

Cellular Systems Biology
or
Biological Network Analysis

Joel S. Bader
Department of Biomedical Engineering
Johns Hopkins University
(c) 2012

December 4, 2012

Preface

Cells are systems. Standard engineering and mathematics texts should provide an excellent introduction to understanding how cells behave, mapping inputs to outputs. Unfortunately, cells are not linear, time-independent systems. Saturation and cooperative response break linearity. Cellular states change with time. Cells are not even deterministic, violating the assumptions of non-linear systems analysis.

This book provides a self-contained introduction to cells as non-linear, time-dependent, stochastic, spatial systems. Each major section is motivated by a canonical biological pathway or phenomenon that requires the introduction of new concepts. All the required mathematical techniques are developed from the motivating examples.

The book is designed as a text for advanced undergraduate or graduate students. Prerequisites are univariate calculus, linear algebra, basic molecular biology, and rudimentary facility with a programming language for computational experiments. Linear systems and Laplace transforms are helpful, but are also reviewed in the initial chapters. Each chapter is designed to be covered in an hour lecture, and problems are provided in an Appendix.

This book is developed from course notes for “Systems Bioengineering III: Genes to Cells,” taught by me since 2007 as a required course for our B.S. in Biomedical Engineering.

Joel S. Bader, Baltimore, MD

Contents

Preface	ii
I Cells as Linear Systems	1
1 Cellular Signal Transduction	2
2 Linear Systems Analysis	3
3 The Laplace Transform and Complex Variables	4
4 Signal Transduction Cascades and MAPK Signaling	5
5 Generating Functions for Pharmacokinetics and Pharmacodynamics	6
6 Positive Feedback and Caffeine Response	7
II Cells as Non-linear Systems	8
7 Saturation and Cooperative Response	9
8 Joint Models of Transcription and Translation	10
9 Positive and Negative Auto-Regulation	11
10 Combinatorial Regulation	12
11 Non-Linear Cascades and Logic Gates	13

III Cells as Stochastic Systems	14
12 Delta-Notch Signaling	15
13 Stochastic Dynamics	16
14 Noise in Gene and Protein Expression	17
15 Stochastic Simulations and the Gillespie Algorithm	18
16 Stability Analysis	19
IV Cells as Spatial Systems	20
17 Morphogen Gradient Patterning	21
18 Diffusion	22
19 Solving the Diffusion Equation	23
20 Patterning and Noise	24
V Cellular Networks	25
21 Diffusion on a Network	26
22 Network Topology, Motifs, and Clustering	27
23 The Giant Component	28
24 Network Partitioning and Spectral Clustering	29
25 Metabolic Networks and Flux Balance Analysis	30
A Problems	31

Part I

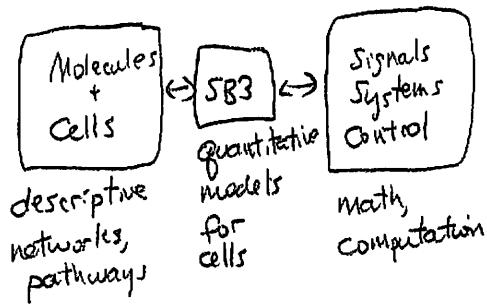
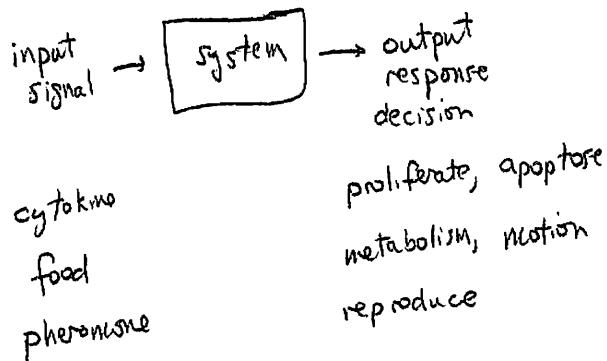
Cells as Linear Systems

Chapter 1

Cellular Signal Transduction

Linear Models for Cell Signaling

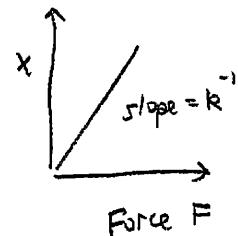
①



Review: Linear Time-Independent Systems

Static: Spring

$$\begin{array}{c} \text{Force } F \downarrow \\ \text{Spring } \xrightarrow{\text{---}} \\ \uparrow x \downarrow \end{array} \quad \begin{array}{l} F = kx \\ \text{or} \\ x = \frac{1}{k} F \end{array} \quad \begin{array}{l} \text{response fn} \rightarrow \text{linear} \\ \uparrow \quad \uparrow \\ \text{response} \quad \text{signal, input} \end{array}$$



Time-dependent

$$F(t) \uparrow \quad x(t) \quad \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array} \quad t \rightarrow$$

How does $x(t)$ depend on $F(t)$?

Simplest possibility: linear response

$$x(t) = \int_0^t H(t-t') F(t') dt'$$

↑
response fn
transfer fn

$$\begin{aligned} \text{Laplace Transform} &= LT = \mathcal{L} \\ \Rightarrow \tilde{x}(s) &= \tilde{H}(s) \tilde{F}(s) \end{aligned}$$

time-independent: $H(t, t')$ depends on $t-t'$

causal: $H(t-t') = 0$ for $t-t' < 0$

Does this model work for cell signals?

Can we calculate $\tilde{H}(s), H(t)$ from cellular/molecular properties?

Can we design a circuit with a desired $H(t)$?

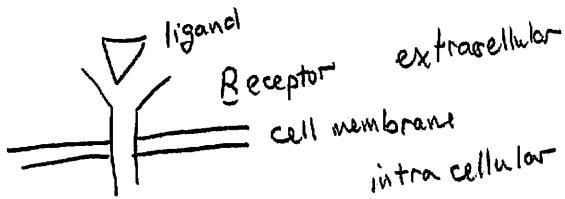
$t-t'$ = time lag

$t > t'$ → past forces

$t < t'$ → future forces

MAPK signaling as a model mitogen-activated protein kinase

(2)

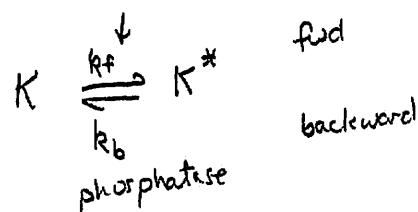


Kinase inactive → kinase-(P) active phosphorylated

more steps in cascade → TF to nucleus

R^* active receptor

steps in cascade → TF to nucleus



$$K^* + K = K_{\text{tot}}$$

weak activation: $K \gg K^*$

$$K = K_{\text{tot}} - K^* \approx K_{\text{tot}}$$

$$\frac{dK^*}{dt} = k_f R^*(t) K - \frac{k_b P}{\alpha} K^*(t)$$

$$\frac{dK^*}{dt} \approx \frac{k_f K_{\text{tot}} R^*(t)}{\beta(t)} - \alpha K^*(t)$$

$$\beta(t)$$

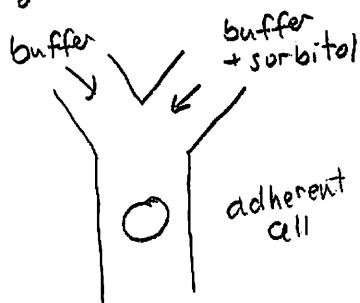
our input signal ↴ more on this next lecture.

$$\frac{dK^*}{dt} = \beta(t) - \alpha K^*(t) \stackrel{LT}{\Rightarrow} s \tilde{K}^*(s) - K(0) = \tilde{\beta}(s) - \alpha \tilde{K}^*(s)$$

$$\tilde{K}^*(s) = \frac{1}{s+\alpha} K(0) + \frac{1}{s+\alpha} \tilde{\beta}(s)$$

$$K^*(t) = e^{-\alpha t} K(0) + \int_0^t e^{-\alpha(t-t')} \tilde{\beta}(t') dt'$$

Probing this circuit: microfluidics



control sorbitol concentration

at cell by adjusting pressure

switch up to 2 Hz



time

What is the response of the MAPK circuit that

responds to high osmolarity?

Does it resemble $K^*(t) = \int dt' H(t-t') \beta(t')$ for a reasonable $H(t)$?

Chapter 2

Linear Systems Analysis

Lecture 2 The Laplace Transform

From last time:

$$\frac{d}{dt} x(t) = \beta(t) - \alpha x(t) \quad \rightarrow \quad s \tilde{x}(s) - x(t=0) = \tilde{\beta}(s) - \alpha \tilde{x}(s)$$

↓

$$x(t) = e^{-\alpha t} x(0) + \int_0^t e^{-\alpha(t-t')} \beta(t') dt' \quad \leftarrow \text{a miracle happens}$$

$$\tilde{x}(s) = \frac{1}{s+\alpha} [x(0) + \tilde{\beta}(s)]$$

You teach yourself BME, we train your minds to think.

Think of $x(t) = \text{vector } \begin{pmatrix} x(t_1) \\ x(t_2) \\ \vdots \end{pmatrix} \quad \beta(t) = \text{vector } \begin{pmatrix} \beta(t_1) \\ \beta(t_2) \\ \vdots \end{pmatrix}$

Linear algebra:

$$A \tilde{x} = \tilde{\beta} - \alpha \tilde{x}$$

Easy if $A \tilde{x} = \lambda \tilde{x}$ eigenvalue, eigenvector

so write $\tilde{x} = \sum_i c_i \phi_i$ where $A \phi_i = \lambda_i \phi_i$

and $\phi_j^+ \phi_i^- = \delta_{ij}$ (orthonormal eigenvectors)

$$\tilde{\beta} = \sum_i \beta_i \phi_i$$

$$A \sum_i c_i \phi_i = \sum_i \beta_i \phi_i - \alpha \sum_i c_i \phi_i$$

$$\sum_i c_i \lambda_i \phi_i = \sum_i \beta_i \phi_i - \alpha \sum_i c_i \phi_i$$

↓ left multiply by ϕ_j^+

$$c_j \lambda_j = \beta_j - \alpha c_j$$

$$c_j = \frac{1}{\lambda_j + \alpha} \beta_j$$

(3)

What is β_j ?

$$\text{If } \beta = \sum_i c_i \phi_i$$

$$\text{then } \beta_j = \phi_j^+ \cdot \beta$$

$$x(t) = \sum_i \frac{1}{\lambda_i + \alpha} \beta_i \phi_i(t)$$

Natural basis set: eigenvectors of $\frac{d}{dt}$

$$\frac{d}{dt} \phi = \lambda \phi$$

$$\frac{d\phi}{\phi} = \lambda dt$$

$$\ln \phi = \lambda t$$

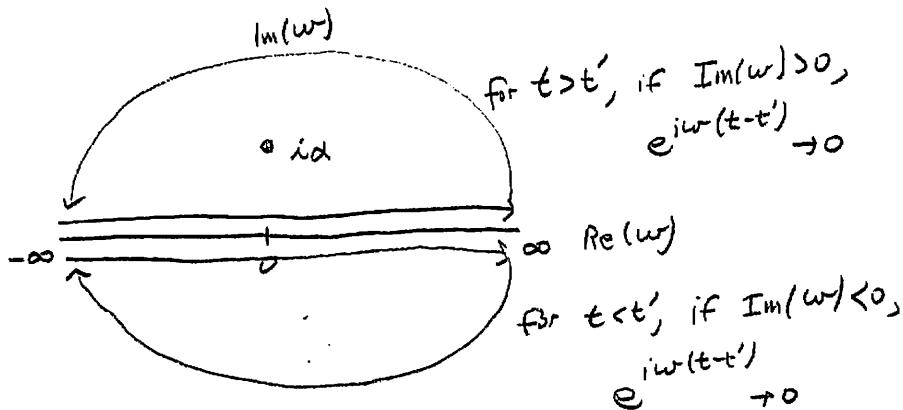
$$\phi = e^{\lambda t} \quad \text{conventional choice: write } \lambda = i\omega$$

$$\sum_i \phi_i(t) \rightarrow \frac{1}{2\pi} \int_{-\infty}^{\infty} dw e^{i\omega t}$$

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dw \frac{e^{i\omega t}}{i\omega + \alpha} \int_{-\infty}^{\infty} dt' e^{-i\omega t'} \beta(t') \quad \hat{\beta}(w) = \int_{-\infty}^{\infty} dt e^{-i\omega t} \beta(t)$$

$$x(t) = \int_{-\infty}^{\infty} dt' \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} dw \frac{e^{i\omega(t-t')}}{i\omega + \alpha} \right] \beta(t')$$

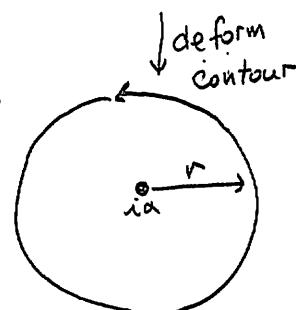
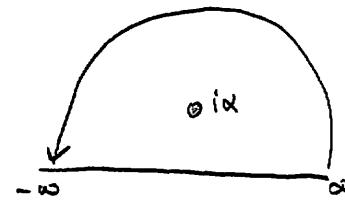
↓ what is this integral?



(3)

Contour Integrals.

$$\begin{aligned}
 \frac{1}{2\pi} \int_{-\infty}^{\infty} dw \frac{e^{i\omega t}}{i\omega + a} &= \frac{1}{2\pi} \oint dw \frac{e^{i\omega t}}{i\omega + a} \\
 &= \frac{1}{2\pi} \int_0^{2\pi} d\theta i r e^{i\theta} \frac{e^{i(i\alpha + re^{i\theta})t}}{i(i\alpha + re^{i\theta}) + a} \\
 &= \frac{1}{2\pi} e^{-at} \int_0^{2\pi} d\theta e^{ire^{i\theta} t} \\
 &\quad \downarrow \text{collapse contour, } r \rightarrow 0 \\
 &= e^{-at} \quad \text{for } t > 0
 \end{aligned}$$

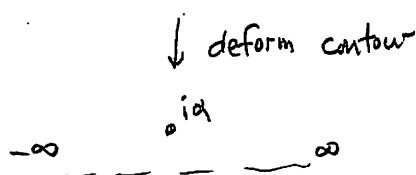
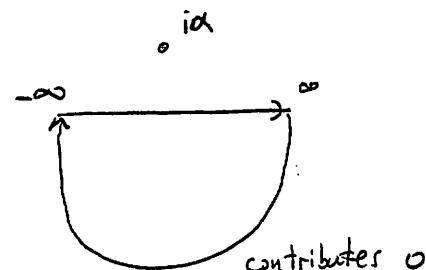


$$\omega = i\alpha + re^{i\theta}$$

$$\theta = 0 \rightarrow 2\pi$$

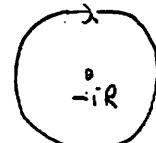
$$dw = ire^{i\theta} d\theta$$

$$\begin{aligned}
 \frac{1}{2\pi} \int_{-\infty}^{\infty} dw \frac{e^{i\omega t}}{i\omega + a} &\stackrel{t < 0}{=} \frac{1}{2\pi} \oint dw \frac{e^{-i\omega t}}{i\omega + a} \\
 &= \frac{1}{2\pi} \int_0^{2\pi} d\theta i r e^{i\theta} \frac{e^{-i(-iR + re^{i\theta})/t}}{i(-iR + re^{i\theta}) + a} \\
 &= -\frac{1}{2\pi} \int_0^{2\pi} d\theta i r e^{i\theta} \frac{e^{-(\infty)t}}{\infty} \rightarrow 0
 \end{aligned}$$



so:

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} dw \frac{e^{i\omega(t-t')}}{i\omega + a} = \begin{cases} e^{-a(t-t')}, & t > t' \\ 0, & t < t' \end{cases}$$



$$\omega = -iR + re^{i\theta}$$

$$\theta = 2\pi \rightarrow 0$$

$$\begin{aligned}
 x(t) &= \int_{-\infty}^{\infty} dt' H(t-t') \beta(t') = \int_{-\infty}^t dt' H(t-t') \beta(t') \\
 &\quad \downarrow \begin{bmatrix} e^{-a(t-t')}, & t > t' \\ 0, & t < t' \end{bmatrix} \quad \text{causality: } x(t) \text{ depends on past} \\
 &\quad \qquad \qquad \qquad \text{not future}
 \end{aligned}$$

Chapter 3

The Laplace Transform and Complex Variables

(3A)

Think about this weirdness.

Analytical fn with a Taylor series:

$$f(x) = c_0 + c_1 x + c_2 x^2 + \dots$$



$$\oint dz f(z) = \oint dz \sum_j c_j z^j = \sum_j c_j \frac{z^{j+1}}{x_0} \Big|_{x_0} = c_j (x_0^{j+1} - x_0^{j+1}) = 0$$

Function with a singularity at 0: $f(x) = \underbrace{\sum_{j=0}^{\infty} c_j x^j}_{\text{non-singular part}} + \frac{b}{x} + \sum_{k=2}^{\infty} \frac{b_k}{x^k}$

$$\oint dz f(z) = \sum_{j=0}^{\infty} c_j \frac{x^{j+1} - x_0^{j+1}}{j+1} \Big|_{x_0} \rightarrow 0$$

$$+ \sum_{k=2}^{\infty} \frac{b_k}{-k+1} x^{-k+1} \Big|_{x_0} \rightarrow 0$$

$$+ b \ln x \Big|_{x_0} \rightarrow ?$$

In complex plane, $x_0 = r e^{i\theta}$ at start

$$x_0 = r e^{i\theta + 2\pi i} \text{ at end}$$

$$b \ln x \Big|_{r e^{i\theta}}^{r e^{i\theta + 2\pi i}} = b \ln \left[\frac{r e^{i\theta + 2\pi i}}{r e^{i\theta}} \right] = b \cdot 2\pi i$$

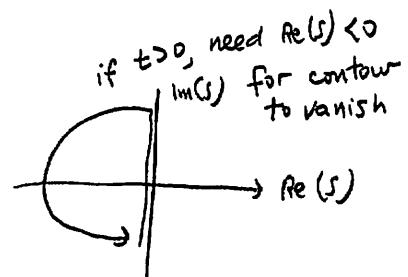
Mechanical rewriting

$$x(t) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi i} e^{i\omega t} \hat{x}(\omega) \quad \text{define } s = i\omega$$

$$ds = i d\omega$$

↓

$$x(t) = \int_{-\infty}^{i\infty} \frac{ds}{2\pi i} e^{st} \tilde{x}(s) \stackrel{t>0}{=} \int_{-\infty}^{i\infty} \frac{ds}{2\pi i} e^{st} \tilde{x}(s)$$



$$\hat{x}(\omega) = \int_{-\infty}^{\infty} dt e^{-i\omega t} x(t)$$

$$\downarrow s = i\omega$$

$$\tilde{x}(s) = \int_{-\infty}^{\infty} dt e^{-st} x(t) \quad \text{causality, pretend } x(t) = 0 \text{ for } t < 0$$

$$\int_0^{\infty} dt e^{-st} x(t)$$

No more lookup tables.

$$LT(1) = \int_0^{\infty} dt e^{-st} = \frac{1}{s}$$

$$\begin{aligned} LT\left(\frac{d}{dt} f(t)\right) &= \int_0^{\infty} dt e^{-st} \frac{d}{dt} f(t) \\ &= \int_0^{\infty} dt \left\{ \frac{d}{dt} \left[e^{-st} f(t) \right] - \left(\frac{d}{dt} e^{-st} \right) f(t) \right\} \\ &= \underbrace{-e^{-st} f(t)}_0 + s \tilde{f}(s) \\ &\sim -f(t=0) \end{aligned}$$

$$So L\left(\frac{d}{dt}\right) \approx s$$

(5)

Similarly, for time displacement.

$$f(t+a) = f(t) + a \frac{d}{dt} f(t) + \frac{1}{2} a^2 \frac{d^2}{dt^2} f(t) + \frac{1}{3!} a^3 \frac{d^3}{dt^3} f(t) + \dots$$

$$= \left[1 + a \frac{d}{dt} + \frac{a^2}{2} \frac{d^2}{dt^2} + \dots \right] f(t)$$

$$f(t+a) = \left[e^{a \frac{d}{dt}} \right] f(t)$$

$$\begin{aligned} LT[f(t+a)] &= \int_0^\infty dt e^{-st} f(t+a) e^{-sa} e^{+sa} \\ &= e^{as} \underbrace{\int_0^\infty dt e^{-(t+a)s} f(t+a)}_{\tilde{f}(s)} \end{aligned}$$

$$\text{Again: } e^{a \frac{d}{dt}} \Leftrightarrow e^{as}$$

time
domain Laplace
domain

$$\frac{d}{dt} \Leftrightarrow s$$

Back to our problem.

$$\frac{d}{dt} x(t) = \beta(t) - \alpha x(t)$$

$$s \tilde{x}(s) - x(0) = \tilde{\beta}(s) - \alpha \tilde{x}(s)$$

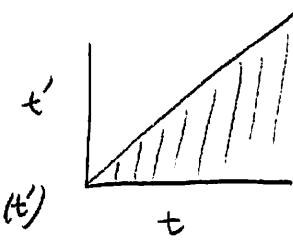
$$\tilde{x}(s) = \frac{1}{s+\alpha} \left[x(0) + \tilde{\beta}(s) \right]$$

$$x(t) = e^{-\alpha t} x(0) + \int_0^t e^{-\alpha(t-t')} \beta(t') dt'$$

(6)

Convolution:

$$\begin{aligned}
 \text{LT } f * g(t) &= \int_0^\infty dt e^{-st} \int_0^t dt' f(t-t') g(t') \\
 &\quad \downarrow \text{causality, define } f(t)=g(t)=0 \text{ for } t<0 \\
 &= \int_0^\infty dt \int_0^t dt' e^{-st} f(t-t') g(t') \\
 &= \int_0^\infty dt' \int_0^\infty dt e^{-st} e^{-st'} f(t-t') g(t') \\
 &= \int_0^\infty dt' e^{-st'} g(t') \int_0^\infty dt e^{-st} f(t-t') \\
 &= \tilde{f}(s) \tilde{g}(s)
 \end{aligned}$$



Multiple poles

$$\frac{1}{2\pi i} \oint_{-\infty}^{i\infty} ds \frac{e^{st} f(s)}{(s+\alpha)(s+\beta) \dots} = \frac{1}{2\pi i} \cdot 2\pi i \left[\frac{e^{-\alpha t} f(-\alpha)}{(-\alpha+\beta) \dots} + \frac{e^{-\beta t} f(-\beta)}{(-\beta+\alpha) \dots} + \dots \right]$$

Chapter 4

Signal Transduction Cascades and MAPK Signaling

Linear cascade

Illustrates

- convolution
- description of response, real time: amplitude gain, t_{\max} , $t_{1/2}$
- description of response, LT: AUC, t_{mean} , σ_t^2

$$\Rightarrow \begin{array}{l} R \\ \hline \end{array} \quad t=0 \Rightarrow \text{all off}, \quad X_1 = X_2 = \dots = X_n = 0$$

$\overset{\text{s(t) input signal}}{\underset{\text{R}_1, R(t)}{\overbrace{X_1}}} - \alpha X_1(t)$

$$\overset{\text{X}_1 \Leftrightarrow X_1}{\downarrow} \quad \overset{\text{X}_2 \Leftrightarrow X_2}{\downarrow} \quad \vdots \quad \overset{\text{X}_n \Leftrightarrow X_n}{\uparrow \quad \uparrow}$$

$$\overset{\text{off state}}{\text{X}_1} \quad \overset{\text{on state}}{\text{X}_2} \quad \vdots \quad \overset{\text{inactive}}{\text{X}_n}$$

$$\dot{X}_1(t) = \underbrace{R_1}_{\beta_1} X_1^0 - \alpha X_1(t) \quad \tilde{X}_1 = \frac{\tilde{S}}{s+\alpha} = \frac{\beta_1}{s+\alpha}$$

$$\dot{X}_2(t) = \underbrace{R_2}_{\beta_2} X_2^0 X_1(t) - \alpha X_2(t) \quad \tilde{X}_{n-1} = \frac{\beta_{n-1}}{s+\alpha} \tilde{X}_{n-2}$$

$$\vdots$$

$$\dot{X}_n(t) = \underbrace{R_n}_{\beta_n} X_n^0 X_{n-1}(t) - \alpha X_n(t) \rightarrow s \tilde{X}_n = \beta_n \tilde{X}_{n-1} - \alpha$$

$$\tilde{X}_n = \frac{\beta_n}{s+\alpha} \tilde{X}_{n-1}$$

non-saturating / weak

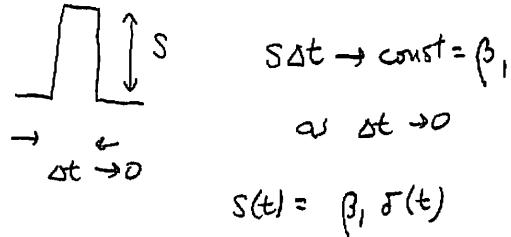
approximation:

$$X_1^0 = X_1^{\text{tot}} - X_1 \approx X_1^{\text{tot}}$$

↑ ↑ ↑

inactive total active

Input signal: impulse



$$\tilde{S}(s) = \int_0^\infty dt e^{-st} \beta_1 \delta(t) = \beta_1$$

$$\tilde{X}_n(s) = \frac{\beta_n \beta_{n-1} \cdots \beta_1}{(s+\alpha)^n} \approx \frac{\beta^n}{(s+\alpha)^n} \quad \text{if rate const., concentrations & eqbal}$$

$$X_n(t) = \int_{-\infty}^{\infty} \frac{ds}{2\pi i} e^{st} \cdot \frac{\beta^n}{(s+\alpha)^n}$$

↑ n^{th} order pole

(2)

Remember: $\oint dz \frac{c}{(z-a)^n} = 0$ for $n \neq 1$ and $c = \text{const.}$
 contour around
 a

$\oint ds \frac{e^{st}}{(s+a)^n} = 0$ for $n=1$? No! $e^{st} \neq \text{cont.}$

Have to expand e^{st} in a Taylor around a .

Trick: $1 = e^{-at} e^{at}$

$$x_n(t) = \frac{1}{2\pi i} \oint ds \frac{\beta^n}{(s+a)^n} e^{st} e^{at} e^{-at} = \frac{1}{2\pi i} \beta^n e^{-at} \underbrace{\oint ds \frac{1}{(s+a)^n} e^{(s+a)t}}_{1 + (s+a)t + \frac{1}{2}(s+a)^2 t^2 + \dots + \frac{1}{(n-1)!} (s+a)^{n-1} t^{n-1} + \dots}$$

$$= \beta^n e^{-at} \cdot \frac{1}{2\pi i} \cdot \frac{1}{(n-1)!} t^{n-1}$$

$$x_n(t) = \beta_1 \cdot (\beta t)^{n-1} \cdot \frac{1}{(n-1)!} e^{-at} \quad \begin{matrix} \text{check units: } \beta_1 = \text{const.} \\ \beta_2 \dots \beta_n = \frac{1}{t^{n-1}} \end{matrix} \quad \begin{matrix} \text{only surviving} \\ \text{term} \end{matrix}$$

Time domain: max amplitude
 as long as $x_n(t) \neq 0$

$$0 = \frac{d}{dt} x_n(t) = \frac{1}{x_n(t)} \frac{d}{dt} x_n(t) = \frac{d}{dt} \ln x_n(t) \quad x_n(t) \text{ and } \log x_n(t) \text{ have max. at same value}$$

$$0 = \frac{d}{dt} (n-1) \ln t - at = \frac{(n-1)}{t} - \alpha \Rightarrow t_{\max} = \frac{(n-1)}{\alpha}$$

Time domain: amplitude gain = $\frac{x_n(t_{\max})}{\beta_1}$
 input amplitude

each step

brings delay α

(3)

$$x_n(t_{\max}) = \beta_1 \cdot \left[\beta \cdot \frac{(n-1)}{\alpha} \right]^{n-1} \cdot \frac{1}{(n-1)!} e^{-\alpha(n-1)/\alpha}$$

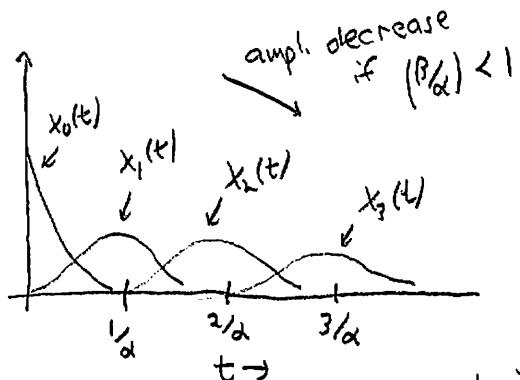
$$= \beta_1 \cdot \left(\frac{\beta}{\alpha} \right)^{n-1} \cdot \frac{(n-1)^{n-1} e^{-(n-1)}}{(n-1)!}$$

Stirling's approx: $n! \approx \sqrt{2\pi n} n^{\frac{n}{e}}$

$$x_n(t_{\max}) \approx \beta_1 \cdot \left(\frac{\beta}{\alpha} \right)^{n-1} \cdot \frac{1}{\sqrt{2\pi(n-1)}}$$

$$\text{amplitude gain} = \frac{x_n(t_{\max})}{\text{input } \beta_1} = \left(\frac{\beta}{\alpha} \right)^{n-1} \cdot \frac{1}{\sqrt{2\pi(n-1)}}$$

↑
amp.



