

Introduction to Stochastic Processes



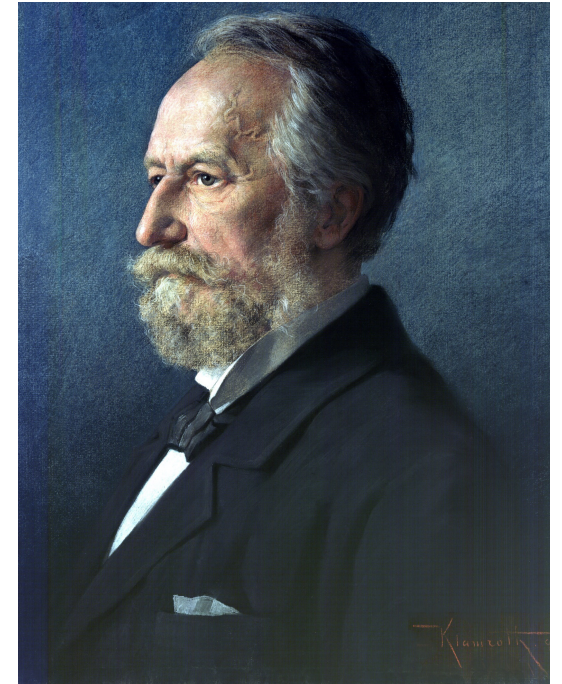
Dr. Sudipta Kumar Sinha
Department of Chemistry, IIT Ropar

Diffusion:

The laws of diffusion were first investigated by Graham. Later by Loschmidt, and that of liquids by Fick and Voit with great accuracy

The laws of diffusion were first formulated by Fick and he proposed following two laws of diffusion

$$J = -D\nabla\rho , \quad \frac{\partial\rho}{\partial t} = -\nabla\cdot J$$



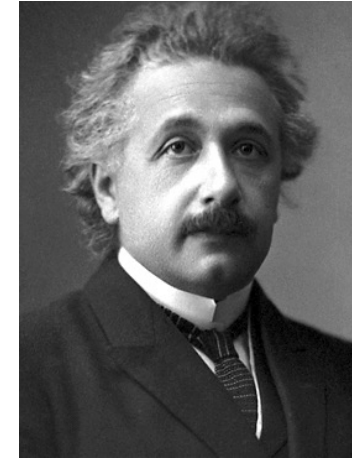
Adolf Fick

The above two equations combine to give the diffusion equation

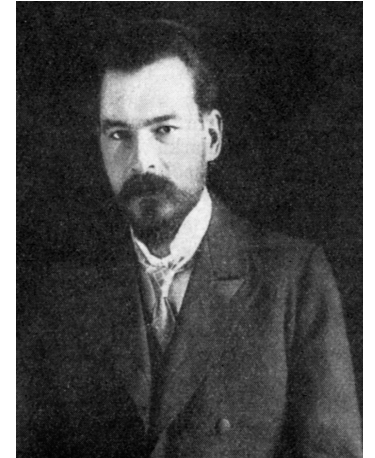
$$\frac{\partial\rho}{\partial t} = \nabla\cdot D\nabla\rho$$

Brownian Motion

They used kinetic theory and demonstrated, theoretically, that the phenomenon of diffusion is the result of Brownian motion.



Einstein (1905)



Smoluchowski (1906)

Einstein's Theory:

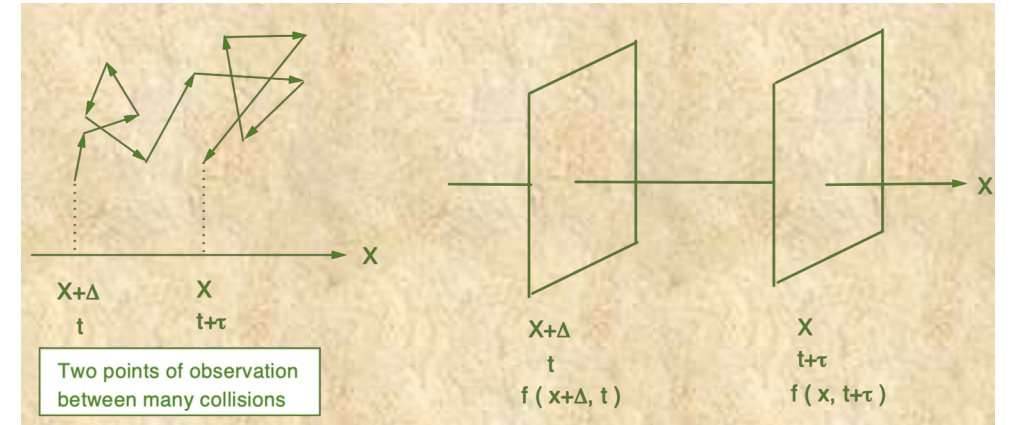
- ❖ The motion of each particle is independent of the others
- ❖ The movements of one and the same particle after different intervals of time must be considered as mutually independent processes

Let dn be the number of particles which experience a displacement between Δ and $\Delta + d\Delta$ in time interval τ

$$dn = n\phi(\Delta) d\Delta, \quad \text{where} \quad \int_{-\infty}^{\infty} \phi(\Delta) d\Delta = 1, \quad \phi(\Delta) = \phi(-\Delta),$$

The value of the concentration $\rho(x, t)$ after time has elapsed can be computed from the values of $\rho(x + \Delta, t)$ for all possible values of Δ , weighted by $\phi(\Delta)$

$$\rho(x, t + \tau) = \int_{-\infty}^{\infty} \rho(x + \Delta, t) \phi(\Delta) d\Delta.$$



Expanding ρ in Taylor's series for small τ and Δ , the equation

$$\rho(x, t) + \frac{\partial \rho(x, t)}{\partial t} \tau + \dots = \rho(x, t) \int_{-\infty}^{\infty} \phi(\Delta) d\Delta + \frac{\partial \rho(x, t)}{\partial x} \int_{-\infty}^{\infty} \Delta \phi(\Delta) d\Delta + \frac{\partial^2 \rho(x, t)}{\partial x^2} \int_{-\infty}^{\infty} \frac{\Delta^2}{2} \phi(\Delta) d\Delta + \dots$$

Neglecting terms of higher order, the above equation reduces to

$$\frac{\partial \rho(x, t)}{\partial t} = D \frac{\partial^2 \rho(x, t)}{\partial x^2}, \quad \text{where} \quad D = \frac{1}{\tau} \int_{-\infty}^{\infty} \frac{\Delta^2}{2} \phi(\Delta) d\Delta.$$

And the solution of the above equation is

$$\rho(x, t) = \frac{n}{\sqrt{4\pi Dt}} \exp\left\{-\frac{x^2}{4Dt}\right\} \quad \text{Or} \quad p(x, t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left\{-\frac{x^2}{4Dt}\right\}$$

It follows that the mean and variance value

$$\mathbb{E}x(t) = 0, \quad \mathbb{E}x^2(t) = 2Dt.$$

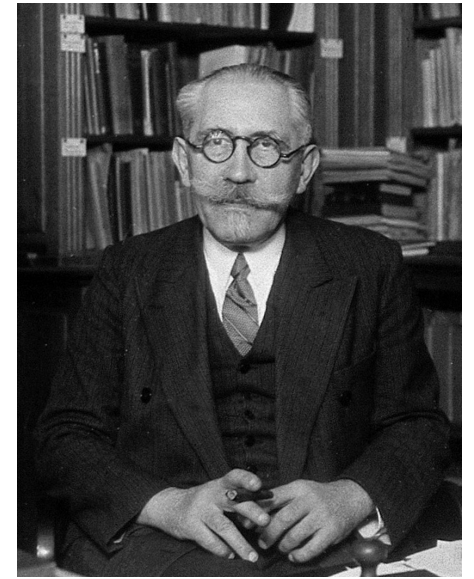
Langevin's Approach (1908):

The dynamics of a free Brownian particle is governed by the frictional force $-6\pi a\eta v$ and by a fluctuational force X , that results from the random collisions of the Brownian particle with the molecules of the surrounding fluid, after the frictional force is subtracted.

$$m\ddot{x} = -6\pi a\eta\dot{x} + F(t)$$

If additional fields of force act on the diffusing particles

$$m\ddot{x} + \Gamma\dot{x} - F_{ext}(t) = F(t) \quad \text{where} \quad F_{ext} = -\nabla U(x)$$



P. Langevin

The above equations are stochastic differential equations, because those are driven by a random force $F(t)$

The main mathematical difference between the two approaches is that Einstein assumes that the displacements Δ are independent, whereas Langevin assumes that the random force $F(t)$ and the displacement x are independent.

