

Extracting moments from Master Equations.

Bahram Houchmandzadeh

CNRS & Grenoble Universités, Lab. Spectrométrie Physique, BP87, 38402 St-Martin d'Hères Cedex, France.

(Dated: Last update : November 27, 2009)

This short tutorial is a first approach to discrete Master Equation manipulation. Only the simplest cases are considered, mostly when moment closure is exact, *i.e.* moments of order k only depend on moments of lower order.

I. INTRODUCTION.

It is not easy to solve a discrete Master Equation (ME), even for the simplest cases. It is however possible to extract various moments from the ME, and write directly a differential equations for them. We are mostly interested in first and second moments (variance and correlations). We'll show here, through few examples, how this is done, going from simple, zero dimensional systems to fully spatial ones.

II. 1-COMPONENT SYSTEM.

Suppose we have a system which can be found in various states n at time t with probabilities $P(n, t)$. What we know about this system is the jump probabilities, *i.e.* the *rate of transitions* between different states, which we denotes by $W^+(n)$ and $W^-(n)$. We will suppose, for simplicity, that these jumps are independent of time, but this constraint can be relaxed easily. The meaning of these rates is the following : If we prepare N (think one billion, or one million billion, a big number) independent system in the state n , after a (short time) dt , we'll find that a proportion $W^\pm(n)dt$ of them have jump into state $n \pm 1$. Again, for the sake of simplicity, we suppose that there is no jump to states other than $n \pm 1$.

Now, we are in the position of formulating a master equation by counting the proportion of systems coming to state n and the proportion leaving it :

$$\frac{dP(n, t)}{dt} = W^+(n-1)P(n-1) - W^+(n)P(n) + W^-(n+1)P(n+1) - W^-(n)P(n) \quad (1)$$

The above equation is in fact an (enumerable) infinite number of coupled equations, one for each n .

A. zero-th moment : \sum_n

Now, suppose that we sum all these equations. We'll alternatively say that we apply the \sum_n operator to eq.(1). On the left hand side, we can change the order of \sum_n and (d/dt) operators, and P being a probability, we get

$$\begin{aligned} (d/dt) \sum_n P(n, t) &= d1/dt \\ &= 0 \end{aligned}$$

On the right hand side, let us look at the first term, containing the $n-1$:

$$\sum_{n=-\infty}^{n=\infty} W^+(n-1)P(n-1) = \sum_{n=-\infty}^{n=\infty} W^+(n)P(n) \quad (2)$$

To get this equality, we have changed our summation variable from n to $n-1$ (*i.e.* replaced all n 's by $n+1$). The sum being between minus and plus infinity, the domain of summation does not change. This is just to say that for example, $\sum_n n$ and $\sum_n (n-1)$ are the same thing, because both are short hand notation for $\dots(-2) + (-1) + 0 + 1 + 2$. Thus, the right hand side of eq.(1) is also zero. We have got, by applying the \sum_n operator, $0 = 0$, which we already knew.

B. first-moment : $\sum_n n$

This exercise of index summation can be pushed forward to gain more information about the dynamic of the system. Suppose that before summing all the equations, we multiply each by its index, n and then sum them up, thus applying the $\sum_n n$ operator. For the lhs of eq.(1), we get

$$\frac{d \sum_n n P(n,t)}{dt} = \frac{d \langle n \rangle}{dt} \quad (3)$$

where $\langle n \rangle$ denotes “average n ”. In fact, you should, whenever you see a symbol like $\langle f(n) \rangle$, think of replacing it with the words “average of function $f(n)$ ”, or with the mathematical operation $\sum_n f(n)P(n)$. If $E(n)$ is the energy of a hydrogen atom in quantum state n , and $P(n)$ the probability of finding an atom in state n , then $\langle E(n) \rangle$ is the average energy of the hydrogen atom present in our cavity.

What about the right hand side ? Repeating the operation of eq.(2) for the first term, we get

$$\sum_{n=-\infty}^{n=\infty} n W^+(n-1) P(n-1) = \sum_{n=-\infty}^{n=\infty} (n+1) W^+(n) P(n)$$

and when adding the second term, what remains is just $\sum_n W^+(n) P(n)$ which, as we said, we denote $\langle W^+(n) \rangle$. Applying the same trick to the third and fourth term gives us $\langle -W^-(n) \rangle$, so finally, we get a full equation for the evolution of the average :

$$\frac{d \langle n \rangle}{dt} = \langle W^+(n) - W^-(n) \rangle \quad (4)$$

a. *Example 1 : brownian movement.* Consider a particle on a 1d lattice , hopping to the left or right side with the same rate β , independent of its position n on the lattice. Then, $W^+(n) = W^-(n)$ and thus, $(d/dt) \langle n \rangle = 0$. Thus, the average position of the particle remains constant.

b. *Example 2 : birth-death.* Consider an ecosystem composed of n_0 organism at time t_0 . Birth and death events are random phenomena which occur by rates λ and μ : the probability for one birth(death) to occur at time t , during a short interval dt , if there are n individuals present at this time is λndt . Thus, in our notation, $\langle W^+(n) \rangle = \lambda n$ and $\langle W^-(n) \rangle = \mu n$. The average number of the system is thus given by

$$\frac{d \langle n \rangle}{dt} = \langle \lambda n - \mu n \rangle = (\lambda - \mu) \langle n \rangle$$

This is just a first order differential equation for the average, which has the following solution :

$$\langle n \rangle = n_0 e^{(\lambda - \mu)t}$$

The average number in the ecosystem grow or decrease exponentially, depending on the relative value of birth and death rate.

