The origin and consequences of noise in biochemical systems

Surprising things happen when we take the discreteness of molecule number seriously, abandoning the notion that chemical concentrations may be treated as continuous variables. Here we will show how ideas of discreteness force us into dealing with issues of noise and randomness. We will arrive at a probabilistic description of reaction kinetics which, in the limit of large numbers, will reproduce the familiar reaction-rate description. We will show how probabilistic systems may be treated numerically using Monte Carlo simulations, and use such simulations to investigate how frequently large random fluctuations can change the state of a bistable system. This will help us understand the stability and spontaneous switching rates of actual biological switches such as those which underlie cell fate determination during development, or long-term potentiation in neurons.

1. Introduction

Consider a simple chemical system in which a molecule X is created at some constant rate k and destroyed in a first-order reaction with rate γ. If the total number n of molecules is large, as is the case in standard chemical systems (Avogadro's number is about 10^{24}), we can ignore the fact that n is an integer, but treat it instead as a continuous variable (Fig. 1a, dashed line), writing

$$\frac{dn}{dt} = k - \gamma n \equiv f_n - g_n . \qquad (1)$$

However, this is certainly not applicable in living cells, where there are typically thousands of molecules of a given protein, hundreds of free ribosomes and RNA polymerases, tens of mRNA molecules of each kind, and one or two copies of most genes. Moving, then, to a discrete description, we might guess that the actual time evolution of n would be somewhat coarse, but still completely predictable. For example, we could imagine a situation in which creation events occurred at time intervals $\Delta t = \Delta n/(k\text{-}\gamma n)$, with Δn = 1 (Fig. 1a, solid line). However, this is physically impossible: it would require the system to somehow

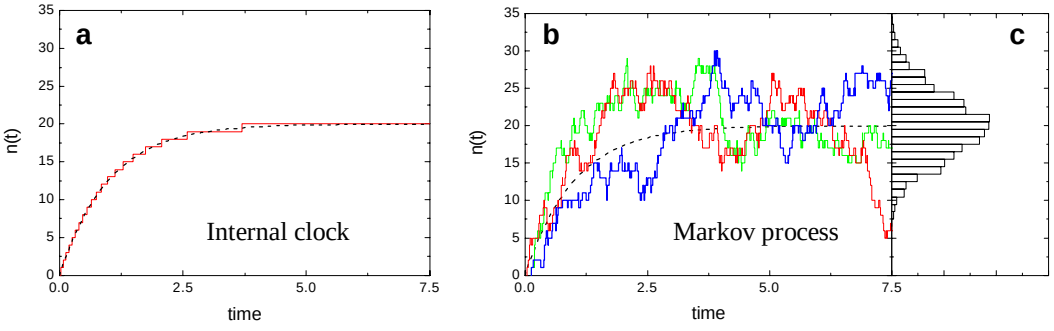

Figure 1: The implications of discreteness

keep track of the time that elapsed between event occurrences, as a clock does between successive ticks. But there is no internal clock in our simple system – there are only molecules which collide into one another. The system has no memory of the past, so its response can only depend on present conditions. (This is the defining property of a Markov process.) What therefore happens is that creation and destruction reactions occur with some *probability* per unit time, proportional to the reaction rates. This means that each time the reaction is run with fixed initial conditions, it will proceed somewhat differently:

the system will be stochastic (Fig. 1b). If we ran several experiments of this kind, recording the number of molecules present after some fixed time had elapsed, we would find a distribution of possible values (Fig. 1c).

2. Probabilistic formulation of reaction kinetics: the Master Equation

If a reaction occurs at some rate r, then in a large time interval T it will occur, on average, rT times. If this interval is divided into N smaller sub-intervals, the chance that the reaction occurred in any one of those sub-intervals is rT/N. Writing $dt = T/N$, this shows that the probability of reaction with rate r occurring in a small time interval dt is just rdt. (To be careful, we must eliminate the possiblity that more than one reaction occurred in this interval; however, for small enough dt, that outcome is negligible.)

