6/2/04

BP297
Stochastics & Fluctuations

The biochemical reactions considered so far are deterministic.

$$A \underset{k_r}{\overset{k_f}{\rightleftharpoons}} B$$

$$\frac{dA}{dt} = -k_f A + k_r B$$

$$\frac{dB}{dt} = k_f A - k_r B$$

For given initial conditions, at a given time t, the results are always exactly the same. But in real biological systems, we have 2 problems: fluctuations, and the small numbers problem.

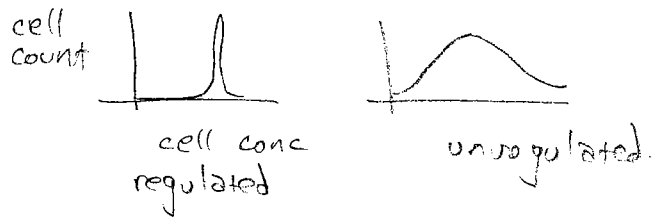
Small numbers problem: Compute the number of H^+ atoms in an E coli: it's about 100. How many chromosomes in an E coli? 1.

Fluctuations matter. Examples

(1) In a race, it's not the average rate that matters. It's who gets there first. Also, activated kinetics is not about average molecules; it's about the highest-energy molecules.

(2) Biology works hard to beat the fluctuations

A Becskei & L Serrano Nature 405 590 (2000)

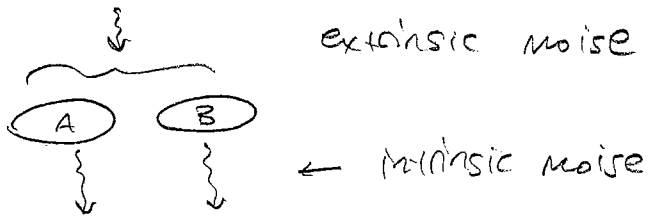


(3) Stochastics determine biologically important variability

- Noise generates errors in DNA replica (Rao et al Nature 420 231 (2002)).
- Noise causes divergence of cell fates, creates cell individuality: lytic-lysogeny path, for example (J Hasty et al PNAS 97 2075 (2000)).
- Aggregation is very sensitive to flucs in prot conc: one person gets Alzheimers'; another doesn't. Irrev.

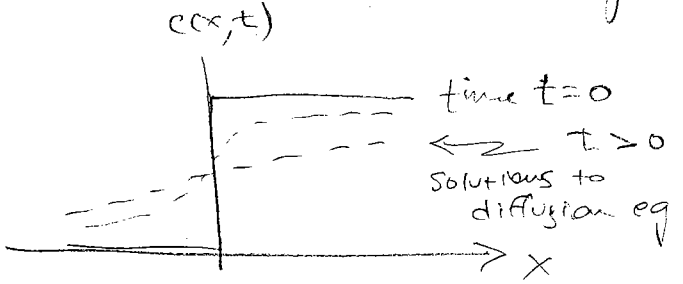
(4). Recent work: extrinsic vs intrinsic noise:

(a) Swain et al PNAS 99 12,795 (2002), (b) Elowitz et al Science 297 1183 (2002), (c) Erin O'Shea.

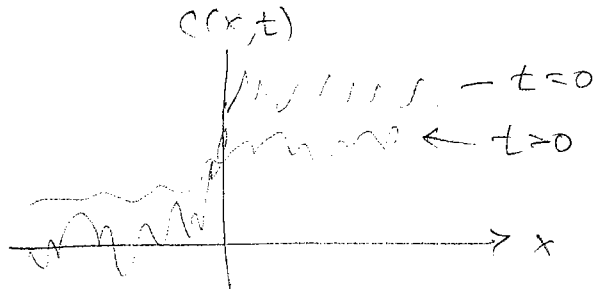


First, Diffusion

The diffusion equation does not describe fluctuations



Diffn Eq.



"Real" Systems

Why? Because Fick's law $J = -D\frac{\partial c}{\partial x}$

is about <u>averages</u>, $\langle J\rangle = -D\frac{\partial \langle c\rangle}{\partial x}$, not about distributions. Here's where it comes from

(1) $\langle J\rangle = \langle c\rangle\langle v\rangle$

(2) force $\langle f\rangle = \zeta\langle v\rangle$

(3) force $\langle f\rangle = -\frac{\partial \mu}{\partial x} = -kT\frac{\partial \ln c(x)}{\partial x}$

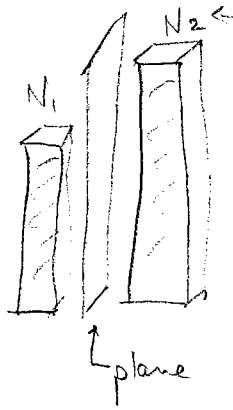
Combine:

$$\langle J\rangle = -\frac{ckT}{\zeta}\frac{\partial \ln c}{\partial x} = -\frac{kT}{\zeta}\frac{\partial c}{\partial x} = -D\frac{\partial c}{\partial x}$$

where $D = \frac{kT}{\zeta}$.