Time domain: $t \downarrow$

$$\text{If } x_n(t) \approx \text{prefac. } e^{-\frac{1}{2} \frac{(t-t_{\max})^2}{\sigma^2}},$$

$$\text{then } t_{1/2} \approx \sigma$$

$$\text{write } \ln x_n(t) \approx -\frac{1}{2} \frac{(t-t_{\max})^2}{\sigma^2}$$

$$\frac{d \ln x_n(t)}{dt} = -\frac{(t-t_{\max})}{\sigma^2} = 0 \text{ at } t=t_{\max}$$

$$= \frac{(n-1)}{t} - \alpha$$

Pattern:

time delay \leftrightarrow relaxn
time width \rightarrow rate \propto

$$\frac{d^2 \ln x_n(t)}{dt^2} = -\frac{1}{\sigma^2} = -\frac{(n-1)}{t^2}$$

$$\text{so } \sigma^2 \approx \frac{t_{\max}^2}{(n-1)}$$

gain $\leftrightarrow \left(\frac{\beta}{\alpha} \right) \frac{\text{production}}{\text{decay}}$

$$T \approx \sqrt{\frac{t_{\max}}{(n-1)}} = \sqrt{\frac{1}{n-1}} \cdot \frac{(n-1)}{\alpha} = \frac{\sqrt{n-1}}{\alpha} \quad \text{width grows.}$$

Chapter 5

Generating Functions for Pharmacokinetics and Pharmacodynamics

What about 1st signal, $x_1(t)$? $\sigma = \sqrt{\frac{1-\alpha}{\alpha}}$?

(4)

This is from Stirling's approx.

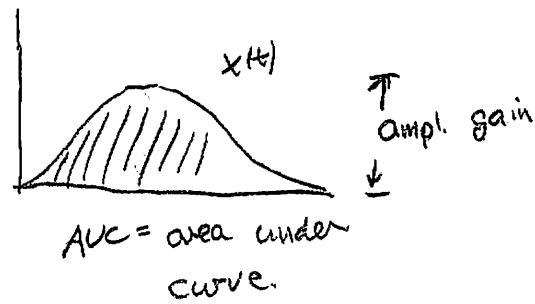
Go back to $x_1(t) = \beta_1 \underbrace{\frac{(\beta t)^{n-1}}{(n-1)!}}_{1 \text{ for } n=1} e^{-\alpha t}$

at $t_{1/2}$, $x_1(t) = \frac{\beta_1}{2} = \beta_1 e^{-\alpha t} \Rightarrow -\ln 2 = -\alpha t \Rightarrow t_{1/2} = \frac{1}{\alpha} \ln 2$

Similar quantities from Laplace domain \rightarrow LT as generating fn.

Gain \rightarrow AUC

$AUC(x_n(t)) = \int_0^\infty dt x_n(t)$ for a pos. signal.



$$= \lim_{s \rightarrow 0} \int_0^\infty dt e^{-st} x_n(t) = \tilde{x}_n(s)$$

$$AUC(x_n(t)) = \lim_{s \rightarrow 0} \frac{\beta^n}{(s+\alpha)^n} = \left(\frac{\beta}{\alpha}\right)^n \quad \begin{matrix} \text{same ratio as ampl. gain!} \\ \text{almost: } \left(\frac{\beta}{\alpha}\right) \cdot \text{ampl. gain.} \end{matrix}$$

$t_{max} \rightarrow t_{mean}$

$$t_{mean} = \frac{\int_0^\infty dt t \cdot x_n(t)}{\int_0^\infty dt x_n(t)} \leftarrow AUC$$

Generating fn trick!

(5)

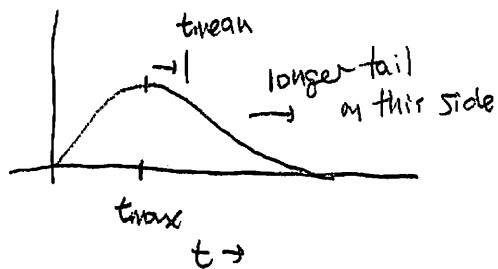
$$\begin{aligned} \frac{d}{ds} \ln \tilde{X}(s) &= \frac{1}{\tilde{X}(s)} \frac{d}{ds} \tilde{X}(s) = \frac{1}{\tilde{X}(s)} \frac{d}{ds} \int_0^{\infty} dt e^{-st} x(t) \\ &= \frac{1}{\tilde{X}(s)} \int_0^{\infty} dt (-t) e^{-st} x(t) \stackrel{s \rightarrow 0}{=} \frac{\int_0^{\infty} dt (-t) x(t)}{\int_0^{\infty} dt x(t)} = -t_{\text{mean}} \end{aligned}$$

so $t_{\text{mean}} = - \frac{d}{ds} \ln \tilde{X}(s) \Big|_{s=0}$

\uparrow
always take \ln first!

$$t_{\text{mean}}(x_n) = - \frac{d}{ds} \ln \left(\frac{P^n}{(s+\alpha)^n} \right) \Big|_{s=0} = - \frac{d}{ds} -n \ln(s+\alpha) = \frac{n}{s+\alpha} \Big|_{s=0} = \frac{n}{\alpha}$$

vs. $t_{\text{max}} = \frac{(n-1)}{\alpha}$



Width:

$$\sigma^2 = \frac{\int_0^{\infty} dt (t - \bar{t})^2 x(t)}{\int_0^{\infty} dt x(t)} = \frac{\int_0^{\infty} dt t^2 x(t)}{\int_0^{\infty} dt x(t)} - \left[\frac{\int_0^{\infty} dt t x(t)}{\int_0^{\infty} dt x(t)} \right]^2$$

$$= \frac{d^2}{ds^2} \ln \tilde{X}(s) \Big|_{s=0}$$

future Hw

$$\hat{\sigma}^2(x_n) = \frac{d^2}{ds^2} \ln \frac{\beta^n}{(s+\alpha)^n} = \frac{d^2}{ds^2} -n \ln(s+\alpha) = \frac{d}{ds} - \frac{n}{(s+\alpha)} = \frac{n}{(s+\alpha)^2} \Big|_{s=0} = \frac{n}{\alpha^2}$$

$$\hat{\sigma}^2(x_n) = \frac{n}{\alpha^2}$$

$$f(x_n) = \frac{\sqrt{n}}{\alpha} \quad \text{vs. time domain } \sigma(x_n) \approx \frac{\sqrt{n-1}}{\alpha}$$

Summary:

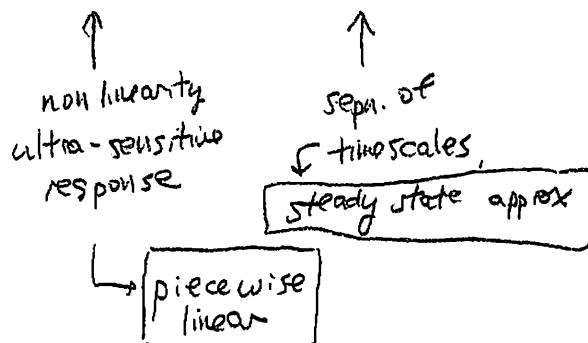
time domain	LT gen. fn
<hr/>	
$t_{max} = \frac{(n-1)}{\alpha}$	$\bar{t} = \frac{n}{\alpha}$
$\text{ampl. gain} = \left(\frac{\beta}{\alpha}\right)^{n-1}$	$AUC = \left(\frac{\beta}{\alpha}\right)^n$
$\sigma \approx \frac{\sqrt{n-1}}{\alpha}$	$\sigma = \frac{\sqrt{n}}{\alpha}$

Time domain depends on α

gain, depends on ampl., area (β/α)

Next time:

Signaling \rightarrow TF binding \rightarrow transcription \rightarrow translation

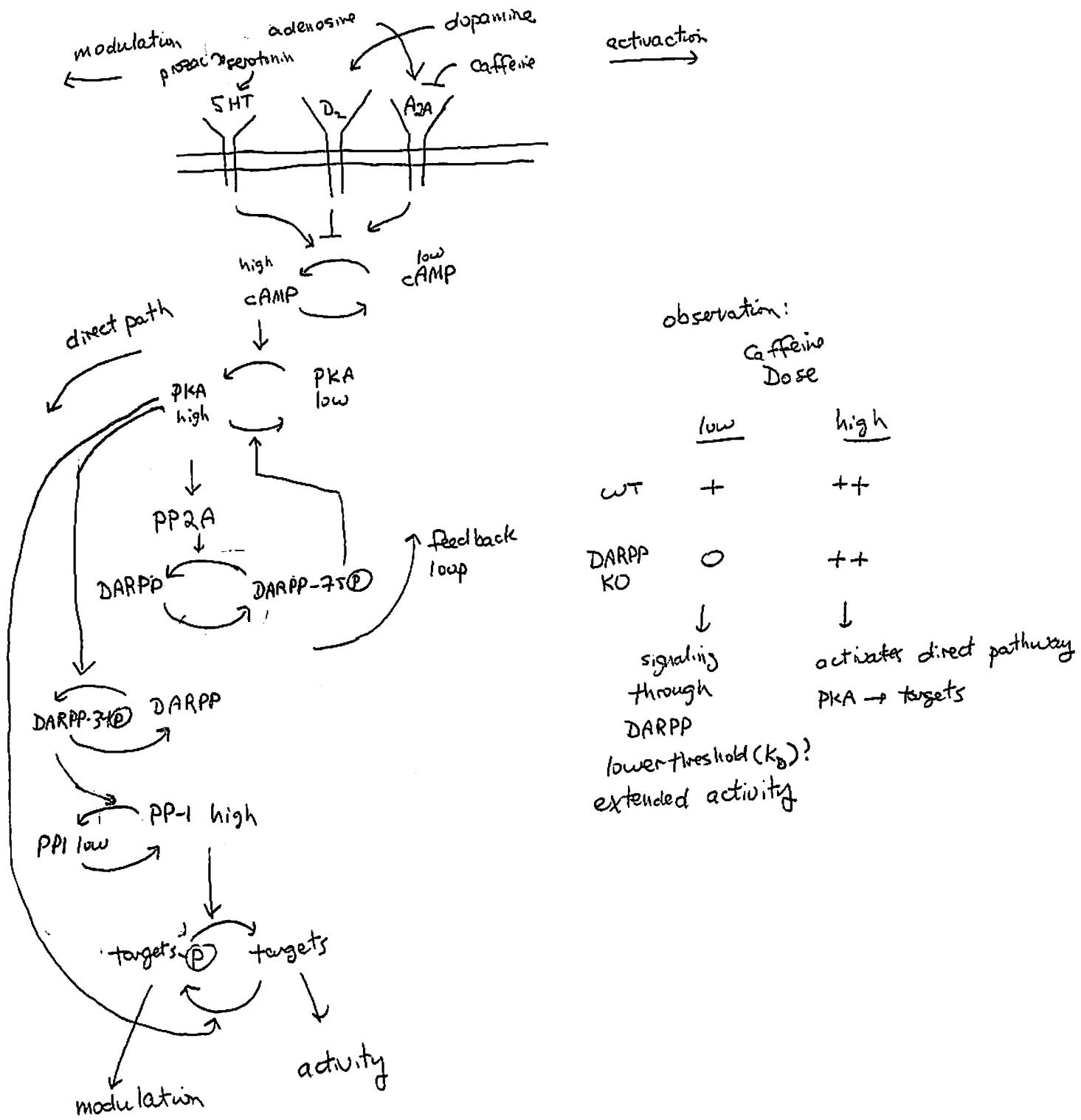


Chapter 6

Positive Feedback and Caffeine Response

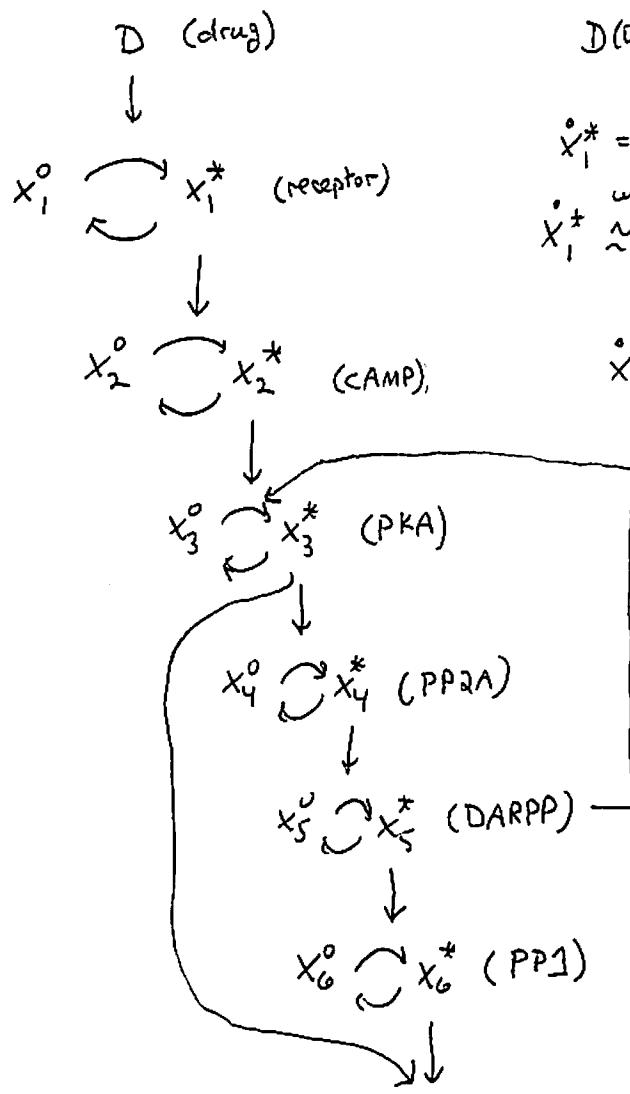
(2)

Biology of Caffeine - see two Science articles



(3)

stripped-down model

abbrev: $\tilde{x}_i = \frac{k_i}{(S+d_i)}$

$$\tilde{x}_5 = \frac{\frac{k_5}{(S+d_5)}}{1 - \frac{k_5 k_4 k_3^f}{(S+\alpha)^3}} \tilde{D}$$

$$\tilde{x}_6 = \frac{\frac{k_6}{(S+\alpha)^6}}{1 - \frac{\tilde{x}_5 k_4 k_3^f}{(S+\alpha)^3}} \tilde{D}$$

<u>ODE</u>	<u>LT</u>
$D(D) = 0$	$x_1^* = \frac{k_1 D(t)}{(S+d_1)} - \alpha_1 x_1^*$ weak activation
	$x_1^* \approx (k_1 D(t)) \tilde{D} - \alpha_1 x_1^*$ $\downarrow x_1^* = \text{active form}$
	$\dot{x}_1^* = k_1 D(t) - \alpha_1 x_1^* \rightarrow \tilde{x}_1^* = \frac{k_1}{S+d_1} \tilde{D}$
	$\dot{x}_2 = k_2 x_1^* - \alpha_2 x_2 \quad \tilde{x}_2 = \frac{k_2}{S+d_2} \frac{k_1}{S+d_1} \tilde{D}$
	$\dot{x}_3 = k_3 x_2 + k_3^f x_5 - \alpha_3 x_3 \quad \tilde{x}_3 = \frac{k_3 k_2 k_1}{(S+\alpha_3)(S+d_2)} \frac{1}{S+\alpha_1} \tilde{D}$
	$\dot{x}_4 = k_4 x_3 - \alpha_4 x_4 + \frac{k_3^f}{S+\alpha_3} \tilde{x}_5$
	$\dot{x}_5 = k_5 x_4 - \alpha_5 x_5$
	$\tilde{x}_4 = \frac{k_4}{S+\alpha_4} \tilde{x}_3 \quad \text{keep in terms of } \tilde{x}_3$
	$\tilde{x}_5 = \frac{k_5 k_4}{(S+\alpha_5)(S+\alpha_4)} \tilde{x}_3$
	Aha! substitute back
	$\tilde{x}_3 = \frac{k_3 k_2 k_1}{(S+d_3)(S+d_2)(S+\alpha_1)} \tilde{D} + \frac{k_5 k_4 k_3^f}{(S+d_5)(S+\alpha_4)(S+d_2)} \tilde{x}_3$
	$\tilde{x}_3 = \frac{\frac{k_3 k_2 k_1}{(S+d_3)(S+d_2)(S+\alpha_1)} \tilde{D}}{1 - \frac{k_5 k_4 k_3^f}{(S+\alpha_5)(S+\alpha_4)(S+d_2)}} \tilde{x}_3$

(4)

Finally

direct path feedback path

$$\dot{T} = R_T x_3 + k_T^f x_6 - \alpha_T T$$

$$\tilde{T} = \frac{\frac{R_T R_3 R_2 R_T}{(s+\alpha_T)(s+d_3)(s+d_2)(s+d_1)}}{1 - \frac{R_3 k_T k_T^f}{(s+d_5)(s+d_4)(s+d_3)}} \tilde{D} + \frac{\frac{k_T^f}{(s+\alpha)^3}}{1 - \frac{k_T^3}{(s+\alpha)^3}} \tilde{D}$$

DARPP knock out, $k_f = 0$

$$\tilde{T}_{KO} = \frac{k^4}{(s+\alpha)^4} \tilde{D}$$

What does it all mean?

Drug exposure = $\int_0^\infty dt D(t) = AUC$

$$D(t) = -\alpha_D t D(t)$$

$$\frac{1}{2} = e^{-\alpha_D t_{1/2}} = e^{-\ln 2}$$

$$D(t=0) = D_0$$

$$\alpha_D = \frac{\ln 2}{t_{1/2}} \approx \frac{\ln 2}{4 \text{ hrs}}$$

$$D(s) = \frac{D_0}{s + \alpha_D}$$

$$AUC(D) = \frac{D_0}{\alpha_D}$$

$$AUC(T_{WT}) = \frac{(k/\alpha)^4}{1 - (k/\alpha)^3} AUC(D) + \frac{(k/\alpha)^7}{1 - (k/\alpha)^3} AUC(D)$$

$$AUC(T_{KO}) = (k/\alpha)^4 AUC(D)$$

from extended signaling of pKA

$$\frac{\text{effective dose (WT)}}{\text{effective dose (KO)}} = \frac{1}{1 - (k/\alpha)^3} + \frac{(k/\alpha)^3}{1 - (k/\alpha)^3}$$

from route through PP-1

gain from feedback

(5)

What about duration?

$$\bar{t}_D = - \frac{d}{ds} \ln \tilde{D}(s) \Big|_{s=0} = - \frac{d}{ds} \ln \left(\frac{D}{s+\alpha_D} \right) = \frac{d}{ds} \ln(s+\alpha_D) = \frac{1}{s+\alpha_D} \Big|_{s=0} = \frac{1}{\alpha_D}$$

$$\bar{t}_T = - \frac{d}{ds} \left\{ \ln \left[\frac{k^4}{(s+\alpha)^4} \tilde{D} \right] - \ln \left[1 - \frac{k^3}{(s+\alpha)^3} \right] \right\}$$

ignore PPL route
for simplicity

$$= \frac{d}{ds} \ln(s+\alpha)^4 - \frac{1}{1 - \frac{k^3}{(s+\alpha)^3}} \frac{d}{ds} \left(1 - \frac{k^3}{(s+\alpha)^3} \right) - \frac{d}{ds} \ln \tilde{D}$$

$$= \frac{4}{\alpha} - \left[\frac{1}{1 - \frac{k^3}{\alpha^3}} \cdot \frac{-3k^3}{\alpha^4} \right] + \frac{1}{\alpha_D}$$

\downarrow
 G_{feedback}

$$= \underbrace{\frac{4}{\alpha} + \frac{1}{\alpha} \cdot \frac{3 \left(\frac{k^3}{\alpha^3} \right)}{\left[1 - \frac{1}{G_f} \right]} \cdot G_f}_{\text{extra duration from feedback}} + \frac{1}{\alpha_D} \quad G_f = \frac{1}{1 - \frac{k^3}{\alpha^3}}$$

$$1 - \frac{k^3}{\alpha^3} = \frac{1}{G_f}$$

$$\bar{t}_{T_{\text{wr}}} = \underbrace{\frac{4}{\alpha} + \frac{1}{\alpha_D}}_{\bar{t}_{T_{\text{KU}}}} + \underbrace{\frac{3}{\alpha} \left[G_f - 1 \right]}_{\text{extra duration from feedback.}}$$

$$\frac{k^3}{\alpha^3} = 1 - \frac{1}{G_f}$$

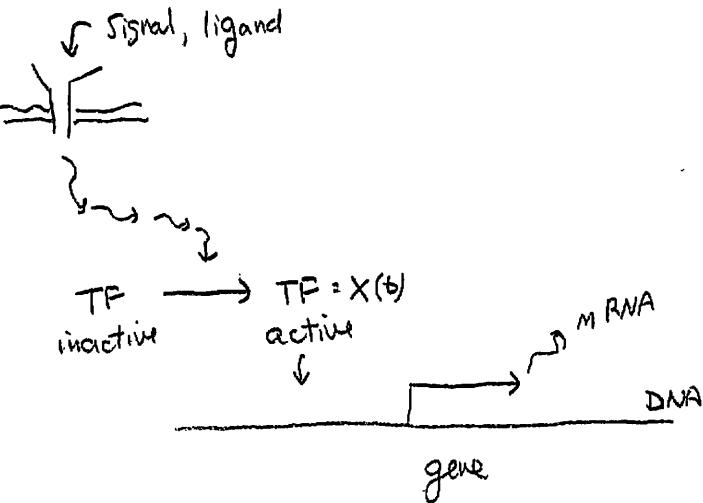
Part II

Cells as Non-linear Systems

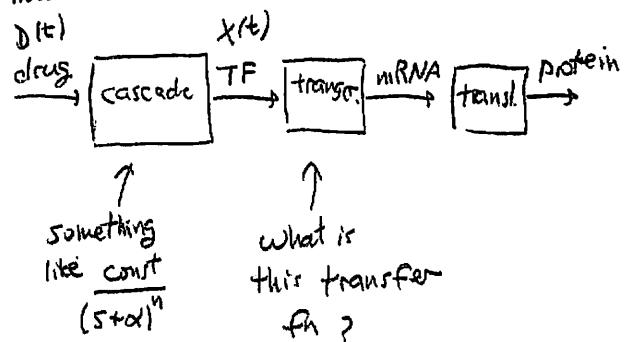
Chapter 7

Saturation and Cooperative Response

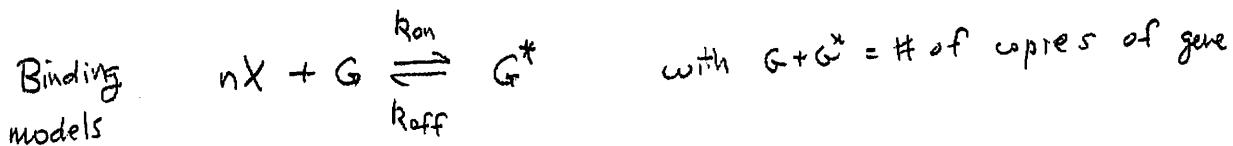
Lecture 4 Cooperativity, Logic Response, Piecewise Linear



block model:



Rate of mRNA production = $\begin{cases} 0, & \text{promoter unbound} \\ \beta_m, & \text{promoter bound} \end{cases}$



$$\frac{dG^*}{dt} = -k_{\text{off}} G^* + k_{\text{on}} X^n [G_{\text{tot}} - G^*]$$

$$s \tilde{G}^* - G^*(0) = -k_{\text{off}} \tilde{G}^* - k_{\text{on}} X^n \tilde{G}^* + k_{\text{on}} X^n \frac{G_{\text{tot}}}{s}$$

$$[s + k_{\text{off}} + k_{\text{on}} X^n] \tilde{G}^* = \frac{k_{\text{on}} X^n G_{\text{tot}}}{s}$$

$$\tilde{G}^*(s) = \frac{\frac{k_{\text{on}} X^n G_{\text{tot}}}{s}}{s(s + k_{\text{off}} + k_{\text{on}} X^n)}$$

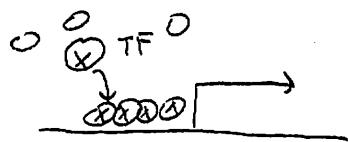
fraction active

$$\frac{G^*(t)}{G_{\text{tot}}} = \frac{\frac{k_{\text{on}} X^n}{s}}{k_{\text{off}} + \frac{k_{\text{on}} X^n}{s}} \left(1 - e^{-\frac{(k_{\text{off}} + k_{\text{on}} X^n)t}{s}} \right)$$

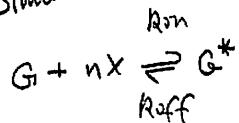
Lecture: Saturation + Cooperativity = Logistic response

①

Timescale separation



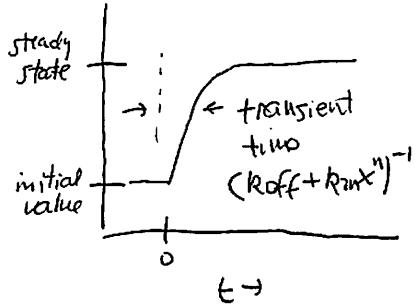
Simultaneous binding model



Gene

$$\frac{dG^*}{dt} = k_{\text{on}} GX^n - k_{\text{off}} G^*$$

$$SG^* - G^*(0) = k_{\text{on}} X^n \left[\frac{G_{\text{tot}}}{S} - G^* \right] - k_{\text{off}} G^*$$



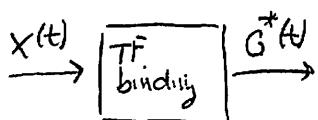
$$[S + k_{\text{on}} X^n + k_{\text{off}}] G^* = k_{\text{on}} X^n G_{\text{tot}} + G_0$$

$$G^* = \frac{1}{S + k_{\text{on}} X^n + k_{\text{off}}} G_0 + \frac{k_{\text{on}} X^n}{S(S + k_{\text{on}} X^n + k_{\text{off}})} G_{\text{tot}}$$

$$\frac{G^*(t)}{G_{\text{tot}}} = e^{- (k_{\text{on}} X^n + k_{\text{off}}) t} \left(\frac{G_0}{G_{\text{tot}}} \right) + \frac{(X_K)^n}{1 + (X_K)^n} \left(1 - e^{- (k_{\text{on}} X^n + k_{\text{off}}) t} \right)$$

Fast transient \rightarrow quickly reach steady-state value

$$K^n = k_{\text{off}} / k_{\text{on}}$$



non-linear transform

$$\frac{G^*(t)}{G_{\text{tot}}} = \frac{(X_K)^n}{1 + (X_K)^n}$$

activator

$$= \frac{1}{1 + (X_K)^n}$$

repressor

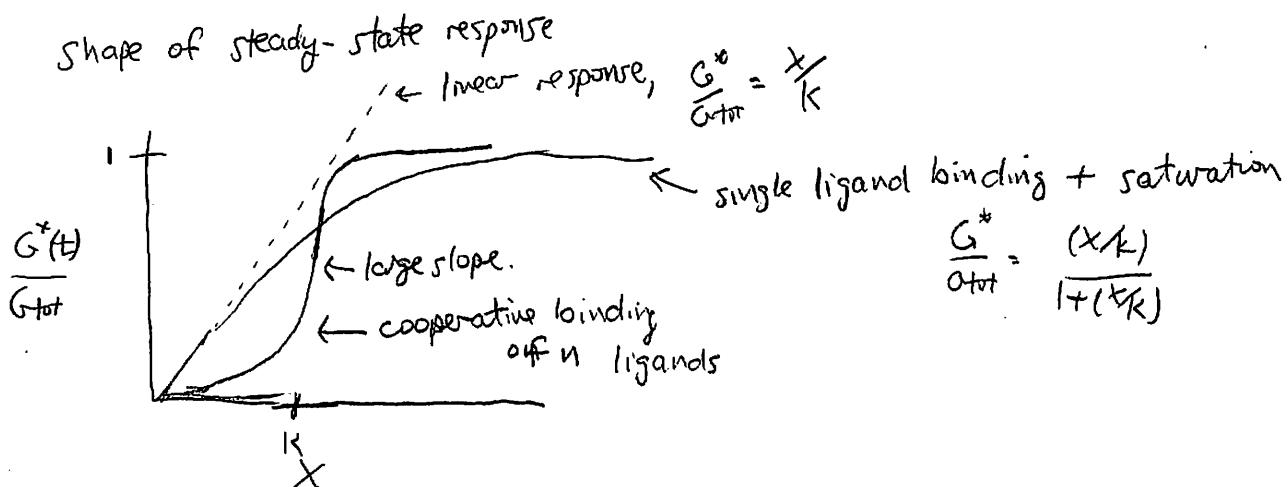
Approximation FF:

Separation of timescales: ignore transient, use immediate response

$$\frac{dG^*}{dt} = 0 \Rightarrow k_{\text{on}} GX^n - k_{\text{off}} G^* = 0$$

$$\Rightarrow \frac{G^*}{G_{\text{tot}}} = \frac{(X_K)^n}{1 + (X_K)^n}$$

②



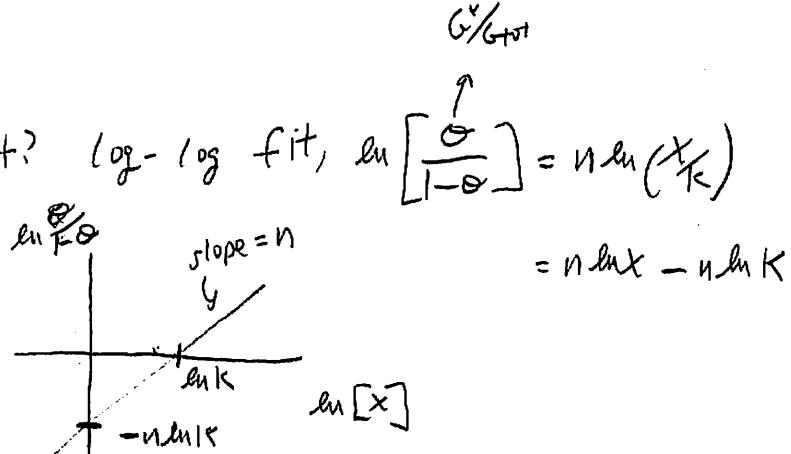
ultrasensitivity = faster than linear response.