Denoting $v = \dot{x}$ and multiplying the Langevin equation by x

$$\frac{m}{2} \frac{d^2}{dt^2} x^2 - mv^2 = -3\pi a \eta \frac{d}{dt} x^2 + \Xi x.$$

Averaging under the assumption that the fluctuational force and the displacement of the particle x are independent, we obtain

$$\frac{m}{2} \frac{d^2}{dt^2} \langle x^2 \rangle + 3\pi a \eta \frac{d}{dt} \langle x^2 \rangle = kT, \quad \text{since} \quad \frac{m}{2} \bar{v}_{x,y,z}^2 = \frac{kT}{2},$$

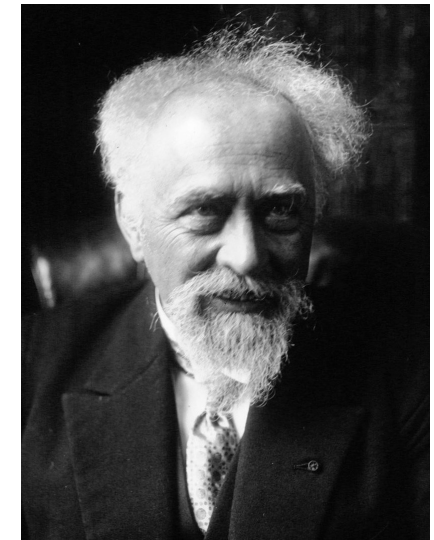
$$d\langle x^2 \rangle / dt = kT / 3\pi a \eta + C e^{-6\pi a \eta t / m} \quad \text{Where } C \text{ is a constant}$$

If times well in excess of the frictional relaxation time, $m/6\pi a \eta$

It follows that $\langle x^2 \rangle - \langle x_0^2 \rangle = (kT/3\pi a \eta)t$ **Einstein's relation!**

$$\sqrt{\overline{(\Delta x)^2}} = \sqrt{\frac{RTt}{3\pi\eta aN}} \quad \text{Where } N=R/k$$

In 1908, Perrin computed the Avogadro number from observations of the Brownian movement, obtaining $N = 6.85 \times 10^{23} \text{ mol}^{-1}$



Perrin

Thus, Langevin's equation of motion, according to Newton's second law of motion, is for a particle of mass m

$$m \frac{d^2 x(t)}{dt^2} = -\zeta \frac{dx(t)}{dt} + F(t) \quad \text{where} \quad \zeta \dot{x} = 6\pi\eta a \dot{x}$$

(i) $F(t)$ is independent of x .

(ii) $F(t)$ varies extremely rapidly compared to the variation of $x(t)$.

$$\overline{F(t)} = 0 \quad \overline{F(t)F(t')} = 2\zeta kT \delta(t - t')$$

We note that the noise force $F(t)$ in the Langevin equation may be related to the drag coefficient ζ as follows.

$$\int_0^\infty \overline{F(t)F(t+\tau)} d\tau = \frac{1}{2} \int_{-\infty}^\infty \overline{F(t)F(t+\tau)} d\tau = \zeta kT \int_{-\infty}^\infty \delta(t) dt$$

Hence

$$\zeta = \frac{1}{kT} \int_0^\infty \overline{F(t)F(t+\tau)} d\tau \quad \text{F-D theorem}$$

Criticism of Langevin's Equation:

Doob show how that equation should properly be interpreted as an integral equation and not as a differential equation



We write the Langevin equation $du(t) = -\beta u(t)dt + dB(t)$

The distribution of $[B(s+t) - B(t)]$ is Gaussian with mean zero, variance $c^2|t|$ and if $t_1 < \dots < t_n$, $B(t_2) - B(t_1)$, ..., $B(t_n) - B(t_{n-1})$ are mutually independent random variables and $B(t)$ follows the definition of Weiner process Doob

Let us integrate both sides of the above equation after multiplying a continuous function $f(t)$

$$\int_{t=a}^b f(t) du(t) = -\beta \int_a^b f(t) u(t) dt + \int_a^b f(t) dB(t)$$

Setting $f(t) = e^{\beta t}$ and since integration of by parts is permissible, we obtain

$$u(t) = u(0)e^{-\beta t} + \int_{s=0}^t e^{-\beta(t-s)} dB(s)$$

Wiener Process:

Let $X(t)$ be a random variable and let $X(0) = 0$. We shall assume the following.

1. The displacement $X(t) - X(s)$ of the Brownian particle over (s, t) is the sum

$$\sum_{k=1}^n [X(t_k) - X(t_{k-1})].$$

2. The increments are independent and they follow Markov process

$$X(t_1) - X(s), X(t_2) - X(t_1), \dots, X(t_n) - X(t_{n-1})$$

3. We have $\langle X(t_1) - X(s) \rangle = \langle X(t_k) - X(t_{k-1}) \rangle = 0$

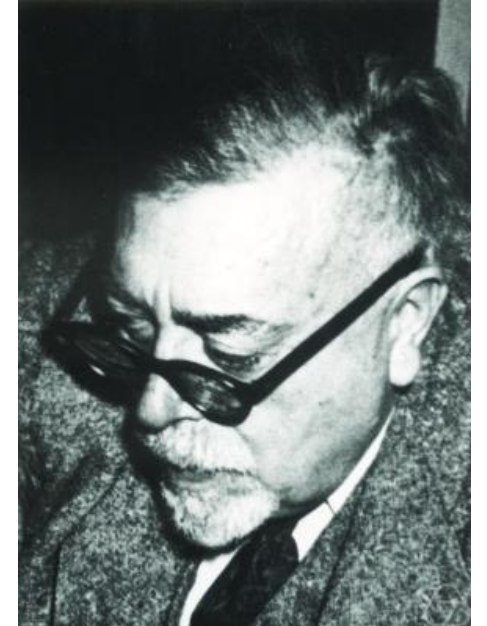
4. Since we have assumed that $X(t) - X(s)$ is the sum of a large number of independent random variables, each having an arbitrary distribution, it follows from the

central limit theorem

$$\phi(u)_{X(t)-x(s)} = e^{\{-\frac{1}{2}u^2 \langle [X(t)-x(s)]^2 \rangle\}}$$

It follows that

$$\langle [X(t) - x(s)]^2 \rangle = c^2 |t - s| \quad \text{and} \quad \langle X(s)X(t) \rangle = c^2 \min(s, t)$$



Wiener

OU theory of Brownian Motion:

Every process, which is MARKOVian, stationary and GAUSSian is (by DOBB's theorem) the ORNSTEIN-UHLENBECK process.

The relation obtained by Einstein and Langevin has fundamental flaw that it is not root mean square differentiable at $t = 0$

$$\langle (\Delta x)^2 \rangle = \frac{2kT}{\zeta} |t|$$

(ignoring the inertia of the particles)

In 1930, Uhlenbeck and Ornstein by including the inertia of the particles

$$\langle (\Delta x)^2 \rangle = \frac{2kT}{\zeta^2} \left(\frac{\zeta}{m} |t| - 1 + e^{-\zeta|t|/m} \right)$$

Here, the root mean square differentiable at $t = 0$

$$\langle (\Delta x)^2 \rangle = \frac{kT}{m} t^2$$

We write the Langevin equation in phase space (x,v)

$$\dot{x}(t) = v(t)$$

$$m\dot{v}(t) = -\zeta v(t) + \lambda(t)$$



Uhlenbeck

The white noise $\lambda(t)$ follows the definition

$$\overline{\lambda(t_1)\lambda(t_2)} = 2D\delta(t_1 - t_2), \quad D = \zeta kT$$

It follows that

$$v(t) = \dot{x}(t) = v_0 e^{-\beta t} + \frac{1}{m} \int_0^t e^{-\beta(t-t')} \lambda(t') dt'$$

$$\Delta x = x(t) - x_0 = \frac{v_0}{\beta} (1 - e^{-\beta t}) + \frac{1}{m\beta} \int_0^t [1 - e^{-\beta(t-t')}] \lambda(t') dt'$$

Therefore, we obtain

$$\overline{\Delta x} = \frac{v_0}{\beta} (1 - e^{-\beta t}) \quad \text{where } \beta = \zeta/m$$

and

$$\overline{(\Delta x)^2} = \left[\frac{v_0}{\beta} (1 - e^{-\beta t}) \right]^2 + \frac{2Dt}{(m\beta)^2} + \frac{D}{m^2\beta^3} [-3 + 4e^{-\beta t} - e^{-2\beta t}]$$

If we have a Maxwell-Boltzmann distribution of initial velocities v_0

$$\langle (\Delta x)^2 \rangle = \frac{2kT}{m\beta^2} \left(\beta t - 1 + e^{-\beta t} \right), \quad (t > 0)$$

inertia is included, however, $\langle (\Delta x)^2 \rangle$ differentiable and the velocity exists.