Let's use a 2-layer model instead to capture fluctuations

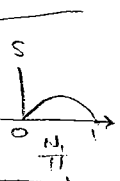
The basic question: What is the physical driving force for flow of particles?



Thermo says the number of configs in this system is

$$\Omega = \frac{(N_1 + N_2)!}{N_1!\, N_2!}$$

$$= -N_1 \ln\left(\frac{N_1}{N}\right) - N_2 \ln\left(\frac{N_2}{N}\right)$$



The entropy $S = k \ln \Omega$ is maximal when $N_1^* = N_2$, so that's the state of equilibrium.

Our question is how it gets there. A particle can jump from either column to the other in time Δt with probability p. It stays, with probability $q = 1 - p$.

Unlike thermo, where we define *states*, for dynamics we define a *process* that takes place between time t and $t + \Delta t$, in which m_1 particles jump from column 1 and m_2 particles jump from column 2. Particles are independent and jumps are too.

Principle of Maximum Caliber

We write a dynamical partition function

$$W(n_1 m_1 n_2 m_2) = \left[p^{m_1} q^{N_1 - m_1} \frac{N_1!}{m_1!\,(N_1-m_1)!} \right] \left[p^{m_2} q^{N_2 - m_2} \frac{N_2!}{m_2!\,(N_2-m_2)!} \right]$$

that applies between times t and $t + \Delta t$ when $N_1 = N_1(t)$ and $N_2 = N_2(t)$.

In the same spirit that the Second Law says to maximize S to predict equilibrium, now too for dynamics, we maximize W over all possible processes (ie over m_1 and m_2), to predict the flux. Since they are independent, we maximize W_1 and W_2 separately.

$$\left. \frac{\partial \ln W_1}{\partial m_1} \right|_{N_1,\, m_1 = m_1^*} = 0$$

$$\ln W = m_1 \ln p + (N_1 - m_1) \ln q + N_1 \ln N_1 - m_1 \ln m_1 - (N_1 - m_1) \ln (N_1 - m_1)$$

We call $\mathcal{C} = \ln W$ the <u>caliber</u>.

So

$$\frac{\partial \ln W_1}{\partial m_1} = \ln p - \ln q - \ln m_1 - 1 + \ln(N_1 - m_1) + 1 = 0$$

$$\Rightarrow \ln\left(\frac{m^*}{N - m^*}\right) = \ln\left(\frac{p}{1-p}\right) = \ln \alpha$$

$$\Rightarrow m^* = \alpha N - \alpha m^*$$

$$\Rightarrow \frac{m^*}{N} = \frac{\alpha}{1+\alpha} = p$$

$$\Rightarrow \boxed{m^* = pN}$$

So, the flux is

$$\langle J \rangle = m_1^* - m_2^* = p\,(N_1(t) - N_2(t))$$

which is equivalent to Fick's Law

$$\langle J \rangle = -D\,\frac{\Delta C}{\Delta x} = -D\,\Delta C \quad \text{if } \Delta x = 1 \text{ unit.}$$

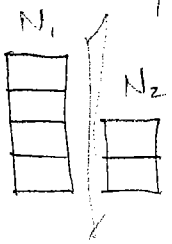
* The principle of Maximum Caliber recovers for us Fick's Law for the mean flux at time t.
Now, let's look at the fluctuations

---

* Note that this derivation of Fick's Law does not require an equilibrium assumption $f = -\frac{\partial \mu}{\partial x}$, where $\mu = kT\ln c$, or averaging $\langle J \rangle = \langle c \rangle v$.

Now, the Fluctuations

Simple example $N_1 = 4$; $N_2 = 2$



For simplicity $p = q = \frac{1}{2}$: Multiplicity table

m_1 \ m_2	0	1	2
0	1	2	1
1	4	8	4
2	6	12	6
3	4	8	4
4	1	2	1

Maxwell's Demon. (backflux).

$m_1^* = 2$ $m_2^* = 1$ $\Rightarrow$

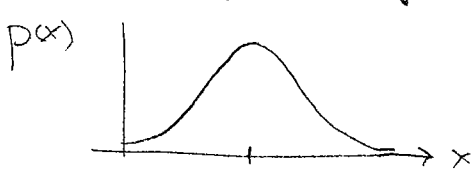
$\langle J \rangle = 1$

superflux

You can use W to compute the relative populations of all the possible fluxes.