Quantitative measure: look at $\frac{G^*/G_{tot}}{1 - G^*/G_{tot}} = \left(\frac{x}{K}\right)^n$

single ligand binding: $n = 1$

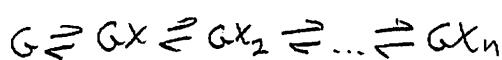
multiple binding: $n > 1$

How to extract the exponent? log-log fit, $\ln \left[\frac{\sigma}{1-\sigma} \right] = n \ln \left(\frac{x}{K} \right)$



Other models:

sequential binding:



independent binding:

n sites, $K = 1 \dots n$

prob that site k is bound = $\frac{(x/K)}{1 + (x/K)} = \sigma_i$

prob that k of n sites are bound = $\binom{n}{k} \sigma_i^k (1-\sigma_i)^{n-k} = \binom{n}{k} \frac{(x/K)^k}{(1+x/K)^n}$

Are these models cooperative?

$$\text{Is } \frac{d \ln \frac{\sigma}{1-\sigma}}{d \ln x} > 1 ?$$

(3)

Sequential binding

$$G_0 \rightleftharpoons G_1 \rightleftharpoons G_2 \rightleftharpoons \dots \rightleftharpoons G_n \quad \text{with} \quad \sum_k G_k = 1$$

$$G_0 + X \rightleftharpoons G_1$$

$$k_{on} G_0 X = k_{off} G_1 \Rightarrow \frac{G_1}{G_0} = \frac{k_{on} X}{k_{off}} = \frac{X}{K} \quad K = \frac{k_{off}}{k_{on}} \quad \text{here}$$

$$\frac{G_2}{G_1} = \frac{X}{K} \Rightarrow G_2 = \left(\frac{X}{K}\right)^2 G_0$$

$$\vdots \\ G_n = \left(\frac{X}{K}\right)^n G_0$$

$$\text{Saturation: } G_0 + G_1 + \dots + G_n = \left[1 + \left(\frac{X}{K}\right) + \dots + \left(\frac{X}{K}\right)^n\right] G_0 = G_{tot}$$

$$\frac{G_0}{G_{tot}} = \frac{1}{1 + \left(\frac{X}{K}\right) + \dots + \left(\frac{X}{K}\right)^n}$$

$$\frac{G^*}{G_{tot}} = \frac{\left(\frac{X}{K}\right)^n}{1 + \left(\frac{X}{K}\right) + \dots + \left(\frac{X}{K}\right)^n}$$

cooperativity:

$$\frac{\Theta}{1-\Theta} = \frac{G^*/G_{tot}}{1-G^*/G_{tot}} = \frac{\left(\frac{X}{K}\right)^n}{1 + \dots + \left(\frac{X}{K}\right)^{n-1}} = \begin{cases} \text{small } X, \sim \left(\frac{X}{K}\right)^n & \xrightarrow{\text{looks cooperative at the beginning}} \\ \text{large } X, \sim \left(\frac{X}{K}\right) & \xrightarrow{\text{not so cooperative at the end.}} \end{cases}$$

Independent binding:

$$\frac{G^*}{G_{tot}} = \Theta_1^n \quad \text{with} \quad \Theta_1 = \frac{\left(\frac{X}{K}\right)}{1 + \left(\frac{X}{K}\right)}$$

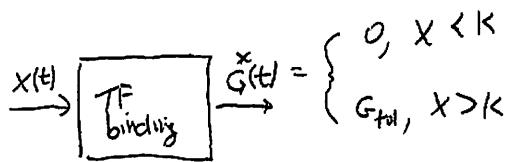
small X , looks cooperative
 large X , $\sim \frac{1}{n-1} \left(\frac{X}{K}\right)$ not so cooperative

$$\frac{G^*/G_{tot}}{1 - G^*/G_{tot}} = \frac{\Theta_1^n}{1 - \Theta_1^n} = \frac{\left(\frac{X}{K}\right)^n / \left(1 + \left(\frac{X}{K}\right)^{-1}\right)^n}{1 - \frac{\left(\frac{X}{K}\right)^n}{\left(1 + \left(\frac{X}{K}\right)\right)^n}} = \frac{\left(\frac{X}{K}\right)^n}{1 + n\left(\frac{X}{K}\right) + \frac{n(n-1)}{2} \left(\frac{X}{K}\right)^2 + \dots + (n-1) \left(\frac{X}{K}\right)^{n-1}}$$

Chapter 8

Joint Models of Transcription and Translation

(4)



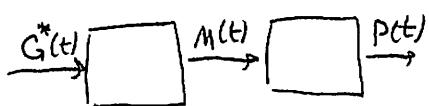
approx 1: steady-state

approx 2: continuous, ultrasensitive \rightarrow binary response logic \rightarrow piecewise linear

$$\frac{G^*(t)}{G_{\text{tot}}} = \Theta(X > K)$$

$H(X > K) \leftarrow$ also used as a truth fn.

Next steps: transcription, translation



$$\dot{M}(t) = \beta_m G^*(t) - \alpha_m M(t)$$

mRNA decay rate, $\sim \frac{1}{\text{lifetime}} \sim \frac{1}{5-10 \text{ min}}$

transcription rate, $\sim 1 \text{ transcript}/5 \text{ min}$

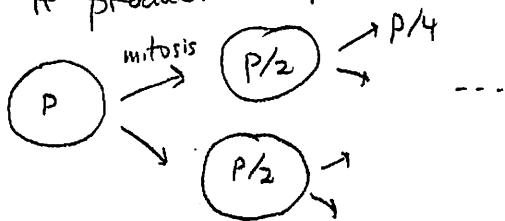
$$\dot{P}(t) = \beta_p M(t) - \alpha_p P(t)$$

protein decay rate $\rightarrow ?$

translation rate, $\sim 1 \text{ protein}/5 \text{ min}$

Active degradation vs. dilution

If production stops:

Mammalian: $\gamma \sim \text{days}$ Yeast, bacteria: $\gamma \sim \text{hours}$

$$\gamma_{\text{protein}} \gg \gamma_{\text{mRNA}}$$

hrs-days min

$$P(t) = P_0 2^{-\frac{t}{\gamma}} = P_0 e^{-\frac{t}{\gamma} \ln 2}$$

$$\dot{P}(t) = -\frac{\ln 2}{\gamma} P(t)$$

$\downarrow \alpha_{\text{dilution}} \sim \frac{\ln 2}{\gamma} \sim \frac{1}{\gamma}$

Starting from $M(0) = P(0) = 0$, $G^* = 1$:

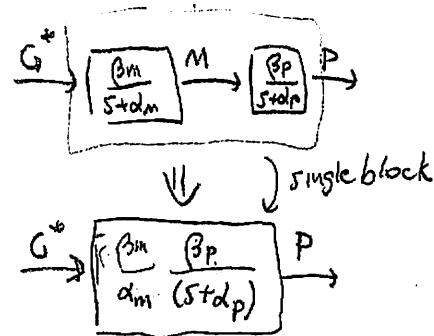
$$\dot{M} = \beta_M - \alpha_M M \Rightarrow s\tilde{M} = \frac{\beta_M}{s} - \alpha_M \tilde{M} \Rightarrow \tilde{M} = \frac{\beta_M}{s(s + \alpha_M)} \Rightarrow M(t) = \boxed{\frac{\beta_M}{\alpha_M}} (1 - e^{-\alpha_M t})$$

$$\dot{P} = \beta_P M - \alpha_P P \Rightarrow s\tilde{P} = \frac{\beta_P \tilde{M}}{s} - \alpha_P \tilde{P} \Rightarrow \tilde{P} = \frac{\beta_P \beta_M}{s(s + \alpha_M)(s + \alpha_P)} \quad \gamma_M \ll \gamma_P \\ \alpha_M \gg \alpha_P$$

$$P(t) = \frac{\beta_P \beta_M}{\alpha_P (\alpha_M - \alpha_P)} (1 - e^{-\alpha_P t}) + \underbrace{\frac{\beta_P \beta_M}{(-\alpha_M)(-\alpha_M + \alpha_P)} e^{-\alpha_M t}}_{\text{fast transient}}$$

$$\downarrow \alpha_M \gg \alpha_P$$

$$P(t) \approx \frac{\beta_P \beta_M}{\alpha_P} \left(1 - e^{-\alpha_P t}\right) \quad \text{effective production rate}$$



$$\dot{P}(t) = \boxed{\frac{\beta_P \beta_M}{\alpha_M}} G^*(t) - \alpha_P P(t)$$

$$\text{effective } \beta = \beta_P \cdot \left[\frac{\frac{1 \text{ copy}}{5 \text{ min}}}{\frac{1}{100} \text{ min}} \right] \approx \frac{2 \text{ copies}}{5 \text{ min}}$$

$$\frac{1}{5 \text{ min}}$$

$$\# \text{ copies protein per cell} \approx \frac{\beta}{\alpha_P} \approx \frac{\frac{2 \text{ copies}}{5 \text{ min}}}{\frac{1}{100} \text{ min}} \approx 40 \text{ copies reasonable.}$$

Chapter 9

Positive and Negative Auto-Regulation

Lecture: Positive and negative auto regulation

①

Theme: physical meaning of pathway diagrams

Motifs: units in a network that we think we understand
resistors, capacitors, logic gates

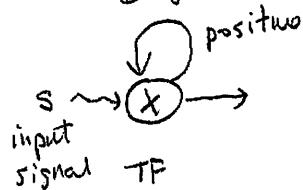
Bottom-up: interactions \rightarrow motifs \rightarrow network \rightarrow function

Top-down:
genes \leftarrow motif \leftarrow network \leftarrow function

Survey of important motifs responsible for cell fate, patterning, development
(bottom-up)

Identifying motifs in a large network (top-down)

Single-gene motifs: auto regulation
Diagram



$$\dot{x} = \beta \Theta(s > k) + \beta x \frac{(x/k_x)^n}{1 + (x/k_x)^n} - \alpha x$$

Function?

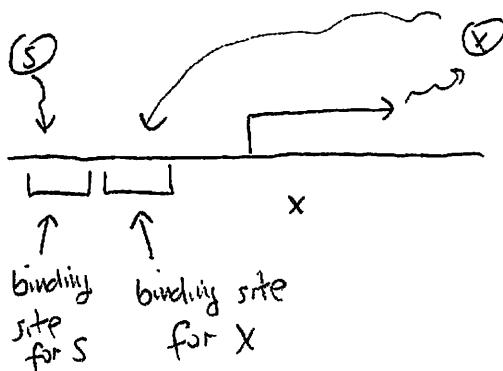
slow response
ultimately slow = memory



$$\dot{x} = \beta [\Theta(s > k) \text{ and } x < k] - \alpha x$$

fast response

Physical mechanism



standard approach: piecewise linear

positive

$$\dot{x} = \beta \Theta(s > k) + \beta_x \frac{(x/k_x)^n}{1 + (x/k_x)^n} - \alpha x$$

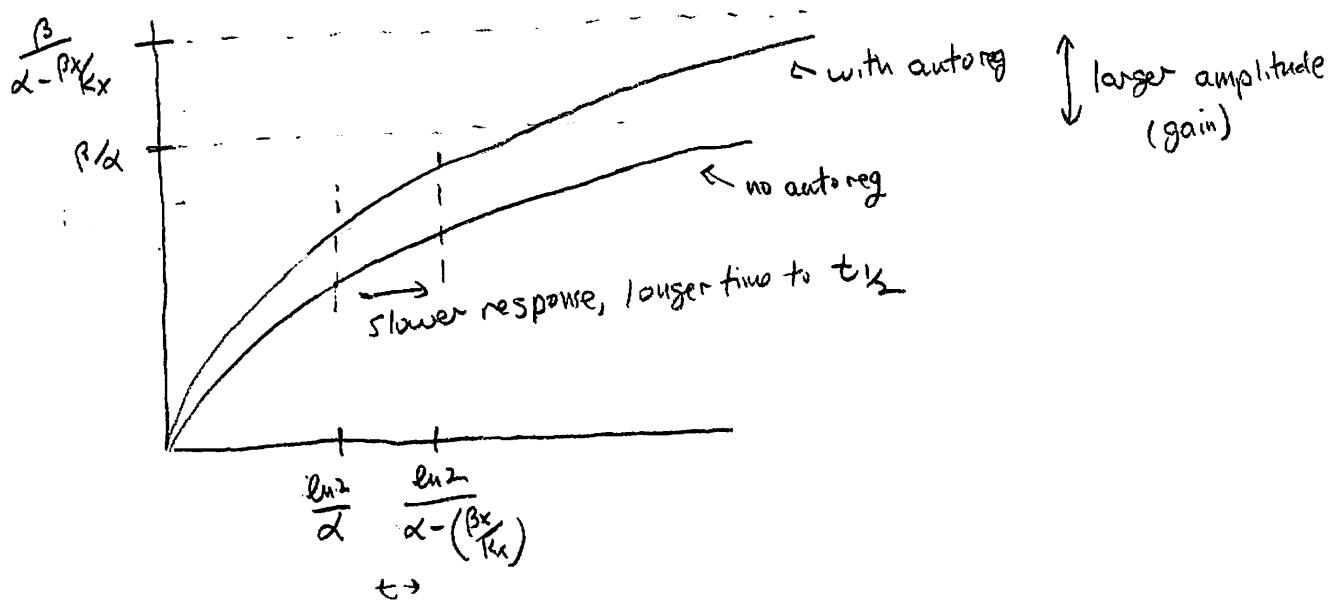
(2)

signal: $S(t) = \begin{cases} 0, & t < 0 \\ \text{large}, & t > 0 \end{cases}$

weak limit: $n=1$ (no cooperativity), $x \ll k_x \Rightarrow \frac{(x/k_x)}{1 + (x/k_x)} \approx \frac{(x/k_x)}{1} \approx \frac{x}{k_x}$

$$\dot{x} = \beta + \frac{\beta_x}{k_x} x - \alpha x = \beta - (\alpha - \frac{\beta_x}{k_x}) x$$

looks just like simple switch, but with slowed response



only valid when $\alpha - \frac{\beta_x}{k_x} > 0$

$$\frac{\beta_x}{\alpha} < k_x$$

↑ steady-state level from self-activation ↑ threshold for self activation

strong limit:

$$\dot{x} = \beta \Theta(s > k) + \beta_x \frac{(x/K_x)^n}{1 + (x/K_x)^n} - \alpha x$$

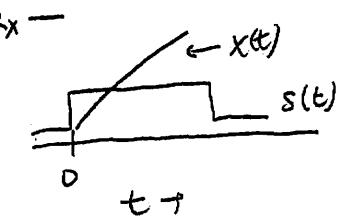
$\frac{\partial x}{\partial s} > K_x$ strong self-activator
 $n \gg 1$ cooperative

$$\dot{x} = \beta \Theta(s > k) + \beta_x \Theta(x > K_x) - \alpha x$$

Signal: square wave, $s(t) = \begin{cases} 0, & t < 0 \\ \text{large}, & 0 < t < T \\ 0, & t > T \end{cases}$

output: depends on whether s can trigger $x > K_x$

when $x < K_x$, $\dot{x} = \beta s(t) - \alpha x$



critical time t^* : $x(t^*) = K_x$

If $T < t^*$, x never makes it above threshold.

$$x(t) = \begin{cases} t < T, & \frac{\beta}{\alpha} (1 - e^{-\alpha t}) \end{cases}$$

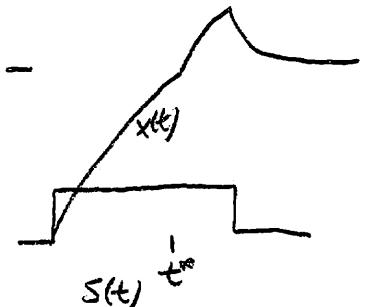
$$x(t) = \begin{cases} t < T, & \frac{\beta}{\alpha} (1 - e^{-\alpha t}) \\ t > T, & \underbrace{\frac{\beta}{\alpha} (1 - e^{-\alpha T})}_{\substack{\uparrow \\ \text{fixed value} \\ \text{when } s \text{ is turned off}}} \cdot \underbrace{e^{-\alpha(t-T)}}_{\substack{\uparrow \\ \text{decay}}} \end{cases}$$



If $T > t^*$

$$x(t) = \begin{cases} t < t^*, & \frac{\beta}{\alpha} (1 - e^{-\alpha t}) \\ t^* < t < T, & \underbrace{K_x e^{-\alpha t}}_{\substack{\uparrow \\ \text{initial cond.} \\ \text{with decay}}} + \underbrace{\left(\frac{\beta + \beta_x}{\alpha}\right) (1 - e^{-\alpha t})}_{\substack{\uparrow \\ \text{final steady state transient}}} \\ t > T, & \left(\frac{\beta + \beta_x}{\alpha}\right) e^{-\alpha t} + \frac{\beta_x}{\alpha} e^{-\alpha t} \end{cases}$$

with $x(t^*) \approx K_x$
previous HW problem



(4)

Negative auto-reg:

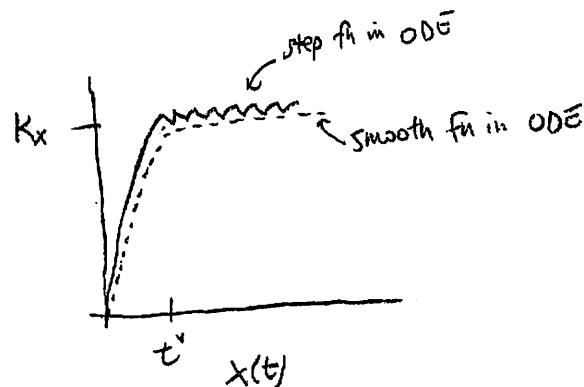
$$\dot{x} = \beta \Theta(s > K \text{ and } x < K_x) - \alpha x$$

$$s(t) = \text{Step fn} = \begin{cases} 0, & t < 0 \\ \text{large}, & t > 0 \end{cases}$$

$$\text{critical time } t^*: x(t^*) = K$$

$$\text{for } t < t^*, \dot{x}(t) = \beta - \alpha x$$

$$x(t) = \frac{\beta}{\alpha} (1 - e^{-\alpha t})$$



$$\text{for } t > t^*, \dot{x} = -\alpha x$$

but then x drops below K_x ,

turns on production

ODE model chokes because \dot{x} is discontinuous at $x = K$

what is $t_{1/2}$?

$$\frac{\beta}{\alpha} (1 - e^{-\alpha t}) = \frac{K}{2}$$

$$1 - e^{-\alpha t} = \frac{K/2}{\beta/\alpha}$$

$$e^{-\alpha t} = 1 - \frac{K/2}{\beta/\alpha}$$

$$t_{1/2} = -\frac{1}{\alpha} \ln \left(1 - \frac{K/2}{\beta/\alpha} \right) \quad t^* = -\frac{1}{\alpha} \ln \left(1 - \frac{K}{\beta/\alpha} \right)$$

↓ tight binding, $K_x \ll \beta/\alpha$, $\ln(1-\epsilon) \approx -\epsilon$

$$t_{1/2} \approx -\frac{1}{\alpha} \cdot -\frac{K/2}{\beta/\alpha} = \frac{1}{2} \frac{K_x}{\beta}$$

finally - response speed scales with production rate!

$$t_{1/2} \approx \frac{\text{conc}}{(\text{conc/time})} \quad \text{correct units}$$

5

Better approach: linearize Hill fn close to $x = Kx$

$$\dot{x} = \beta \frac{1}{1+(x/K)^n} - \alpha x$$

$$\dot{x} \approx \beta \left[\frac{1}{2} - \frac{n}{4K} (x-K) \right] - \alpha \underbrace{x}_{K+(x-K)}$$

$$\dot{x} \approx \left[\frac{\beta}{2} - \alpha K \right] - \left[\frac{\beta n}{4K} + \alpha \right] (x-K)$$

$$\frac{1}{1+(x/K)^n} \approx \frac{1}{1+(\frac{K}{K})^n} + (x-K) \frac{d}{dx} \left(\frac{1}{1+(x/K)^n} \right)$$

$$= \frac{1}{2} + (x-K) \cdot (-1) \frac{1}{(1+(\frac{x}{K})^n)^2} \cdot \frac{nx^{n-1}}{K^n}$$

$$= \frac{1}{2} + (x-K) \cdot \frac{1}{4} \cdot \frac{n}{K}$$

$$= \frac{1}{2} - \frac{n}{4K} (x-K)$$

steady state: $x-K = \frac{\frac{\beta}{2} - \alpha K}{\frac{\beta n}{4K} + \alpha} \xrightarrow{\alpha \gg K} \frac{\beta/2}{\beta n/4K} = \frac{2K}{n}, x \approx K + \frac{2K}{n} \approx K$

response rate = $\frac{\beta n}{4K} + \alpha \gg \alpha$

proportional control: $\dot{x} \propto \text{const.} (x-K)$

Chapter 10

Combinatorial Regulation

Lecture 17. Combinatorial Regulation

①

Objectives

- gloves with Cathy Jancuk
- secret ingredients: rum, nutmeg
- Information to regulate a gene
- Information as # of bp
- How TFs read DNA
- Solutions: multi-TF, multi-domain
- Back to Feed-forward Loops (FFLs)

Information to regulate a gene: how much?

$$S = \text{entropy} = k \ln \frac{\Omega}{\Omega_{\text{after}}}$$

↑ ↑
 convenient # of possibilities
 constant

Information = loss of entropy

or
information content (IC)

$$IC = S_{\text{before}} - S_{\text{after}} = k \ln \left[\frac{\Omega_{\text{before}}}{\Omega_{\text{after}}} \right]$$

Regulating a gene \Leftrightarrow positioning RNAp along DNA

$$\Omega_{\text{before}} = \# \text{ of } \text{bps} \text{ in DNA} = 3 \times 10^9 \text{ (or } 6 \times 10^9 \text{ for W vs C, top vs bottom strand)}$$

$$\left. \begin{aligned} IC_{\text{in}} &= \log_2 \frac{\Omega_{\text{before}}}{\Omega_{\text{after}}} \\ \log_2 \Omega &= \log_2 e^{k \ln \Omega} = [k \ln \Omega] + \ln_2 e \\ \text{this choice: } k &= \frac{1}{\ln_2 e} \end{aligned} \right\}$$

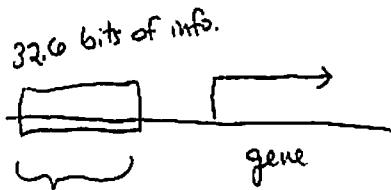
$$\Omega_{\text{after}} = 1$$

$$IC = \log_2 \left[\frac{\Omega_{\text{before}} = 6 \times 10^9}{\Omega_{\text{after}} = 1} \right] = \log_2(1) + \log_2(3) + 9 \log_2 10 = 9 \times \frac{10}{3} = 30$$

IC = 32.6 bits of information to choose a gene (starting point)

(2)

DNA must encode this information using a (unique) regulatory tag/code



IC encoded by nucleotide sequence # of segs of width w

$$IC(\text{sequence of length } w) = \log_2 \left[\frac{s_{\text{before}}}{s_{\text{after}}} \right] = \log_2 \left[\frac{4^w}{1} \right] = w \log_2 4 = 2w$$

↗ sequence
 selected by evolution

General Rule:

$$IC(\text{position}) = IC(\text{regulatory code})$$

$$\downarrow \qquad \downarrow$$

$$\log_2 [\text{genome in bp}] = w \log_2 4$$

$$w = \frac{\log_2 [\text{genome in bp}]}{\log_2 [4]} = \frac{1}{2} \log_2 [\text{genome in bp}]$$