The inertia induced shift as obtained by 2-D stochastic motion in phase space, which is equivalent to a colored noise driven dynamics in configuration space $\left(\frac{-1 + e^{-\beta t}}{\beta} \right), \quad (t > 0)$

The velocity autocorrelation is exponential $\langle v(t)v(t') \rangle_{\text{time}} = \frac{kT}{m} e^{-\zeta|t-t'|/m}$

Formal Description of Stochastic Process:

Definition of B_t

Let us consider B is a Brownian Motion (BM) which is stochastic process (SP) defined in the following way

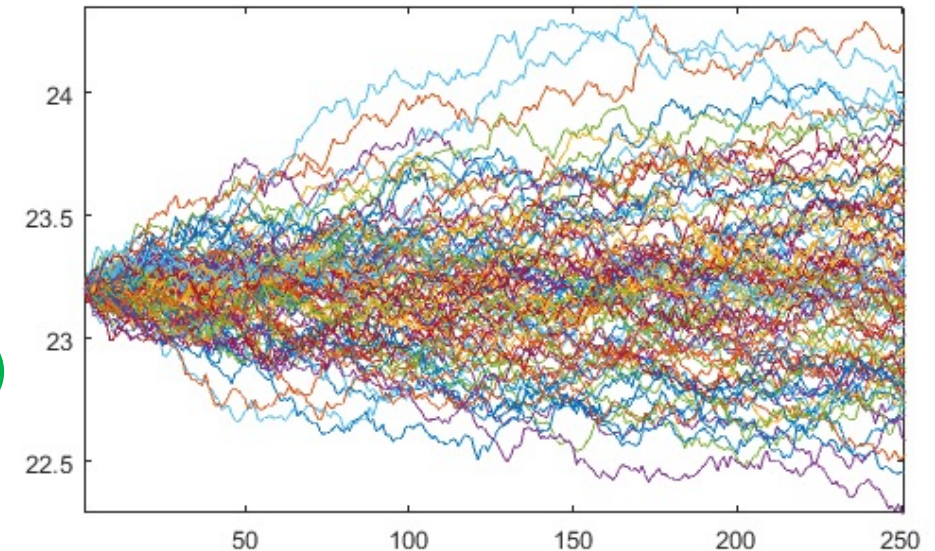
1. For $s < t$, the increment $B_t - B_s \sim N(0, t-s)$
2. $B_t - B_s$ is independent of $\{B_u, 0 \leq u \leq s\}$
3. B_t is continuous in 't' (almost surely pathwise)

B_t is the value of B at time t

Nondifferentiability of B_t

$$\lim_{h \rightarrow 0} \frac{B_{t+h} - B_t}{h} \rightarrow N\left(0, \frac{h}{h^2}\right) \rightarrow N\left(0, \frac{1}{h}\right)$$

Distance travel from one point to another point is ∞



Stochastic Integrals:

Consider the following integral

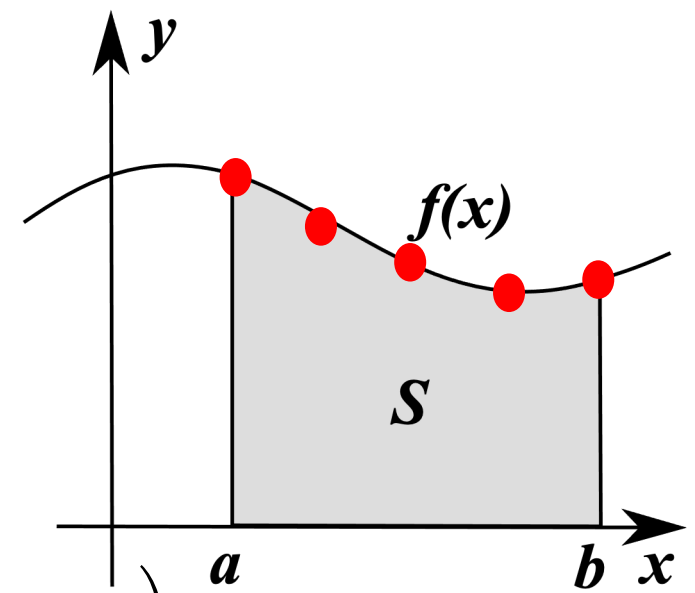
$$\int_0^x x dx = \frac{x^2}{2}$$

But for the stochastic integral

$$\int_0^T B_t dB_t \neq \frac{B_T^2}{2} \rightarrow \sum_i B_{t_i} (B_{t_{i+1}} - B_{t_i})$$

$B_{t_{i+1}}$ B_{t_i} $B_{\frac{t_i+t_{i+1}}{2}}$

Each of them gives different answer!



Hard to calculate this integral, take the expectation, $E\left(\sum_i B_{t_i} (B_{t_{i+1}} - B_{t_i})\right)$

$$\begin{aligned} E\left(\sum_i B_{t_i} (B_{t_{i+1}} - B_{t_i})\right) &= \sum_i E\left(B_{t_i} (B_{t_{i+1}} - B_{t_i})\right) \quad \text{Linearity of } E \\ &= \sum_i E\left(E\left(B_{t_i} (B_{t_{i+1}} - B_{t_i}) \mid F_{t_i}\right)\right) \quad \text{Tower property of } E: E(E(Y|X)) = E(Y) \end{aligned}$$

Where $F_{t_i} = \{B_t : 0 \leq t \leq t_i\}$: All the information up to time t_i

$$\begin{aligned} &= \sum_i E\left(B_{t_i} E\left(B_{t_{i+1}} - B_{t_i} \mid F_{t_i}\right)\right) \quad \text{Since } E(XY|X) = XE(Y|X) \text{ and } B_{t_i} \text{ is } F_{t_i} \text{ measurable} \\ &= \sum_i E\left(B_{t_i} E\left(B_{t_{i+1}} - B_{t_i}\right)\right) = 0 \quad \text{As } B_{t_{i+1}} - B_{t_i} \text{ is independent of } F_{t_i} \text{ (second property of BM)} \end{aligned}$$

Therefore, it appears that $\int_0^T B_t dB_t \approx \sum_i B_{t_i} (B_{t_{i+1}} - B_{t_i})$ has zero mean

Let us guess the expectation of integral $E(\int_0^T B_t dB_t) \approx E\left(\frac{B_T^2}{2}\right) = \frac{T}{2} \neq 0$ since $B_T \sim N(0, T)$

From the above argument it is clear that $\int_0^T B_t dB_t \neq \frac{B_T^2}{2}$

The actual value of the $\int_0^T B_t dB_t = \frac{B_T^2}{2} - \frac{T}{2}$ (Ito integral)

$$\int_0^T B_t dB_t \approx \sum_i B_{\frac{t_i+t_{i+1}}{2}}(B_{t_{i+1}} - B_{t_i}) = \frac{B_T^2}{2} \quad (\text{Stratonovich integral})$$

$$\int_0^T B_t dB_t \approx \sum_i B_{t_{i+1}}(B_{t_{i+1}} - B_{t_i}) = \frac{B_T^2}{2} + \frac{T}{2} \quad (\text{Backward Ito integral})$$

Ito integral:

Consider the following Taylor Expansion of a function $f(x)$

$$df(x) = f'(x)dx + \frac{f''(x)}{2!}(dx)^2 + \frac{f'''(x)}{3!}(dx)^3 + \dots$$

$$\text{If } f(x)=x^2 \text{ and } x=B_t \text{ then } df(B_t) = f'(B_t)dB_t + \frac{f''(B_t)}{2!}(dB_t)^2 + \frac{f'''(B_t)}{3!}(dB_t)^3 + \dots$$

And since $(dB_t)^2 \approx dt$ for BM; Substitute it and take the integral

$$f(B_t) - f(B_0) = \int_0^T f'(B_t)dB_t + \frac{1}{2} \int_0^T f''(B_t)dt$$

Upon substitution of the $f(x)$

$$B_T^2 - B_0^2 = \int_0^T 2B_t dB_t + \frac{1}{2} \int_0^T 2dt \rightarrow \int_0^T B_t dB_t = \frac{B_T^2}{2} - \frac{T}{2}$$

More on Ito integral:

In general a stochastic integral is written as $\int_0^T X_t dB_t$ Where X_t is a stochastic process, which is F_t measurable

Then $\int_0^T X_t^2 dB_t < \infty$ (Square integrable)

If $E \int_0^T X_t^2 dB_t < \infty$ then $E \int_0^T X_t dB_t = 0$ and $E \left(\int_0^T X_t dB_t \right)^2 = E \int_0^T X_t^2 dB_t$

If X_t is deterministic then $\int_0^T X_t dB_t \sim N \left(0, \int_0^T X_t^2 dB_t \right)$

Martingales:

A process M is called Martingale if

- 1. $E|M| < \infty$
- 2. $E(M_t | F_s) = M_s$ for $s \leq t$

The BM is a Martingale $E(B_t | F_s) = B_s$ for $s \leq t$

If $E \int_0^T X_t^2 dB_t < \infty$ then $M_t = \int_0^t X_s dB_s$ is a Martingale

Solution of SDE Using Ito formula:

ODE: $dX_t = \sigma X_t dt \rightarrow \frac{dX_t}{dt} = \sigma X_t \rightarrow X_t - X_0 = \int_0^t \sigma X_s ds \rightarrow X_t = X_0 e^{\sigma t}$

SDE: $dX_t = \underbrace{\mu X_t dt}_{\text{drift}} + \underbrace{\sigma X_t dB_t}_{\text{diffusion}} \neq \frac{dX_t}{dt} = \mu X_t + \sigma X_t \frac{dB_t}{dt}$ **Not differentiable**