We now consider an ensemble of identical systems, each having the same initial conditions, and define $p_n(t)$ as the number of these systems which have precisely n molecules at time t. This number can increase if a molecule of X is created in some system having n-1 molecules, or if a molecule of X is destroyed in some system having n+1 molecules; it can decrease if a molecule of X is created or destroyed in some system having n molecules. These four possibilities are shown in Fig. 2; for clarity, consider just one of these for the moment. Suppose there are p_{n-1} systems having n-1 molecules at some time t. In a small time interval dt, the probability that there will be a molecule created in any one of these systems is $f_{n-1}\,dt$. Therefore, the total number of systems in which a molecule is created will be given by $p_{n-1}\,f_{n-1}\,dt$. Each of these systems will then enter the pool of systems which have n molecules, adding to the number p_n that were there to begin with. Thus, $p_n(t+dt) = p_n(t) + p_{n-1}\,f_{n-1}dt$, or $dp_n/dt = f_{n-1}\,p_{n-1}$. If we now include all four fluxes, we obtain the Master Equation:

$$\frac{dp_n}{dt} = -(f_n + g_n)p_n + f_{n-1}p_{n-1} + g_{n+1}p_{n+1}. \qquad (2)$$

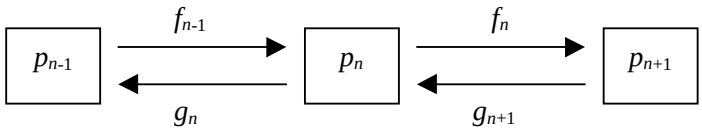

Figure 2: Derivation of the Master Equation

Note that this is actually an infinite set of equations, one for each n. The Master Equation is linear in the quantities p_n, so it remains unchanged when divide the number of systems in a given state n by the fixed total number of systems. In that case, $\Sigma_n\, p_n = 1$, and we can think of $p_n(t)$ as the probability for any given system to be in state n. To connect with experiments: the ensemble of systems could be a population of cells, and p_n would represent the fraction of cells having n copies of some protein.

3. Emergence of the deterministic law

It is possible to obtain all the moments of the probability distribution $p_n(t)$ without explicitly solving the Master Equation. For example, the mean number of molecules, calculated by averaging over all the systems at a given time, is:

$$\langle n \rangle = \sum_n n p_n \,. \tag{3}$$

Summing Eq. 2 over n, and using f_n and g_n from Eq. 1, we obtain

$$\begin{aligned}
\frac{d}{dt}\langle n \rangle &= \frac{d}{dt}\sum n p_n = -k\sum n p_n + k\sum n p_{n-1} - \gamma\sum n^2 p_n + \gamma\sum n(n+1)p_{n+1} \\
&= -k\sum n p_n + k\sum (n-1)p_{n-1} + k\sum p_{n-1} - \gamma\sum n^2 p_n + \gamma\sum (n+1)^2 p_{n+1} - \gamma\sum (n+1)p_{n+1} \\
&= k\sum p_{n-1} - \gamma\sum (n+1)p_{n+1} = k - \gamma\langle n \rangle \,,
\end{aligned} \tag{4}$$

where we have used the fact that $\Sigma\, h(n) = \Sigma\, h(n\pm 1)$ if the sum is carried over all n. In short,

$$\frac{d}{dt}\langle n \rangle = k - \gamma\langle n \rangle \,. \tag{5}$$

That is, the mean molecule number still obeys the deterministic equation. (This result is true whenever the rates f_n and g_n are linear functions of n.)