Some Properties of the Fluctuations

Note that a binomial distribution resembles a gaussian function, for large enough x.



$$p(x) \propto e^{-\frac{(x-\langle x\rangle)^2}{2\sigma^2}} \approx e^{-\frac{(m_1 - N_1 p)^2}{2N_1 pq}}$$

where the mean is $\langle x_1\rangle = N_1 p = m_1^*$

and standard dev is $\sigma_1 = \sqrt{N_1 pq}$

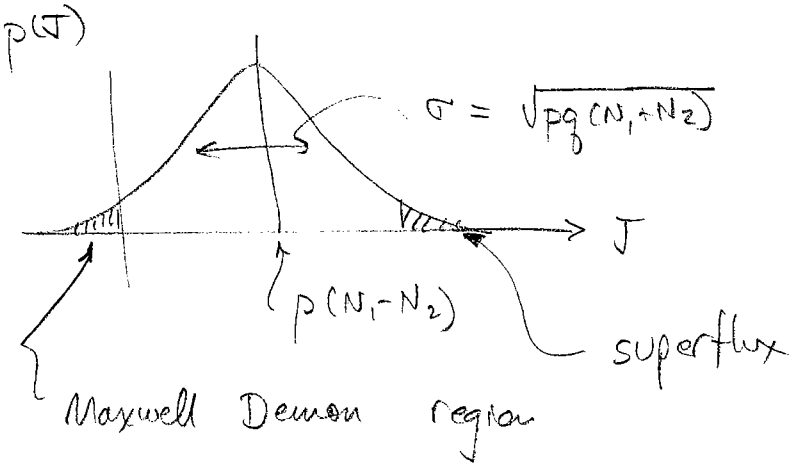
We are interested in the distribution of flux $P(J)$, not the distributions of m_1 & m_2 individually.

$$J = m_1 - m_2. \qquad (1)$$

A law of gaussian functions: If you have a gaussian in m_1 with mean m_1^* and STD σ_1 and a gaussian in m_2 with mean m_2^* and " σ_2, then the distribution of the difference, Eq (1) (or sum) is also a gaussian, with mean $(m_1^* - m_2^*)$ and STD $\sigma^2 = \sigma_1^2 + \sigma_2^2$, so

$$p(J) \propto e^{-\frac{(J - pAN)^2}{2pqN}}$$

where $\Delta N = N_1 - N_2$ and $N = N_1 + N_2$



Some properties of p(J).

① The $\frac{\text{STD}}{\text{mean}}$ indicates the size of backflux

$$= \frac{\sqrt{pqN}}{pfN} \sim \frac{1}{N^{1/2}} \quad \text{which} \to 0 \text{ as } N \uparrow.$$

where $f = \frac{N_1 - N_2}{N_1 + N_2}$. $\Rightarrow$ Small numbers problem: noise decreases as system size gets larger.

② Even if N gets large, if $\Delta N \to 0$, there's substantial backflux.

Now, Chemistry & Fluctuations

$A \underset{k_r}{\overset{k_f}{\rightleftharpoons}} B$ The rate coefficients k_f and k_r are for average rates. Sometimes an A converts faster, and sometimes slower, than this.

Best-known example. Radioactive decay

$A \xrightarrow{k_f} B$ ← radon

A: radium

B: α-particle

If you have N radium atoms, then in a time interval t, the probability is

$$P(m) = \frac{a^m e^{-a}}{m!} \qquad m = 0, 1, 2, \ldots$$

that m of them will decay in that time interval. This is a Poisson distribution $a = m^+ = Np$.

Now, we show that the Poisson distribution is just a manifestation of the model above when

(1) $p \ll q$
(2) $N \to \infty$ such that $a = pN$ is finite.

Again, start with

$$P(m,N) = p^m q^{N-m} \frac{N!}{m!(N-m)!} \qquad (2)$$

We will use

$$\lim_{N\to\infty} \left(1 - \frac{a}{N}\right)^N = e^{-a}.$$

So, re-write Eq (2) as

$$P(m,N) = p^m \frac{(1-p)^N}{(1-p)^m} \frac{N!}{(N-m)!} \frac{1}{m!}$$

$$= p^m \frac{\left(1-\frac{a}{N}\right)^N}{(1-p)^m} N^m \frac{1}{m!}$$

$$= \frac{p^m e^{-a}}{m!}$$

where $(1-p)^m \approx 1$ since m is small and $p \ll 1$, and $N(N-1)\cdots(N-m+1) \approx N^m$.