$X_t - X_0 = \int_0^t \mu X_s ds + \int_0^t \sigma X_s dB_s$ **Hard to solve**

Solution Using Ito Formula:

$df(X_t) = f'(X_t)dX_t + \frac{f''(X_t)}{2!} (dX_t)^2$

Rearrange the SDE $\frac{dX_t}{X_t} = d \log X_t = \mu dt + \sigma dB_t$ Thus $f(X_t) = \log X_t; f'(X_t) = \frac{1}{X_t}; f''(X_t) = -\frac{1}{X_t^2}$

$d \log X_t = \mu dt + \sigma dB_t - \frac{1}{2} (\mu dt + \sigma dB_t)^2$
 $= \mu dt + \sigma dB_t - \frac{1}{2} \sigma^2 dt$ Since $dt^2 \rightarrow 0; dt dB_t \rightarrow 0; dB_t^2 \rightarrow dt$

$\log X_t - \log X_0 = \int_0^t \left(\mu - \frac{1}{2} \sigma^2 \right) ds + \int_0^t \sigma dB_s = \left(\mu - \frac{1}{2} \sigma^2 \right) t + \sigma dB_t$

$X_t = X_0 e^{-\left(\mu - \frac{1}{2} \sigma^2\right)t + \sigma dB_t}$

Fokker Planck Equation:

Time evolution of the probability distribution function of the Brownian Particle

$$\frac{d\mathbf{a}}{dt} = \mathbf{v}(\mathbf{a}) + \mathbf{F}(\mathbf{t}) \quad \langle \mathbf{F}(\mathbf{t}) \rangle = \mathbf{0} \quad \langle \mathbf{F}(\mathbf{t})\mathbf{F}(\mathbf{t}') \rangle = 2\mathbf{B}\delta(\mathbf{t} - \mathbf{t}')$$

where $\mathbf{v}(\mathbf{a})$ is some given function of the variables \mathbf{a} . Our aim is to obtain a noise average probability distribution function, $f(\mathbf{a}, t)$ of the values of \mathbf{a} at time t . To do that one can recognize, $f(\mathbf{a}, t)$ is conserved i.e.

$$\int d\mathbf{a} f(\mathbf{a}, t) = 1$$

Immediately, it suggest that the time derivative of the density, $f(\mathbf{a}, t)$ is balanced by the divergence of a flux, a velocity times that density

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

Where a noise term, $\mathbf{F}(\mathbf{t})$ is there. Upon taking the average over noise, the Fokker-Planck equation is obtained

$$\frac{\partial \langle f(\mathbf{a}, t) \rangle}{\partial t} = -\frac{\partial}{\partial \mathbf{a}} \cdot \mathbf{v}(\mathbf{a}) \langle f(\mathbf{a}, t) \rangle + \frac{\partial}{\partial \mathbf{a}} \cdot \mathbf{B} \cdot \frac{\partial}{\partial \mathbf{a}} \langle f(\mathbf{a}, t) \rangle$$

Illustrations:

Start with a Langevin equation, where the relaxation time, $\tau = m/\zeta$ is very much shorter than any natural time scale associated with motion in the potential $U(x)$. The inertial term,

$$m \frac{d^2 x(t)}{dt^2}, \text{ can be ignored} \quad \frac{dx(t)}{dt} = -\frac{1}{\zeta} U(x)' + \frac{1}{\zeta} F(t)$$

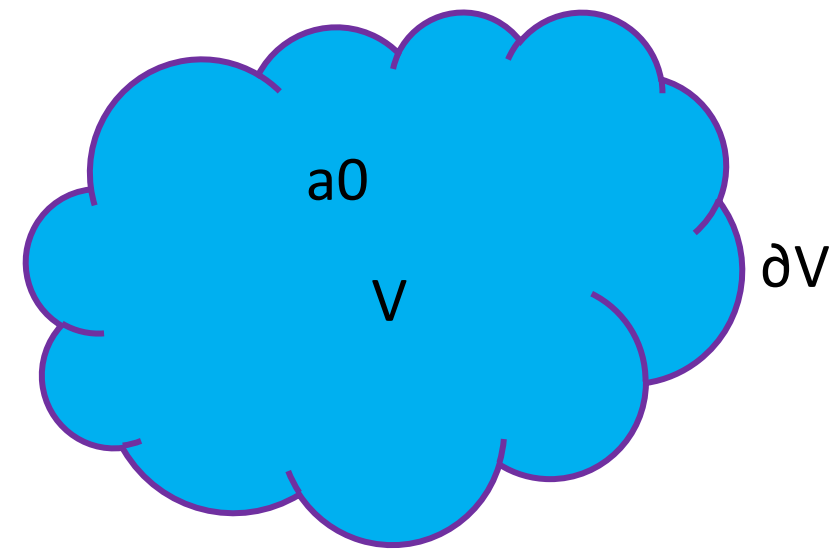
It leads to a F-P equation which is commonly known as *Smoluchowski equation*,

$$\begin{aligned} \frac{\partial f}{\partial t} &= \frac{1}{\zeta} \frac{\partial}{\partial x} U(x)' f + \frac{kT}{\zeta} \frac{\partial^2}{\partial x^2} f && \text{Where } D = kT/\zeta \\ &= D \frac{\partial}{\partial x} e^{-U(x)/kT} \frac{\partial}{\partial x} e^{U(x)/kT} f \end{aligned}$$

It describes the diffusion over a barrier. Moreover, one can make it self-adjoint equation by a trick. Upon substitution $f = \sqrt{f_{eq}} g$ leads to a Schrodinger like equation

$$-\frac{\partial g}{\partial t} = D \left(-\frac{\partial^2}{\partial x^2} + U_{eff}(x) \right) g \quad \text{Where} \quad U_{eff}(x) = \left\{ \left(\frac{1}{2kT} \frac{\partial U}{\partial x} \right)^2 - \frac{1}{2kT} \frac{\partial^2 U}{\partial x^2} \right\}$$

First Passage Time:



- ✓ Evolution ($\mathbf{a}(t)$) of a set of variables \mathbf{a} is governed by a Langevin equation
- ✓ The initial point \mathbf{a}_0 starts out somewhere in a "volume" V in this space, bounded by a "surface" ∂V
- ✓ The first passage time is the first time that the point leaves V

The motion of a cloud of initial points satisfies the Fokker-Planck equation. An absorbing boundary condition is imposed on ∂V that removes all paths that have crossed the boundary of V before time t and focus on only those points that have not left V by time t . Then the distribution of points that have not left by time t is $P(\mathbf{a}, t)$, and satisfies

$$\frac{\partial P}{\partial t} = -\nabla_{\mathbf{a}} \cdot (\mathbf{v}(\mathbf{a})P) + \nabla_{\mathbf{a}} \cdot \mathbf{B} \cdot \nabla_{\mathbf{a}} P = DP$$

$$\text{Thus } P(\mathbf{a}, t) = e^{tD} \delta(\mathbf{a} - \mathbf{a}_0)$$

$$P(\mathbf{a}, 0) = \delta(\mathbf{a} - \mathbf{a}_0) \quad P(\mathbf{a}, t) = 0 \quad \text{On } \partial V$$

The integral of P over all \mathbf{a} in the volume V is the number of all starting points that are still in V at time t ; it depends on the initial location \mathbf{a}_0 .

$$S(t, \mathbf{a}_0) = \int_V d\mathbf{a} P(\mathbf{a}, t)$$

The difference $S(t) - S(t + dt)$ is the number of initial points that have not left before time t but have left during the time interval dt following t and therefore determines the distribution of first passage times $\rho(t, \mathbf{a}_0)$,

$$S(t, \mathbf{a}_0) - S(t + dt, \mathbf{a}_0) = \rho(t, \mathbf{a}_0)dt \quad \text{Or} \quad \rho(t, \mathbf{a}_0) = -\frac{dS(t, \mathbf{a}_0)}{dt}$$

The mean first passage time is the first moment of t ,

$$\tau(\mathbf{a}_0) = \int_0^t d\tau \tau \rho(\tau, \mathbf{a}_0) = \int_0^t \tau dS(\tau, \mathbf{a}_0) = \int_0^{t \rightarrow \infty} d\tau S(\tau, \mathbf{a}_0) = \int_0^{\infty} dt S(t, \mathbf{a}_0)$$

There is more direct way to calculate

$$\begin{aligned}\tau(\mathbf{a}_0) &= \int_0^\infty dt S(t, \mathbf{a}_0) = \int_0^\infty dt \int_V d\mathbf{a} P(\mathbf{a}, t) = \int_0^\infty dt \int_V d\mathbf{a} e^{tD} \delta(\mathbf{a} - \mathbf{a}_0) \\ &= \int_0^\infty dt \int_V d\mathbf{a} \delta(\mathbf{a} - \mathbf{a}_0) \left(e^{tD^\dagger} \mathbf{1} \right)\end{aligned}$$