C. Variance : $\sum_n n^2$

The above game generalizes to moment of order k : to get the evolution of $\langle n^k \rangle$, apply the operator $\sum_n n^k$ to the master equation. We seldom go beyond the second moment in the every day life, but this quantity is of prime importance. Playing the index game for the second moment thus gives us

$$\frac{d \langle n^2 \rangle}{dt} = 2 \langle n(W^+(n) - W^-(n)) \rangle + \langle W^+(n) + W^-(n) \rangle \quad (5)$$

What is even more interesting is the centered second moment, called the variance $V = \langle n^2 \rangle - \langle n \rangle^2$. Using eqs.(4,5), one get

$$\begin{aligned} \frac{dV}{dt} &= \frac{d \langle n^2 \rangle}{dt} - 2 \langle n \rangle \frac{d \langle n \rangle}{dt} \\ &= 2 \langle (n - \langle n \rangle)(W^+(n) - W^-(n)) \rangle + \langle W^+(n) + W^-(n) \rangle \end{aligned}$$

c. Example 1 : brownian movement. We get

$$\frac{d\langle n^2 \rangle}{dt} = 2\beta$$

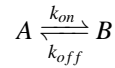
and thus, $\langle n^2 \rangle = 2\beta t$ (assuming the particle is at $n = 0$ at time 0). This is a well known result of brownian motion. Moreover, if we call ℓ the lattice size, the position $x = \ell n$ and $\langle x^2 \rangle = 2\beta \ell^2 t$. The coefficient $D = 2\beta \ell$ is called the diffusion coefficient.

d. Example 2 : birth-death. Let us suppose that the birth and death rate are equal ($\lambda = \mu$) (if not, the variance will also grow or decrease exponentially). There is no change in the average then : $\langle n \rangle = n_0$. The second moment reads,

$$\frac{d\langle n^2 \rangle}{dt} = (\lambda + \mu)n_0 t + n_0^2$$

The variance thus grows (linearly) in time, even though the average remains constant ! This is a particular case where the variance can become much larger than the square of the average.

e. Exercise 1. For the chemical equation



obtain the mean and variance of species A, supposing that all molecules are A at $t = 0$.

D. The characteristic function : $\sum_n \exp(ins)$

Instead of painfully computing the moments one after the other, there is way to compute them all together at once. At least formally. Consider a function $\phi(s, t)$ defined on the interval $[0, 2\pi]$, of which $P(n, t)$ are the Fourier coefficients at time t :

$$\phi(s, t) = \sum_{n=-\infty}^{+\infty} P(n, t) \exp(ins)$$

If ϕ is known, it is straight forward to get P :

$$P(n, t) = \frac{1}{2\pi} \int_0^{2\pi} \phi(s, t) \exp(-ins) ds$$

Even without computing the above integral, we can get all the moments by differentiating ϕ . For example, the first two moments are given by

$$\begin{aligned} \langle n \rangle &= \sum_n n P(n, t) = \frac{1}{i} \frac{\partial \phi}{\partial s} \Big|_{s=0} \\ \langle n^2 \rangle &= \sum_n n^2 P(n, t) = - \frac{\partial^2 \phi}{\partial s^2} \Big|_{s=0} \end{aligned}$$

and so on : the k -th moment is equal to the k -th derivative of ϕ at point $s = 0$, scaled by $(1/i)^k$.

Now, let us apply the operator $\sum_n \exp(ins)$ to the master equation (1). Playing the dummy index game again and exchanging summation and derivation when ought to bring us

$$\frac{\partial \phi}{\partial t} = (e^{is} - 1) \langle e^{ins} W^+(n) \rangle + (e^{-is} - 1) \langle e^{ins} W^-(n) \rangle \quad (6)$$

This doesn't seem very helpful until we notice that the term inside the $\langle \rangle$'s can, sometimes, be related to the function ϕ and its derivatives.

f. *Example 1 : brownian motion.* There, the jump probabilities are constant and independent of n : $W^+(n) = W^-(n) = \beta$. Thus, each bracket of eq.(6) gives back exactly the function ϕ and

$$\frac{\partial \phi}{\partial t} = \beta(e^{is} + e^{-is} - 2)\phi = 2\beta(\cos(s) - 1)\phi \quad (7)$$

As the particle was at $n = 0$ at time $t = 0$, $P(n, 0) = \delta_{n,0}$ and $\phi(s, 0) = 1$. The solution of eq.(7) is thus

$$\phi(s, t) = \exp[2\beta(\cos(s) - 1)t]$$

We can now take the various derivatives and obtain the corresponding moments. We can do a little more here and get the probabilities, because the integration of the above function is well known :

$$P(n, t) = e^{-2\beta t} I_n(2\beta t) \quad (8)$$

where I_n is the Bessel I function of order n . For long times, i.e. $2\beta t > n^2/4$, a very good approximation of the above function is

$$P(n, t) = \frac{1}{\sqrt{2\pi(2\beta t)}} e^{-n^2/2(2\beta t)}$$

The density probability of finding a particle between $x = \ell n$ and $x = \ell(n + 1)$ is $P(x, t)\ell = P(n, t)$ and therefore,

$$P(x, t) = \frac{1}{\sqrt{2\pi D t}} e^{-x^2/2Dt} \quad (9)$$

where we have defined the diffusion coefficient $D = 2\beta\ell^2$. This is the usual gaussian of brownian motion.

g. *Example 2 : birth-death.* We will again take the special case of equal birth and death rate $\lambda = \mu = 1$. Therefore, $W^+(n) = W^-(n) = n$, and $\langle \exp(ins)n \rangle = -i\partial_s \phi$. The characteristic differential equation reads :

$$\partial_t \phi = 2(\cos(s) - 1)\partial_s \phi$$

which can be solved :

$$\phi(s, t) = \left[\frac{(t-1)e^{is} - t}{te^{is} - (t+1)} \right]^{n_0} \quad (10)$$

For $n_0 = 1$ (exactly one organism at the initial time)

$$\begin{aligned} P(0, t) &= t/(1+t) \\ P(n, t) &= t^{n-1}/(1+t)^{n+1} \quad n > 0 \end{aligned}$$

E. Generalization.

What we said above about three specific cases can obviously be extended to any function $f(n)$:

$$\frac{d \langle f(n) \rangle}{dt} = \langle (f(n+1) - f(n))W^+(n) \rangle - \langle (f(n) - f(n-1))W^-(n) \rangle$$

Note that for the moments $f(n) = n^k$, $f(n) - f(n-1)$ is a polynomial of order $k-1$; if the jump rates are no more than linear in n , both sides of the above rate equation will be polynomials of the same degree. For many realistic situations, this will not be the case and some kind of approximation and cutting will be involved. The various scheme for moments closure approximations are beyond the scope of this document¹.

¹ The great majority of them lack mathematical rigor ; the difficult task of estimating the errors introduced by such schemes is often done through numerical simulations.