4. Steady state: the Poisson distribution

Assume now that we are in steady state, so $dp_n/dt = 0$. Then,

$$0 = -(k + \gamma n)p_n + k p_{n-1} + \gamma(n+1)p_{n+1} \,, \tag{6}$$

or, setting $\bar{n} = k/\gamma$,

$$(n+1)p_{n+1} - \bar{n}p_n = n p_n - \bar{n} p_{n-1} \,. \tag{7}$$

Since this is true for all n, both sides must be equal to a constant; and because p_n must be normalizable, it can be shown that the constant is simply zero. Therefore,

$$p_n = \frac{\bar{n}}{n} p_{n-1} = \frac{\bar{n}^2}{n(n-1)} p_{n-2} = \ldots = \frac{\bar{n}^n}{n!} p_0 \qquad \Rightarrow \qquad \Sigma\, p_n = p_0 \Sigma \frac{\bar{n}^n}{n!} = p_0 e^{\bar{n}} \,.$$

Setting $\Sigma p_n = 1$ gives $p_0 = e^{-\bar{n}}$. The final steady state result is known as the Poisson distribution:

$$p_n = \frac{\bar{n}^n}{n!} e^{-\bar{n}} \qquad\qquad \bar{n} = k/\gamma \tag{8}$$

5. The limit of large numbers

The mean and variance of the Poisson distribution are given by:

$$\langle n \rangle = \langle \delta n^2 \rangle = \bar{n} \,. \tag{9}$$

The relative standard deviation is therefore

$$\frac{\sqrt{\langle \delta n^2 \rangle}}{\langle n \rangle} = \frac{1}{\sqrt{\langle n \rangle}} \,. \tag{10}$$

This gives us a precise notion of what it means to have a 'large number' of molecules in our system: we can expect deviations from deterministic behavior of the order of the inverse square root of the number of molecules involved. Therefore, an ensemble of systems with an average number of 20 molecules will show a spread of 22% about this value (Fig. 3a), while one with 500 molecules will show a spread of just 4% (Fig. 3b).

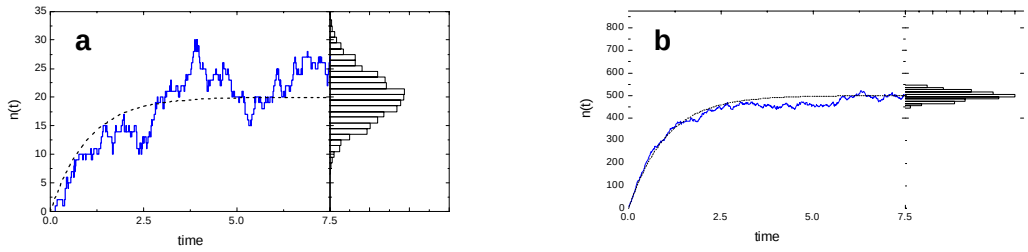

Figure 3: The large number limit

6. The Fokker-Planck equation

For intermediate molecule numbers, when the difference between n and n+1 may be neglected but fluctuations must still be taken into account, there is a useful approximation to the Master Equation which provides some physical insight. We replace n by a continuous variable, and use the notation $h(n)$ in place of the h_n used previously. Any function of n can be Taylor expanded:

$$h(n+\Delta n) = h(n) + \frac{\partial h}{\partial n}\Delta n + \frac{1}{2}\frac{\partial^2 h}{\partial n^2}\Delta n^2 + \vartheta(\Delta n^3) \,. \tag{11}$$

Thus,

$$f(n-1)p(n-1) = f(n)p(n) - \frac{\partial}{\partial n} f(n)p(n) + \frac{1}{2}\frac{\partial^2}{\partial n^2} f(n)p(n) + \ldots$$

$$g(n+1)p(n+1) = g(n)p(n) + \frac{\partial}{\partial n} g(n)p(n) + \frac{1}{2}\frac{\partial^2}{\partial n^2} g(n)p(n) + \ldots \quad . \tag{12}$$