Now the integration over \mathbf{a} and drop the subscript 0

$$\tau(\mathbf{a}) = \int_0^\infty dt \left(e^{tD^\dagger} \mathbf{1} \right)$$

$$D^\dagger \tau(\mathbf{a}) = \int_0^\infty dt D^\dagger e^{tD^\dagger} \mathbf{1} = \int_0^\infty dt \frac{d}{dt} e^{tD^\dagger} \mathbf{1} = -1$$

The lower limit survives and the upper limit vanishes due to absorbing boundary condition. The MFPT problem is determined by solving inhomogeneous adjoint equation

$$D^\dagger \tau(\mathbf{a}) = -1 \quad \tau(\mathbf{a}) = 0 \text{ on } \partial V$$

Application to the Smoluchowski equation:

The Smoluchowski equation is written as
$$\frac{\partial f}{\partial t} = D \frac{\partial}{\partial x} e^{-U(x)/kT} \frac{\partial}{\partial x} e^{U(x)/kT} f$$

The corresponding adjoint equation is

$$D e^{\frac{U(x)}{kT}} \frac{\partial}{\partial x} e^{-\frac{U(x)}{kT}} \frac{\partial}{\partial x} \tau(x) = -1$$

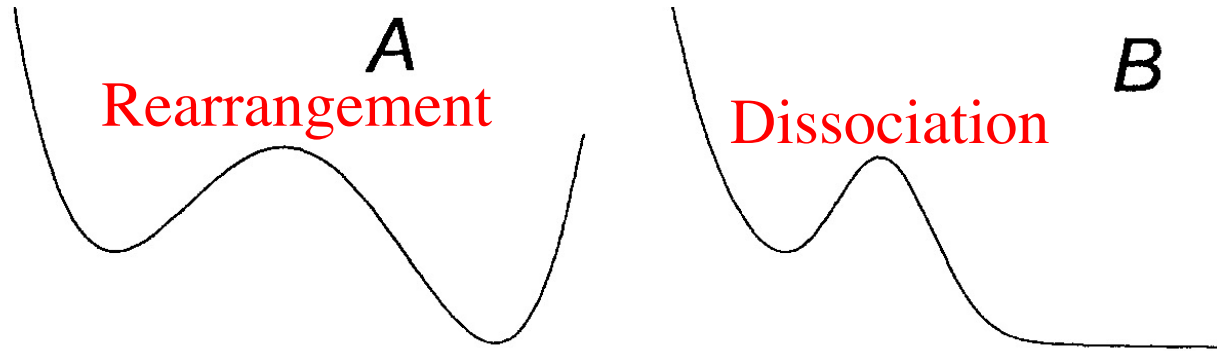
- ❖ The coordinate x is the starting position of the Brownian particle.
- ❖ The absorbing barrier is located at b , and we assume that there is a reflecting barrier at a , with $a < x < b$.

$$e^{-\frac{U(x)}{kT}} \frac{\partial}{\partial x} \tau(x) = -\frac{1}{D} \int_a^x e^{-\frac{U(z)}{kT}} dz \quad \text{Integrate once over } x$$

$$\tau(x) = \frac{1}{D} \int_x^b dy e^{\frac{U(y)}{kT}} \int_a^y dz e^{-\frac{U(z)}{kT}} \quad \text{integrate once more over } x, \text{ using the boundary conditions at the two limits } a \text{ and } b$$

Kramer's Problem:

Determine the rate at which the Brownian particle escapes from a potential well



Assume the motion is purely diffusive, allows to use Smoluchowski equation, and the barrier is sufficiently high: The rate of arrival at the barrier is estimated by taking the reciprocal of the first passage time to the barrier.

Place an absorbing barrier at x_{\max} , $U(x_{\max})=U_{\max}$ and the reflecting barrier at $x = a$, provided by a repelling potential at $x \rightarrow \infty$ (initial position)

$$\tau(x) = \frac{1}{D} \int_x^{x_{\max}} dy e^{U(y)/kT} \int_{-\infty}^y dz e^{-U(z)/kT}$$

At low temperature (small kT), the integral over z is dominated by the potential near the minimum

$$U(z) = U_{min} + \frac{1}{2} \omega_{min}^2 (z - x_{min})^2 + \dots$$

Then the upper limit of integration can be replaced by infinity, and the integral is

$$\int_{-\infty}^y dz \exp\left(-\frac{U(z)}{kT}\right) \cong \int_{-\infty}^{\infty} dz \exp\left(-\frac{U_{min}}{kT}\right) \exp\left(-\frac{\omega_{min}^2}{2kT} (z - x_{min})^2\right) = \exp\left(-\frac{U_{min}}{kT}\right) \sqrt{\frac{2\pi kT}{\omega_{min}^2}}$$

The integral over y is dominated by the potential near the barrier and has the quadratic expansion,

$$U(y) = U_{max} - \frac{1}{2} \omega_{max}^2 (y - x_{max})^2 + \dots$$

The integral over y is practically independent of x as long as x is near the potential minimum, so the lower limit can be replaced by minus infinity,

$$\int_x^{x_{max}} dy \exp\left(-\frac{U(y)}{kT}\right) \cong \int_{-\infty}^{x_{max}} dy \exp\left(\frac{U_{max}}{kT}\right) \exp\left(-\frac{\omega_{max}^2}{2kT} (y - x_{max})^2\right)$$

$$= \frac{1}{2} \exp\left(-\frac{U_{max}}{kT}\right) \sqrt{\frac{2\pi kT}{\omega_{max}^2}} \quad \begin{array}{l} 1/2 \text{ appears because only half} \\ \text{of the Gaussian is included} \end{array}$$

The mean first passage time (MFPT) (in the high barrier limit) is

$$\tau(x) \cong \frac{1}{2D} \frac{2\pi kT}{\omega_{max}\omega_{min}} \exp\left(\frac{U_{max}-U_{min}}{kT}\right) \quad \text{Where } D=kT/\zeta$$

The rate of arrival is $1/\tau$, and the rate of crossing, k_{κ} , is half of that, so that

$$k_{\kappa} \cong \frac{\omega_{max}\omega_{min}}{2\pi\zeta} \exp\left(-\frac{U_{max}-U_{min}}{kT}\right)$$

The escape rate is correct only in the high friction limit of Brownian motion

- i) The rate of activation has a typical Arrhenius form $Ae^{-E_0/kBT}$
- ii) The rate is thus inversely proportional to the friction coefficient of the medium.

Master Equation:

Master equation is a typical probability balance equation

Recall the evolution of probability of a Brownian particle, given by Einstein

$$\rho(x, t + \tau) = \int_{-\infty}^{\infty} \rho(x + \Delta, t) \phi(\Delta) d\Delta.$$

probability of a jump
of magnitude

This equation relates the probability distribution function of a Brownian particle at \mathbf{x} and time $t + \tau$, to that for the particle at a previous position $\mathbf{x} + \Delta$ at an earlier time t

Let $y = \mathbf{x} + \Delta$; it follows $dy = d\Delta$. We also can write $\phi(\Delta) = \phi(y \rightarrow x)$ and $P(x, t) = \rho(x, t)$

$$P(x, t + \tau) = \int \phi(y \rightarrow x) P(y, t) dy$$

Upon expansion $P(x, t) + \tau \frac{\partial P(x, t)}{\partial t} = \int \phi(y \rightarrow x) P(y, t) dy$

$$\frac{\partial P(x, t)}{\partial t} = \int \phi(y \rightarrow x) P(y, t) dy - P(x, t)$$

Consider the fact $\int \phi(x \rightarrow y) dy = \int \phi(-\Delta) d\Delta = 1$ and replace ϕ/τ by W

$$\frac{\partial P(x, t)}{\partial t} = \int W(y \rightarrow x) P(y, t) dy - \int W(x \rightarrow y) P(x, t) dy$$

The above equation is the continuous version of Master equation, and W is the probability jump per unit time. Immediately, one can immediately write its discrete version

Gain of m state due to transition
from the other n states

$$\frac{dP_m(t)}{dt} = \sum_n W_{nm} P_n(t) - \sum_n W_{mn} P_m(t)$$

Loss due to transition from
the m to all other n states

At equilibrium $\frac{dP_m(t)}{dt} = 0$ implies $\sum_n W_{nm} P_n^{eq}(t) = \sum_n W_{mn} P_m^{eq}(t)$

For each pair of n and m , separately
the transitions must balance

$$W_{nm} P_n^{eq} = W_{mn} P_m^{eq}$$

Principle of detailed balance!