F. Problem.

Suppose that we have a large pool of individuals forms of S different species, equally represented. We choose N individuals from the pool, what is the probability $P(s, N)$ of obtaining s different species $P(s, N)$? [Help : imagine a process by which you pick individuals one by one. Then, $W(s \rightarrow s) = s/S$ and $W(s \rightarrow s+1) = 1 - s/S$, where s is the number of different species already present in your bag. Form the master equation and solve exactly for $P(s, N)$. You can also extract the moments and show that

$$\langle s(N) \rangle = S(1 - (1 - 1/S)^N)$$

Push the computation to the variances, and evaluated the number of species missed after 10 picking if you have obtained 4 species.

III. 2-COMPONENTS SYSTEM.

What we'll do here is merely a repetition of the above exercises. However, as we'll encounter for the first time correlations, we'll do it in some details. It is also a good first step before tracking the spatially extended systems. Suppose now that our system has a probability $P(n, m; t)$ to be in state (n, m) at time t . The master equation reads

$$\begin{aligned} \frac{dP(n, m)}{dt} = & W_1^+(n-1, m)P(n-1, m) - W_1^+(n, m)P(n, m) + \\ & W_1^-(n+1, m)P(n+1, m) - W_1^-(n, m)P(n, m) + \\ & W_2^+(n, m-1)P(n, m-1) - W_2^+(n, m)P(n, m) + \\ & W_2^-(n, m+1)P(n, m+1) - W_2^-(n, m)P(n, m) \end{aligned} \quad (11)$$

Applying $\sum_{n, m} n$ will give the evolution of $\langle n \rangle$ (note that the third and fourth lines of eq.(11) do not contribute to this sum) : $d\langle n \rangle / dt = \langle W_1^+(n, m) - W_1^-(n, m) \rangle$. We would get a similar result for $\langle m \rangle$ if we had applied $\sum_{n, m} m$.

The next step is to compute the evolution of $\langle n^2 \rangle$ by applying $\sum_{n, m} n^2$. The formal expression we get is similar to the one of the last section. The main difference is that if $W_1^-(n, m)$ contains m explicitly (as it often does, we'll see a basic example below), then $d\langle n^2 \rangle / dt$ will depend on correlations $\langle nm \rangle$. We can get an equation for this latter by applying $\sum_{n, m} nm$ to the ME :

$$\frac{d\langle nm \rangle}{dt} = \langle m(W_1^+(n, m) - W_1^-(n, m)) \rangle + \langle n(W_2^+(n, m) - W_2^-(n, m)) \rangle \quad (12)$$

So the task will be to solve two coupled differential equations.

h. Example 1 : brownian motion again ! We suppose this time that our brownian particle jumps around a $2d$ -lattice, and the jump rate is, as before, β : the probability for the particle to jump, in the short interval dt , to one of its neighbor is βdt , and hence, its probability to leave the cell is $4\beta dt$. As all jumps probabilities are equal, we'll get $d\langle n \rangle / dt = d\langle m \rangle / dt = 0$. The second moments do not contain correlations and as before, $d\langle n^2 \rangle / dt = d\langle m^2 \rangle / dt = 2\beta$. By putting $x = n\ell$, $y = m\ell$ and $r^2 = x^2 + y^2$, we obtain $\langle r^2 \rangle = 4\beta\ell^2 t = Dt$. In general, diffusion coefficient at d dimension is defined as $D_d = 2d\beta\ell^2$ and $\langle r^2 \rangle = Dt$.

i. Example 2 : protein production from a non-regulated gene. This is a two step process : first, RNA is produced from DNA and then RNA is read by ribosomes to produce proteins. Both RNA and protein are subject to degradation (with rates μ_r and μ_p). Let's call r and p the number of RNAs and proteins inside the cell. RNA production from the gene is a poissonian stochastic process with rate λ_r and thus,

$$W_1^+(r, p) = \lambda_r \quad (13)$$

$$W_1^-(r, p) = \mu_r r \quad (14)$$

As you notice, none of the above equations contain p , so we can treat RNA as a one component chemical reaction $\emptyset \xrightarrow{\lambda_r} R \xrightarrow{\mu_r} R+1$:

$$\frac{d\langle r \rangle}{dt} = \lambda_r - \mu_r \langle r \rangle$$

and the stationary level is given by $\langle r \rangle = \lambda_r / \mu_r$. For the second moment,

$$\frac{d\langle r^2 \rangle}{dt} = 2\lambda_r \langle r \rangle - 2\mu_r \langle r^2 \rangle + \lambda_r + \mu_r \langle r \rangle$$

A short massage of this expression will show that the stationary variance is $Var(r) = \langle r \rangle$: the probability distribution of RNAs is thus poissonian.

The production of proteins depends on the amount of available RNAs and the efficiency of this reaction λ_p :

$$W_2^+(r, p) = \lambda_p r \quad (15)$$

$$W_2^-(r, p) = \mu_p p \quad (16)$$

The stationary average number of proteins is easy to compute and reads $\langle p \rangle = \lambda_p r / \mu_p$. The second moment however contains correlations :

$$\frac{d\langle p^2 \rangle}{dt} = 2\lambda_p \langle rp \rangle - 2\mu_p \langle p^2 \rangle + \lambda_p \langle r \rangle + \mu_p \langle p \rangle$$

So we have to write them down :

$$\begin{aligned} \frac{d\langle rp \rangle}{dt} &= \langle p(\lambda_r - \mu_r r) \rangle + \langle r(\lambda_p r - \mu_p p) \rangle \\ &= -(\mu_r + \mu_p) \langle rp \rangle + \lambda_r \langle p \rangle + \lambda_p \langle r^2 \rangle \end{aligned}$$

As $\langle r \rangle$, $\langle r^2 \rangle$ and $\langle p \rangle$ are known, the above coupled equations are easily solved if enough care is taken of not loosing a parameter along the road. The interested reader can show that the steady state variance is given by $Var(p) / \langle p \rangle = 1 + b$ where b is the average number of proteins produced per RNA. Protein distribution is therefore non-poissonian.

j. Example 3. RNA burst. In a beautiful paper, Golding *et al.*² measured the RNA production in single cells and in real time. They came with the (apparently) strange result that for the random variable RNA, $V / \langle r \rangle = 4$. The RNA production seems simili-poissonian, but with a ratio of 4 instead of one between the variance and the average. In order to explain this result, they imagined the RNA production as a two step process : a gene can be turned ON or OFF. During the ON period, the gene can transcribe RNA.