Substituting this into Eq. 2, we obtain the Fokker Planck equation:

$$\frac{\partial p(n,t)}{\partial t} = -\frac{\partial}{\partial n}\left[(f-g)p - \frac{1}{2}\frac{\partial}{\partial n}(f+g)p\right] = -\frac{\partial J}{\partial n}, \tag{13}$$

where J represents a probability flux. This can be thought of as a diffusion equation: every particle represents a system in our ensemble; the particle position is analogous to the number of molecules in the system; and J is the flux of particles across any boundary. In steady state, J must be a constant; however, the flux at $n = 0$ must be zero (no system can pass to having negative particle number), so the flux must be zero everywhere. This gives us the following equation:

$$(f-g)p = \frac{1}{2}\frac{\partial}{\partial n}(f+g)p. \tag{14}$$

Setting $q = (f + g)p$,

$$\frac{(f-g)}{(f+g)}q = \frac{1}{2}\frac{\partial q}{\partial n} \quad \Rightarrow \quad \frac{1}{q}\frac{\partial q}{\partial n} = \frac{\partial \ln(q)}{\partial n} = 2\frac{(f-g)}{(f+g)} \quad \Rightarrow \quad q = Ae^{\int 2\frac{(f-g)}{(f+g)}dn}$$

Therefore,

$$p(n) = \frac{A}{(f+g)}e^{-\varphi(n)}, \qquad \text{with} \qquad \varphi(n) = -\int 2\frac{(f-g)}{(f+g)}dn'. \tag{15}$$

That is, the reaction is analogous to a thermodynamic system in some potential φ.

7. Steady state of a bistable system

Consider the autocatalytic genetic system introduced in problem set 1:

$$\frac{dx}{dt} = \underbrace{\frac{v_0 + v_1 K_1 K_2 x^2}{1 + K_1 K_2 x^2}}_{f(x)} - \underbrace{\gamma \cdot x}_{g(x)} \tag{16}$$

In the problem set, x represented a concentration, and the decay term arose due to dilution. We use a different interpretation here, assuming that the cell volume is fixed, and that the decay is due to some active process. We can then change parameter units so that x represents molecule number. If $\gamma = 1$, $K_1K_2 = 10^{-4}$, $v_0 = 12.5$ and $v_1 = 200$, the system is bistable. This results in a double well potential, with the two stable states separated by some energy barrier. For this case, $(f - g)$, the potential $\varphi(x)$, and the steady state distribution $p(x)$ are shown in Fig. 4.

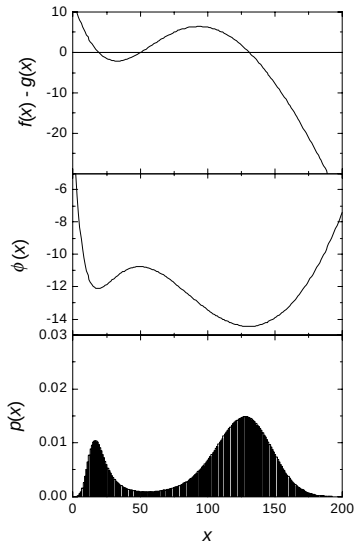

Figure 4:
A stochastic bistable system

8. Waiting times between reaction events

Suppose a chemical reaction occurs with rate r. What is the time interval between successive occurrences of the reaction? The probability that the reaction occurs in some time interval dt is rdt; the probability that it does not occur is therefore $1 - rdt$. The probability that it occurs only after some time τ can be calculated as follows:

$$P(\tau) \equiv \wp(\text{next occurence is in the interval } \tau \text{ to } \tau+d\tau) = \wp(\text{does not occur for } t < \tau)\ \wp(\text{occurs in } \tau \text{ to } \tau+d\tau) \quad (17)$$

But

$$\wp(\text{does not occur for } t < \tau) = \wp(\text{does not occur for } t < \tau-d\tau)\ \wp(\text{does not occur in } \tau-d\tau \text{ to } \tau) \quad (18)$$