Waiting Times

Single-molecule experiments often measure a quantity related to the Poisson dist; called the waiting time distribution. The WTD is the distribution of <u>times you wait between events</u>, rather than the number of events per unit time. For situations that give the Poisson dist for the latter, the WTD is exponential, ie

WTD $$q(t) = \lambda e^{-\lambda t} \qquad (3)$$

PD if $$p(m,t) = \frac{(\lambda t)^m e^{-\lambda t}}{m!}$$

We can derive $q(t)$ from $p(t)$, or use MaxCal again in a simple way.

$$S = \sum q_t \ln q_t \qquad \text{subject to}$$

$$\sum q_t = 1$$

$$\sum t q_t = \langle t \rangle = \tau$$

$$dC = -\sum (-1 - \ln q_t^* - \alpha - \lambda t)\, dq_t = 0$$

$$\Rightarrow \quad q_t^* = \frac{e^{-\lambda t}}{\sum_t e^{-\lambda t}} = \frac{1}{Z} e^{-\lambda t}$$

To normalize $q_t^* = \lambda e^{-\lambda t} \quad \Rightarrow \quad Z = \frac{1}{\lambda}$

so $\langle t \rangle = \tau = \sum t q_t^* = \sum t e^{-\lambda t} = -\frac{\partial \ln Z}{\partial \lambda}$

$$= + \frac{1}{\lambda}$$

so $$\boxed{q_t^* = \lambda e^{-\lambda t} = \frac{1}{\tau} e^{-t/\tau}}$$

This is the waiting time used also in the Gillespie algorithm.

Two-Step Rxns.

$$A \xrightarrow{\tau_1} B \xrightarrow{\tau_2} C$$

$$g_A(t_1) = \frac{1}{\tau_1} e^{-t/\tau_1}$$

$$g_B(t_2) = \frac{1}{\tau_2} e^{-t/\tau_2}$$

From A to C, we have

$$g(t) = \int_0^t g_A(t_1)\, g_B(t-t_1)\, dt_1$$

$$= \int \left(\frac{1}{\tau_1} e^{-t_1/\tau_1}\right)\left(\frac{1}{\tau_2} e^{-(t-t_1)/\tau_2}\right) dt_1$$

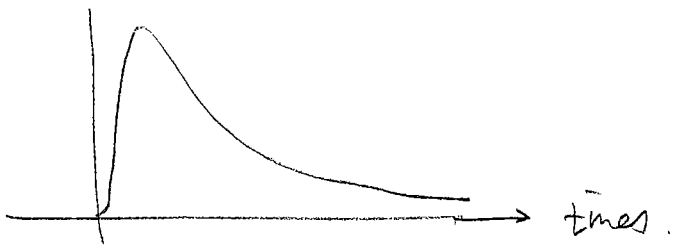
$$= \left(\frac{1}{\tau_1 \tau_2}\right) e^{-t/\tau_2} \underbrace{\int_0^t e^{-t_1\left(\frac{1}{\tau_1}-\frac{1}{\tau_2}\right)} dt_1}$$

$$\frac{e^{-t\left(\frac{1}{\tau_1}-\frac{1}{\tau_2}\right)} - 1}{-\left(\frac{1}{\tau_1}-\frac{1}{\tau_2}\right)}$$

$$= \frac{1/(\tau_1\tau_2)}{\left(\frac{1}{\tau_1} - \frac{1}{\tau_2}\right)} \left(e^{-t/\tau_2} - e^{-t/\tau_1}\right)$$

$$= \frac{1}{\tau_2 - \tau_1} \left(e^{-t/\tau_2} - e^{-t/\tau_1}\right)$$

which looks like this



Langevin Eqn / Gillespie / Stochsim

$$m\frac{dv}{dt} = -\xi v + f(t)$$ Langevin

$f(t)$ fluctuating force with props

$\langle f(t) \rangle = 0$

$\langle f(t) f(t') \rangle = A\delta(t-t')$

What you can do with this is to get various time correlation fcns. For instance, multiply by $v(0)$ and take the avg:

$$m \frac{d}{dt} \langle v(0)\, v(t) \rangle = -\xi \langle v(0)\, v(t) \rangle + \underbrace{\langle v(0) f(t) \rangle}_{=0 \text{ uncorrelated}}$$

Solving gives

$$\langle v(0)\, v(t) \rangle = A e^{-\frac{\xi}{m} t}$$

Since $\langle v(0)^2 \rangle = \frac{kT}{m}$, it means $A = \frac{kT}{m}$

For complex systems, this can be done by the Gillespie algorithm or StochSim:

- D Gillespie J Phys Chem 81 2340 (1977)
- M Gibson & J Bruck J Phys Chem 104 1876 (2000)
- Rao et al Nature 420 231 (2002).