Let us call n the state of the gene (0 or 1) and r the number of RNA molecule. The state of the gene is a binomial process

$$W_1^+(n) = k(1 - n)$$

$$W_1^-(n) = k'n$$

where k, k' are ON-OFF switch rates. the equilibrium state is given by

$$\langle n \rangle = \frac{k}{k + k'} = \tau$$

where τ is the fraction of time spent by the gene in the ON state and

$$\langle n^2 \rangle = \tau$$

As an honorable binomial process, the variance of n is given by $\tau(1 - \tau)$. Production of RNA can only take place during the ON state of the gene, $\emptyset \xleftarrow{\mu} R \xrightarrow{\lambda n} R + 1$:

$$W_2^+(n, r) = \lambda n$$

$$W_2^-(n, r) = \mu r$$

where λ is the synthesis rate of RNA and μ its degradation as before. Processing these rates as before, we get

$$\langle r \rangle = (\lambda / \mu) \tau$$

² Golding, Paulsson, Zawilski & Cox, Cell, **123**:1025(2005)

the actual number of RNA is as before the ratio of production over degradation, but this time multiplied by the fraction of time when the gene is ON. After processing the correlations, we get

$$\langle r^2 \rangle = (\lambda/\mu)\tau(1 + \alpha)$$

where

$$\alpha = \frac{\lambda(k + \mu)}{\mu(k + k' + \mu)}$$

First, note that if $\tau \ll 1$, then $\langle r \rangle^2 \ll \langle r^2 \rangle$ and thus,

$$\frac{V}{\langle r \rangle} \approx (1 + \alpha)$$

More over, if $k, k' \ll \mu$, then an increase in τ will increase $\langle r \rangle$, but will not affect noticeably the ratio $V/\langle r \rangle$. This is what is observed.

IV. SPATIALLY EXTENDED SYSTEM.

We could also have called this section “infinite components system”. The smart term for this, used in quantum mechanics, is “second quantification”. As an example, suppose that you have placed n_0 particle on *each* cell of an array (which we’ll suppose $1d$ for the moment) ; particles are subject to stochastic evolution such as diffusion. It will be useless to track the spatial probability distribution of each particle, specially as they are indistinguishable. A much more useful information is the probability of having n_i particles on cell i .

This will be our task below. Notations however will become cumbersome very soon, so we should spend some time on them. The state of the system will be referred to by the vector $\mathbf{n} = (\dots, n_i, n_{i+1}, \dots)$ and $P(\mathbf{n}; t)$ will be the probability of being in state \mathbf{n} at time t , *i.e.* having n_i particles on site i . We will frequently have to manipulate the state $(\dots, n_{i-1}, n_i \pm 1, n_{i+1}, \dots)$, *i.e.* the state \mathbf{n}' where one particle has been added (or removed) at site i of state \mathbf{n} . We will note these states by $a_i^+ \mathbf{n}$ and $a_i \mathbf{n}$. The a ’s are often referred to as creation and annihilation operator. When one particle jumps from site i to site $i + 1$, the state \mathbf{n} becomes the state $a_{i+1}^+ a_i \mathbf{n}$. Let us not forget that this is just a short hand notation for : “state $(\dots, n_{i-1}, n_i, n_{i+1}, \dots)$ becomes $(\dots, n_{i-1}, n_i - 1, n_{i+1} + 1, \dots)$ ”.

We will also need a notation for “number of particle on site i ”. When there is no possible confusion on the state \mathbf{n} , we will just use n_i ; if there is a risk, we will use $N_i(\mathbf{n})$. N_i is the counting operator of site i .

Fundamental example : diffusion.

As for before, the jump (to a neighboring site) probability for a given particle is, independent of all others, β . What is then the probability for *one* particle migrating from site i to site $i + 1$ during a short interval dt ? Of course, it is proportional to how many particles are already there on i :

$$W(i \rightarrow i + 1; \mathbf{n}) = \beta N_i(\mathbf{n}) = \beta n_i \quad (17)$$

To write a master equation, *i.e.* write an evolution equation for $P(\mathbf{n}; t)$, we have to count all the possibilities of coming to and leaving the state \mathbf{n} . Leaving state \mathbf{n} is easy : it is enough to be in this state and have one particle jumping from site i to site $i \pm 1$ (whatever i). On the other hand, to come to this state, it is enough to be in the state $a_i^+ a_{i \pm 1} \mathbf{n}$ and have one particle on site i to migrate to site $i \pm 1$. Summing up all these possibilities

$$\frac{dP(\mathbf{n}; t)}{dt} = \sum_i W(i \rightarrow i + 1; a_i^+ a_{i+1} \mathbf{n}) P(a_i^+ a_{i+1} \mathbf{n}; t) - W(i \rightarrow i + 1; \mathbf{n}) P(\mathbf{n}; t) \quad (18)$$

$$+ W(i \rightarrow i - 1; a_i^+ a_{i-1} \mathbf{n}) P(a_i^+ a_{i-1} \mathbf{n}; t) - W(i \rightarrow i - 1; \mathbf{n}) P(\mathbf{n}; t) \quad (19)$$

You could think that we have forgotten some events, like being in state $a_{i+1}^+ a_i \mathbf{n}$ and having one particle leaving $i + 1$ for i . But because we are summing on all sites i , this event is indeed accounted for in the above master equation. As we said, notations are a little cumbersome, but nothing is really complicated. For example, $W(i \rightarrow i + 1; a_i^+ a_{i+1} \mathbf{n}) = \beta(n_i + 1)$ and $W(i \rightarrow i + 1; \mathbf{n}) = \beta n_i$.

Let us just recall now the definition of various moments. The average number of particle at site k is

$$\langle n_k \rangle = \sum_{\{n_i\}} n_k P(\mathbf{n})$$

which is just a generalisation of what we have seen at section III. All other moments are computed alike. For example, the correlation between site k and site ℓ is

$$\langle n_k n_\ell \rangle = \sum_{\{n_i\}} n_k n_\ell P(\mathbf{n})$$

Evolution of the average.

Following the same lines of arguments developed above, to get the average number of particle at site k , we apply the $\sum_{\{n_j\}} n_k$ to the master equation (18). The left hand side will give, as usual, $d\langle n_k \rangle / dt$.