1-D Random Walk:

n_r : Total no of leftward jump
 n_l : Total no of rightward jump

The master equation for this walk

$$\frac{dP_{n_l n_r}(t)}{dt} = -(k_r + k_l)P_{n_l n_r}(t) + k_r P_{n_l n_r - 1}(t) + k_l P_{n_l - 1 n_r}(t)$$

This is an infinite set of equations because the number of possible jumps are unlimited

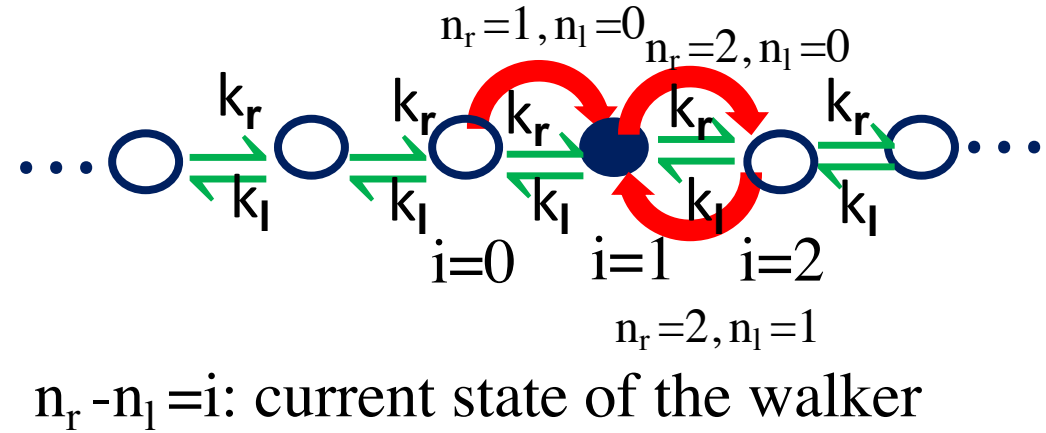
Introducing two auxiliary variables s_r and s_l and associated generating function, G and substitute it into above equation

$$G(s_l, s_r, t) = \sum_{s_l, s_r=0}^{\infty} P_{n_l n_r}(t) s_l^{n_l} s_r^{n_r}$$

$$\frac{dG}{dt} = [-(k_r + k_l) + k_r s_r + k_l s_l]G \quad \text{And its solution} \quad G = e^{\{-k_l(1-s_l)t\}} e^{\{-k_r(1-s_r)t\}}$$

The solution reflects the fact that the joint probability distribution for the left and right jumps reflects two statistically independent Poisson distributions

All the sites are equivalent to each other



Expand the equation, $G = e^{\{-k_l(1-s_l)t\}} e^{\{-k_r(1-s_r)t\}}$ in a Taylor series in s_r and s_l

$$G(s_l, s_r, t) = \sum_{s_l, s_r=0}^{\infty} e^{-k_l t} \frac{(k_l t)^{n_l}}{n_l!} e^{-k_r t} \frac{(k_r t)^{n_r}}{n_r!} s_l^{n_l} s_r^{n_r}$$

Upon comparing with its definitions

$$P_{n_l n_r}(t) = e^{-k_l t} \frac{(k_l t)^{n_l}}{n_l!} e^{-k_r t} \frac{(k_r t)^{n_r}}{n_r!} s_l^{n_l} s_r^{n_r}$$

The distribution of walker's position along the line follows immediately

$$P_i(t) = \sum_{s_l, s_r=0}^{\infty} P_{n_l n_r}(t) \delta_{n_r - n_l, i} = e^{-(k_l + k_r)t} \sum_{n_l=0}^{\infty} \frac{(k_l t)^{n_l}}{n_l!} \frac{(k_r t)^{n_l+i}}{(n_l+i)!}$$

This sum can be shown to approach a Gaussian distribution as time approaches infinity

The mean and variances can be calculated by using the following property of $G(s_r, s_l, t)$

$$\frac{\partial^a \partial^b}{\partial s_l^a \partial s_r^b} G(s_l, s_r, t) |_{s_l=s_r=1} = \langle n_l(n_l - 1) \dots (n_l - a + 1) n_r(n_r - 1) \dots (n_r - b + 1) \rangle(t)$$

Therefore, the mean will be $\langle i \rangle(t) = \langle n_r - n_l \rangle(t) = (k_r - k_l)t$

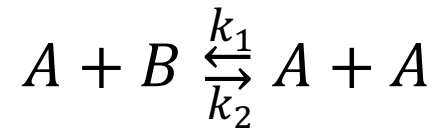
and the variance will be $[\langle i^2 \rangle - \langle i \rangle^2](t) = [\langle (n_r - n_l)^2 \rangle - \langle i \rangle^2](t) = (k_r + k_l)t$

The distribution function,
$$P(i, t) \approx \frac{1}{\sqrt{2\pi(k_r + k_l)t}} \exp \left[-\frac{(i - (k_r - k_l)t)^2}{2(k_r + k_l)t} \right]$$

- ✓ The GF reduces infinite number of coupled equation to a finite number of dynamic equation at the expense of new continuous auxiliary variables
- ✓ The whole exercise is relatively simple, involving only strategic placement of auxiliary variables within the ordinary eom
- ✓ The method provides probability and statistical moments for the occurrence of events under investigation
- ✓ Possible to track multiple types of events within a single GF eom, each type is associated with unique auxiliary variable

Chemical Kinetics:

Consider the following bimolecular reaction



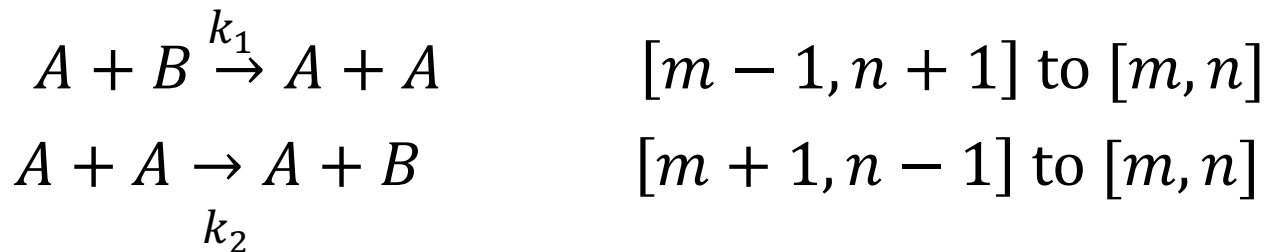
$$m \rightarrow A$$

$$n \rightarrow B$$

$$m + n = N \text{ (Conservation)}$$

A state of the system is completely specified by $[m, n]$

With the above prescription, the following transitions will happen



Then the Master equation takes the following form

$$\frac{dP_{m,n}}{dt} = \frac{k_1}{V} (m - 1)(n + 1)P_{m-1,n+1}(t) - \frac{k_1}{V} mnP_{m,n}(t) + \frac{k_2}{V} (m + 1)^2 P_{m+1,n-1}(t) - \frac{k_2}{V} m^2 P_{m,n}(t)$$

Because of the conservation law, one can eliminate one of the variables

$$\frac{dP_m}{dt} = \frac{k_1}{V} (m - 1)(N - m + 1)P_{m-1}(t) - \frac{k_1}{V} m(N - m)P_m(t) + \frac{k_2}{V} (m + 1)^2 P_{m+1}(t) - \frac{k_2}{V} m^2 P_m(t)$$

The above equation can be written in the form of shift operator as

$$\frac{dP_m}{dt} = \frac{k_1}{V} (E^{-1} - 1)m(N - m)P_m(t) + \frac{k_2}{V} (E - 1)m^2P_m(t)$$

Where the $EP_m(t) = P_{m+1}(t)$ $E^{-1}P_m(t) = P_{m-1}(t)$

Rather than attempting a complete solution of the master equation, we give a mean field description of a chemical reaction. Let us introduce a concentration variable

$$C = \frac{m}{V}; C_0 = \frac{N}{V}; \rho(C, t) = P_m(t)$$

In the limit of large V, we can expand in powers of 1/V

$$EP_m(t) = P_{m+1}(t) = P\left(\frac{m+1}{V}, t\right) = \rho\left(C + \frac{1}{V}, t\right) = \left[1 + \frac{1}{V}\frac{\partial}{\partial C} + \frac{1}{2V^2}\frac{\partial^2}{\partial C^2} + \dots\right]\rho$$

$$E^{-1}P_m(t) = P_{m-1}(t) = P\left(\frac{m-1}{V}, t\right) = \rho\left(C - \frac{1}{V}, t\right) = \left[1 - \frac{1}{V}\frac{\partial}{\partial C} + \frac{1}{2V^2}\frac{\partial^2}{\partial C^2} + \dots\right]\rho$$

$$E = 1 + \frac{1}{V}\frac{\partial}{\partial C} + \frac{1}{2V^2}\frac{\partial^2}{\partial C^2} + \dots \quad \text{And} \quad E^{-1} = 1 - \frac{1}{V}\frac{\partial}{\partial C} + \frac{1}{2V^2}\frac{\partial^2}{\partial C^2} + \dots$$

Upon substitution of these relations into the master equation in all aspects

$$\begin{aligned}\frac{\partial \rho}{\partial t} &= \frac{k_1}{V} V^2 \left[-\frac{1}{V} \frac{\partial}{\partial C} + \frac{1}{2V^2} \frac{\partial^2}{\partial C^2} + \dots \right] C(C_0 - C)\rho + \frac{k_2}{V} V^2 \left[\frac{1}{V} \frac{\partial}{\partial C} + \frac{1}{2V^2} \frac{\partial^2}{\partial C^2} + \dots \right] C^2 \rho \\ &= -\frac{\partial}{\partial C} (k_1 C(C_0 - C) - k_2 C^2) \rho + \frac{1}{2V} \frac{\partial^2}{\partial C^2} (k_1 C(C_0 - C) + k_2 C^2) \rho\end{aligned}$$

