles, are being struck by many much smaller particles, atoms, or molecules that make up the liquid or gas, and the Brownian motion is determined by the pattern of the strikes from these relatively invisible particles.
  - It is named after the Scottish botanist Robert Brown who, in 1827, was supposedly studying pollen particles floating in water under a microscope. With the mathematical analysis of Einstein (1905) and Smoluchowski (1906), scientific observation of this motion gave strong indirect evidence of the existence of atoms and molecules. (Lucretius, in his scientific poem *On the Nature of Things*, circa 60 BCE, had used Brownian motion as evidence for the atomic nature of matter.)
  - **Brownian object** – an object that exhibits Brownian motion – could be called a Brownian entity.

    Examples: Brownian particle, Brownian vortex loop, Brownian stock price

B Taylor Series

B.1 One-dimensional Taylor Series

A one-dimensional function $f : \mathbb{R} \to \mathbb{R}$ that is analytic equals its Taylor series in some interval. We can write the Taylor series in what I would call the "derivative series form" like so,

$$f(x+h) \quad = \quad \sum_n \frac{h^n}{n!}\, \mathrm{d}_x^n f(x),$$

where d_x^n is shorthand for $(\mathrm{d}/\mathrm{d}x)^n = \mathrm{d}^n/\mathrm{d}x^n$.

B.2 Multi-dimensional Taylor Series in Multi-index Notation

A multi-dimensional Taylor series of an analytic function $f : \mathbb{R}^n \to \mathbb{R}$ can be nicely written in what is called multi-index notation. We have $\mathbf{x} \in \mathbb{R}^n$ and $f(\mathbf{x}) \in \mathbb{R}$, and the multi-index α is an n-tuple of non-negative-integer-valued indices in $\mathbb{N}^n$. Here are the notational properties and definitions:

- $\alpha \in \mathbb{N}^n$
- $\alpha = \langle \alpha_1, \ldots, \alpha_n \rangle$
- Magnitude of α: $|\alpha| \equiv \alpha_1 + \cdots + \alpha_n$
- Factorial of α: $\alpha! \equiv \alpha_1! \cdots \alpha_n!$
- $x^\alpha \equiv x_1^{\alpha_1} \cdots x_n^{\alpha_n}$
- $\partial_x^\alpha = \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n} = (\partial/\partial x_1)^{\alpha_1} \cdots (\partial/\partial x_n)^{\alpha_n} = \dfrac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}$

Given these definitions the Taylor series of f about $\mathbf{x}$ in terms of elements of $\mathbf{h}$ is

$$f(\mathbf{x}+\mathbf{h}) = \sum_{|\alpha|\geq 0} \frac{h^\alpha}{\alpha!}\, \partial_x^\alpha f(\mathbf{x}).$$

Expanding this out, one gets what one expects:

$$\begin{aligned}
f(\mathbf{x}+\mathbf{h}) \quad &= \quad \frac{h^{\langle 0,0,\ldots,0\rangle}}{\langle 0,0,\ldots,0\rangle!}\, \partial_x^{\langle 0,0,\ldots,0\rangle} f(\mathbf{x}) \\
&\quad + \left(\frac{h^{\langle 1,0,\ldots,0\rangle}}{\langle 1,0,\ldots,0\rangle!}\, \partial_x^{\langle 1,0,\ldots,0\rangle} f(\mathbf{x}) + \frac{h^{\langle 0,1,\ldots,0\rangle}}{\langle 0,1,\ldots,0\rangle!}\, \partial_x^{\langle 0,1,\ldots,0\rangle} f(\mathbf{x}) + \cdots + \frac{h^{\langle 0,0,\ldots,1\rangle}}{\langle 0,0,\ldots,1\rangle!}\, \partial_x^{\langle 0,0,\ldots,1\rangle} f(\mathbf{x}) \right) \\
&\quad + \left(\frac{h^{\langle 2,0,\ldots,0\rangle}}{\langle 2,0,\ldots,0\rangle!}\, \partial_x^{\langle 2,0,\ldots,0\rangle} f(\mathbf{x}) + \frac{h^{\langle 1,1,\ldots,0\rangle}}{\langle 1,1,\ldots,0\rangle!}\, \partial_x^{\langle 1,1,\ldots,0\rangle} f(\mathbf{x}) + \cdots + \frac{h^{\langle 1,0,\ldots,1\rangle}}{\langle 1,0,\ldots,1\rangle!}\, \partial_x^{\langle 1,0,\ldots,1\rangle} f(\mathbf{x}) \right. \\
&\quad \left. + \frac{h^{\langle 0,2,\ldots,0\rangle}}{\langle 1,0,\ldots,0\rangle!}\, \partial_x^{\langle 1,0,\ldots,0\rangle} f(\mathbf{x}) + \cdots + \frac{h^{\langle 0,1,\ldots,1\rangle}}{\langle 0,1,\ldots,1\rangle!}\, \partial_x^{\langle 0,1,\ldots,1\rangle} f(\mathbf{x}) + \cdots + \frac{h^{\langle 0,0,\ldots,2\rangle}}{\langle 0,0,\ldots,2\rangle!}\, \partial_x^{\langle 0,0,\ldots,2\rangle} f(\mathbf{x}) \right) \\
&\quad + \cdots \\
&= \quad f(\mathbf{x}) + h_i\, \partial_x^i f(\mathbf{x}) + \frac{1}{2} h_j h_k\, \partial_x^j \partial_x^k f(\mathbf{x}) + \cdots
\end{aligned}$$