For the right hand side, little care should be taken as we have to perform a double sum $\sum_{\{n_j\}} n_k \sum_i \{\dots\}$. Note that terms in the second summation which do not contain n_k explicitly will vanish after the summation. The only sites with non-vanishing contributions are then $i = k \pm 1$ and $i = k$.

k. (i) Let us look at $i = k - 1$ in full details. There are four terms in the summation. We have to perform first

$$\sum_{\{n_j\}} n_k W(k-1 \rightarrow k; a_{k-1}^+ a_k \mathbf{n}) P(a_{k-1}^+ a_k \mathbf{n}; t) \quad (20)$$

We will play again the ‘‘dummy index’’ game. Performing variables change $n_k \rightarrow n_k + 1$ and $n_{k-1} \rightarrow n_{k-1} - 1$ will transform state $a_{k-1}^+ a_k \mathbf{n}$ into state \mathbf{n} . The boundaries of summation being minus and plus infinity, equation (20) is then rewritten :

$$\sum_{\{n_j\}} (n_k + 1) W(k-1 \rightarrow k; \mathbf{n}) P(\mathbf{n}; t)$$

The second term is just $-\beta \sum_{\{n_j\}} n_k W(k-1 \rightarrow k; \mathbf{n}) P(\mathbf{n}; t)$. The next two terms do not contain n_k explicitly in their transition rates (which involve sites $k-1$ and $k-1$) and their contribution is zero. Therefore, the contribution of the site $i = k - 1$ to the average evolution is

$$\langle W(k-1 \rightarrow k; \mathbf{n}) \rangle$$

By the same token, the contribution of site $i = k + 1$ to the average evolution will be $\langle W(k+1 \rightarrow k; \mathbf{n}) \rangle$.

l. (ii) It remains to compute the contribution of site $i = k$. For

$$\sum_{\{n_j\}} n_k W(k \rightarrow k \pm 1; a_k^+ a_{k \pm 1} \mathbf{n}) P(a_k^+ a_{k \pm 1} \mathbf{n}; t)$$

Changing variables $n_k \rightarrow n_k - 1$ and $n_{k \pm 1} \rightarrow n_{k \pm 1} + 1$, the sum transforms into

$$\sum_{\{n_j\}} (n_k - 1) W(k \rightarrow k \pm 1; \mathbf{n}) P(\mathbf{n}; t)$$

and therefore, the total contribution of site $i = k$ to the average evolution is $-\langle W(k \rightarrow k-1; \mathbf{n}) \rangle - \langle W(k \rightarrow k+1; \mathbf{n}) \rangle$. The evolution equation takes the symmetric form

$$\begin{aligned} \frac{d\langle n_k \rangle}{dt} &= \langle W(k-1 \rightarrow k; \mathbf{n}) \rangle - \langle W(k \rightarrow k-1; \mathbf{n}) \rangle \\ &+ \langle W(k+1 \rightarrow k; \mathbf{n}) \rangle - \langle W(k \rightarrow k+1; \mathbf{n}) \rangle \end{aligned} \quad (21)$$

Remembering the definition of transition rates (17),

$$\frac{d\langle n_k \rangle}{dt} = \beta (\langle n_{k-1} \rangle + \langle n_{k+1} \rangle - 2\langle n_k \rangle) \quad (22)$$

which will become the usual diffusion equation $\partial_t c = D \partial_x^2 c$ when we will take the continuous limit. More on this later.

Evolution of second moments.

Let us recall what we did in the last section. We saw that the contribution of site $i = k \pm 1$ to the average evolution is

$$\langle [(n_k + 1) - n_k] W(k \pm 1 \rightarrow k; \mathbf{n}) \rangle$$

The contribution of site $i = k$ is

$$\langle [(n_k - 1) - n_k] W(k \rightarrow k \pm 1; \mathbf{n}) \rangle$$

A little exercise will show that this generalises naturally to moments of order p , when operator $\sum_{\{n_j\}} n_k^p$ is applied to the diffusion master equation. The contribution of these sites to the evolution of the second moment of site k will be respectively

$$\langle [(n_k + 1)^2 - n_k^2] W(k \pm 1 \rightarrow k; \mathbf{n}) \rangle = \langle (2n_k + 1) W(k \pm 1 \rightarrow k; \mathbf{n}) \rangle$$

and

$$\langle [(n_k - 1)^2 - n_k^2] W(k \rightarrow k \pm 1; \mathbf{n}) \rangle = \langle (-2n_k + 1) W(k \rightarrow k \pm 1; \mathbf{n}) \rangle$$

We therefore get a nice symmetric form for the second moment as before

$$\begin{aligned} \frac{d \langle n_k^2 \rangle}{dt} &= \langle 2n_k W(k+1 \rightarrow k; \mathbf{n}) \rangle - \langle 2n_k W(k \rightarrow k+1; \mathbf{n}) \rangle \\ &+ \langle 2n_k W(k-1 \rightarrow k; \mathbf{n}) \rangle - \langle 2n_k W(k \rightarrow k-1; \mathbf{n}) \rangle \\ &+ \langle W(k+1 \rightarrow k; \mathbf{n}) + W(k-1 \rightarrow k; \mathbf{n}) + 2W(k \rightarrow k-1; \mathbf{n}) \rangle \end{aligned}$$

There is however a small problem here if we recall the definition of transition rates (17)

$$\begin{aligned} \frac{1}{\beta} \frac{d \langle n_k^2 \rangle}{dt} &= -4 \langle n_k^2 \rangle + 2 \langle n_k n_{k+1} \rangle + 2 \langle n_k n_{k-1} \rangle \\ &+ 2 \langle n_k \rangle + \langle n_{k+1} \rangle + \langle n_{k-1} \rangle \end{aligned} \quad (23)$$

The evolution of the second moment at site k depends on the correlations between this site and its neighbor $\langle n_k n_{k \pm 1} \rangle$, which are unknown quantities.

We had already encountered this problem when handling two-component systems. To compute $\langle n_k n_{k+1} \rangle$, we just have to apply $\sum_{\{n_j\}} n_k n_{k+1}$ to the master equation (18). Sites $i = k-1, \dots, k+2$ will contribute to the sum. We will drop writing the state \mathbf{n} in transition rates W .