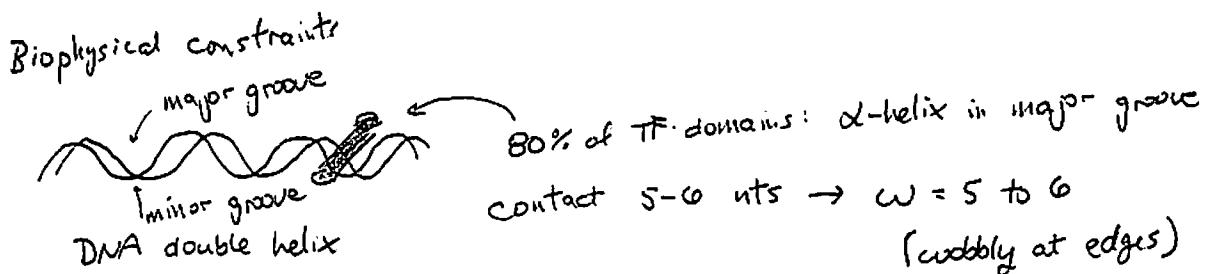
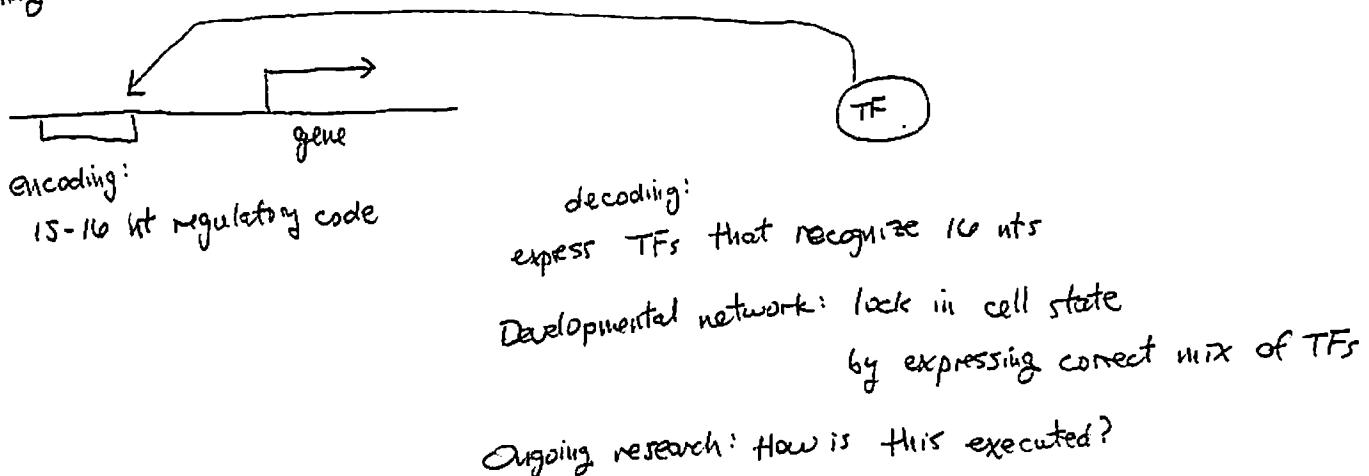
<u>organism</u>	<u>genome</u>	<u>$\log_2 \text{Genome}$</u>	<u>w</u>
H1N1	13 kb	$\underbrace{\log_2(13)}_{3.7} + \underbrace{3 \log_2 10}_{10} \approx 14$	7

$E. coli$	5 Mb	$\underbrace{\log_2(5)}_{2.3} + \underbrace{6 \log_2 10}_{20} \approx 22.3$	11-12
-----------	------	-----------------------------------------------------------------------------	-------

Yeast	12 Mb	$\underbrace{\log_2(4 \cdot 3)}_{2+1.6} + 6 \log_2 10 \approx 23.6$	12-13
-------	-------	---------------------------------------------------------------------	-------

Human	3 E9	$\underbrace{\log_2(3)}_{1.5} + \underbrace{9 \log_2 10}_{30} \approx 31.5$	15-16
-------	------	-----------------------------------------------------------------------------	-------

Summary



Problem: 5-6 nt recognition per TF domain
15-16 nt IC

Solution: multidomain recognition

Solution!
multiple TFs combini
3-4 TFs \rightarrow 15-20 nt recognition

Suppose: a complete library of all 5-bp TFs

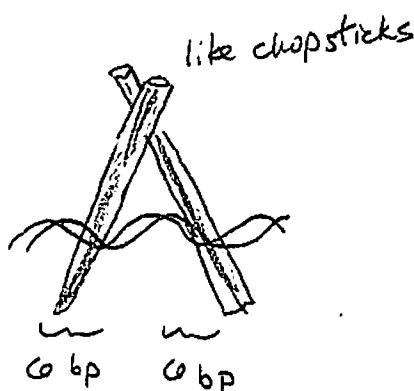
Then: can select subsets of 4 TFs to cover all 20 nt codes

Feasible?

$$\text{Need } 4^5 \text{ proteins} = 2^{10} = 1024$$

Human: 10% TFs, ~ 2000

See Newman Keating 2003 Science
bZIP (basic leucine Zippers)



④

Solution 2: multi-domain TF



Zinc finger TFs: 3 α -helix domains + linkers

15-18 nts

benefit: tighter binding (used for DNA surgery)

drawback: need many more

Human genome:

\sim 2000 TFs of which

\sim 1000 Zn fingers

Hierarchical combinatorial regulation

Chapter 11

Non-Linear Cascades and Logic Gates

Lecture 16

①

$\rightarrow \overset{\leftarrow}{X}$ negative feedback

$$\dot{X} = \beta_0 \Theta(X < K) - \alpha X$$

$\rightarrow \overset{\leftarrow}{X}$ positive feedback

$$\dot{X} = \beta_0 + \beta_1 \frac{(X/K)^n}{1+(X/K)^n} - \alpha X$$

$X \ll K$, with $n \approx 1$

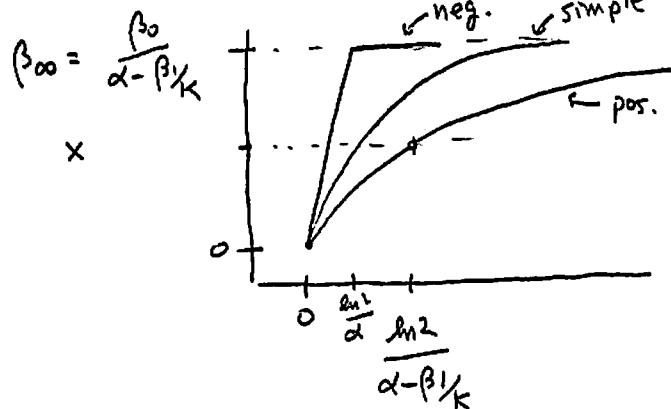
$X \gg K$

$$\dot{X} = \beta_0 + \beta_1 \frac{X}{K} - \alpha X$$

$$\dot{X} = \beta_0 + \beta_1 - \alpha X$$

$$\dot{X} = \beta_0 - (\alpha - \frac{\beta_1}{K}) X$$

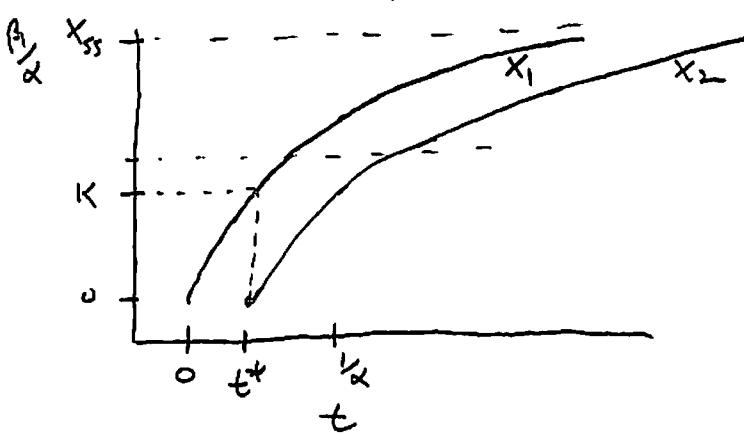
$$\frac{\beta_1}{K} > \alpha$$



$$\dot{X}_1 = \beta_1 u(t) - \alpha X_1(t)$$

$$\dot{X}_2 = \beta_2 \Theta(X_1 > K) - \alpha X_2$$

$u(t-t^*)$



method: walk through in time
key points: t^* , t_K

$$X_1(t^*) = \frac{x_{ss}}{\alpha} (1 - e^{-\alpha t^*}) = K$$

conc: x_{ss}

$$1 - e^{-\alpha t^*} = \frac{K}{x_{ss}}$$

$$1 - \frac{K}{x_{ss}} = e^{-\alpha t^*}$$

$$t^* = -\frac{1}{\alpha} \ln \left[1 - \frac{K}{x_{ss}} \right]$$

low threshold $K \ll x_{ss}$ high threshold $K \approx x_{ss}/2$

$$t^* \approx \frac{1}{\alpha} \cdot \frac{K}{\beta/\alpha} \approx \frac{K}{\beta}$$

$$t^* \approx \frac{\ln 2}{\alpha}$$

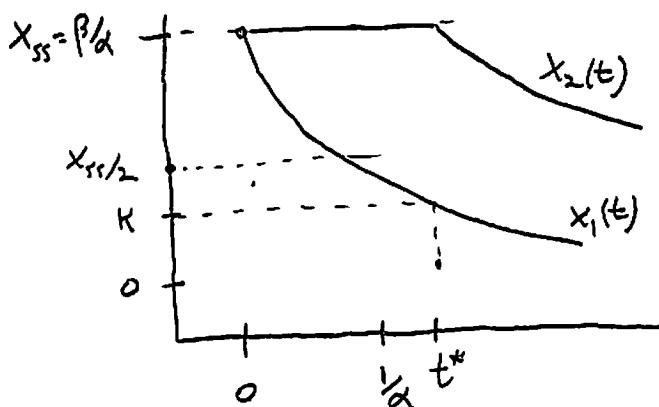
(2)

off step:

$$\dot{x}_1 = -\alpha x_1$$

$$\dot{x}_2 = \underbrace{\beta \Theta(x_1 > K) - \alpha x_2}_{\dot{x}_2 = 0 \text{ until } x_1 > K} \quad \text{with } x_2(0) = x_{ss} = \frac{\beta}{\alpha}$$

$$\dot{x}_2 = -\alpha x_2 u(t - t^*)$$



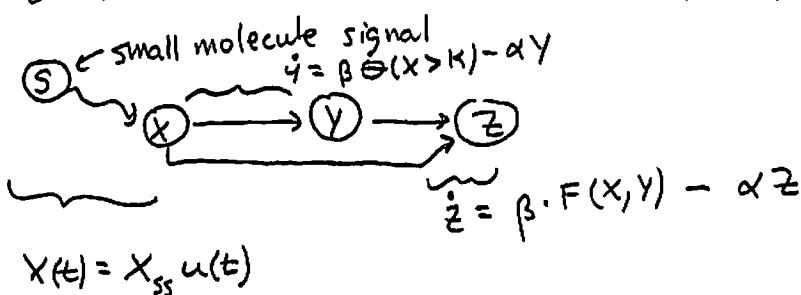
$$x_1(t^*) = x_{ss} e^{-\alpha t^*} = K$$

$$t^* = \frac{1}{\alpha} \ln\left(\frac{x_{ss}}{K}\right)$$



on step: $x(t) = x_{ss} u(t)$, $x(0) = y(0) = z(0) = 0$
 $z \text{ delay} \approx \frac{K}{\beta} \text{ to } \frac{1}{\alpha}$

3-TF circuits

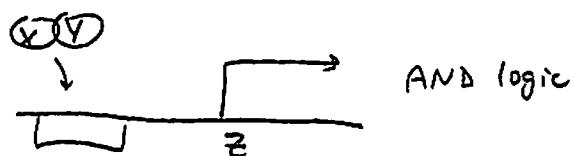
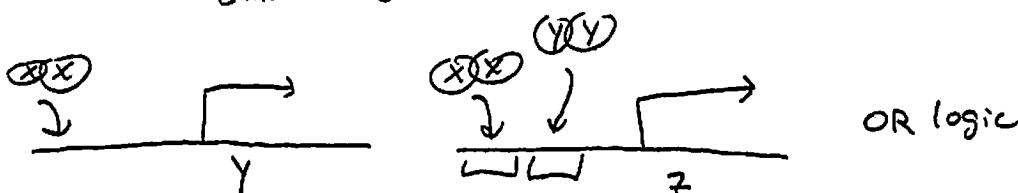


off step: $x(t) = 0$, $y(0) = z(0) = \beta/\alpha$

$$z \text{ delay} \approx \frac{1}{\alpha} \ln\left(\frac{ss}{K}\right)$$

Example: leucine zipper TFs

bind as dimers



AND logic:

on step:

$$\dot{y} = \beta - \alpha y \quad y(t) = Y_{ss} (1 - e^{-\alpha t})$$

off step:

$$\dot{y} = -\alpha y$$

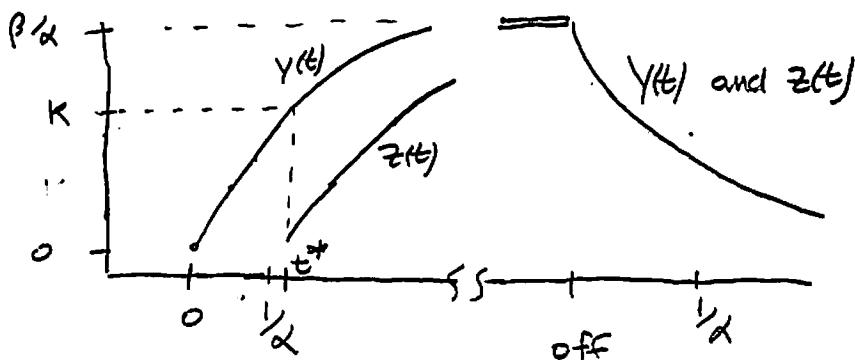
$$\dot{z} = \beta \Theta(y > K) - \alpha z$$

$$\dot{z} = -\alpha z$$

$$\dot{z} = \beta u(t - t^*) - \alpha z$$

no delay

$$\text{delay of } t^* = -\frac{1}{\alpha} \ln \left[1 - \frac{K}{Y_{ss}} \right]$$

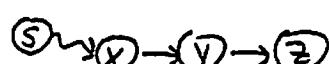


OR logic:

on step: no delay

$$\text{off step: delay of } t^* = \frac{1}{\alpha} \ln \left(\frac{Y_{ss}}{K} \right)$$

Summary:

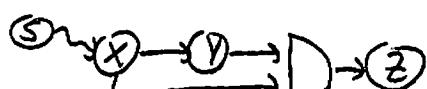


$$\text{delay off } -\frac{1}{\alpha} \ln \left(1 - \frac{K}{Y_{ss}} \right)$$

on

off

$$\text{delay of } \frac{1}{\alpha} \ln \left(\frac{Y_{ss}}{K} \right)$$



delay

no delay



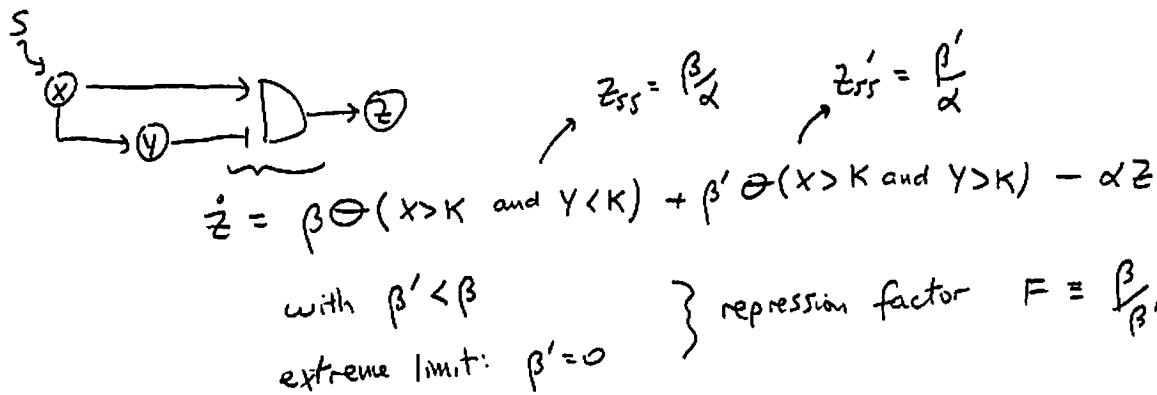
no delay

delay

Lecture 17: Incoherent FFL

Facebook: talk face-to-face once / week. Exam key today

①



standard approach

① divide into piecewise linear regimes, write as ODE

② for each regime, get t_{start} , trend, z_{start} , z_{end}

and $t_{1/2} = \text{time when } z = z_{start} + \frac{1}{2}(z_{end} - z_{start})$

$$x(t) = (\text{large}) \cdot u(t)$$

$$y(0) = 0$$

$$z(0) = 0$$

On step:

$$\text{regime 1: } \begin{array}{ll} x > K & \dot{z} = \beta - \alpha z \\ y < K & \dot{y} = \beta_y - \alpha y \end{array}$$

$$z(t) = \frac{\beta}{\alpha} (1 - e^{-\alpha t}) \quad y(t) = \frac{\beta_y}{\alpha} (1 - e^{-\alpha t})$$

ends at

$$y(t^*) = K = \underbrace{\frac{\beta_y}{\alpha}}_{Y_{ss}} (1 - e^{-\alpha t^*})$$

$$e^{-\alpha t^*} = 1 - \frac{K}{Y_{ss}}$$

$$t^* = -\frac{1}{\alpha} \ln \left(1 - \frac{K}{Y_{ss}} \right)$$

Regime 2:

note: I will not try to trick you about $t^* + \Delta t$ vs t
for offsets between regimes

$$y(t^*) = K$$

$$z(t^*) = z_{ss} \left[1 - e^{-\alpha t^*} \right] \quad e^{-\alpha t^*} = 1 - \frac{K}{Y_{ss}}$$

$$z(t^*) = z_{ss} \cdot \frac{K}{Y_{ss}}$$

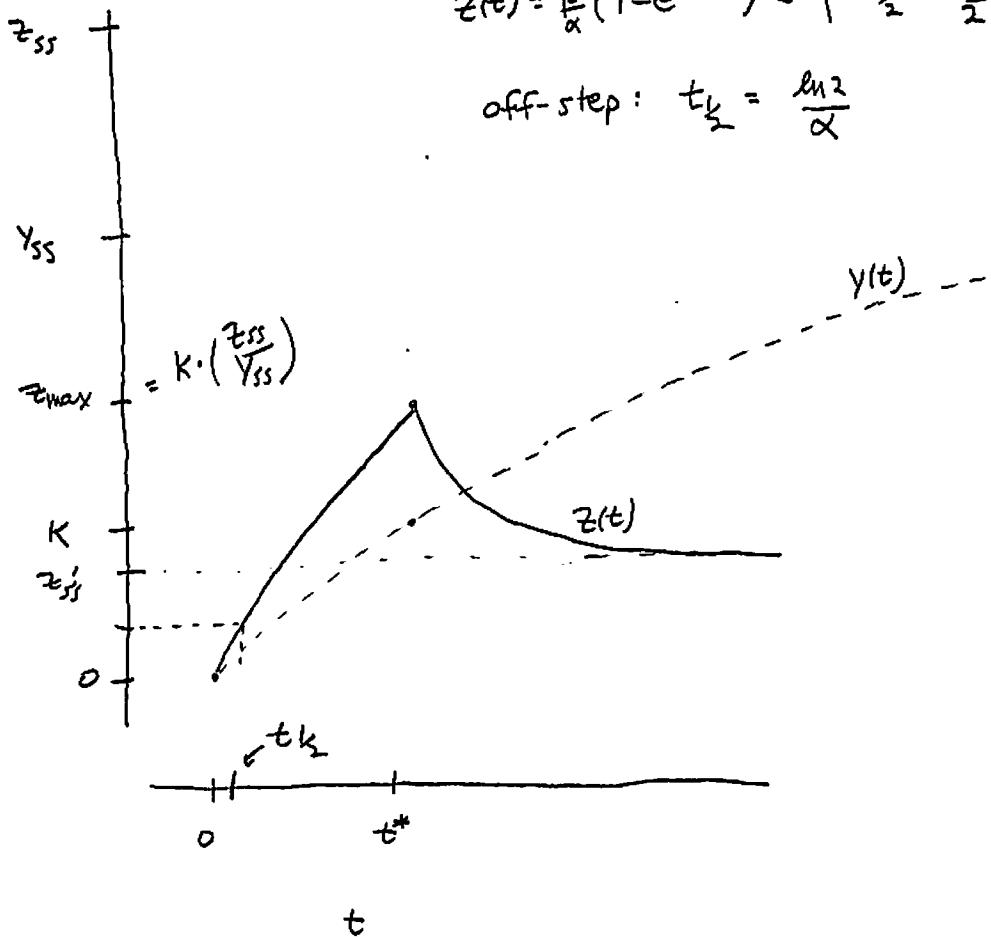
here t is offset from t^*

$$z(t) = \frac{\beta'}{\alpha} + (z(t^*) - \frac{\beta'}{\alpha}) e^{-\alpha(t-t^*)}$$

$$\dot{z}(t+t^*) = \beta' - \alpha z \Rightarrow z(t) = z'_{ss} + (\frac{z_{ss}}{Y_{ss}} K - z'_{ss}) e^{-\alpha(t-t^*)}$$

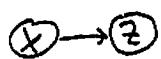
$$t_{1/2} \text{ calculated from long-time } z(t) = z'_{ss} = \frac{\beta'}{\alpha}$$

$$z(t) = \frac{\beta'}{\alpha} (1 - e^{-\alpha t_{1/2}}) \approx \beta t_{1/2} = \frac{1}{2} \frac{\beta'}{\alpha} \Rightarrow t_{1/2} \approx \frac{1}{2\alpha} \left(\frac{\beta'}{\beta} \right)$$



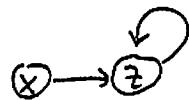
(3)

Big Picture

onoff

$$\frac{\ln 2}{\alpha}$$

$$\frac{\ln 2}{\alpha}$$



slower

slower, $\rightarrow \infty$ for
strong feedback
(bistable switch)



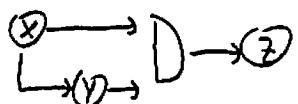
faster

$$\frac{\ln 2}{\alpha}$$



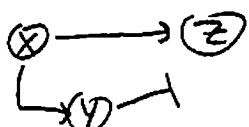
$$\frac{\ln 2}{\alpha}$$

slower



slower

$$\frac{\ln 2}{\alpha}$$



faster

$$\frac{\ln 2}{\alpha}$$

positive \rightarrow slower, bistable
region

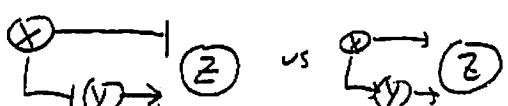
negative \rightarrow faster, overshoot \rightarrow oscillations
region (feedback)

coherent: paths have same sign

incoherent: paths have different sign

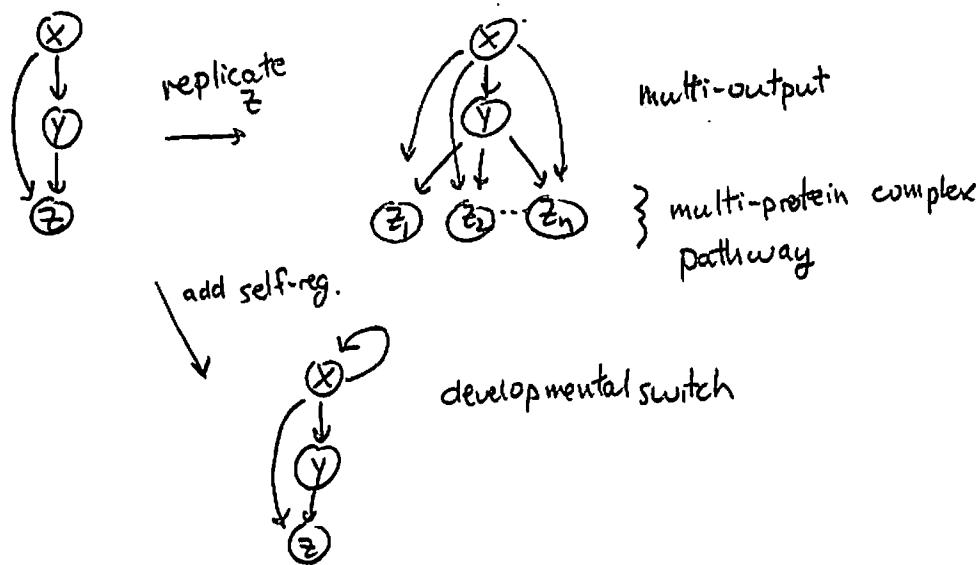
X	0	1	0	1
Y	1	0	0	1
Z	1	0	0	1

steady states

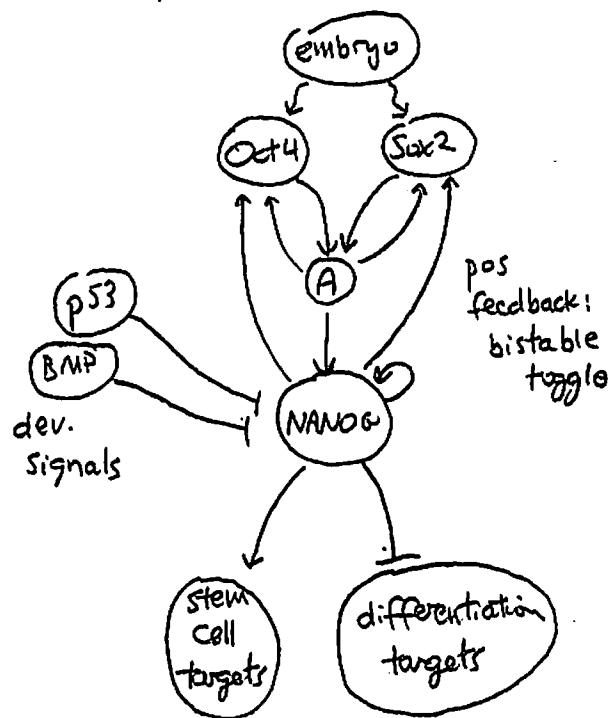


on by default vs off by default

Generalizations



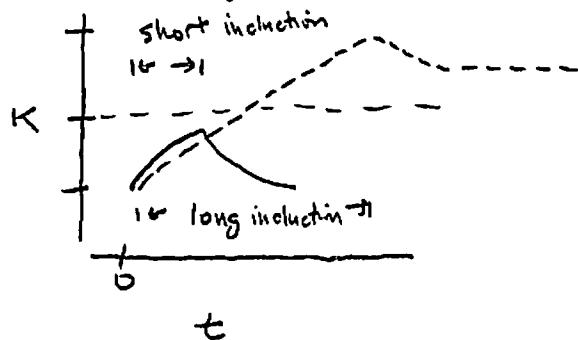
Example: Resetting a stem cell



jump-start:
 $\text{Oct4}, \text{Sox2}, + \text{Lhx28}$ (wise.)
 Nanog

$\text{Oct4}, \text{Sox2}, + \text{klf4}, \text{c-Myc}$ (kyoto)

Express long enough to toggle Nanog



Part III

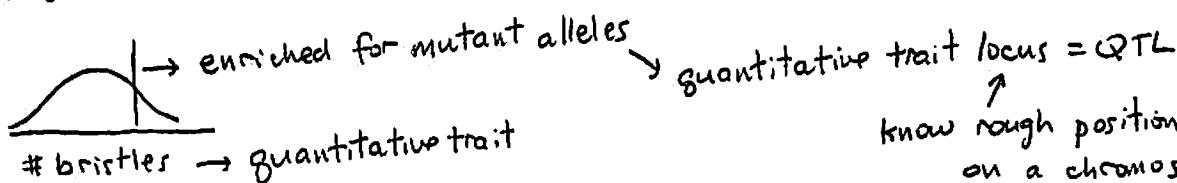
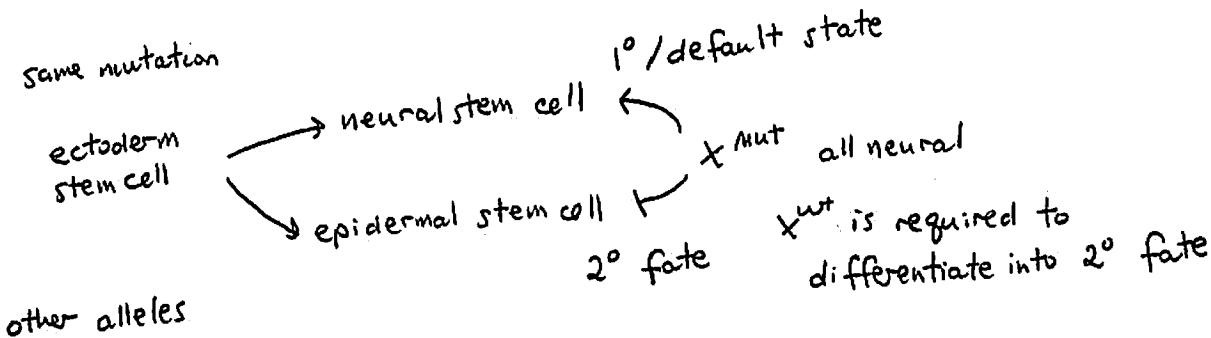
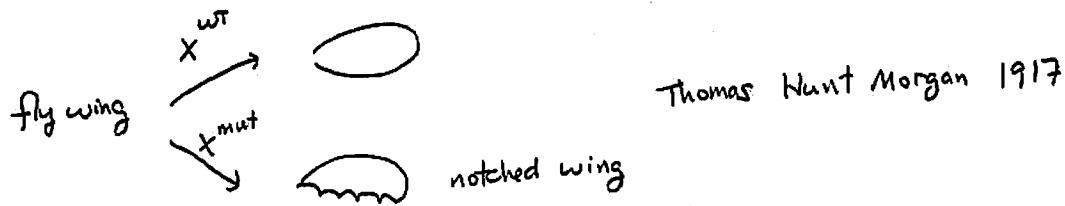
Cells as Stochastic Systems

Chapter 12

Delta-Notch Signaling

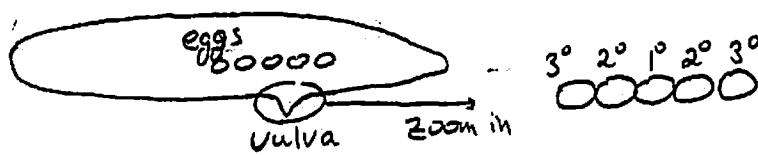
Lecture 20 Delta-Notch Signaling.