B.3 1D Taylor Series of a Product of Functions

The product of two Taylor series is the Taylor series of the product:

$$f(x+h)\, g(x+h) \quad = \quad (fg)(x+h)$$

and

$$\left[\sum_{n=0}^{\infty} \frac{h^n}{n!}\, \partial_x^n f(x) \right] \left[\sum_{m=0}^{\infty} \frac{h^m}{m!}\, \partial_x^m g(x) \right] \quad = \quad \sum_{\ell=0}^{\infty} \frac{h^\ell}{\ell!}\, \partial_x^\ell (fg)(x) \quad \equiv \quad \sum_{\ell=0}^{\infty} \frac{h^\ell}{\ell!}\, \partial_x^\ell \{ f(x)\, g(x) \}$$

Here's a proof using combinatorics:

$$\begin{aligned}
f(x+h)\,g(x+h) &= \left[\sum_{n=0}^{\infty} \frac{h^n}{n!}\,\partial_x^n f(x)\right]\left[\sum_{m=0}^{\infty} \frac{h^m}{m!}\,\partial_x^m g(x)\right] \\
&= \sum_{\ell=0}^{\infty}\sum_{m=0}^{\ell} \frac{h^\ell}{(\ell-m)!m!}\,\partial_x^{\ell-m} f(x)\,\partial_x^m g(x) \qquad (\ell \equiv n+m,\text{ so } n=\ell-m) \\
&= \sum_{\ell=0}^{\infty} \frac{h^\ell}{\ell!}\left(\sum_{m=0}^{\ell} \frac{\ell!}{m!(\ell-m)!}\partial_x^{\ell-m} f(x)\,\partial_x^m g(x)\right) \\
&= \sum_{\ell=0}^{\infty} \frac{h^\ell}{\ell!}\left(\sum_{m=0}^{\ell} \binom{\ell}{m}\partial_x^{\ell-m} f(x)\,\partial_x^m g(x)\right) \\
&= \sum_{\ell=0}^{\infty} \frac{h^\ell}{\ell!}\,\partial_x^\ell\{f(x)\,g(x)\} \\
&= \sum_{\ell=0}^{\infty} \frac{h^\ell}{\ell!}\,\partial_x^\ell(fg)(x) \\
&= (fg)(x+h)
\end{aligned}$$

B.4 Multi-index Taylor Series of a Product of Functions

The product of two Taylor series is the Taylor series of the product:

$$f(\mathbf{x}+\mathbf{h})\,g(\mathbf{x}+\mathbf{h}) \quad = \quad (fg)(\mathbf{x}+\mathbf{h})$$

and

$$\left[\sum_{|\alpha|=0}^{\infty} \frac{h^\alpha}{\alpha!}\,\partial_x^\alpha f(\mathbf{x})\right]\left[\sum_{|\beta|=0}^{\infty} \frac{h^\beta}{\beta!}\,\partial_x^\beta g(\mathbf{x})\right] \quad = \quad \sum_{|\gamma|=0}^{\infty} \frac{h^\gamma}{\gamma!}\,\partial_x^\gamma(fg)(\mathbf{x}) \quad \equiv \quad \sum_{|\gamma|=0}^{\infty} \frac{h^\gamma}{\gamma!}\,\partial_x^\gamma\{f(\mathbf{x})\,g(\mathbf{x})\}$$

Here's a proof using combinatorics:

$$\begin{aligned}
f(\mathbf{x}+\mathbf{h})\,g(\mathbf{x}+\mathbf{h}) &= \left[\sum_{|\alpha|=0}^{\infty} \frac{h^\alpha}{\alpha!}\,\partial_x^\alpha f(\mathbf{x})\right]\left[\sum_{|\beta|=0}^{\infty} \frac{h^\beta}{\beta!}\,\partial_x^\beta g(\mathbf{x})\right] \\
&= \sum_{|\gamma|=0}^{\infty}\sum_{|\beta|=0}^{|\gamma|} \frac{h^\gamma}{(\gamma-\beta)!\beta!}\,\partial_x^{\gamma-\beta} f(\mathbf{x})\,\partial_x^\beta g(\mathbf{x}) \qquad (\gamma \equiv \alpha+\beta,\text{ so } \alpha=\gamma-\beta) \\
&= \sum_{|\gamma|=0}^{\infty} \frac{h^\gamma}{\gamma!}\left(\sum_{|\beta|=0}^{|\gamma|} \frac{\gamma!}{\beta!(\gamma-\beta)!}\partial_x^{\gamma-\beta} f(\mathbf{x})\,\partial_x^\beta g(\mathbf{x})\right) \\
&= \sum_{|\gamma|=0}^{\infty} \frac{h^\gamma}{\gamma!}\left(\sum_{|\beta|=0}^{|\gamma|} \binom{\gamma}{\beta}\partial_x^{\gamma-\beta} f(\mathbf{x})\,\partial_x^\beta g(\mathbf{x})\right) \\
&= \sum_{|\gamma|=0}^{\infty} \frac{h^\gamma}{\gamma!}\,\partial_x^\gamma\{f(\mathbf{x})\,g(\mathbf{x})\} \\
&= \sum_{|\gamma|=0}^{\infty} \frac{h^\gamma}{\gamma!}\,\partial_x^\gamma(fg)(\mathbf{x})
\end{aligned}$$

