Stochastic processes

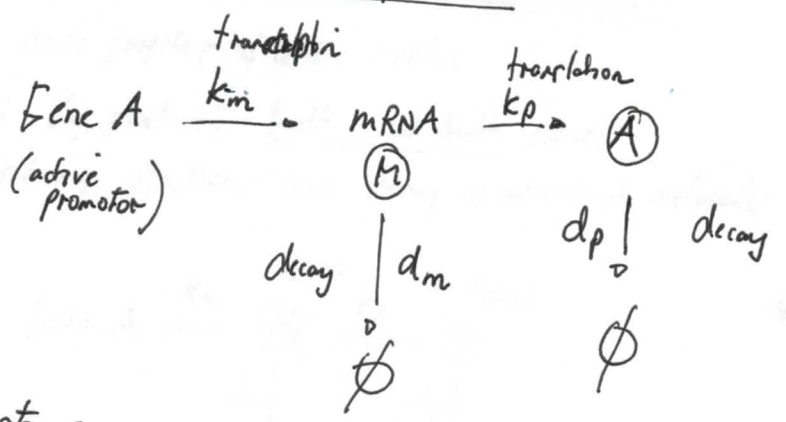
traditional approach: deterministic system described by ODE
rate equations

$$\frac{d\vec{x}}{dt} = f(\vec{x}, p)$$

only approximation:

- large molecule number, well stirred compartments
  ⇒ bulk reactions

Simple model of gene expression



Concentrations:

$$\frac{d[M]}{dt} = k_m - d_m M$$

$$\frac{dA}{dt} = k_p M - d_p A$$

Linear model; linear rates

Steady state

M_s, A_s $\quad \frac{dM}{dt} = 0 = \frac{dA}{dt}$

$$k_m - d_m M_s = 0$$

$$M_s = \frac{k_m}{d_m}$$

$$A_s = \frac{k_p}{d_p} M_s = \frac{k_p}{d_p}\frac{k_m}{d_m}$$

$$\boxed{A_s = \frac{k_m}{d_p} \cdot b}$$

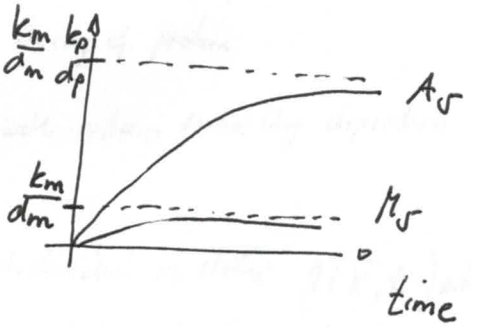
$$b = \frac{k_p}{d_m}$$ burst size

average number of proteins translated per mRNA

Simple solution:

$$M(t) = \frac{k_m}{d_m}\left(1 - e^{-d_m t}\right)$$

$$A(t) = \frac{k_m}{d_m}\frac{k_p}{d_p}\left(1 - e^{-d_p t}\right)$$



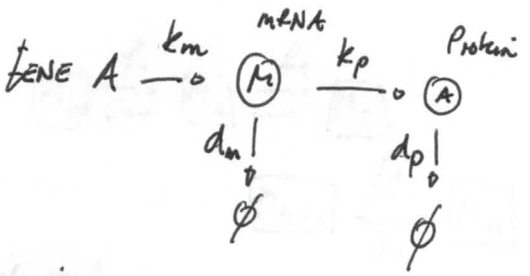
but: cells are intrinsically noisy bioreactors

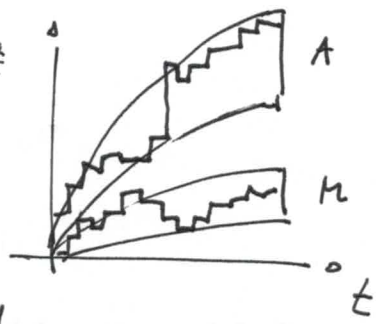
- low copy number (especially mRNA) & many proteins can result in large fluctuations in
  - molecule numbers
  - reaction rates

$\Rightarrow$ description by ODE not appropriate

Modelling gene expression as stochastic process

- emergent noise properties of genetic system
- description by stochastic birth-and death process
  considering synthesis and decay of individual molecules





- system specified at any time by total number of mRNA and protein molecules

$$\vec{x}(t) = (M, A)(t) \qquad M, A \in \mathbb{Z}_0^+$$

- all reactions are assumed to be Poisson; i.e the probability of a reaction with rate k happening in a time dt is

$$k \cdot dt$$

- time evolution is given by the transitions between different states

$$\vec{x} \dashrightarrow \vec{x}'$$

transitions:

$(M, A) \dashrightarrow (M+1, A)$ transcription of mRNA
$(M, A) \dashrightarrow (M-1, A)$ decay of mRNA
$(M, A) \dashrightarrow (M, A+1)$ translation of protein
$(M, A) \dashrightarrow (M, A-1)$ decay of protein

- transitions are stochastic; i.e occur with certain probability depending on the current state of the system
- instead of single trajectory: probability distribution of states $p(\vec{x}, t)$ at time t
- the time evolution of the probability distribution is described by the

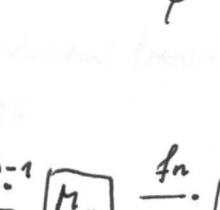
master equation

$$\frac{d\vec{p}(\vec{x}, t)}{dt} = \sum_{\vec{x}'} \left[W(\vec{x}' \dashrightarrow \vec{x})\, p(\vec{x}', t) - W(\vec{x} \dashrightarrow \vec{x}')\, p(\vec{x}, t) \right]$$