①

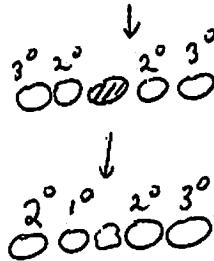


↑
 know rough position
 on a chromosome
 but can cover 10-100 genes

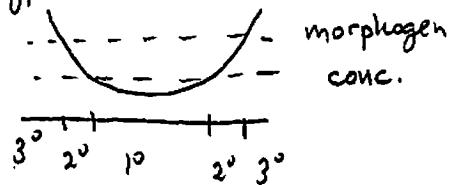
C. elegans (nematode/worm)
 hermaphrodite



experiment:
 ablate 1° cell



3 cell fates
 hypothesis: gradient?



mutant alleles → all 1° → Muv genes



vertebrates

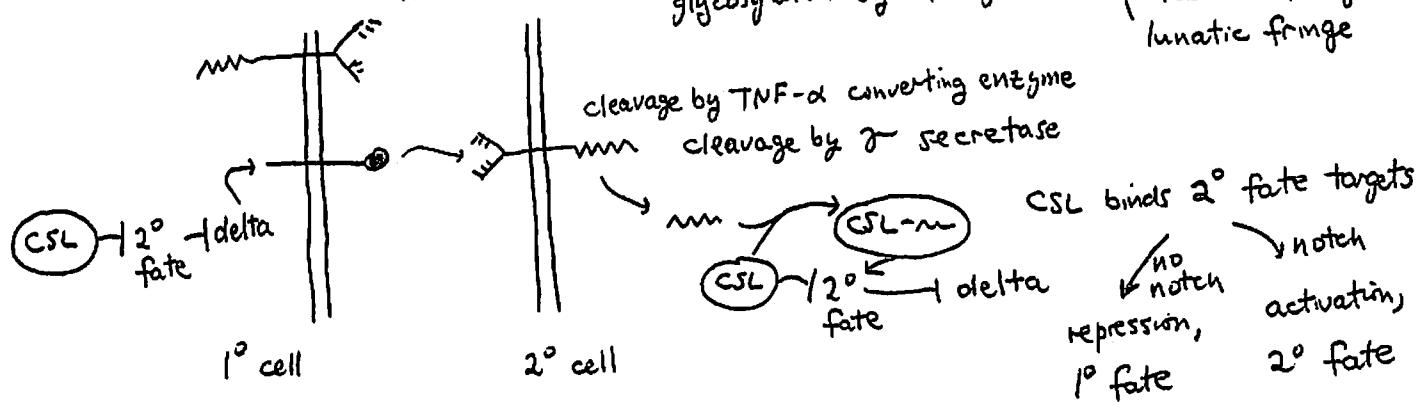
specification of stem cells/progenitor cells
 neural, cardiac, endocrine, bone

developmental disorders
 leukemia
 protein cleavage disorders
 presenilin, Alzheimer's

(2)

Molecular biology: 1980s - 1990s

notch = receptor, expressed on 1° & 2°
 glycosylated by fringe family
 manic fringe
 radical fringe
 lunatic fringe



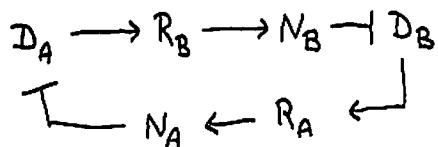
Model: Two cells, A and B (keep separate from 1° 2° for 1° + 2° fate)

Variables $R_{A,B}$ = notch receptor surface concentration $\approx R_{tot}$
 expressed on 1° + 2° cells, high conc.

$D_{A,B}$ = concentration of Delta

$N_{A,B}$ = concentration of cleaved notch

diagram



stable states:

D_A, N_B hi, D_B, N_A low \rightarrow A is 1°

D_A, N_B low, D_B, N_A hi \rightarrow B is 1°

Coupled ODEs



↓

D_B

$$\frac{dN_B}{dt} = \underbrace{k R_B D_A}_{\beta_N} - \alpha_N N_B$$

$$\frac{dN_A}{dt} = \beta_N D_B - \alpha_N N_A$$

$$\frac{dD_B}{dt} = \beta_D \frac{1}{1 + (N_B/K)^n} - \alpha_D D_B$$

$$\frac{dD_A}{dt} = \beta_D \frac{1}{1 + (N_A/K)^n} - \alpha_D D_A$$

(3)

Simplification: protein cleavage much faster than synthesis

N_A, N_B in steady-state compared to D_A, D_B

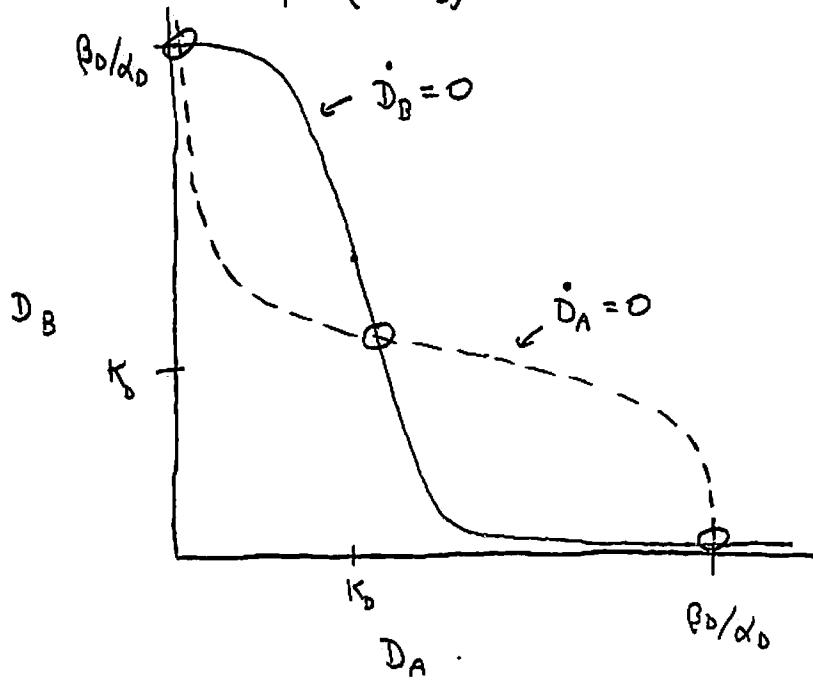
$$\frac{dN_A}{dt} = 0 = \beta_N D_B - \alpha_N N_A \Rightarrow N_A \approx \left(\frac{\beta_N}{\alpha_N}\right) D_B$$

$$\frac{dD_A}{dt} = \beta_D \frac{1}{1 + \left(\frac{\beta_N D_B / \alpha_N}{K}\right)^n} - \alpha_D D_A \quad \text{define } K_D = \frac{K}{(\beta_N / \alpha_N)}$$

$$\dot{D}_A = \beta_D \frac{1}{1 + \left(\frac{D_B / K_D}{K}\right)^n} - \alpha_D D_A$$



$$\dot{D}_B = \beta_D \frac{1}{1 + \left(\frac{D_A / K_D}{K}\right)^n} - \alpha_D D_B$$



nullclines: $\dot{D}_A = 0$
 $\dot{D}_B = 0$

fixed points: $\dot{D}_A = \dot{D}_B$

3 fixed points:

<u>D_A</u>	<u>D_B</u>	<u>Fate</u>
$\approx K$	$\approx K$	undetermined
β/α	0	$1^\circ, 2^\circ$
0	β/α	$2^\circ, 1^\circ$

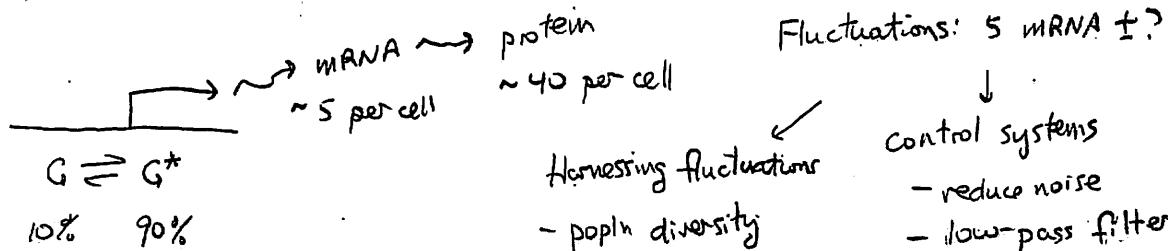
Chapter 13

Stochastic Dynamics

Lecture:

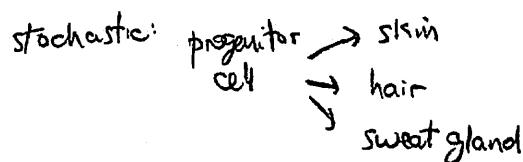
Stochastic Dynamics

Basic switch:



Non-ODE dynamics

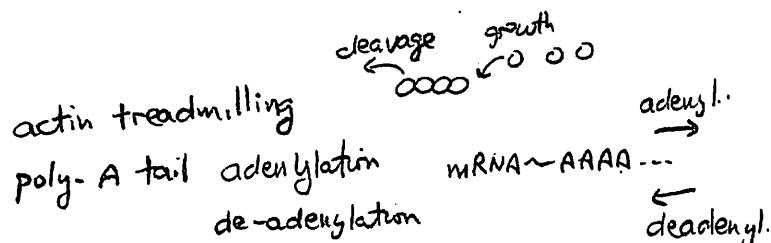
cell fate



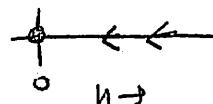
ODE: progenitor cell \rightarrow Frankenstein
each cell part skin, hair, sweat

particle number

$$0 \xrightleftharpoons[\alpha]{\beta} 1 \xrightleftharpoons[\alpha]{\beta} 2 \xrightleftharpoons[\alpha]{\beta} 3 \dots$$



$$\text{ODE: } \frac{dn}{dt} = \beta - \alpha \Rightarrow \text{if } \beta - \alpha < 0, \text{ fixed pt. at } 0$$



stochastic: Call P_0 = prob. of state 0

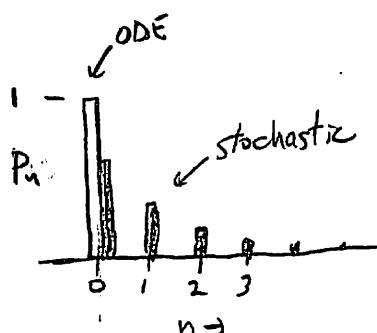
detailed balance:

$$\beta P_0 = \alpha P_1 \Rightarrow P_1 = (\frac{\beta}{\alpha}) P_0$$

$$\beta P_1 = \alpha P_2 \Rightarrow P_2 = (\frac{\beta}{\alpha}) P_1 = (\frac{\beta}{\alpha})^2 P_0$$

$$P_n = (\frac{\beta}{\alpha})^n P_0 = (1 - \frac{\beta}{\alpha}) (\frac{\beta}{\alpha})^n$$

$$1 = \sum_{n=0}^{\infty} P_n = P_0 [1 + (\frac{\beta}{\alpha}) + (\frac{\beta}{\alpha})^2 + \dots] = \frac{P_0}{1 - \frac{\beta}{\alpha}} \Rightarrow P_0 = (1 - \frac{\beta}{\alpha})$$



(2)

Fluctuations at equilibrium

$$\dot{P} = \beta - \alpha P \quad P = \# \text{ protein molecules}$$

$P_n = \underbrace{\text{prob. pf}}_{\text{state with } n \text{ molecules}}$

$$P_0 \xrightleftharpoons[\alpha]{\beta} P_1 \xrightleftharpoons[\alpha]{\beta} P_2 \xrightleftharpoons[\alpha]{\beta} \dots$$

$$\alpha P_1 = \beta P_0 \Rightarrow P_1 = \frac{\beta}{\alpha} P_0$$

$$2\alpha P_2 = \beta P_1 \Rightarrow P_2 = \frac{1}{2} \left(\frac{\beta}{\alpha}\right)^2 P_0$$

$$3\alpha P_3 = \beta P_2 \Rightarrow P_3 = \frac{1}{3!} \left(\frac{\beta}{\alpha}\right)^3 P_0$$

$$P_n = \frac{1}{n!} \left(\frac{\beta}{\alpha}\right)^n P_0 \Rightarrow P_n = \frac{\left(\frac{\beta}{\alpha}\right)^n}{n!} e^{-\frac{\beta}{\alpha}} \quad \text{Poisson}$$

$$1 = \sum_{n=0}^{\infty} P_n = P_0 \underbrace{\sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{\beta}{\alpha}\right)^n}_{e^{\frac{\beta}{\alpha}}}$$

most likely value: $O = \frac{d}{dn} P_n$ or, since P_n monotonic with $\ln P_n$,

$$O = \frac{d}{dn} \ln P_n = \frac{d}{dn} \left[n \ln \left(\frac{\beta}{\alpha} \right) - \frac{\beta}{\alpha} - \ln n! \right] \quad \ln n! \sim n \ln n$$

$$O = \ln \left(\frac{\beta}{\alpha} \right) - \ln n \Rightarrow n = \frac{\beta}{\alpha} \text{ most likely}$$

↓
ODE steady-state

Chapter 14

Noise in Gene and Protein Expression

(3)

Mean, Variance: generating fn method

$$P_n = \text{prob of state } n = \frac{1}{n!} (\beta/\alpha)^n e^{-(\beta/\alpha)}$$

Define discrete version of LT:

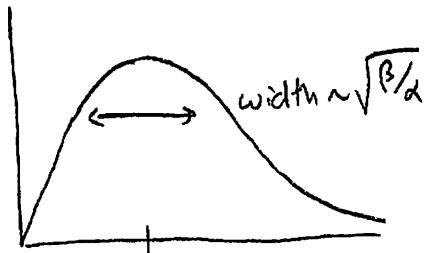
$$\tilde{P}(s) = \sum_{n=0}^{\infty} e^{-sn} P_n$$

$$\text{Then } \langle n \rangle = -\frac{d}{ds} \ln \tilde{P}(s) \Big|_{s=0}$$

$$\text{Var}(n) = \frac{d^2}{ds^2} \ln \tilde{P}(s) \Big|_{s=0}$$

$$\text{AUC} = \tilde{P}(s) \Big|_{s=0} = 1 \quad \text{for normalized distribution}$$

$$1 = \sum_{n=0}^{\infty} P(n)$$



$$\bar{n} = \beta/\alpha$$

$$\begin{aligned} \tilde{P}(s) &= \sum_{n=0}^{\infty} e^{-sn} \frac{1}{n!} (\beta/\alpha)^n e^{-\beta/\alpha} \\ &= e^{-\beta/\alpha} \sum_{n=0}^{\infty} \frac{1}{n!} \left[(\beta/\alpha) e^{-s} \right]^n \\ &= e^{-\beta/\alpha} \exp \left[\frac{\beta}{\alpha} e^{-s} \right] = \exp \left[\frac{\beta}{\alpha} (e^{-s} - 1) \right] \end{aligned}$$

$$\text{AUC} = \tilde{P}(s) \Big|_{s=0} = 1 \quad \checkmark$$

$$\begin{aligned} \langle n \rangle &= -\frac{d}{ds} \ln \tilde{P}(s) = -\frac{d}{ds} \left[\frac{\beta}{\alpha} (e^{-s} - 1) \right] \\ &= \frac{\beta}{\alpha} e^{-s} \stackrel{s=0}{=} \frac{\beta}{\alpha} \quad \text{same as most likely value} \end{aligned}$$

$$\text{Var}(n) = \frac{d^2}{ds^2} \left[\frac{\beta}{\alpha} (e^{-s} - 1) \right] \stackrel{s=0}{=} \frac{\beta}{\alpha} e^{-s} \stackrel{s=0}{=} \frac{\beta}{\alpha}$$

which you know from prob stats

Typical result for independent events:
variance \approx mean

(4)

Time correlations:

How long do fluctuations last?

Duration long enough to trigger a downstream event?

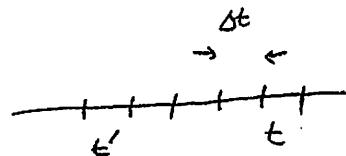
Quantify: Autocorrelation $C(t) = \langle P(t)P(0) \rangle - \langle P(t) \rangle \langle P(0) \rangle$

$$C(t) = \langle P(t+\Delta t)P(t) \rangle - \langle P \rangle^2$$

at equilibrium

$$\text{Already know } C(0) = \langle P^2 \rangle - \langle P \rangle^2 = \beta/\alpha$$

$$\text{Model: } P(t) = \sum_{t' \leq t} \text{Prob(protein made at } t' \text{ and survives to } t)$$



↑
time intervals
width Δt

↓
random variable $\pi(t')$ ← small
with value 1 or 0

$$\text{Prob}(1) = [\beta \Delta t] e^{-\alpha(t-t')} \rightarrow \langle \pi(t') \rangle = \beta \Delta t e^{-\alpha(t-t')}$$

$$\text{Prob}(0) = 1 - \beta \Delta t e^{-\alpha(t-t')}$$

$$\text{so } P(t) = \sum_{t' \leq t} \pi(t')$$

$$\langle P(t) \rangle = \sum_{t' \leq t} \langle \pi(t') \rangle = \sum_{t' \leq t} \beta \Delta t e^{-\alpha(t-t')} \stackrel{\Delta t \rightarrow 0}{=} \int_{-\infty}^t dt' \beta e^{-\alpha(t-t')}$$

$$= \beta e^{-\alpha t} \frac{1}{\alpha} e^{\frac{\alpha t'}{\alpha}} \int_{-\infty}^t = \frac{\beta}{\alpha} \quad \checkmark$$

$$C(t) = \langle P(t)P(0) \rangle - \langle P \rangle^2 \quad (5)$$

$$= \left\langle \sum_{t_1 \leq t} \pi(t_1) \sum_{t_2 \leq 0} \pi(t_2) \right\rangle - \left\langle \sum_{t_1 \leq t} \pi(t_1) \right\rangle \left\langle \sum_{t_2 \leq 0} \pi(t_2) \right\rangle$$

$$= \sum_{t_1=t_2 \leq 0} \langle \pi(t_1)^2 \rangle - \langle \pi(t_1) \rangle^2 + \sum_{t_1 \neq t_2} \underbrace{\langle \pi(t_1) \pi(t_2) \rangle}_{\substack{\text{events are independent}}} - \langle \pi(t_1) \rangle \langle \pi(t_2) \rangle$$

$$= \sum_{t_1=t_2 \leq 0} \langle \pi(t_1) \rangle - \langle \pi(t_1) \rangle^2 \quad \langle \pi(t_1) \pi(t_2) \rangle = \langle \pi(t_1) \rangle \langle \pi(t_2) \rangle$$

$$= \int_{-\infty}^0 dt' \beta e^{-\alpha(t-t')} \left[1 - \underbrace{\beta \Delta t e^{-\alpha(t-t')}}_{\substack{\text{no contribution} \\ \text{as } \Delta t \rightarrow 0}} \right]$$

$$= \beta e^{-\alpha t} \frac{1}{\alpha} e^{\alpha t'} \Big|_{-\infty}^0 = \frac{\beta}{\alpha} e^{-\alpha t}$$

$$C(t) = C(0) e^{-\alpha t}$$

↑ same as the response fn! $P(t) = \beta(t) - \alpha P(t)$

$$\tilde{P}(s) = \frac{1}{s+\alpha} \tilde{P}(s)$$

Fluctuation-Dissipation Theorem:

$$\frac{P(t) - P(0)}{P(t \rightarrow \infty) - P(0)} = \frac{C(t)}{C(t)}$$

↑ decay from a prepared state ↑ decay of equilibrium fluctuations

Chapter 15

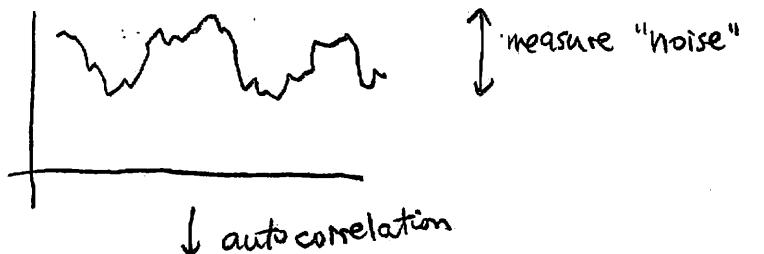
Stochastic Simulations and the Gillespie Algorithm

Experimental investigations

(6)

expression of GFP, YFP proteins

single cell:



↓ autocorrelation

$$C(t) = \frac{1}{T} \int_0^T dt' x(t') x(t'+t) - \left[\frac{1}{T} \int_0^T dt x(t) \right]^2$$

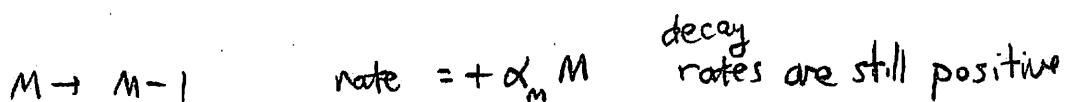
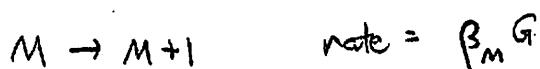
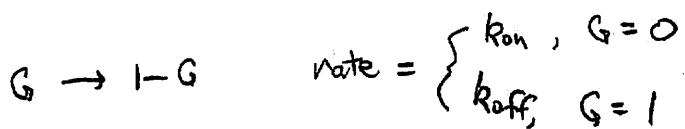
Simulation: Gillespie Algorithm

Developed for nuclear weapons research stochastic nuclear decay

Problem: Discrete system, state (vector) = $\begin{pmatrix} G \\ M \\ P \end{pmatrix}$

- # active promoters
- # mRNA
- # proteins

Reaction channels



When is the next event?

$$\sum = k_{tot}$$

As random variable,

$$\text{total rate to a new state} = k_{tot} = \sum_{\text{channels}} k_{\text{channel}}$$

what is the prob. that the system has no transition in time t ?

$$\text{Prob (no transition)} = e^{-k_{\text{tot}} t}$$

$$\Downarrow \quad \text{Prob}(\text{transition between } 0+t) = 1 - \text{Prob}(\text{no transition}) = 1 - e^{-k_{\text{off}} t}$$

$$\Downarrow$$

$$\text{Prob (transition at } t) = \frac{d}{dt} \text{Prob (transition betw. } 0 \text{ to } t) = \frac{1}{k_{tot} t} e^{-k_{tot} t}$$

\downarrow

prob. density

\downarrow

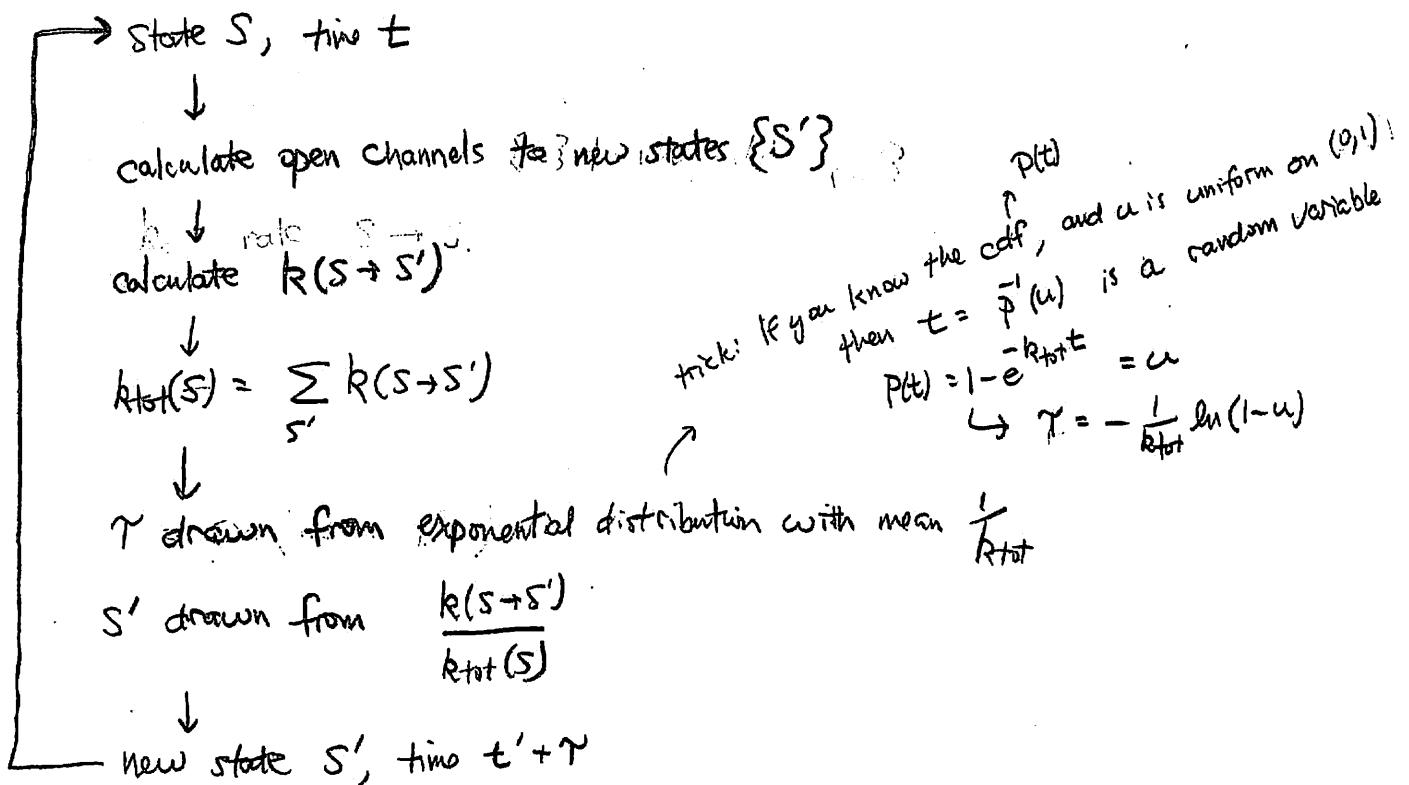
cum. prob. distribution

$$\text{prob. density} \quad \text{cum. prob. distribution}$$

Average waiting time = $\frac{1}{k_{\text{tot}}}$ which new state? $\text{Prob}(S') = \frac{k(S \rightarrow S')}{\sum_{S'} k(S \rightarrow S')} \leftarrow b_{\text{tot}}(S)$

Fancier: select γ from exponential distribution, $P_\lambda(\gamma) = \frac{1}{\bar{\tau}_{tot}} e^{-\bar{\tau}_{tot}\gamma}$

Algorithm:



Correlation Function by FT

$$C(\tau) = \frac{1}{T} \int_{-T/2}^{T/2} dt_1 x(t_1) x(t_1 + \tau)$$

$$x(t) = \int_{-\frac{\pi}{\Delta t}}^{\frac{\pi}{\Delta t}} d\omega \frac{e^{-i\omega t}}{2\pi} \hat{x}(\omega)$$

$$\begin{aligned} C(\tau) &= \frac{1}{T} \int_{-T/2}^{T/2} dt_1 \int_{-\frac{\pi}{\Delta t}}^{\frac{\pi}{\Delta t}} d\omega \frac{e^{i\omega t_1}}{2\pi} \hat{x}(\omega) x(t_1 + \tau) \cdot e^{i\omega \tau} e^{-i\omega \tau} \\ &= \frac{1}{T} \int_{-\frac{\pi}{\Delta t}}^{\frac{\pi}{\Delta t}} d\omega \frac{e^{-i\omega \tau}}{2\pi} \hat{x}(\omega) \int_{-T/2}^{T/2} dt_1 e^{-i(-\omega)(t_1 + \tau)} x(t_1 + \tau) \end{aligned}$$

$$C(\tau) = \frac{1}{T} \int_{-\frac{\pi}{\Delta t}}^{\frac{\pi}{\Delta t}} d\omega \frac{e^{-i\omega \tau}}{2\pi} \hat{x}(\omega) \hat{x}(-\omega)$$

why? Computational efficiency, $n \ln n$ vs n^2 . filtering.
 ↓ ↑
 FT direct

Why not? Tricky to set up, dealiasing, wrap-around.