Fokker Planck equation

Immediately one can recognize

$$v(C) = (k_1 C(C_0 - C) - k_2 C^2)$$

$$B(C) = (k_1 C(C_0 - C) + k_2 C^2)$$

Stochastic Modelling and Simulation of Biochemical Network

Motivation

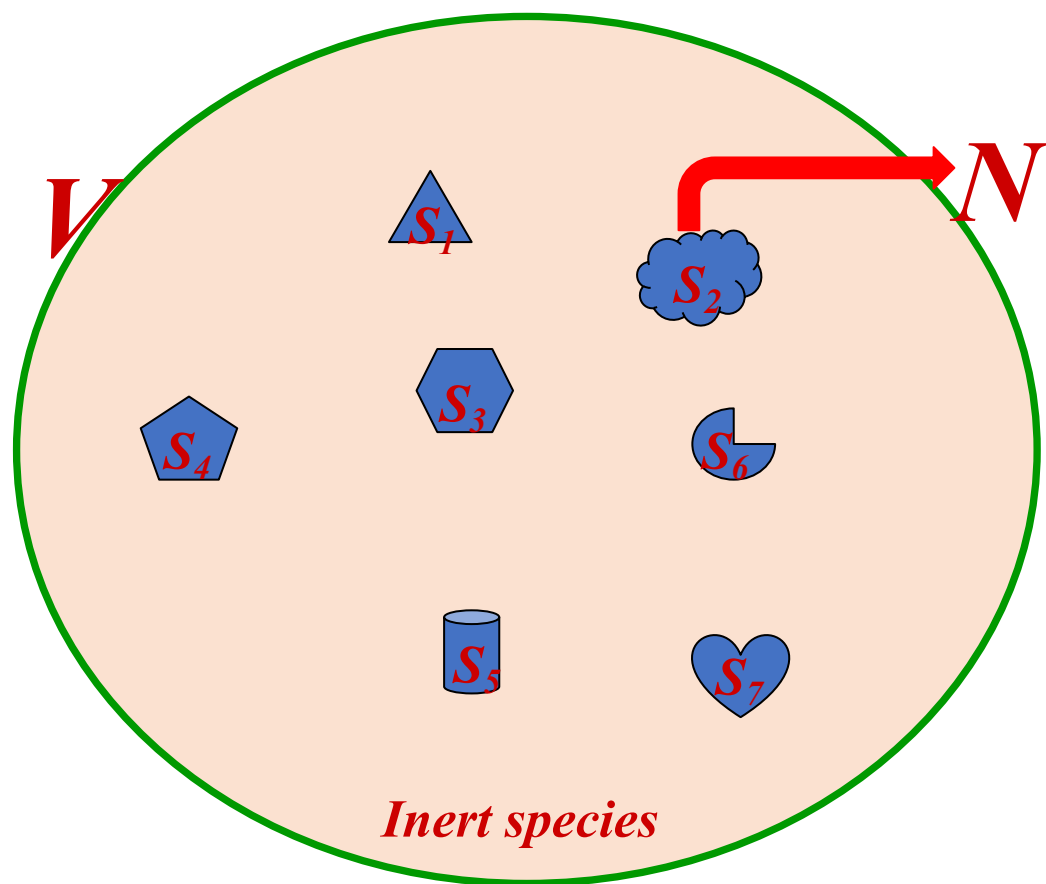
- ◆ Realistic description of chemical system
- ◆ Numerical solution of stochastic master equation
- ◆ Consider the fluctuations and correlations of the chemical system

Advantages over Deterministic

- ◆ Stochastic formulation is always valid where deterministic valid; in fact stochastic formulation reduces to deterministic one at thermodynamic limit
- ◆ Deterministic approach is applicable when only few species are present in the system (both by analytically or numerically); stochastic is applicable for any number of species

Specific Aim for Developing the Algorithm

Numerically simulate the Markov process that master equation describes analytically; even though the master equation never explicitly used during formulation



$X_j \cong$ current number of molecules of chemical species S_j in V

There will be M chemical reactions $\{R_\mu, \mu=1\dots M\}$ and each of the chemical reaction characterized by a reaction parameter c_μ

$*$ \rightarrow reaction products,

$S_j \rightarrow$ reaction products,

$S_j + S_k \rightarrow$ reaction products $(j \neq k),$

$2S_j \rightarrow$ reaction products,

$S_i + S_j + S_k \rightarrow$ reaction products $(i \neq j \neq k \neq i)$

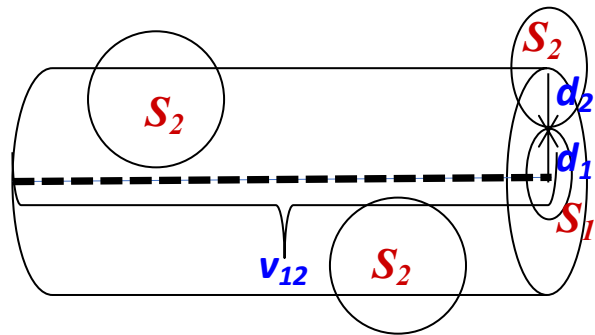
$S_j + 2S_k \rightarrow$ reaction products $(j \neq k),$

$3S_j \rightarrow$ reaction products.

Fundamental Hypothesis

$c_\mu \delta t \equiv$ average probability, to first order in δt , that a particular combination of R_μ reactant molecules will react accordingly in the next time interval δt .

Under what conditions, the above hypothesis has a legitimate physical basis



Consider this reaction $R_\mu : S_1 + S_2 \rightarrow 2S_3$

Center to center distance between two molecules $d_{12} \equiv (d_1 + d_2)/2$

And the relative velocity, v_{12}

“collision volume” $\delta V_{\text{coll}} = \pi d_{12}^2 \cdot v_{12} \delta t$

average probability $\langle \delta V_{\text{coll}}/V \rangle = \pi d_{12}^2 \cdot \langle v_{12} \rangle \delta t/V$

For ideal gas

$$\langle \delta V_{\text{coll}}/V \rangle = V^{-1} \pi d_{12}^2 (8kT/\pi m_{12})^{1/2} \delta t$$

Thus, the reaction parameter

$$c_\mu = V^{-1} \pi d_{12}^2 (8kT/\pi m_{12})^{1/2} \exp(-u_\mu^*/kT)$$

Relation Between c_μ And k_μ

Consider this reaction $R_\mu : S_1 + S_2 \rightarrow 2S_3$

X_1 number of molecule S_1 and X_2 number of molecule S_2 , then there will be distinct $X_1 X_2$ combination of reactant molecules inside V and the addition theorem of probability that $X_1 X_2 c_\mu dt$ gives the probability that the R_μ reaction occur in volume V

Thus the average rate at which R_μ reaction occur in volume V , $\langle X_1 X_2 c_\mu \rangle = \langle X_1 X_2 \rangle c_\mu$

Thus the average rate per unit volume, $\langle X_1 X_2 \rangle c_\mu / V = \langle x_1 x_2 \rangle V c_\mu$, $x_i \equiv X_i / V$,

The k_μ is conventionally defined as the average rate per unit volume divided by the product of the average densities of the reactants

$$k_\mu = \langle x_1 x_2 \rangle V c_\mu / \langle x_1 \rangle \langle x_2 \rangle$$

Correlations and fluctuations automatically come in stochastic formulations!

Reaction Probability Density Function

$\mathcal{P}(X_1, X_2, \dots, X_N; t) \equiv$ the probability that there will be X_1 molecules of S_1 , and X_2 molecules of S_2, \dots , and X_N molecules of S_N , in V at time t ,

and its moments

$$X_i^{(k)}(t) \equiv \sum_{X_1=0}^{\infty} \cdots \sum_{X_N=0}^{\infty} X_i^k \mathcal{P}(X_1, \dots, X_N; t) \quad (i = 1, \dots, N; k = 1, 2, \dots).$$

runs. Particularly useful are the $k = 1$ and $k = 2$ averages; this is because the quantities

$$X_i^{(1)}(t)$$

$$\text{and } \Delta_i(t) \equiv \{X_i^{(2)}(t) - [X_i^{(1)}(t)]^2\}^{1/2}$$

measure, respectively, the average number of S_i molecules in V at time t , and the magnitude of the root-mean-square fluctuations about this average.

Continued...

$$h_{\mu} = 1,$$

reaction probability dens $h_{\mu} = X_j,$

$$h_{\mu} = X_j X_k,$$

$P(\tau, \mu) d\tau \equiv$ probability at time t

reaction in V will occur

time interval $(t + \tau, t + \tau + d\tau)$

$$h_{\mu} = X_j(X_j - 1)/2,$$

$$h_{\mu} = X_i X_j X_k,$$

$P(\tau, \mu)$ is a joint probability density

variable τ ($0 \leq \tau < \infty$) and the differential

$$h_{\mu} = X_j X_k (X_k - 1)/2,$$

$h_{\mu} c_{\mu} \delta t =$ probability, to first

$$h_{\mu} = X_j(X_j - 1)(X_j - 2)/6,$$

occur in V in the next time interval δt .