$W(x' \dashrightarrow x)$ describes probability for transition
$\sum_{\vec{x}'}$: considers all possible states

steady state distribution of mRNA

$$\frac{dM}{dt} = k_m - d_m M$$



differential equations

for all different states: M_n: probability of being in state with n mRNA

$$\frac{dM_n}{dt} = f_{n-1} M_{n-1} + g_{n+1} M_{n+1} - (f_n + g_n) M_n$$

in equilibrium

$$M_n f_n = g_{n+1} M_{n+1} \qquad f_n = k_m$$

$$g_n = d_m n \qquad g_{n+1} = d_m (n+1)$$

$$M_n k_m = d_m (n+1) M_{n+1}$$

$$M_{n+1} = \frac{k_m}{d_m (n+1)} M_n$$

$n=0$: $M_1 = \frac{k_m}{d_m} M_0$

$n=1$: $M_2 = \frac{k_m}{d_m \cdot 2} M_1 = \left(\frac{k_m}{d_m}\right)^2 \frac{1}{2} M_0$

$n=2$: $M_3 = \left(\frac{k_m}{d_m}\right)^3 \frac{1}{2 \cdot 3} M_0$

$\vdots$

$$M_n = \left(\frac{k_m}{d_m}\right)^n \frac{1}{n!} M_0$$

with $\lambda = \frac{k_m}{d_m}$

probabilities sum to 1

$$\sum_{n=0}^{\infty} M_n = M_0 \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} = M_0 e^{\lambda} = 1$$

$$\Rightarrow M_0 = e^{-\lambda}$$

steady state values of all state probabilities

$$p(n) = \frac{\lambda^n}{n!} e^{-\lambda}$$ poisson distribution for mRNA in steady state

$$\langle M \rangle = \lambda = \frac{k_m}{d_m}$$ identical to deterministic ode

$$\delta M^2 = \lambda = \frac{k_m}{d_m}$$

Coefficient of variation $\frac{\delta M^2}{\langle M \rangle} = 1$

Solving the master equation (Gillespie)

- by solving the master equation the time evolution of the probability distribution can be solved
- very difficult: analytically, but also numerically
- strategy: simulate trajectories of individual transitions which are consistent with the ME

Important to know all rates for all the processes: For example

$(n, A) \xrightarrow{w_1 = k_m} \circ \quad (n+1, A)$
$(n, A) \xrightarrow{w_2 = d_m n} \circ \quad (n-1, A)$
$(n, A) \xrightarrow{w_3 = k_f n} \circ \quad (n, A+1)$
$(n, A) \xrightarrow{w_4 = d_f A} \circ \quad (n, A-1)$

Gillespie algorithm

1. System is in $\vec{x}$ at time t: (n, A) (t)
2. estimate probability ~~for~~ w_i for all feasible transitions $\vec{x} \dashrightarrow \vec{x}'$
3. estimate the time Δt until which the transition happens

$$\Delta t = \frac{1}{w_0} \log(\varepsilon_1) \quad \text{with } w_0 = \sum w_i$$
$$\varepsilon_1 \in [0, 1] \text{ random number}$$

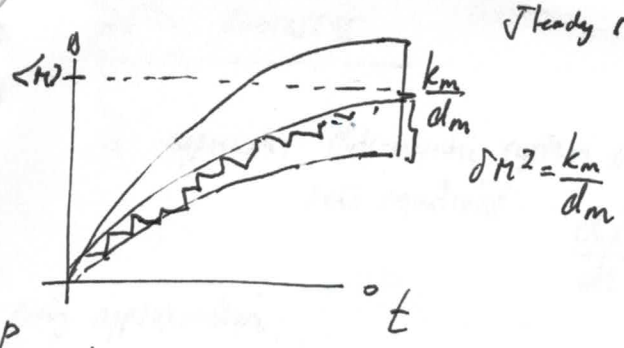
4. estimate which transition w_α happens
   probability of individual transitions is proportional to w_i

$$\sum_{i=1}^{\alpha-1} w_i \leq \varepsilon_2 w_0 \leq \sum_{i=1}^{\alpha} w_i \qquad \varepsilon_2 \in [0, 1]$$

5. update state $\vec{x} \xrightarrow{w_\alpha} \vec{x}'$ and time $t \to t + \Delta t$

Example

steady state



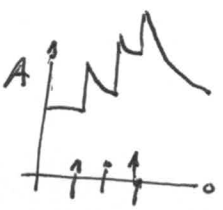
for the protein amount in steady state

$$\langle A \rangle = \frac{k_m}{d_m} \cdot \frac{k_p}{d_p} = \frac{k_m}{d_p} \cdot b \qquad b = \frac{k_p}{d_m}$$ protein per mRNA burst

$$\frac{\delta A^2}{\langle A \rangle} = \left(\frac{b}{1 + \frac{d_p}{d_m}} \right) + 1 \cong b + 1$$

for $d_m \gg d_p$ degradation of mRNA much faster than protein

- between synthesis and degradation of mRNA it is transcribed by ribosomes releasing a burst of proteins
- noise width of protein distribution is determined primarily via average burst size b



- <u>intrinsic noise</u> is <u>controlled</u> on <u>translational level</u> the larger the burst, the larger the noise
- <u>mean</u> an <u>variance</u> on protein level can be independently controlled

b: lac I = 40
lac A = 5

time dependency can be calculated:

$$\langle A \rangle (t) = \frac{k_m}{d_p} \cdot b \left(1 - e^{-d_p t}\right)$$

$$\frac{\delta A^2}{\langle A \rangle}(t) = \frac{\left(1 - e^{-2 d_p t}\right) \cdot b}{\left(1 - e^{-d_p t}\right)} b + 1$$

noise out of equilibrium is stronger than in equilibrium

variance δA^2 relaxes to steady state at a rate $2 d_p t$ twice as fast as mean