Notation

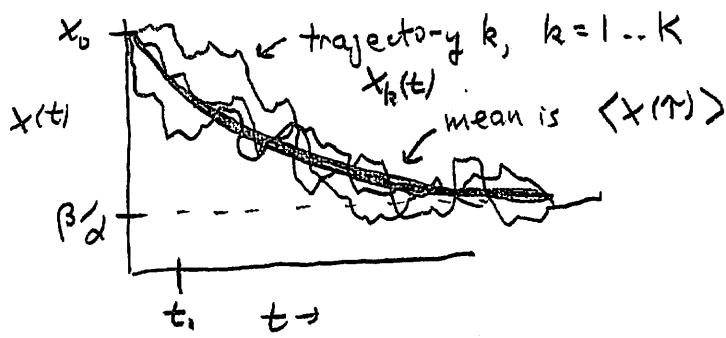
$$C(\tau) = \frac{1}{T} \int_0^T dt_1 X(t_1) X(t_1 + \tau) - \left[\frac{1}{T} \int_0^T dt_1 X(t_1) \right]^2$$

$$\langle x(t_1) x(t_2) \rangle = C(|t_1 - t_2|)$$

Non-equilibrium: $X(0)$ fixed at x_0 , breaks time independence.

$\langle x(t) \rangle$ means $\langle x(t) \rangle$ given $x(0) = x_0$, or $\langle x(t) \rangle_{x(0)=x_0}$

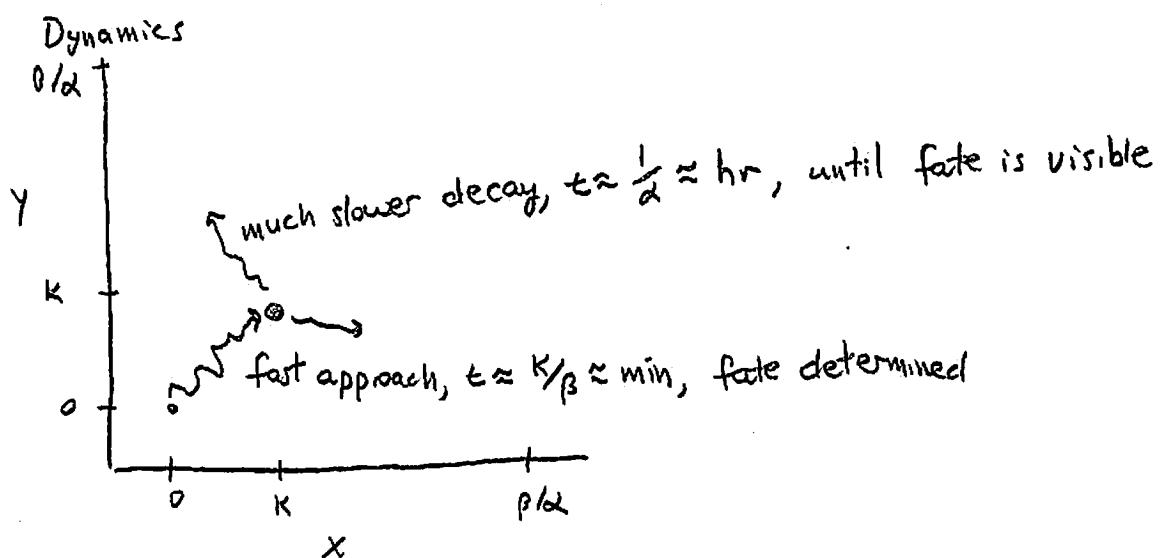
obtain $\langle x(t) \rangle_{x(0)=x_0}$ as an average over trajectories, not over time.



$$\left\langle x(t) \right\rangle = -\pi \sum_{k=-\infty}^{\infty} x_k(t)$$

$$\langle x(0) \rangle|_{x(0)=x_0} = x_0$$

$$\langle x(t) \rangle_{x(0)=x_0} \rightarrow \frac{\beta}{\alpha} \quad \text{as } t \rightarrow \infty$$

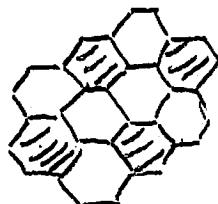


stable states

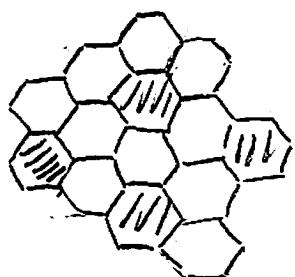
1D: ... $\frac{1}{2}$ filled

... $\frac{1}{3}$ filled

2D:



$\frac{1}{3}$ filled



$\frac{1}{4}$ filled

Chapter 16

Stability Analysis

(4)

Stability analysis

$$\begin{pmatrix} \dot{D}_A \\ \dot{D}_B \end{pmatrix} = \begin{pmatrix} F(D_A, D_B) \\ F(D_B, D_A) \end{pmatrix} \stackrel{\text{Taylor series}}{\approx} \begin{pmatrix} F(D_A^*, D_B^*) \\ F(D_B^*, D_A^*) \end{pmatrix} + \underbrace{\begin{pmatrix} \frac{\partial}{\partial D_A} F(D_A, D_B) & \frac{\partial}{\partial D_B} F(D_A, D_B) \\ \frac{\partial}{\partial D_B} F(D_B, D_A) & \frac{\partial}{\partial D_A} F(D_B, D_A) \end{pmatrix}}_{\text{called the Jacobian evaluated at each fixed pt.}} \begin{pmatrix} D_A - D_A^* \\ D_B - D_B^* \end{pmatrix}$$

\downarrow

$$F(x, y) = \beta \frac{1}{1 + (y/k)^n} - \alpha x$$

$$F(x, y) = 0 = \beta \frac{1}{1 + (y/k)^n} - \alpha x \Rightarrow \frac{1}{1 + (y^*/k)^n} = \frac{\alpha}{\beta} x^*$$

$$\frac{\partial F(x, y)}{\partial x} = -\alpha \quad \frac{\partial F(x, y)}{\partial y} = -\frac{\beta \cdot n y^{n-1}/k^n}{(1 + (y/k)^n)^2} = -\frac{\beta n y^{n-1}}{k^n} \frac{\alpha^2}{\beta^2} x^{*2}$$

$$\frac{\partial F(x, y)}{\partial y} = -n \alpha \frac{y^{n-1} x^{*2}}{k^n (\beta/\alpha)}$$

Quick reminder about stability:

$$\dot{x} = Mx \Rightarrow x(t) = e^{Mt} x(0)$$

matrix $M \rightarrow$ diagonalize, $M = UDU^+$ with $D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$

$$e^{Mt} = 1 + Mt + \sum_{i=1}^{\infty} t^i \frac{M^i}{i!} + \dots \quad U^+ U = UU^+ = I$$

$$= 1 + t \cdot UDU^+ + \frac{t^2}{2} \underbrace{UDU^+}_{1} UDU^+ + \dots$$

$$= U \left[1 + tD + \frac{t^2}{2} D^2 + \dots \right] U^+ = U e^{tD} U^+$$

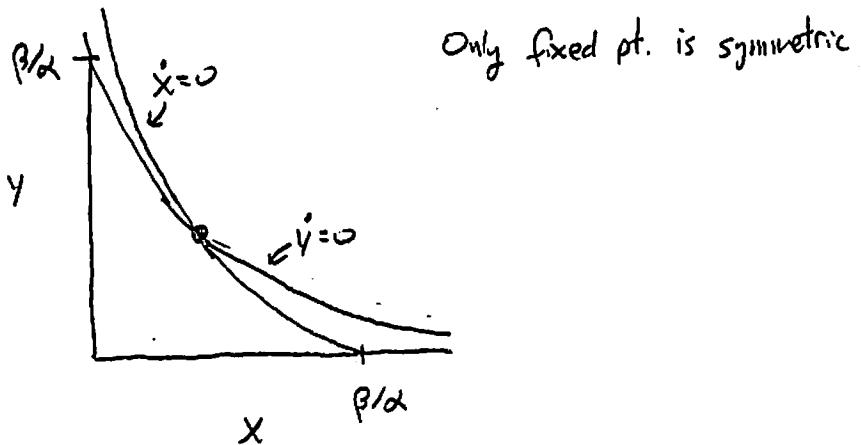
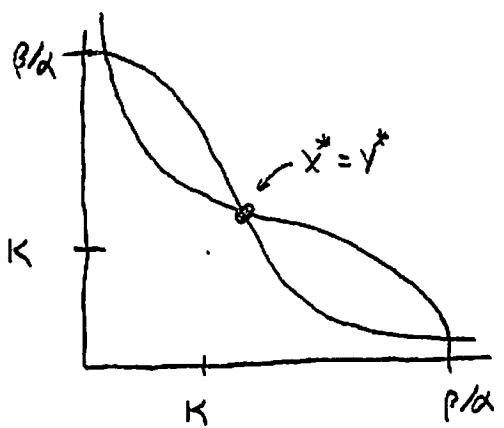
Grows if any $\operatorname{Re}(\lambda) > 0 \rightarrow$ unstable

Decays if all $\operatorname{Re}(\lambda) < 0 \rightarrow$ stable

(SA)

Stability analysis

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} -\alpha & -n\alpha \frac{y^{n-1} x^n}{K^n S} \\ -n\alpha \frac{x^{n-1} y^n}{K^n S} & -\alpha \end{pmatrix} \begin{pmatrix} \delta x \\ \delta y \end{pmatrix}$$

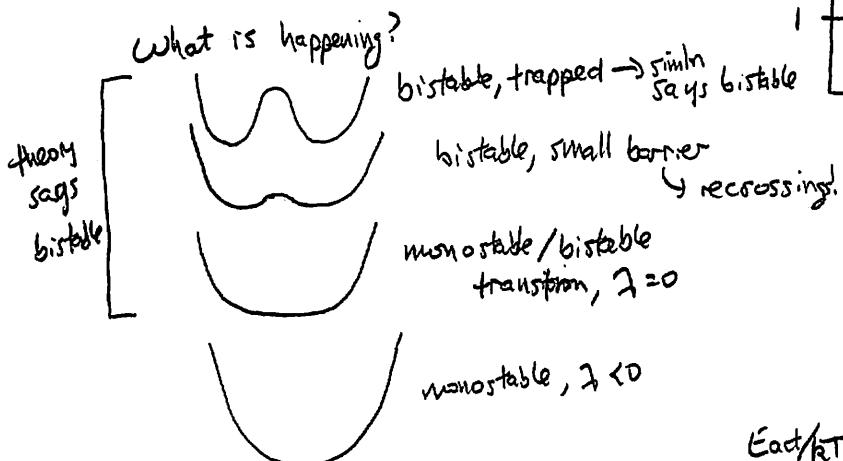
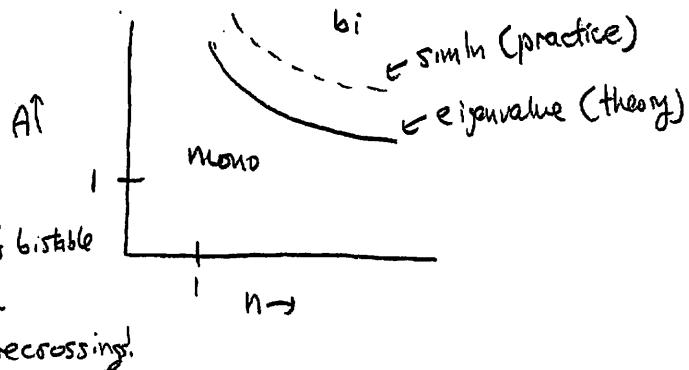
with x^*, y^* = fixed pt S = steady-state max = $\frac{\beta}{\alpha}$ First, some intuition. For $n=1$, nullclines only intersect once:Second, for $n>1$, multiple intersections.At the intersection, $x^* = y^*$ and both equal K as $n \rightarrow \infty$.At finite n , however, if $y^* = K$,then $\dot{x} = 0$ gives $\frac{\beta}{1 + (\frac{y^*}{K})^n} - \alpha x^* = 0$ or $x^* = \frac{S}{n}$ instead of $x^* = K$.Setting $x^* = y^* = K$ is good when $n \rightarrow \infty$
or when $K \approx S/2$

but may be bad otherwise.

we expect $K \leq x^* < S/2$ in general. Even if $(\frac{x^*}{K}) \approx 1$,
 $(\frac{x^*}{K})^n$ will amplify the error.

Bistability

Difference between theory + practice
↓
simln

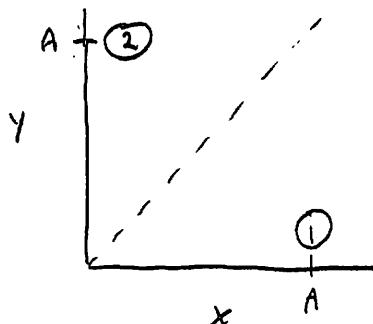


Remember chemical kinetics, rate $\propto e^{-kt}$

Let's calculate!

$$\dot{X} = \frac{A}{1+Y^n} - X$$

$$\dot{Y} = \frac{A}{1+X^n} - Y$$



Imagine we start in state 1.

Calculate rate of $1 \rightarrow 2$ transitions, or state lifetime = 1/kat
(state) > organism lifetime \Rightarrow bistable

(state) < organism lifetime \Rightarrow mono stable.

Prepare in state ①. Decay is $\frac{d}{dt} H(t) = -k \overset{\text{prob. state 1, } X > Y}{\underset{\text{prob. state 2, } Y > X}{\begin{cases} \text{if } t \\ \text{if } t \end{cases}}} \Rightarrow H(t) = e^{-kt}$
 $\frac{d}{dt} H(t) = k e^{-kt} \approx k$ for short t.

Use fluctuation-dissipation, $H(t)$ for prepared state $\approx \langle H(0)H(t) \rangle - \underbrace{\langle H \rangle^2}_0$ for equil.

$$k = -\frac{d}{dt} \langle H(0)H(t) \rangle = -\frac{d}{dt} \langle H(-t)H(0) \rangle = \langle \dot{H}(-t)H(0) \rangle = \langle \dot{H}(0)H(t) \rangle$$

(2)

$\langle \dot{H}(0) H(t) \rangle$ contribution when $x=y$ just before $t=0$
 $x \rightarrow x+1$ just after $t=0$
 $x > y$ for later t .

$$\langle \dot{H}(0) H(t) \rangle \approx \langle |\dot{H}(0)| \rangle_{x=y} \cdot \frac{\Pr(x=y)}{\Pr(\text{all states})}$$

$\hookrightarrow 1$

$$\langle |\dot{H}(0)| \rangle \approx \frac{1}{T_{\text{cell div.}}}$$

Guess for $\Pr(x,y)$: $\Pr(x,y) \approx \Pr(x|y) \Pr(y|x)$

$$\Pr(x|y) = \text{Poisson}\left(\frac{A}{1+y^n}\right)$$

$$\Pr(y|x) = \text{Poisson}\left(\frac{A}{1+x^n}\right)$$

$$\Pr(x=y) = \sum_{z=0}^{\infty} \left[\frac{1}{z!} \left(\frac{A}{1+z^n} \right)^z e^{-\left(\frac{A}{1+z^n} \right)} \right]^2$$

Most likely value: $z = \frac{A}{1+z^n}$ not tractable. oh well.

stability

$$\dot{X} = \beta_x \frac{1}{1 + (Y/K)^n} - \alpha X$$

change of variables: $X = \frac{x}{K}$

$$Y = \frac{y}{K}$$

$$\dot{Y} = \beta_y \frac{1}{1 + (X/K)^m} - \alpha Y$$

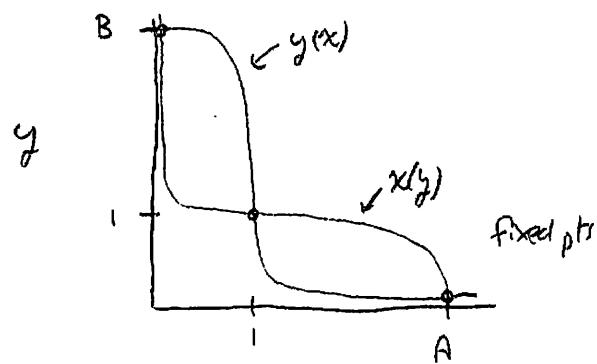
$$T = \alpha t$$

$$\frac{d(Kx)}{dT/\alpha} = \beta_x \frac{1}{1 + y^n} - \alpha Kx$$

$$\frac{dx}{dt} = \overbrace{\left(\frac{\beta_x/\alpha}{K} \right)}^A \frac{1}{1 + y^n} - x$$

↑ unitless ratio
of steady-state
threshold

$$\frac{dy}{T} = \overbrace{\left(\frac{\beta_y/\alpha}{K} \right)}^B \frac{1}{1 + x^m} - y$$



stability:

$$\begin{pmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{pmatrix} = \begin{pmatrix} \dot{x}(x, y) \\ \dot{y}(x, y) \end{pmatrix} \underset{\text{near fixed pt}}{\approx} \begin{pmatrix} \frac{\partial \dot{x}}{\partial x} & \frac{\partial \dot{x}}{\partial y} \\ \frac{\partial \dot{y}}{\partial x} & \frac{\partial \dot{y}}{\partial y} \end{pmatrix} \begin{pmatrix} \delta x \\ \delta y \end{pmatrix}$$

$$\hookrightarrow \begin{pmatrix} \dot{x}(x^*, y^*) + \delta \dot{x} \\ \dot{y}(x^*, y^*) + \delta \dot{y} \end{pmatrix}$$

↓
0 at fixed pt

$$\text{Define } \begin{pmatrix} x(\tau) \\ y(\tau) \end{pmatrix} = \begin{pmatrix} x^* \\ y^* \end{pmatrix} + \underline{r}(\tau)$$

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = M \begin{pmatrix} \delta x \\ \delta y \end{pmatrix} \rightarrow \dot{\underline{r}}(\tau) = M \underline{r}(\tau) \rightarrow \underline{r}(\tau) = e^{M\tau} \underline{r}(0) \quad M = U^* D U$$

$$\underline{r}(\tau) = e^{M\tau} \underline{r}(0) = e^{U^* D U \tau} \underline{r}(0) = U^* e^{D\tau} U \underline{r}(0)$$

stable if $e^{D\tau} \rightarrow 0$

(2)

$$U^T e^{DT} U = U^T \underbrace{e^{\begin{pmatrix} \lambda_1 T & 0 \\ 0 & \lambda_2 T \end{pmatrix}}}_U$$

stable if all $\lambda < 0$ unstable if at least one $\lambda > 0$

simplify $x = \frac{A}{1+y^n} \Rightarrow \frac{A}{(1+y^n)^2} = \frac{x^2}{A}$

$$\dot{x} = A \frac{1}{1+y^n} - x$$

$$\frac{\partial \dot{x}}{\partial x} = -1$$

$$\frac{\partial \dot{x}}{\partial y} = -A \frac{ny^{n-1}}{(1+y^n)^2}$$

$$\frac{\partial \dot{y}}{\partial x} = -B \frac{mx^{m-1}}{(1+x^m)^2}$$

$$\frac{\partial \dot{y}}{\partial y} = -1$$

$$\frac{B}{(1+x^m)^2} = \frac{y^2}{B}$$

$$Mr = 2r$$

$$(M - 2I)r = 0$$

$$O = |M - 2I|$$

$$O = \det(M - 2I) = (-1-\lambda)^2 - \frac{ABmnx^{m-1}y^{n-1}}{(1+x^m)^2(1+y^n)^2}$$

$$M = \begin{pmatrix} -1 & \frac{nx^2y^{n-1}}{A} \\ my^2x^{m-1} & -1 \end{pmatrix}$$

$$(-1-\lambda)^2 = \frac{ABmnx^{m-1}y^{n-1}}{(1+x^m)^2(1+y^n)^2}$$

$$\lambda = -1 \pm \sqrt{\frac{ABmnx^{m-1}y^{n-1}}{(1+x^m)^2(1+y^n)^2}}$$

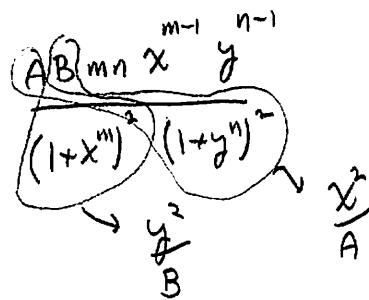
always < 0

$$\text{stable: } -1 + \sqrt{\frac{ABmnx^{m-1}y^{n-1}}{(1+x^m)^2(1+y^n)^2}} < 0$$

$$(-1-\lambda)^2 - \frac{mnx^{m+1}y^{n+1}}{AB} = 0$$

$$\lambda = -1 \pm \sqrt{\frac{mnx^{m+1}y^{n+1}}{AB}}$$

$$\frac{ABmnx^{m-1}y^{n-1}}{(1+x^m)^2(1+y^n)^2} < 1$$



$$= \frac{mn}{AB} x^{m+1} y^{n+1}$$

fixed pt 2: $x, y \approx 1$

unstable (2 states) when

$$\frac{mn}{AB} x^{m+1} y^{n+1} \gtrsim 1$$

$$mn \gtrsim AB$$

large hill coeff \rightarrow unstable

fixed pt 1:

$$x \approx A \quad y \approx \frac{B}{A^m}$$

$$\text{stable when } \frac{mn}{AB} \left(A \right)^{m+1} \left(\frac{B}{A^m} \right)^{n+1} \lesssim 1$$

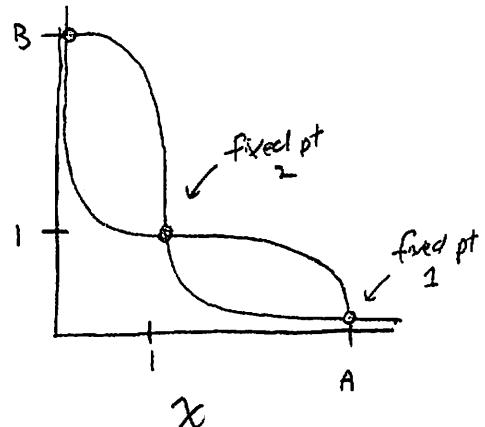
$$mn \frac{A^m B^n}{A^{m(n+1)}} \lesssim 1$$

$$mn \lesssim \frac{A^{mn}}{B^{mn}}$$

large amplitude

$$y = \frac{B}{(1+x^m)}$$

$$x = \frac{A}{1+y^n}$$



for $A=B, m=n$,
want

$$A^2 \lesssim m^2 \lesssim \frac{A^{n^2-n+1}}{A^n}$$

$$A \lesssim n \lesssim A^{\frac{(n^2-n)/2}{n(n-1)/2}}$$

since $A >$ threshold

$$1 < A \lesssim n \lesssim A$$

$n=1: 1 < A \lesssim 1 < A^0$ no solution, ring state

$$n=2: A \lesssim 2 \lesssim A^{2(1)/2} \quad A=2$$

$$n=3: A \lesssim 3 \lesssim A^3 \quad A = 3^{\frac{1}{3}} \rightarrow 3$$

Part IV

Cells as Spatial Systems

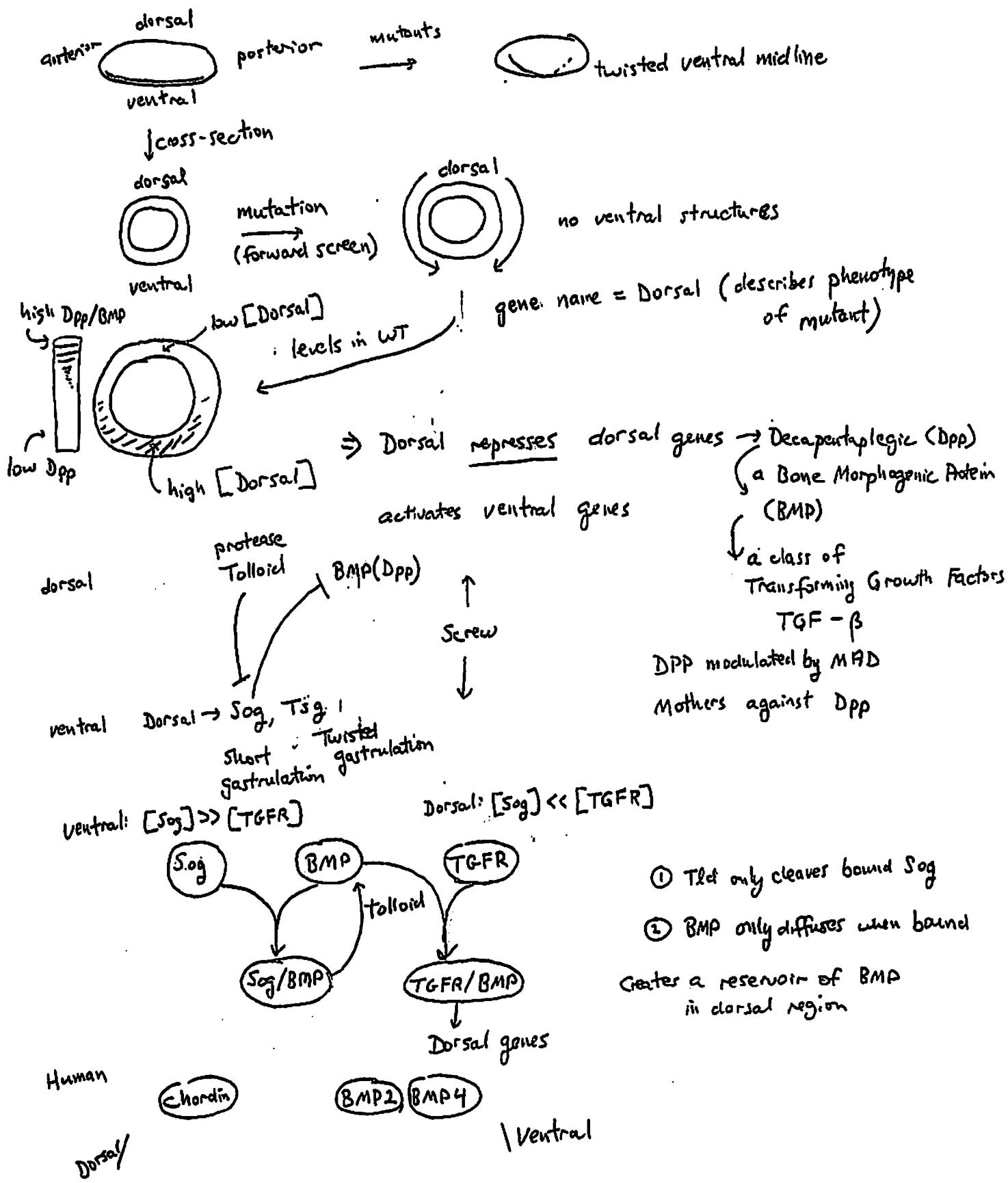
Chapter 17

Morphogen Gradient Patterning

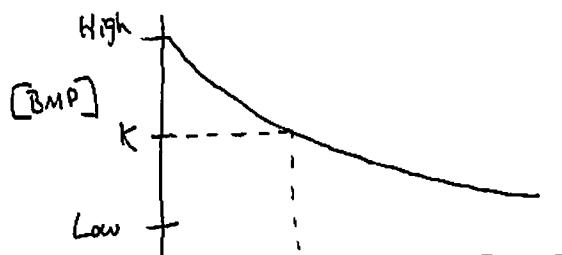
Lecture 19

④

morphogen gradients



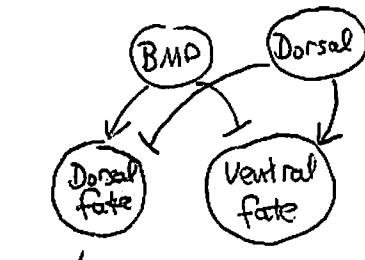
European flag model



distance from top →

| on | off |

✓ multiple morphogens



Imagine an ODE

$$\dot{T} = \frac{\beta \left(\frac{BMP}{K} \right)^n}{1 + \left(\frac{BMP}{K} \right)^n} - \alpha T$$

$BMP > K \rightarrow$ expressed

$BMP < K \rightarrow$ repressed



Bl W R France

Black
Red
Yellow

R
W

Hungary

GW Or Ireland

Black Y R Belgium

Bl Y R Romania

R
W

Austria

Y G Lithuania

G W R Italy

Pur
W Pur

Latvia

Blu
Black Estonia

W

Blue
W
R

Serbia

Chapter 18

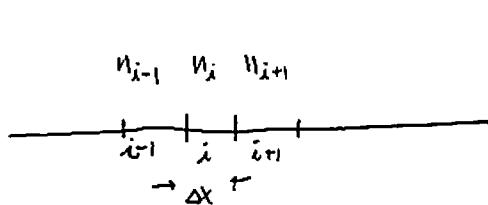
Diffusion

Lecture 21

①

objectives

- PDE model for diffusion
- Morphogen gradients, steady state
- Diffusion to capture, mean first passage time



hopping rate k

time $\Delta t \rightarrow$ probability of hop is $k\Delta t$

prob of hop to right

$$n_i(t+\Delta t) = n_i(t) - (k\Delta t)n_i(t) + (k\Delta t)p_r n_{i-1}(t) + (k\Delta t)p_l n_{i+1}(t)$$

$$\underbrace{\frac{n_i(t+\Delta t) - n_i(t)}{\Delta t}}_{\dot{n}_i(t)} = k \left[p_r n_{i-1}(t) + p_l n_{i+1}(t) - n_i(t) \right]$$

↓ isotropic, $p_r = p_l = \frac{1}{2}$

$$\dot{n}_i(t) = \frac{1}{2} k \left[n_{i-1}(t) + n_{i+1}(t) - 2n_i(t) \right]$$

$$\downarrow n_{i\pm 1} = \Delta x \cdot \rho(x \pm \Delta x)$$

$$\Delta x \cdot \dot{\rho}(x, t) = \frac{k}{2} \cdot \Delta x \left[\underbrace{\rho(x-\Delta x, t) + \rho(x+\Delta x, t)}_{\text{↓ Taylor series}} - 2\rho(x, t) \right]$$

$$\begin{aligned} & \rho(x, t) - \cancel{\nabla \rho(x, t) \cdot \Delta x} + \frac{1}{2} \Delta x^2 \cancel{\nabla^2 \rho(x, t)} \\ & \rho(x, t) + \cancel{\Delta x \nabla \rho(x, t)} + \frac{1}{2} \Delta x^2 \cancel{\nabla^2 \rho(x, t)} - 2\rho(x, t) \end{aligned}$$

$$\begin{aligned} \dot{\rho}(x, t) &= \underbrace{\frac{1}{2} k \Delta x^2}_{D} \nabla^2 \rho(x, t) \\ D &= \frac{1}{2} k \Delta x^2 \end{aligned}$$

Chapter 19

Solving the Diffusion Equation

Suppose $p(x,t) = \delta(x)$ at $t=0$

what is $p(x,t)$ for $t > 0$?