- Contribution of site $i = k-1$ is $\langle [(n_k + 1)n_{k+1} - n_k n_{k+1}] W(k-1 \rightarrow k) \rangle$.
- Contribution of site $i = k$ is $\langle [(n_k - 1)(n_{k+1} + 1) - n_k n_{k+1}] W(k \rightarrow k+1) \rangle + \langle [(n_k - 1)n_{k+1} - n_k n_{k+1}] W(k \rightarrow k-1) \rangle$
- Contribution of site $i = k+1$ is $\langle [n_k(n_{k+1} - 1) - n_k n_{k+1}] W(k+1 \rightarrow k+2) \rangle + \langle [(n_k + 1)(n_{k+1} - 1) - n_k n_{k+1}] W(k+1 \rightarrow k) \rangle$.
- Contribution of site $i = k+2$ is $\langle [n_k(n_{k+1} + 1) - n_k n_{k+1}] W(k+2 \rightarrow k+1) \rangle$

The dummy index game is clear : when writing an expression containing $W(i \rightarrow j)$, the prefactor which contain the j index (the receiving end) will get +1 and the one containing i (the giving end) will get -1. Simplifying and summing the above expression,

$$\begin{aligned} \frac{d \langle n_k n_{k+1} \rangle}{dt} &= \langle n_{k+1} [W(k-1 \rightarrow k) - W(k \rightarrow k-1)] \rangle \\ &+ \langle (n_k - n_{k+1}) [W(k \rightarrow k+1) - W(k+1 \rightarrow k)] \rangle \\ &+ \langle n_k [-W(k+1 \rightarrow k+2) + W(k+2 \rightarrow k+1)] \rangle \\ &- \langle W(k \rightarrow k+1) \rangle - \langle W(k+1 \rightarrow k) \rangle \end{aligned}$$

And replacing with the actual transition rates

$$\frac{1}{\beta} \frac{d \langle n_k n_{k+1} \rangle}{dt} = -4 \langle n_k n_{k+1} \rangle + \langle n_{k-1} n_{k+1} \rangle + \langle n_k n_{k+2} \rangle + \langle n_k^2 \rangle + \langle n_{k+1}^2 \rangle - \langle n_k \rangle - \langle n_{k+1} \rangle \quad (24)$$

It will not consume to much energy to show that this formula generalizes to $\langle n_i n_j \rangle$ for $|i - j| > 1$:

$$\frac{1}{\beta} \frac{d \langle n_i n_j \rangle}{dt} = -4 \langle n_i n_j \rangle + \langle n_{i \pm 1} n_j \rangle + \langle n_i n_{j \pm 1} \rangle \quad (25)$$

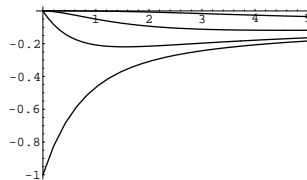


Figure 1 functions $e^{-t}I_n(t)$ for $n = 0, 1, 2, 4$.

Solving the moment equations for homogeneous initial conditions.

To solve eqs(25-23) we need to precise the initial conditions. The condition in which we are interested here is one having translational symmetry. For example, having exactly n_0 particle on each site at $t = 0$. As time flows, particles will jump around and destroy the initial condition ; moreover, each realisation will do that in a different way. What we can expect however is that *average* quantities will keep the symmetry of initial conditions, because the master equation we use is just an ordinary differential equations on *probabilities*. We will thus assume that $\langle n_i \rangle = \langle n_j \rangle = \mu \forall i, j$. Eq. (22) reads :

$$\frac{d\mu}{dt} = -2\mu + \mu + \mu = 0$$

The average number of particles per site remains constant $\mu(t) = n_0$. For the correlation, translation symmetry imposes that $\langle n_i n_j \rangle = \langle n_{i+k} n_{j+k} \rangle \forall i, j, k$: correlations between two sites then can only depend on their distance $|i - j|$. Calling $u_n = \langle n_i n_{i \pm n} \rangle$, eqs (25-23) read (normalizing time by the 2β factor) :

$$\begin{aligned} \frac{du_0}{dt} &= -2u_0 + 2u_1 + 2\mu \\ \frac{du_1}{dt} &= -2u_1 + u_0 + u_2 - \mu \\ \frac{du_n}{dt} &= -2u_n + u_{n-1} + u_{n+1} \end{aligned}$$

It will be wiser to use $v_0 = u_0 - \mu$, $v_n = u_n$ which will give a more symmetric pictures of the above equations :

$$\frac{dv_n}{dt} = -2v_n + v_{n+1} + v_{n-1} \quad (26)$$

with the initial conditions $v_0 = n_0^2 - n_0$ and $v_n = n_0^2$. It would have been even wiser to define $v_0 = u_0 - \mu^2 - \mu$, $v_n = u_n - \mu^2$ which does not change the form of equation (26) but only the initial conditions : $v_0 = -n_0$ and $v_n = 0$. We will come back to why it is necessary to subtract average from the variance in diffusive processes.

Bessel functions I_n have a nice property, linking derivative of the n th order to adjacent ones :

$$\frac{dI_n(t)}{dt} = \frac{1}{2} (I_{n-1}(t) + I_{n+1}(t))$$

and therefore, the natural solution of eqs.(26) is

$$v_n(t) = -n_0 e^{-2t} I_n(2t) \quad (27)$$

. The asymptotic behavior of modified Bessel function is $I_n(t) \approx (1/\sqrt{t}) \exp(t)$, so all correlations will die out eventually. To get to the continuum limit, we use the approximation of eq.(8), keeping in mind that $n_i = \ell c(x)$ (where $x = i/\ell$) and $\langle n_i n_{i+k} \rangle = \ell^2 \langle c(y)c(y+x) \rangle$

$$\begin{aligned} \langle n_i n_{i+k} \rangle &= -n_0 e^{-4\beta t} I_n(4\beta t) + n_0^2 + n_0 \delta_{k,0} \\ \langle c(y)c(y+x) \rangle &= -\frac{c_0}{\sqrt{2\pi(2Dt)}} \exp\left(-\frac{x^2}{2(2Dt)}\right) + c_0^2 + c_0 \delta(x) \end{aligned}$$

m. What is the spatial correlation function ? What do we mean by $f(k) = \langle n_i n_{i+k} \rangle$? This is just a histogram of particles distance : if we find two particles at distance k of each other, they will contribute one unit to $f(k)$. Now, if we have n_i particles in cell i and n_{i+k} particles in cell $i+k$, we will have $n_i \times n_{i+k}$ particles with distance k . Of course, the distance between each particle and itself ; therefore, each particle contributes one unit to $\langle n_i^2 \rangle$, whatever the size of cell array . This is the reason behind terms such as $n_0 \delta_{k,0}$ or $c_0 \delta(x)$.