$$= \quad (fg)(\mathbf{x}+\mathbf{h})$$

$$\binom{\gamma}{\beta} = \binom{\gamma_1}{\beta_1}\binom{\gamma_2}{\beta_2}\cdots\binom{\gamma_n}{\beta_n}$$

C Migration Holors M

Given that $\mathbf{R} \in \mathbb{R}^d$ is a stochastic variable that describes the state of a Brownian object, that $\boldsymbol{\xi} = \Delta\mathbf{R}$ is a stochastic change in the state variable $\mathbf{R}$, that $p[\mathbf{R}, \boldsymbol{\xi}; \Delta t, t]$ is the probability that the Brownian object transitions to a state $\mathbf{R}$ at time t by a change $\boldsymbol{\xi}$ with duration Δt, and that $\alpha \in \mathbb{N}^d$ is a multi-index as defined in Appendix B.2, we define $\langle \xi^\alpha \rangle_{\Delta t}(\mathbf{R}, t)$, which are elements of a mean-transition-increment holor[9] to a state $\mathbf{R}$ at time t over a time Δt, by

$$\langle \xi^\alpha \rangle_{\Delta t}(\mathbf{R}, t) \equiv \int \mathrm{d}^d\xi \; \xi^\alpha \; p[\mathbf{R}, \boldsymbol{\xi}; \Delta t, t].$$

For example, with a 4D ($d = 4$) state-space, if $\alpha = \langle 0, 3, 1, 2\rangle$, then $\xi^\alpha = \xi_1^{\alpha_1}\xi_2^{\alpha_2}\cdots\xi_d^{\alpha_d} = {\xi_2}^3\xi_3{\xi_4}^2$ and

$$\langle \xi^\alpha \rangle_{\Delta t}(\mathbf{R}, t) = \left\langle {\xi_2}^3\xi_3{\xi_4}^2 \right\rangle_{\Delta t}(\mathbf{R}, t) = \int \mathrm{d}\xi_1\,\mathrm{d}\xi_2\,\mathrm{d}\xi_3\,\mathrm{d}\xi_4 \; {\xi_2}^3\xi_3{\xi_4}^2 \; p[\mathbf{R}, \boldsymbol{\xi}; \Delta t, t].$$

Now we define a migration holor $M_{|\alpha|}^{\{\alpha\}}(\mathbf{R}, t)$ by

$$M_{|\alpha|}^{\{\alpha\}}(\mathbf{R}, t) \equiv \lim_{\Delta t \to 0} \frac{\langle \xi^\alpha \rangle_{\Delta t}(\mathbf{R}, t)}{\alpha! \Delta t},$$

where

$$\{\alpha\} \equiv \left\{ \bigcup_j \{j\}_{\alpha_j} \right\} = \{i_1, i_2, \ldots, i_{|\alpha|}\}$$

and $\{j\}_{\alpha_j}$ means a multiset consisting of α_j instances of the number j. For example, again taking a 4D state-space with $\alpha = \langle 0, 3, 1, 2\rangle$, we have $|\alpha| = 6$ and

$$\begin{aligned} \{\alpha\} &= \left\{ \bigcup_j \{j\}_{\alpha_j} \right\} = \big\{\{1\}_0 \cup \{2\}_3 \cup \{3\}_1 \cup \{4\}_2\big\} = \big\{\{\} \cup \{2,2,2\} \cup \{3\} \cup \{4,4\}\big\} \\ &= \{2,2,2,3,4,4\} \\ &= \{i_1, i_2, i_3, i_4, i_5, i_6\}, \end{aligned}$$

thus

$$\begin{aligned} M_{|\alpha|}^{\{\alpha\}}(\mathbf{R}) &= \lim_{\tau \to 0} \frac{\langle \xi^\alpha \rangle_\tau(\mathbf{R})}{\alpha! \tau} = \lim_{\tau \to 0} \frac{\left\langle {\xi_2}^3\xi_3\,{\xi_4}^2 \right\rangle_\tau(\mathbf{R})}{3!1!2!\tau} \\ &= M_6^{\{2,2,2,3,4,4\}}(\mathbf{R}) = M_6^{222344}(\mathbf{R}) = M_6^{422234}(\mathbf{R}) = M_6^{423224}(\mathbf{R}) = \cdots \end{aligned}$$

[9] A holor is a mathematical entity that is made up of one or more independent quantities. A holor may be multiply-indexed, like a tensor, but its transformation properties, under rotation, say, are not necessarily specified.