Recall stochastic model: $\langle x \rangle = 0$ $\langle x^2 \rangle = 2Dt$ $\xrightarrow{\text{in limit that # of jumps is large, } k\tau \gg 1}$
 $\xrightarrow{\text{normal (central limit theorem, Stirling's approx)}}$

$$p(x) \underset{\sim}{\approx} \sqrt{\frac{1}{4\pi Dt}} e^{-\frac{x^2}{2Dt}}$$

Now we can do the integral.

$$\dot{p}(x,t) = D \nabla^2 p(x,t)$$

↓ FT

$$\dot{\hat{p}}(k,t) = D \int_{-\infty}^{\infty} dx \underbrace{e^{-ikx}}_{\downarrow \text{integrate by parts}} \nabla^2 p(x,t)$$

↓ integrate by parts

$$\nabla \left[e^{-ikx} \nabla p(x,t) \right] - (-ik) \underbrace{e^{-ikx} \nabla p(x,t)}_{\nabla \left[e^{-ikx} p(x,t) \right]} - \left[-ik e^{-ikx} p(x,t) \right]$$

$$\dot{\hat{p}}(k,t) = -Dk^2 \hat{p}(k,t)$$

↓ LT

$$s \tilde{\hat{p}}(k,s) - \hat{p}(k,0) = -Dk^2 \tilde{\hat{p}}(k,s)$$

$$\hat{p}(k,0) = \int_{-\infty}^{\infty} dx e^{-ikx} \delta(x) = 1$$

$$\tilde{\hat{p}}(k,s) = \frac{1}{s + Dk^2}$$

$$\hat{p}(k,t) = e^{-Dk^2 t}$$

$$p(x,t) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} e^{-Dk^2 t} e^{-ikx} = \frac{1}{2\pi} \sqrt{\frac{\pi}{Dt}} e^{-\frac{x^2}{4Dt}} = \sqrt{\frac{1}{4\pi Dt}} e^{-\frac{x^2}{4Dt}}$$

Morphogen Diffusion
last time: Density. $\dot{\rho}(x,t) = \rho(x,t) + D \nabla^2 \rho(x,t) - \alpha \rho(x,t)$
from the particle viewpoint

①

$$\begin{array}{c} \pm \Delta x \\ \xrightarrow{k t} \quad \# \text{ of hops} \\ x(t) = \sum_{i=1}^{kt} \Delta x_i \xrightarrow{\pm \Delta x} \end{array}$$

hopping rate k
distance Δx

$$D = \frac{k \Delta x^2}{t}$$

$$\langle x(t) \rangle = \left\langle \sum_{i=1}^{kt} \Delta x_i \right\rangle = \sum_i \langle \Delta x_i \rangle = 0 \quad \begin{matrix} \text{symmetric} \\ \text{no drift} \end{matrix}$$

$$\begin{aligned} \langle x(t)^2 \rangle &= \left\langle \sum_{i=1}^{kt} \Delta x_i \sum_{j=1}^{kt} \Delta x_j \right\rangle = \sum_{i=1}^{kt} \sum_{j=1}^{kt} \langle \Delta x_i \Delta x_j \rangle \quad i \neq j: \langle \Delta x_i \rangle \langle \Delta x_j \rangle = 0 \\ &\quad \text{independent} \quad \downarrow \\ &= kt \cdot \Delta x^2 \quad \text{only } j=i \text{ term} \\ &= 2 \cdot \left(\frac{k \Delta x^2}{2} \right) t = 2 D t \quad \text{contributors} \end{aligned}$$

$$\boxed{\langle x^2 \rangle = 2 D t} \quad \text{or in 3 dimensions, } \langle r^2(t) \rangle = \langle x^2(t) \rangle + \langle y^2(t) \rangle + \langle z^2(t) \rangle = 6 D t$$

Some mechanisms for morphogen patterning

① Burst
 $t=0$, no particles released at origin

cells at position x are activated if $\rho(x,t) > K$ for any time t .

What is the activation boundary?



$$\dot{\rho}(x,t) = n_0 \delta(x) \delta(t) + D \nabla^2 \rho(x,t)$$

↓ FT

$$\begin{aligned} \hat{\rho}(k,t) &= n_0 \delta(t) + D \underbrace{\text{FT}(\nabla^2 \rho(x,t))}_{(ik) \text{FT}(\nabla \rho(x,t))} \\ &\quad (ik) \cdot \text{FT}[\rho(x,t)] \end{aligned}$$

$$\dot{\hat{\rho}}(k,t) = n_0 \delta(t) - D k^2 \hat{\rho}(k,t)$$

↓ ET

$$\hat{s} \hat{\rho}(k,s) = n_0 - D k^2 \hat{\rho}(k,s)$$

$$\begin{aligned} \text{FT}\left(\frac{d}{dx} f(x)\right) &= \int_{-\infty}^{\infty} dx e^{-ikx} \frac{d}{dx} f(x) \\ &= \int_{-\infty}^{\infty} dx \frac{d}{dx} \left[e^{-ikx} f(x) \right] - \int_{-\infty}^{\infty} dx \left[\frac{d}{dx} e^{-ikx} \right] f(x) \\ &\quad \underbrace{e^{-ikx} \Big|_{-\infty}^{\infty} = 0}_{\text{at } \pm \infty} \quad \underbrace{i k \text{FT}[f(x)]}_{i k f(k)} \\ &\quad f(k) \rightarrow 0 \end{aligned}$$

$$s\hat{\rho}(k, s) = n_0 - Dk^2 \hat{\rho}(k, s)$$

$$\hat{\rho}(k, s) = \frac{n_0}{s + Dk^2} e^{-Dk^2 t}$$

$$\hat{\rho}(k, t) = \int_{-\infty}^{\infty} \frac{ds}{2\pi i} \frac{n_0}{s + Dk^2} = n_0 e^{-Dk^2 t}$$

$$\rho(x, t) = n_0 \int_{-\infty}^{\infty} dk e^{ikx} e^{-Dk^2 t} = n_0 e^{-\frac{(ix)^2}{4Dt}} e^{-\frac{x^2}{2\cdot 2Dt}}$$

$$\rho(x, t) = \frac{n_0}{2\pi} \cdot \sqrt{\frac{\pi}{Dt}} \cdot e^{-\frac{1}{4Dt}} = \sqrt{\frac{1}{4\pi Dt}} e^{-\frac{(x)^2}{2\cdot 2Dt}}$$

$$\langle x^2(t) \rangle = \int_{-\infty}^{\infty} dx \frac{x^2}{\int_{-\infty}^{\infty} dx \rho(x, t)}$$

note that:

$$\begin{aligned} \frac{d}{dk} \hat{\rho}(k, t) &= \frac{d}{dk} \int dx e^{-ikx} \rho(x, t) \\ &= \int dx (-ix) e^{-ikx} \rho(x, t) \end{aligned}$$

$$\text{so } \left. \frac{d^2}{dk^2} \hat{\rho}(k, t) \right|_{k=0} = (-i) \int dx x^2 \rho(x, t)$$

$$\left. \frac{d^2}{dk^2} \ln \hat{\rho}(k, t) \right|_{k=0} = -\langle x^2(t) \rangle = -\frac{d^2}{dt^2} \underbrace{\ln [n_0 e^{-Dk^2 t}]}_{\ln n_0 - Dk^2 t} = 2Dt$$

$$\text{Again, } \langle x^2(t) \rangle = 2Dt$$

Activation boundary:

$$\rho_{\max}(x) = \underset{t}{\arg \max} \rho(x, t) = \arg \max_t \frac{n_0}{\sqrt{4\pi Dt}} e^{-\frac{x^2}{4Dt}}$$

$$0 = \frac{d}{dt} \rho(x, t) = \frac{1}{\rho(x, t)} \frac{d}{dt} \rho(x, t) = \frac{d}{dt} \ln \rho(x, t) = \frac{d}{dt} \left(-\frac{1}{2} \ln t - \frac{x^2}{4Dt} \right)$$

$$0 = -\frac{1}{2t} + \frac{x^2}{4Dt}$$

$$t = \frac{x^2}{2Dt} \rightarrow \text{same relationship } x^2 = 2Dt$$

$$\int_{-\infty}^{\infty} du e^{-u^2} = \left[\int_{-\infty}^{\infty} du e^{-u^2} \int_{-\infty}^{\infty} dv e^{-v^2} \right]^{\frac{1}{2}}$$

$$= \left[\int_0^{2\pi} d\theta \int_0^{\infty} dr r e^{-r^2} \right]^{\frac{1}{2}}$$

$$= \left[2\pi \int_0^{\infty} \frac{dz}{2} e^{-z^2} \right] = \sqrt{\pi}$$

$$\int_{-\infty}^{\infty} du e^{-au^2} = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} d(\sqrt{a} \cdot u) e^{-(\sqrt{a} \cdot u)^2} = \sqrt{\frac{\pi}{a}}$$

$$\int_{-\infty}^{\infty} du e^{-au^2 + bu} = \int_{-\infty}^{\infty} du e^{-a(u + \frac{b}{2a})^2 + \frac{b^2}{4a}}$$

$$= e^{\frac{b^2}{4a}} \underbrace{\int_{-\infty}^{\infty} du e^{-a(u + \frac{b}{2a})^2}}_{\sqrt{\frac{\pi}{a}}} \quad v = u + \frac{b}{2a}$$

$$\int_{-\infty}^{\infty} du dv e^{-u^2 - v^2} = \left[\int_{-\infty}^{\infty} du e^{-u^2} \int_{-\infty}^{\infty} dv e^{-v^2} \right]^{\frac{1}{2}}$$

$$(u, v) \rightarrow (r, \theta)$$

$$du dv = r dr d\theta \quad u^2 + v^2 = r^2$$

$$z = r^2$$

$$dz = 2r dr$$

$$z = r^2$$

$$dz = 2r dr$$

$$z = r^2$$

$$dz = 2r dr$$

$$f_{\text{max}}(x) = f(x, t^*) = \frac{n_0}{\sqrt{4\pi D t^*}} e^{-\frac{x^2}{4Dt^*}} \quad \text{with } 2Dt^* = x^2$$

$$= \frac{n_0}{\sqrt{2\pi x^2}} e^{-\frac{1}{2}} = \left(\frac{n_0}{x}\right) \frac{e^{-\frac{1}{2}}}{\sqrt{2\pi}}$$

(in the first unit)

At boundary,

$$\left(\frac{n_0}{x}\right) \frac{e^{-\frac{1}{2}}}{\sqrt{2\pi}} = K$$

$$x = \frac{n_0}{K} \frac{1}{\sqrt{2\pi e}} \quad \begin{matrix} \text{variation in production rate} \\ \rightarrow \end{matrix} \quad \frac{dx}{dn_0} = \frac{1}{K \sqrt{2\pi e}} \quad \begin{matrix} \text{constant variation} \\ \text{in boundary.} \end{matrix}$$

steady-state model, constant spatial decay

$$\dot{\rho}(x, t) = \beta_0 \delta(x) + D \nabla^2 \rho(x, t) - \alpha \rho(x, t)$$

$$\dot{\hat{\rho}}(k, t) = \beta_0 - D k^2 \hat{\rho}(k, t) - \alpha \hat{\rho}(k, t)$$

$$s \hat{\rho}(k, s) - \hat{\rho}(k, t=0) = \frac{\beta_0}{s} - D k^2 \hat{\rho}(k, s) - \alpha \hat{\rho}(k, s)$$

$$(s + D k^2 + \alpha) \hat{\rho}(k, s) = \frac{\beta_0}{s}$$

$$\hat{\rho}(k, s) = \frac{\beta_0}{s(s + D k^2 + \alpha)}$$

$$\hat{\rho}(k, t) = \frac{\beta_0}{D k^2 + \alpha} - \frac{\beta_0}{D k^2 + \alpha} e^{-(D k^2 + \alpha)t}$$

→ ignore transient

$$\hat{\rho}(k, t \rightarrow \infty) = \hat{\rho}(k) = \frac{\beta_0}{D k^2 + \alpha}$$

$$\rho(x) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} \frac{e^{ikx}}{D(k + i\frac{\alpha}{D})(k - i\sqrt{\alpha/D})}$$

$$\begin{matrix} x > 0 & \text{upper } \pm \text{ plane} & ix(i\sqrt{\alpha/D}) & -x\sqrt{\alpha/D} \\ Re(k) > 0 & \beta_0 \frac{1}{2\pi} \cdot 2\pi x & \frac{e}{D \cdot 2ix\sqrt{\alpha/D}} & \frac{\beta_0 e}{2\sqrt{\alpha D}} \end{matrix}$$

$$+ \infty \quad -ix\sqrt{\alpha/D}$$

$$\beta_0 \frac{e}{2\sqrt{\alpha D}} \quad \text{units } \sqrt{x^2/k^2} =$$

$$\rho(x) = \frac{\beta_0}{2\sqrt{\alpha D}} e^{-|x|\sqrt{\alpha/D}}$$

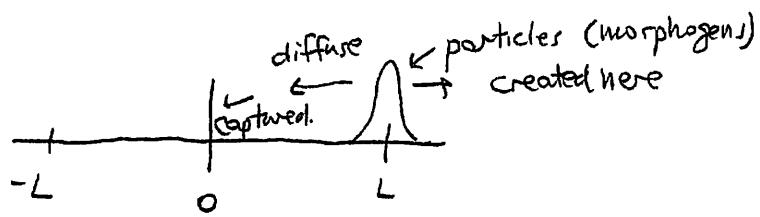
$$\text{At boundary, } \rho(x) = K = \frac{\beta_0}{2\sqrt{\alpha D}} e^{-|x|\sqrt{\alpha/D}}$$

$$|x| = -\sqrt{\frac{D}{\alpha}} \ln \left[\frac{2\sqrt{\alpha D} K}{\beta_0} \right]$$

$$\frac{dx}{d\beta_0} = \frac{d}{d\beta_0} \sqrt{\frac{D}{\alpha}} \ln(\beta_0) = \frac{1}{\beta_0} \sqrt{\frac{D}{\alpha}}$$

Higher β_0 and $\alpha \rightarrow$ less sensitive

Absorbing boundary: method of images.



$$\dot{p}(x,t) = D \frac{\partial^2}{\partial x^2} p(x,t) \quad \text{with} \quad \begin{cases} p(x,0) = \delta(x-L) \\ p(0,t) = 0 \end{cases} \quad \rightarrow \text{if not for this, } p(x,t) = \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{(x-L)^2}{4Dt}}$$

Satisfied by imagining anti-particles emitted at $-L$. \rightarrow method of images.

$$\begin{aligned} p_{\text{tot}}(x,t) &= p_{\text{part}}(x,t) - p_{\text{antipart}}(x,t) \\ &= \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{(x-L)^2}{4Dt}} - \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{(x+L)^2}{4Dt}} \\ &= \frac{1}{\sqrt{4\pi Dt}} \left[e^{-\frac{(x-L)^2}{4Dt}} - e^{-\frac{(x+L)^2}{4Dt}} \right] \end{aligned}$$

For median first passage time, $\int_0^\infty dx p_{\text{tot}}(x,t) = \frac{1}{2}$

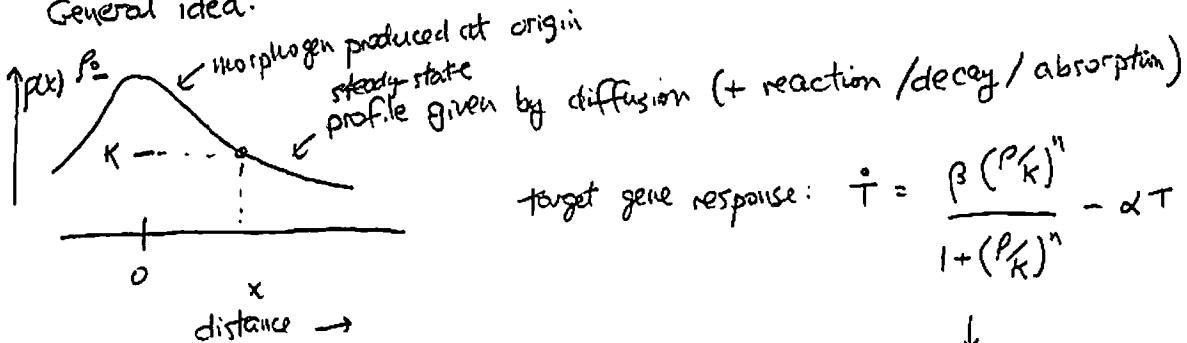
$$\begin{aligned} \frac{1}{2} &= \frac{1}{\sqrt{4\pi Dt}} \int_0^\infty dx e^{-\frac{(x-L)^2}{4Dt}} - e^{-\frac{(x+L)^2}{4Dt}} \\ &= \frac{1}{\sqrt{4\pi Dt}} \left[\int_{-L}^\infty dx e^{-\frac{x^2}{4Dt}} - \int_L^\infty dx e^{-\frac{x^2}{4Dt}} \right] = \frac{1}{\sqrt{4\pi Dt}} \int_{-L}^L dx e^{-\frac{x^2}{4Dt}} \\ (\text{let } z^2 &= \frac{x^2}{2Dt}, \quad z = \frac{x}{\sqrt{2Dt}}, \quad dz = \frac{dx}{\sqrt{2Dt}} \Rightarrow \frac{1}{2} = \frac{1}{\sqrt{2\pi}} \int_{-\frac{L}{\sqrt{2Dt}}}^{\frac{L}{\sqrt{2Dt}}} dz e^{-\frac{z^2}{2}} \right. \end{aligned}$$

Chapter 20

Patterning and Noise

And now for some patterning models.

General idea:



$$\text{target gene response: } \dot{T} = \frac{\beta (\rho/K)^n}{1 + (\rho/K)^n} - \alpha T$$

$$x^* \text{ where } \rho = K$$

$$\begin{array}{ll} x < x^* & x > x^* \\ \text{fate 1} & \text{fate 2} \end{array}$$

$$\text{Sensitivity: } \frac{d \ln x^*}{d \rho_0}$$

\curvearrowleft concen. at origin

Model: Absorbed at distance L

$$\dot{\rho}(x,t) = 0 = D \nabla^2 \rho(x,t) \quad \text{for } x > 0$$

with $\rho(t) = \begin{cases} \rho_0, & x=0 \\ 0, & x=L \end{cases}$

$$\nabla^2 \rho(x,t) = 0$$

irrelevant

$$\downarrow \text{trial form } \rho(x,t) = C_n x^n$$

$$C_n \cdot (n)(n-1) x^{n-2} = 0$$

possibilities:

$$C_n = 0$$

$$n=0 \rightarrow \rho(x,t) = C_0 + C_1 x$$

$$n=1 \downarrow \rho(0)=\rho_0$$

$$\rho(x) = \rho_0 + C_1 x$$

$$\downarrow \rho(L) = 0, C_1 = -\frac{\rho_0}{L}$$

$$\rho(x) = \rho_0 \left[1 - \frac{x}{L} \right]$$

$$\rho(x) = K = \rho_0 \left[1 - \frac{x^*}{L} \right]$$

$$\frac{K}{\rho_0} = 1 - \frac{x^*}{L}$$

$$x^* = L \cdot \left(1 - \frac{K}{\rho_0} \right)$$

$$\frac{d \ln x^*}{d \ln \rho_0} = \rho_0 \frac{d}{d \rho_0} \ln \left(1 - \frac{K}{\rho_0} \right)$$

$$= \rho_0 \cdot \frac{\frac{+K/\rho_0^2}{1-K/\rho_0}}{1-K/\rho_0} = \frac{K/\rho_0}{1-K/\rho_0}$$

Least sensitive for $K \ll \rho_0$.

okay for $\frac{K}{\rho_0} = \frac{1}{2}$, $x^* \approx \frac{L}{2}$

Model 2: Decay everywhere

$$\dot{\rho}(x,t) = D \nabla^2 \rho(x,t) - \alpha \rho(x,t) \quad \text{again } \rho(0) = \rho_0$$

↓ steady state

$$0 = D \nabla^2 \rho(x) - \alpha \rho(x) \quad -x/\lambda \text{ "}" \lambda \text{ for length}$$

$$\downarrow \text{ trial solution } \rho(x) = A e^{-x/\lambda} \rightarrow \rho(0) = \rho_0, A = \rho_0$$

$$0 = D \cdot \rho_0' \frac{1}{\lambda^2} e^{-x/\lambda} - \alpha \rho_0 e^{-x/\lambda}$$

$$\frac{D}{\lambda^2} = \alpha$$

$$\lambda^2 = D/\alpha \quad \rightarrow \text{ think of } \frac{1}{\alpha} \text{ as lifetime.}$$

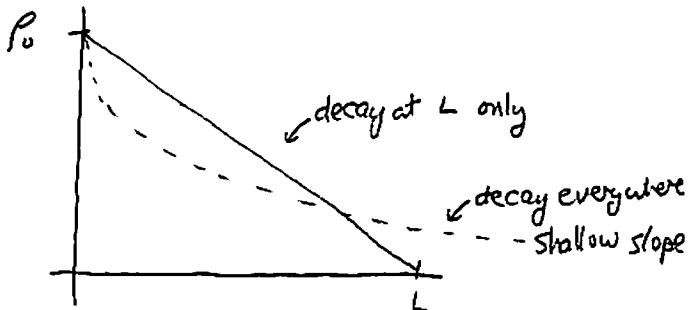
$$\text{mean sq. diffusion length} = R^2(t) = 2Dt = 2D \frac{1}{\alpha} \text{ for particle lifetime!}$$

$$\lambda^2 \approx \text{mean sq. diff. length!}$$

$$\rho(x,t) = \rho_0 e^{-x/\sqrt{D/\alpha}}$$

$$K = \rho(x^*) = \rho_0 e^{-x^*/\sqrt{D/\alpha}}$$

$$x^* = \sqrt{D/\alpha} \cdot \ln(\rho_0/K)$$



sensitivity:

$$\frac{d \ln x^*}{d \ln \rho_0} = \frac{d \ln [\ln \rho_0]}{d \ln \rho_0} = \frac{1}{\ln \rho_0} \quad \rightarrow \begin{array}{l} \text{can make small} \\ \text{when } \rho_0 \text{ is large.} \end{array}$$

x^* is less sensitive

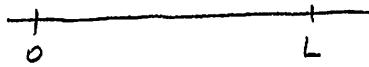
Even better:

$$0 = D \nabla^2 \rho(x) - \alpha [\rho(x)]^n \quad \text{for } n > 1.$$

Transients

Morphogen created here

↙
Time to establish a gradient



Rule of thumb: time to establish gradient ~ time to diffuse across region

$x(t)$ = position of morphogen at $t \rightarrow$ stochastic

$$x(0) = 0$$

Given $x(t)$

$$\langle x(t) \rangle = 0 \quad (\text{symmetry})$$

$$x(t+\Delta t) = \begin{cases} x(t) - \Delta x, & p = \frac{1}{2} \\ x(t) + \Delta x, & p = \frac{1}{2} \end{cases}$$

Suppose $t = n \Delta t = n$ hops

hop every Δt

$$\text{hopping rate} = \frac{1}{\Delta t} = k$$

After n hops,

$$x(t) = \sum_{i=1}^n \Delta x_i \quad \text{where } \Delta x_i = \begin{cases} +\Delta x, & p = \frac{1}{2} \\ -\Delta x, & p = \frac{1}{2} \end{cases}$$

and Δx 's are independent
(iid) identically distributed

$$\langle x(t) \rangle = \sum_i \langle \Delta x_i \rangle = n \langle \Delta x_i \rangle = n \cdot 0 = 0$$

$$\langle \Delta x \rangle = (\pm \Delta x)^2 / 2 = \Delta x^2 / 2$$

$$\begin{aligned} \text{Var}(x(t)) &= \text{Var}\left(\sum_i \Delta x_i\right) = n \text{Var}(\Delta x) = n \left[\langle \Delta x^2 \rangle - \langle \Delta x \rangle^2 \right] \\ &= n \Delta x^2 = \left(\frac{t}{\Delta t}\right) \Delta x^2 \end{aligned}$$

$$\text{Var}(x(t)) = t \cdot \left[\frac{\Delta x^2}{\Delta t}\right]$$

diffusion constant
defined as $2D$

$$\boxed{\text{Var}(x(t)) = 2Dt}$$

→ For diffusion across region, $R^2(t) \approx L^2$

also called $R^2(t)$

$$L^2 = 2Dt$$

$$t = \frac{L^2}{2D}$$

Transients



$$R^2 = 2Dt \text{ is usually a good guess. } D = \frac{1}{2} k \Delta x^2$$

$$\text{water: } k \approx 1/10^{12} \text{ sec}$$

$$t \approx \frac{R^2}{2D} \quad R \approx \text{cell width} \approx 10 \mu\text{m} \\ \approx 10^{-3} \text{ cm}$$

$$\Delta x \approx 1 \text{ \AA} \approx 10^{-8} \text{ cm}$$

$$t \approx \frac{10^{-6} \text{ cm}^2}{2 \cdot 10^{-6} / \text{sec}} \approx 0.5 \text{ sec}$$

about right.

$$D \approx \frac{1}{2} \cdot \frac{10^{-12}}{\text{sec}} \cdot 10^{-16} \text{ cm}^2 = \frac{1}{2} \times 10^{-4} \text{ cm}^2/\text{sec}$$

$$D = 5 \times 10^{-5} \text{ cm}^2/\text{sec} \text{ about right.}$$

For transcription/translation:

need NTP or tRNA to diffuse

$$\text{small molecule: } 10^{-6} \text{ cm}^2/\text{s}$$

$$\text{protein: } 10^{-7} \text{ to } 10^{-8} \text{ cm}^2/\text{s}$$

$$t_{\text{tot}} \approx 100 \left(\frac{R^2}{2D} \right) \approx 100-200 \text{ sec}$$

↑ ↑
 each nt diffusion time
 or tRNA 1-2 sec

$$\text{what is } \sqrt{D/\alpha} ? \quad \frac{1}{\alpha} \approx \text{time to capture} = \frac{R^2}{2D} \quad \text{with } R \approx \text{distance to nearest protease.}$$

$$\text{pattern length} \approx \sqrt{D \cdot \frac{R^2}{2D}} \approx \text{distance to nearest protease}$$

requires low protease concentration \rightarrow large variance in α

$$x^* = \sqrt{D/\alpha} \cdot \ln(s_0/k)$$

$\underbrace{\phantom{x^* = \sqrt{D/\alpha} \cdot \ln(s_0/k)}_{\text{will be } \ln(2) \text{ to } \ln(10)} \sim \text{factor of 1 to 2}}$

\downarrow
diffusion length \approx pattern length

not good.