V. GENERALIZATION OF JUMP PROCESSES.

Lets get back to 1-component systems. Until now, we have only considered jump processes of the type $W(i \rightarrow i \pm 1)$. From a practical point of view, this means the following : considering a small time interval dt , the probability for one event (*i.e.* jump, birth, death,...) to occur in this interval is something like αdt , where α is a characteristic constant of the system under consideration. The probability for *two* events to occur is proportional to dt^2 , so for small dt , we can just neglect these processes. Of course, by *small* dt , we mean time interval small enough to set $\alpha dt \ll 1$.

Now, what if the time resolution of our measurement apparatus is not good enough to be in this approximation ? Suppose for example that we are observing GFP-proteins under fluorescence appearing in average every 10 seconds, but our camera has a low sensitivity and we have to integrate for 30" to see an image. Then, the probability for 3 proteins to appear during this time interval is not anymore negligible.

We can easily generalize our master equation describing "more than one" jumps by taking into account terms like $W(n \rightarrow n \pm k)$; such a term means that, being in the state n , the probability to jump to the state $n+k$ during an interval dt is $W(n \rightarrow n \pm k)dt$. This is a function of two variables n, k . We can measure it experimentally by making a histogram of all jumps observed.

$$\frac{dP(n,t)}{dt} = \sum_{k=-\infty}^{\infty} W(n-k \rightarrow n)P(n-k) - W(n \rightarrow n+k)P(n)$$

Note that this equation does not contain $k = 0$, which is automatically removed in the substraction.

A. The average.

Again, playing the dummy index game, multiplying by n and summing over all n 's, changing the variable of summation from n to $n+k$ when needed, we get

$$\frac{\partial \langle n \rangle}{\partial t} = \sum_n \left(\sum_k kW(n \rightarrow n+k) \right) P(n)$$

The inner sum over k 's is just the "average" jump size when the system is in state n , which we can note \bar{k}_n or $\langle k|n \rangle$. The second summation is the summation of this quantity over n 's, weighted by the probabilities of being in state n . Therefore,

$$\frac{\partial \langle n \rangle}{\partial t} = \langle \langle k|n \rangle_1 \rangle_2 \quad (28)$$

where the inner bracket (subscript 1) is the average over the jumps and the outer one is the average over the states of the system n 's.

n. trivial example. Let us reconsider the simplest birth-date process considered in section II, where $W(n \rightarrow n+1) = \lambda n$, $W(n \rightarrow n+k) = 0$ if $k > 1$, $W(n \rightarrow n-1) = \mu n$ and $W(n \rightarrow n+k) = 0$ if $k < -1$. Then $\bar{k}_n = 1 \times \lambda n + (-1) \times \mu n = (\lambda - \mu)n$ and

$$\frac{\partial \langle n \rangle}{\partial t} = (\lambda - \mu) \langle n \rangle \quad (29)$$

which is what we had. It is of course obvious that from (29) we can recover the simpler case (4).

o. RNA or Protein burst. This is a problem we had fully treated in two component systems, but we can simplify it through generalized jumps. Suppose that during a small time, species A (be it RNA or protein or whatever) arrive packed in bags of random size, so that $W(n \rightarrow n+k) = f(k)$ for $k > 0$ (independent of n). On the other hand, we still continue to consider the degradation process of A 's as a simple poisson one $W(n \rightarrow n-1) = -\mu n$. Then,

$$\begin{aligned} \bar{k}_n &= \sum_{k>0} kf(k) - \mu n \\ &= E - \mu n \end{aligned}$$

where E is the average size of the bags. Then

$$\frac{\partial \langle n \rangle}{\partial t} = E - \mu \langle n \rangle$$

So after a given time ($\gg 1/\mu$), the average reaches the stationnary value

$$\langle n \rangle = E/\mu$$

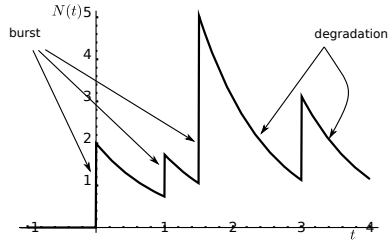


Figure 2 Species A appears in (temporally poissonian) bursts of random size k and degrades with rate μ .

B. The second moment.

We can continue and multiply by n^2 the master equation, and then sum over all n 's. We are now used to this game and we get

$$\begin{aligned} \frac{\partial \langle n^2 \rangle}{\partial t} &= \sum_n \left(\sum_k 2nkW(n \rightarrow n+k) + \sum_k k^2W(n \rightarrow n+k) \right) P(n) \\ &= 2 \langle n \langle k | n \rangle_1 \rangle_2 + \langle \langle k^2 | n \rangle_1 \rangle_2 \end{aligned}$$

p. trivial example : birth-death. We have already computed $\langle k | n \rangle_1 = (\lambda - \mu)n$, so the first term reads

$$2 \langle n \langle k | n \rangle_1 \rangle_2 = 2(\lambda - \mu) \langle n^2 \rangle$$

For the second term, we have $\langle k^2 | n \rangle_1 = (\lambda + \mu)n$; putting all terms together, we have

$$\frac{\partial \langle n^2 \rangle}{\partial t} = 2(\lambda - \mu) \langle n^2 \rangle + (\lambda + \mu) \langle n \rangle$$

Wich we already knew.

q. RNA or Protein burst. Let us call

$$F = \sum_{k>0} k^2 f(k)$$

Computing all the terms accordingly, we have

$$\begin{aligned} \frac{\partial \langle n^2 \rangle}{\partial t} &= 2 \langle n(E - \mu n) \rangle + \langle F + \mu n \rangle \\ &= -2\mu \langle n^2 \rangle + (2E + \mu) \langle n \rangle + F \end{aligned}$$

Again, we can compute the stationary solution $\langle n^2 \rangle$

$$\begin{aligned} \langle n^2 \rangle &= \frac{2E + \mu}{2\mu} \langle n \rangle + \frac{F}{2\mu} \\ &= \frac{E^2}{\mu^2} + \frac{E}{2\mu} + \frac{F}{2\mu} \end{aligned}$$

or the more interesing quantity VMR, the variance $V = \langle n^2 \rangle - \langle n \rangle^2$ over the mean $\langle n \rangle$

$$\frac{V}{\langle n \rangle} = \frac{1 + E}{2} + \frac{F - E^2}{2E}$$

Note that the term term is just the VMR of the bursts. Because we are considering only non-empty bags of species A appearing, $E \geq 1$. So in general, for this process

$$\frac{V}{\langle n \rangle} = 1 + \alpha$$

where the term α has one contribution from the average size of the bags and one contribution from the variance to mean of the size of the bags. For bags of size 1, $VMR = 1$.