We shall calculate the probability in (11) as the product of $P_0(\tau)$, the probability at time t that *no* reaction will occur in the time interval $(t, t + \tau)$, times $h_{\mu} c_{\mu} d\tau$, the subsequent probability that an R_{μ} reaction will occur in the next differential time interval $(t + \tau, t + \tau + d\tau)$:

$$P(\tau, \mu) d\tau = P_0(\tau) \cdot h_{\mu} c_{\mu} d\tau.$$

Continued...

$(t, t + \tau)$ to be divided into K subintervals of equal length $\epsilon = \tau/K$.

The probability that *none* of the reactions R_1, \dots, R_M occurs in the first ϵ subinterval $(t, t + \epsilon)$ is, by (13) and the multiplication theorem for probabilities,

$$\prod_{\nu=1}^M [1 - h_\nu c_\nu \epsilon + o(\epsilon)] = 1 - \sum_{\nu=1}^M h_\nu c_\nu \epsilon + o(\epsilon).$$

This is also the *subsequent* probability that no reaction occurs in $(t + \epsilon, t + 2\epsilon)$, and *then* in $(t + 2\epsilon, t + 3\epsilon)$, and so on. Since there are K such ϵ subintervals between t and $t + \tau$, then $P_0(\tau)$ can be written

$$\begin{aligned} P_0(\tau) &= \left[1 - \sum_{\nu} h_\nu c_\nu \epsilon + o(\epsilon) \right]^K \\ &= \left[1 - \sum h_\nu c_\nu \tau / K + o(K^{-1}) \right]^K. \end{aligned}$$

This is true for any $K > 1$, and in particular it is true in the limit of infinitely large K . Therefore,

$$P_0(\tau) = \lim_{K \rightarrow \infty} \left[1 - \left(\left(\sum h_\nu c_\nu \tau + o(K^{-1}) \right) / K \right) \right]^K,$$

Continued...

or, using the standard limit formula for the exponential function,

$$P_0(\tau) = \exp \left[- \sum_{\nu=1}^M h_{\nu} c_{\nu} \tau \right].$$

reaction probability density function,

$$P(\tau, \mu) = h_{\mu} c_{\mu} \exp \left[- \sum_{\nu=1}^M h_{\nu} c_{\nu} \tau \right]$$

To be precise, this formula gives $P(\tau, \mu)$ for $0 \leq \tau < \infty$ and $1 \leq \mu \leq M$, with τ real and μ integer; for all other values of τ and μ , $P(\tau, \mu)$ is zero.

$$\int_0^{\infty} d\tau \sum_{\mu=1}^M P(\tau, \mu) = \sum_{\mu=1}^M h_{\mu} c_{\mu} \int_0^{\infty} d\tau \exp \left[- \sum_{\nu=1}^M h_{\nu} c_{\nu} \tau \right] = 1$$

Step 0 (initialization). Set the time variable $t = 0$. Specify and store initial values for the N variables X_1, X_2, \dots, X_N , where X_i is the current number of molecules of chemical species S_i . Specify and store the values of the M reaction parameters c_1, c_2, \dots, c_M for the M chemical reactions $\{R_\mu\}$. Using (14), calculate and store the M quantities $h_1c_1, h_2c_2, \dots, h_Mc_M$ which collectively determine the reaction probability density function $P(\tau, \mu)$ in (17). Finally, specify and store a series of "sampling times" $t_1 < t_2 < \dots$, and also a "stopping time" t_{stop} .

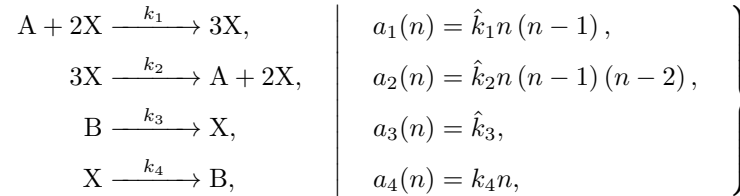
Step 1 By employing suitable Monte Carlo techniques, generate one random pair (τ, μ) according to the joint probability density function $P(\tau, \mu)$ in (17). Explicit methods for doing this are presented in Section 5.

Step 2 Using the numbers τ and μ generated in Step 1, advance t by τ , and change the $\{X_i\}$ values of those species involved in reaction R_μ to reflect the occurrence of one R_μ reaction. Then, recalculate the $h_\nu c_\nu$ quantities for those reactions R_ν whose reactant X_i -values have just been changed. (For example, suppose R_μ is the reaction $S_1 + S_2 \rightarrow 2S_3$. Then after replacing t by $t + \tau$, we would replace X_1, X_2 and X_3 by $X_1 - 1, X_2 - 1$ and $X_3 + 2$, respectively; we would then recalculate $h_\nu c_\nu$ in accordance with (14) for every reaction R_ν in which either S_1 or S_2 or S_3 appears as a *reactant*.)

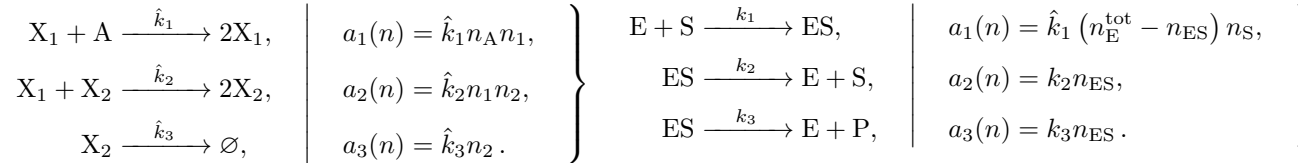
Step 4 If t has just been advanced through one of the sampling times t_i , read out the current molecular population values X_1, X_2, \dots, X_N . If $t > t_{\text{stop}}$, or if no more reactants remain (all $h_\mu = 0$), terminate the calculation; otherwise, return to Step 1.

The Algorithm

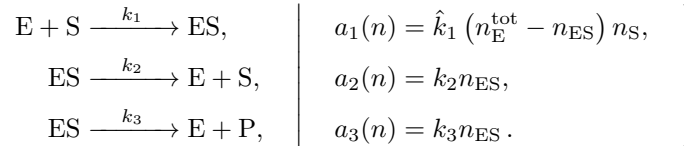
Schlogl model



Lotka-Volterra



Michelis-Menten



1. Initialize the system at $t = 0$ with initial numbers of molecules for each species, n_1, \dots, n_s .
2. For each $j = 1, \dots, r$, calculate $a_j(n)$ based on the current state n .
3. Calculate the exit rate $a_0(n) = \sum_{j=1}^r a_j(n)$. Terminate if $a_0(n) = 0$.
4. Compute a sample τ of the time until the next reaction using (5.21).
5. Update the time $t = t + \tau$.
6. Compute a sample j of the reaction index using (5.23).
7. Update the state n according to R_j . That is, set $n = n + S_{\cdot j}$, where $S_{\cdot j}$ denotes j th column of the stoichiometry matrix S .
8. If $t < t_{\max}$, return to Step 2.

```
function [ssa,ensem] = makeSSA(S,a)
rng('shuffle'); s = size(S,1);
ssa = @gillespie; ensem = @ensemble;
function [tt,nn] = gillespie(n,tmax,dt) % Single run
t = 0; steps = tmax/dt;
tt = zeros(steps,1);
nn = zeros(steps,s);
nn(1,:) = n; idx = 1;
while t<tmax
if all(n==0) % exhaustion check
disp('Reactants_exhausted!');
break;
end
asum = cumsum(a(n));
t = t - (1/asum(end))*log(rand);
j = find(asum>asum(end)*rand,1);
n = n + S(:,j);
if (t - tt(idx)) > dt
idx = idx + 1;
tt(idx) = t;
nn(idx,:) = n;
end
end
tt(idx:end) = []; nn(idx:end,:) = [];
end
function [TT,NN] = ensemble(n0,tmax,dt,runs) % Ensemble
TT = (0:dt:tmax)'; ttmax = zeros(runs,1);
NN = zeros(1+tmax/dt, s, runs);
for i=1:runs
[tt,nn] = gillespie(n0,tmax,dt);
ttmax(i) = tt(end);
NN(:, :, i) = interp1q(tt, nn, TT);
end
idx = (TT > min(ttmax));
TT(idx) = []; NN(idx, :, :) = [];
end
end
```

Suggested Texts and References:

- 1) Stochastic Methods: A Handbook for the Natural and Social Sciences, *Crispin W. Gardiner*
- 2) Theory and Applications of Stochastic Processes: An Analytical Approach, *Zeev Schuss*
- 3) Non equilibrium Statistical Mechanics, *Robert Zwanzig*
- 4) Elements of Non-equilibrium Statistical Mechanics, *V. Balakrishnan*
- 5) Gillespie, Daniel T. "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions." *Journal of computational physics* 22.4 (1976): 403-434.