Setting $Q(\tau) = \wp$ (does not occur for $t < \tau$), this implies $\ln(Q(\tau))-\ln(Q(\tau-d\tau)) = \ln(1-rd\tau) \approx -rd\tau$. Therefore,

$$\frac{d\ln(Q)}{d\tau} = -rd\tau \quad \Rightarrow \quad Q(\tau) = e^{-r\tau}, \quad (19)$$

where we have used $Q(0) = 1$. Inserting this in Eq. 17, we get

$$P(\tau) = e^{-r\tau} rd\tau . \quad (20)$$

The waiting times between successive reactions are therefore exponentially distributed, with mean value $\langle\tau\rangle = 1/r$, and variance $\langle\delta\tau^2\rangle = 1/r$.

9. Stochastic simulation of chemical reactions

If u is a random number drawn from a uniform distribution between zero and one, then the following function of u is distributed precisely as τ is in Eq. 20:

$$\theta = \frac{1}{r}\ln\left(\frac{1}{u}\right) \tag{21}$$

This gives us a very simple prescription for numerically simulating the behavior of a stochastic system. Start with some initial condition for each molecule type. If there are m possible types of reactions (m = 2 for Eq. 2, as there are only creation and destruction events) occuring with rates r_i (i = 1, ..., m), then we can generate m random variables θ_i which are the putative waiting times to the next occurrence of each reaction type. The smallest of these gives the time interval after which the first reaction which will actually occur. At this stage, simply update the variables (for Eq. 2, we would increment n if a creation event occurred, or decrement it if a destruction event occurred), recalculate the rates, and repeat the generation of putative times. Continue this until some convenient time limit is reached. This is how the timecourses in Fig. 1b were generated. To obtain Fig. 1c, simply repeat this procedure several times (2,500 times for Fig. 1b), recording the final state of the system, then calculate the histogram of final states. This procedure is a slight simplification of the Gillespie algorithm.

10. Spontaneous switching rates in a bistable system

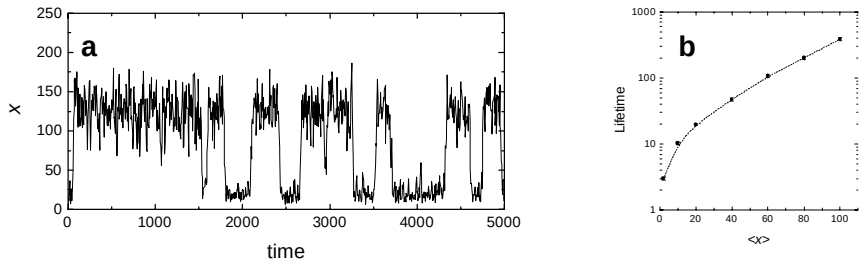

Figure 5: Stochastic transitions and escape times

We can now simulate the time evolution of the system introduced in section 7. We find that an individual cell transitions stochastically between the available steady states (Fig. 5a), while the cell population has a histogram of x values as shown in Fig. 4. An important quantity to investigate is the average escape time from a given state, known as the lifetime of that state. Figure 5b shows the lifetime of the induced state as a function of the average number of X molecules in that state. Once again, we can see the crossover to deterministic behavior: as the number of molecules becomes large, the escape time diverges, and the stable states become truly stable. Note that we are measuring time in units of the degradation time of X, which is typically of the order of a cell lifetime. If we use Eq. 16 as a model of lysis/lysogeny network of phage-λ, then the induced state corresponds to lysogeny, and escape corresponds to spontaneous lysis. Lysogens have been measured to undergo spontaneous lysis at a rate of about 10^{-8} per cell per generation: the switch is extremely stable. Our simple network would require about 1000 molecules to achieve comparable stability, which is consistent with measured repressor concentrations. It is possible, however, to increase the stability of a switch by other means, such as by increasing the cooperativity, or through the use of stabilizing feedback loops.