Appendix A: Alternative formulations of the brownian motion.

1. Basic facts.

Brownian motion plays a central role in physics and therefore many alternative formulations of it coexist ; their usage depends on the taste of the user and the experimental/theoretical problem at hand. In the above section, we used continuous time and discrete lattice approach. On the other hand, an experimentalist recording a brownian motion will take successive pictures at *fixed* time intervals. At each time step i , he will record the displacement d_i : to him d_i 's are *the* (continuous) random variables and he can construct all the above concept of brownian motion by using a “discrete time, continuous space” formulation. Of course, it would be also possible to develop a full “continuous time, continuous space” theory, but we will focus here only on the former formulation.

Let us suppose a one dimensional world where at each time step i , the particle under our microscope makes a random jump d_i . Space and time being homogeneous, we assume the same distribution for all d_i 's ; in particular, $\langle d_i \rangle = 0$ and $\langle d_i^2 \rangle = \sigma^2$. Moreover we assume that the jumps at step i and step j ($i \neq j$) are independent : $\langle d_i d_j \rangle = \langle d_i \rangle \langle d_j \rangle = 0$.

The distance from origin after N time interval is the random variable

$$X_N = \sum_{i=1}^N d_i$$

Obviously, $\langle X_N \rangle = 0$ and $\langle X_N^2 \rangle = N\sigma^2$. Calling τ the duration of time intervals and t the total time of recording, $\langle X^2(t) \rangle = (\sigma^2/\tau)t = Dt$ where the diffusion coefficient D is defined as it should.

The central limit theorem informs us that if random variables d_i are well behaving (*i.e.* possess at least a second order moment σ^2) then for large N , X_N is a gaussian random variable :

$$P(x) = \frac{1}{\sqrt{2\pi s^2}} \exp\left(-\frac{x^2}{2s^2}\right)$$

where $s = \sqrt{N}\sigma$. Using $N = t/\tau$ and $D = \sigma^2/\tau$, we get

$$P(x,t) = \frac{1}{\sqrt{2\pi Dt}} \exp\left(-\frac{x^2}{2Dt}\right)$$

which is the same as eq.(9).

2. The problem of persistence.

As we mentioned above, for large times, the distribution law is gaussian. But how large is large ? Or how should we chose the time interval between two pictures ? The main assumption behind our computation above was that successive steps are independent. This is obviously inaccurate if the time interval between successive pictures is too short : because of its inertia, a 1 micron latex particle in water will (approximately) conserve its direction and speed during one nano-second ; an amoebae will keep moving in the same direction for some time ; a polymer will not bend over a picometer distance. In all these examples, we see that the i step conserve some information from $i - 1$ step. The simplest way of modeling this process is to assume that

$$d_i = p d_{i-1} + \sqrt{1-p^2} \delta_i \tag{A1}$$

p ($\in [0, 1]$) is the amount of conserved information from the last step, and δ_i is the random modification brought to the step $i - 1$. Again, we assume translation invariance in time and space : no moment should depend on the index i : $\langle \delta_i^k \rangle = \langle \delta_j^k \rangle$ and $\langle d_i^k \rangle = \langle d_j^k \rangle$. Moreover, we suppose $\langle \delta_i \rangle = 0$ and $\langle \delta_i^2 \rangle = \sigma^2$. We then have

$$\begin{aligned} \langle d_i^2 \rangle &= p^2 \langle d_{i-1}^2 \rangle + (1-p^2) \langle \delta_i^2 \rangle + 2p\sqrt{1-p^2} \langle d_i \delta_i \rangle \\ &= p^2 \langle d_{i-1}^2 \rangle + (1-p^2) \sigma^2 \end{aligned}$$

because d_i and δ_i are independent. Therefore, $\langle d_i^2 \rangle = \sigma^2$. Now, what about correlation ? Multiplying eq.(A1) by d_{i-1} and taking the average,

$$\langle d_i d_{i-1} \rangle = p \sigma^2$$

More generally,

$$\langle d_i d_{i-k} \rangle = p^k \sigma^2$$

So correlations decrease exponentially and become negligible after $n_c = -1/\log p$ steps, or equivalently, after $t_c = -\tau/\log p$. Sampling the particle at $\gg t_c$, the movement appears truly brownian, while sampling at $\ll t_c$, the movement appears ballistic.

The quantity measured by the experimentalist is the distance from origin after N steps $X_N = \sum_1^N d_i$. Its fluctuations are given by

$$\begin{aligned} \langle X_N^2 \rangle &= \sum_{i=1}^N \sum_{j=1}^N \langle d_i d_j \rangle \\ &= N \langle d_i^2 \rangle + 2 \sum_{k=1}^{N-1} (N-k) \langle d_i d_{i+k} \rangle \\ &= \left(N + 2 \sum_{k=1}^{N-1} (N-k) p^k \right) \sigma^2 \end{aligned}$$

the second term is just a geometric series, so we get

$$\langle X_N^2 \rangle = \frac{1+p}{1-p} \sigma^2 N - 2p \frac{1-p^N}{(1-p)^2} \sigma^2$$

Or, in time unit,

$$\langle X^2(t) \rangle = Dt - A(1 - e^{-t/t_c})$$

where

$$D = \frac{1+p}{1-p} \frac{\sigma^2}{\tau}$$

is the diffusion coefficient and

$$t_c = -\tau/\log p$$

is the persistence time. Note that the amplitude $A = -2p\sigma^2/(1-p)^2$ can be expressed as a function of D , t_c and τ and is not an independent parameter. Experimentally, it is easy to estimate D and A from a linear regression of the data (if long times and diffusive regime prevail) and estimate p through $A/D = 2\tau p/(1-p^2)$ which then give access to t_c . It is also possible to use a non linear curve fitting to estimate A, D and t_c and check then their compatibility.

r. History : Part of this lecture was given in 2006 at Ecole Doctorale de Physique, Grenoble University, as an introduction to fluctuations in biological systems.