Part V

Cellular Networks

Chapter 21

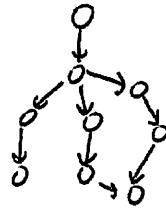
Diffusion on a Network

Network Diffusion

gene/protein network

diffusion / flow from receptor
down multiple paths
to receptor

Master TF \rightarrow subprocess activation



social network

spread of disease
spread of ideas, information

Discrete transition model

step 0: vertex i is active

step 1: vertex i transfers activity to one of its neighbors.

$$\Pr(j \text{ active at step 1} \mid i \text{ active at time } 0) = \frac{\begin{cases} A_{ji} & \text{if } i \sim j \\ 0 & \text{if } i \not\sim j \end{cases}}{d_i} \quad \begin{array}{l} \text{connected to} \\ \text{if } i \sim j \\ \text{not connected} \\ \text{to} \\ d_i \leftarrow \text{degree of } i \end{array}$$

$$P_{ji}(1 \text{ step}) = (A^{-1})_{ji}$$

$$= \sum_j A_{ij} = \sum_j A_{ij} D_i^{-1}$$

step 2: $\Pr(k \text{ active at step 2} \mid i \text{ active at step 0})$

$$= \sum_j P_{kj}(1) P_{ji}(1) = \sum_j A_{kj} D_j^{-1} A_{ji} D_i^{-1} = [A D^{-1} A^{-1}]_{ki}$$

step l : $P_{ji}(l) = [(A D^{-1})^l]_{ji}$ so: $(A D^{-1})^l$ propagates a signal l steps through a network.

Diffusion: think about continuous time random walk

②

$p_j(t)$ = prob. that j is active at time t

with rate $i \leftrightarrow j = A_{ij}$

$$\frac{d}{dt} p_i(t) = \sum_{\text{incoming}} - \sum_{\text{outgoing}} = \sum_j A_{ij} p_j(t) + \left(\sum_j A_{ji} \right) p_i(t)$$

$$\frac{d}{dt} p_j(t) = \sum_i (A_{ij} - D_{ij}) p_j(t)$$

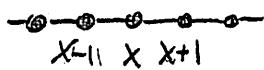
↓ vector $\tilde{p}(t) = \begin{pmatrix} p_1(t) \\ p_2(t) \\ \vdots \\ p_n(t) \end{pmatrix}$

since $D_{ij} = \begin{cases} D_i & \text{for } i=j \\ 0 & \text{for } i \neq j \end{cases}$

$$\frac{d}{dt} \tilde{p}(t) = [A - D] \tilde{p}(t) = - [\underbrace{D - A}_{\text{graph Laplacian}}] \tilde{p}(t)$$

Make this concrete: regular 1D lattice

$\rightarrow \Delta x \leftarrow$



change during one jump

$$\frac{d}{dt} p_x(t) = - \sum_{x'} (D_{xx'} + A_{xx'}) p_{x'}(t) = -2 p_x(t) + p_{x-1}(t) + p_{x+1}(t)$$

when $p_x(t)$ varies smoothly: $p_{x-1}(t) \approx p_x(t) - \frac{\Delta x}{\Delta t} p_x(t) + \frac{(\Delta x)^2}{2} \frac{d^2}{dx^2} p_x(t)$

$$p_{x+1}(t) \approx p_x(t) + \frac{\Delta x}{\Delta t} p_x(t) + \frac{(\Delta x)^2}{2} \frac{d^2}{dx^2} p_x(t)$$

$$\underline{-2 p_x(t) = -2 p_x(t)}$$

$$\frac{d}{dt} p(x,t) \underset{(\Delta t) \frac{d^2}{dx^2}}{\approx} p(x,t) = D \nabla^2 p(x,t)$$

Diffn const. \rightarrow ∇^2 Laplace operator

unfortunate notation:
reuse $D = \begin{cases} \text{degrees matrix} \\ \text{diffn. const.} \end{cases}$

$$\text{D-A} \Leftrightarrow -\nabla^2$$

graph Laplacian negative of Laplace operator

unfortunate sign difference.

Application: signaling.

Receptor is activated,
generates cytokine to maintain concentration M_0 at $x=0$
that is degraded with rate k .

What is concentration profile $M(x)$ at steady-state?

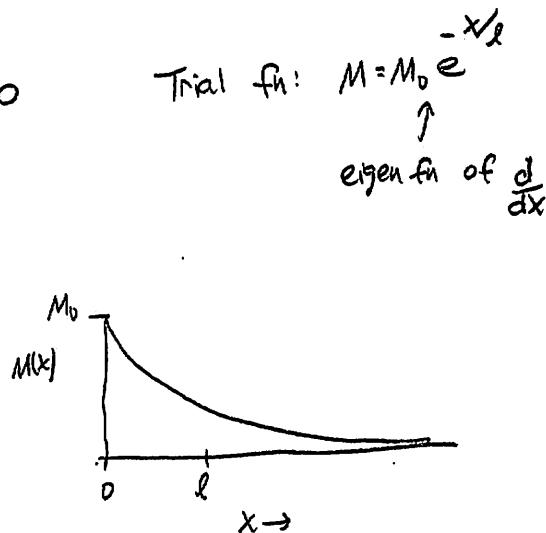
$$\frac{d}{dt} M(x,t) = 0 = D \frac{d^2}{dx^2} M(x,t) - k M(x,t)$$

$$\left[D \frac{d^2}{dx^2} - k \right] M(x) = 0 \quad \text{Trial fn: } M = M_0 e^{-x/l} \quad l \text{ units = length.}$$

$$\left[\frac{D}{l^2} - k \right] M(x) = 0$$

$$l = \sqrt{D/k}$$

$$M(x) = M_0 e^{-x/l}$$



Same idea can be applied to network information diffusion

q_i = rate that signal is injected into node i ($q = \text{query}$)

k = rate that signal is extinguished

what is $\underline{p}_j(t)$ at steady state?

$$\frac{d}{dt} \underline{p}(t) = - [D - A] \underline{p}(t) - k \underline{p}(t) + \underbrace{\underline{q}}_{\text{query vector}} = 0$$

$$[kI + D - A] \underline{p}(t) = \underline{q}$$

$$\underline{p} = [kI + D - A]^{-1} \underline{q} \quad \text{Google PageRank}$$

(4)

what is this doing?

$$\underbrace{p = \left[kI + D - A \right]^{-1} g}_{\downarrow \quad \checkmark}$$

$$p = \underbrace{\left[(1+k)D - A \right]^{-1} g}_{\text{diffusion kernel}}$$

$$= \left[\left[(1+k)D \right] \left[I - \frac{1}{1+k} D^{-1} A \right]^{-1} \right]^{-1} \quad \text{and} \quad (XY)^{-1} = Y^{-1} X^{-1}$$

$$= \left[I - \frac{1}{1+k} D^{-1} A \right]^{-1} \frac{1}{1+k} D^{-1} \quad \frac{1}{1-x} = 1+x+x^2+\dots$$

$$= \left(I + \frac{D^{-1}}{k+1} A + \frac{D^{-1}}{k+1} A \frac{D^{-1}}{k+1} A + \dots \right) \frac{D^{-1}}{k+1}$$

$$= \sum_{l=0}^{\infty} \left(\frac{D^{-1}}{k+1} \right) \left(\frac{1}{k+1} \right)^l \underbrace{\left(AD^{-1} \right)^l}_{\text{propagator for } l \text{ steps}} \quad \begin{matrix} \nearrow \text{symmetric} \\ \downarrow \end{matrix}$$

sum over paths of all length
 longer paths have decreased weight

Google: paths up to $l = 10$ or so.

$$\frac{1}{k+1} \approx 0.8$$

$[D-A]^{-1}$ can be precomputed \rightarrow fast!

(5)

Implementing Friend Finder

Calculate $L = D - A$

Use "Normalized Laplacian" $\hat{L} = \tilde{D}^{-\frac{1}{2}} [L] \tilde{D}^{-\frac{1}{2}} = I + \tilde{D}^{-\frac{1}{2}} A \tilde{D}^{-\frac{1}{2}}$

$K(\lambda) = \text{kernel using decay rate } \lambda = [(I + \lambda)I - \tilde{D}^{-\frac{1}{2}} A \tilde{D}^{-\frac{1}{2}}]^{-1}$

Usually $\lambda=1$ a good choice

Calculate $K(\lambda=1)$

Exploration:

For entire kernel

find top 10 pairs with $A_{ij}=0$ (not friends... yet)

f = index for these predicted friends, $f = 1, 2, 3, \dots$

$K(f)$ = kernel value

How good is the prediction - Relevance?

$P(f) = \# \text{ of known friends in entire network with } K \geq K(f)$

Then Relevance = $\frac{P(f)}{P(f) + f} \Rightarrow \text{prob. that a pair in this list are friends}$

Ignore $i=j$

Remember $i,j = j,i$ so count K_{ij} once, not twice

Personal Friend Finder

Same thing, but just look at your top hits with $A_{ij}=0$

relevance with respect to your friends.

Chapter 22

Network Topology, Motifs, and Clustering

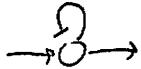
Lecture: Networks

①

Genes

Quantitative models of transcription

Diagrams as ODEs



Genomics: genome-scale discovery of diagrams

TF-DNA binding → gene regulatory networks

protein-protein binding → signal transduction

Problem: Direction, strength of regulation unclear/unknown

$O \rightarrow O$ vs $O \leftarrow O$

weak vs
strong

Opportunity: Analyze large-scale structure

Overall architecture? → (convergent) evolution

Small-scale motifs? → patterns with discrete function
(like auto-regulation)

New problems:

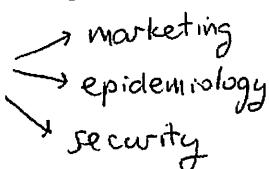
Searching (for candidate genes based on known genes)

clustering / segmenting → pathways, complexes

source detection → disease spread, causality

General problems for large-scale data sets

phone calls, emails, purchasing



Describing a network or graph

Intuition: our social network

N or $V \rightarrow$ # of vertices

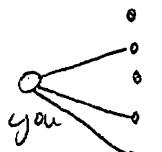
$E \rightarrow$ # of edges (undirected, no self edges)

↳ more complicated models: directed, multi-edges/time stamps, strength or probability, type (EvsP), ...

Intensive properties:

$f =$ fraction of possible edges that are occupied / observed

$J =$ average number of neighbors



$$f(N-1) = J \Rightarrow f = \frac{J}{N-1} \approx \frac{J}{N}$$

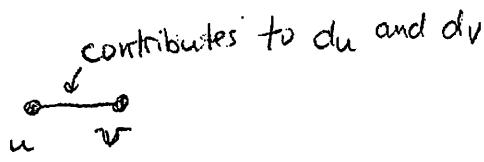
$$f = \frac{E}{\binom{N}{2}} = \frac{2E}{N(N-1)} = \frac{J}{N-1} \Rightarrow J = \frac{2E}{N}$$

Another way to relate $E + J$:

$$\sum_u d_u = N \cdot \langle \text{degree} \rangle = N \cdot J$$

↑ degree of vertex
vertex index

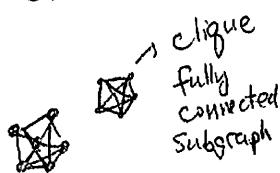
This counts each edge twice



$$NJ = 2E \Rightarrow J = \frac{2E}{N}$$

Many families of graphs have identical (N, E) or (N, f) or (N, J)

Clustered



each cluster:
 $J+1$ vertices
degree = J

$\frac{N}{J+1}$ clusters

Embedded



low-dimensional embedding
each vertex connects to
 J nearest neighbors

Hierarchical



each vertex has J edges
1 parent + $(J-1)$ children
deepest level has all internal connections (clique)

Random



N vertices with
 E edges
selected at random

2A

level	# of vertices in level
0	1
1	J
2	$J(J-1)$
⋮	⋮
ℓ	$J(J-1)^{\ell-1}$

total depth ℓ
with $\ell+1$ strata

$$N = 1 + J + J(J-1) + \dots + J(J-1)^{\ell-1}$$

$$(J-1)N = (J-1) + J(J-1) + \dots + J(J-1)^{\ell-1} + J(J-1)^\ell$$

$$(J-1-1)N = J(J-1)^\ell - 2$$

$$N = \frac{J(J-1)^\ell - 2}{J-2} \quad \text{check: } J=3, \ell=2 \rightarrow \frac{3 \cdot 2^2 - 2}{1} = 10 \quad \checkmark$$

↓ large J

$$N \approx J^\ell$$

$$\ell \approx \log_J N$$

Chapter 23

The Giant Component

What does our network look like?

(3)

Design teams, friends \rightarrow interpenetrating cliques

Classroom seating \rightarrow 2D embedding, probably correlated w/ social interaction or alphabetical

clubs \rightarrow hierarchical president $\xrightarrow{\text{officers}}$ members

No accounting for taste \rightarrow random

Random is often the starting point \rightarrow null model

Baseline for expected amount of clustering, motifs

Calculate distribution under null, compare to observed.

Example: $\Pr(d)$ degree distribution

$$\Pr(d) = \binom{N-1}{d} f^d (1-f)^{N-1-d}$$

Binomial, exact

$$\approx \binom{N}{d} f^d (1-f)^{N-d}$$

since $N-1 \approx N$

\downarrow recall prob stats, $\lim_{\substack{N \rightarrow \infty \\ f \rightarrow 0 \\ Nf \rightarrow J}}$

$$\frac{J^d}{d!} e^{-J}$$

Poisson

$$\binom{N}{d} f^d (1-f)^{N-d} = \frac{N(N-1)\dots(N-d)}{d!} f^d (1-f)^{N-d}$$

$$\approx \frac{N^d}{d!} f^d (1-f)^N$$

since $N \gg d$

$$\approx \left(\frac{Nf}{d!}\right)^d e^{-Nf}$$

since $f = \frac{J}{N} \ll 1$

$$= \frac{J^d}{d!} e^{-J}$$

④

Learned something: How many outcasts
 pariahs
 isolated vertices do we expect?

$$P(d=0) \approx \frac{J^0}{0!} e^{-J} = e^{-J}$$

$$\text{Expected number of outcasts} = \sum_n P(d_n=0) = N e^{-J}$$

when $N e^{-J} < 1$, less than 1 outcast on average
 ↓ take ln

$$\ln N - J < 0$$

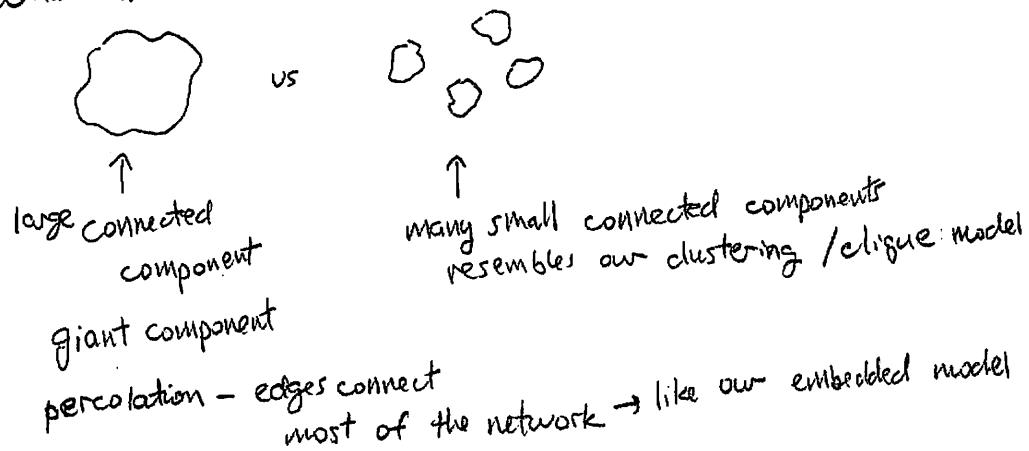
$J > \ln N$ gives a graph with no outcasts

Our class: $N \approx 100$

$$\ln N \approx 2 \underbrace{\ln 10}_{2.3} \approx 4.6$$

If each person has 4-5 friends, no outcasts.

What's about schisms?



What is the size of the largest component?

Is it close to N ? → entire class connected

⑤

p = prob that a vertex is not in the giant component.

For each of its $N-1 \approx N$ neighbors, either

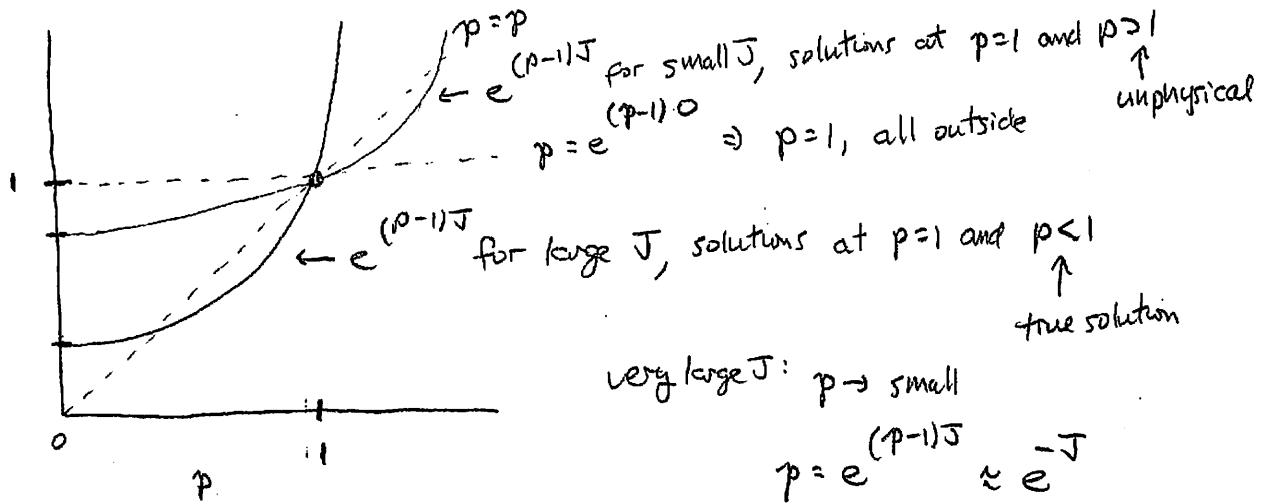
→ no connection prob = $1-f$

→ connected, but other vertex is also outside prob = $f \cdot p$

$$p = \prod_{u=1}^N [(1-f) + (f \cdot p)] = [1 + (p-1)f]^N \underset{f \text{ small}}{\approx} e^{(p-1)fN} = e^{(p-1)J}$$

Self-consistent equation. No closed form solution.

Insight: like nullclines. See where $p=p$ and $p=e^{(p-1)J}$ cross depending on J

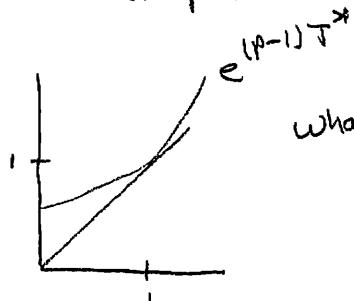


Giant component has $\frac{1}{1-e^{-J}}$ of all vertices

Transition when

$e^{(p-1)J^*}$ is exactly tangent

at $p=1$



what is this special J^* ?

$$\left. \frac{d}{dp} e^{(p-1)J^*} \right|_{p=1} = \left. \frac{d}{dp} p \right|_{p=1}$$

$$J^* = 1$$

→ having 1 neighbor on average gives a giant component.

Small components have e^{-J}

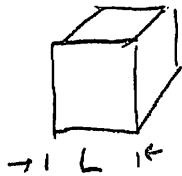
Our class: $e^{-5} \approx 1\%$ might be outside the giant component

(6)

Network diameter - How far on average?

Clustered network: diameter $\rightarrow \infty$ (unconnected components)

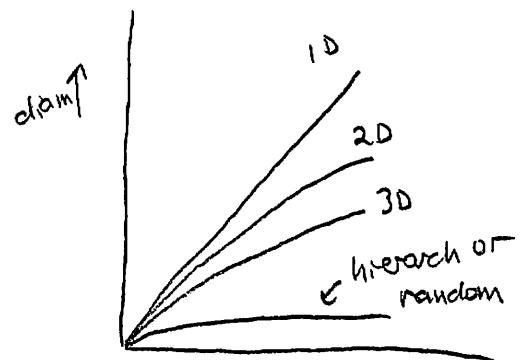
Embedded:



L vertices in each dimension

$N = L^Q$ \leftarrow # dimensions

$$\text{diam} \approx \sqrt{\underbrace{L^2 + L^2 + \dots + L^2}_{Q \text{ terms}}} \approx \sqrt{Q} \cdot L \approx \sqrt{Q} \cdot N^{\frac{1}{Q}}$$



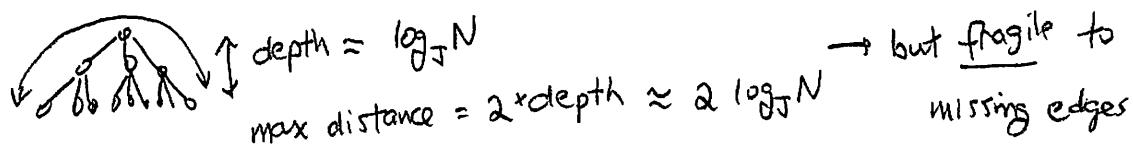
Linear: diam $\sim N$

2D: diam $\sim \sqrt{N}$

3D: diam $\sim \sqrt[3]{N}$

$N \rightarrow \dots$
with J fixed

Hierarchical:

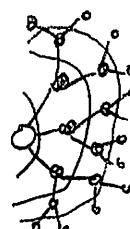


Random:

same idea.

first neighbors = J

second neighbors = $J(J-1)$
 \uparrow
for retracing path



length l neighbors $\approx J(J-1)^{l-1} \approx (J-1)^l = N$ when $l = \text{diameter}$

diameter $\approx \log_{J-1} N$

our class: diam $\approx \log_4 128 \approx \log_4 2^7 \approx 3$ to 4

Humanity: diam $\approx \log_{10} 10^{10} \approx 10$ for good friends

The Giant Component

Infection (application)

p = prob. not infected

f = prob. that an interaction would lead to infection (if the other person is infected)

N = # of interacting people

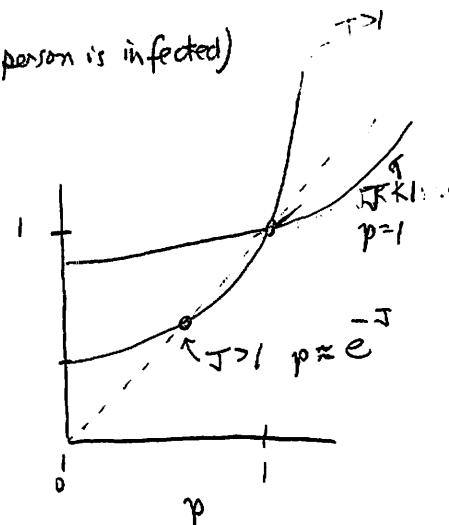
$$p = [(1-f) + fp]^N = [1 + (p-1)f]^N = e^{(p-1)fN}$$

$$p = e^{-(p-1)J}$$

threshold: $J \approx 1$

if each case leads to
→ 1 or more new cases → pandemic!

< 1 new case → no widespread infection



Problems when

$f \rightarrow$ large contagious

$N \rightarrow$ large air travel
computer / phone viruses

Simulation

N = # of vertices

Add edges, keep track of size
of largest component → call this Q

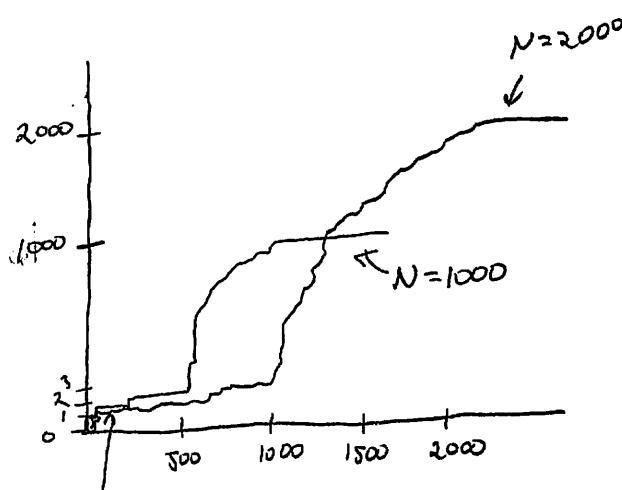
Transition when $J \approx 1$

$$NJ = 2E \Rightarrow \frac{NJ}{2}$$

$$g = \text{size of giant component} \\ = Q/N = 1 - p$$

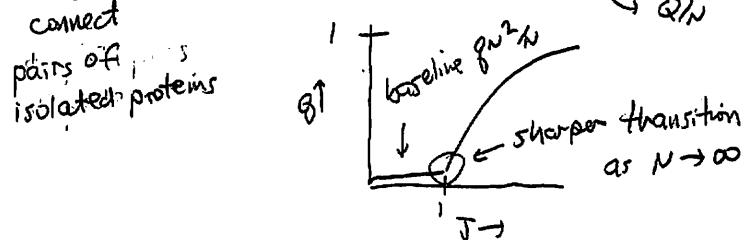
phase transition at $J=1$

$$(1-g) = e^{-gJ} \Rightarrow g = 1 - e^{-gJ}$$



early edges connect pairs of isolated proteins

$E \rightarrow$ collapse onto one plot
as g vs $J \sim 2E/N$
 $\rightarrow Q/N$



Characterizing the critical behavior: What happens near $J=1$? ②

Expand $g = 1 - e^{-qJ}$ near $(\frac{g}{J} = 0)$

Trick: write $J = 1 + x$

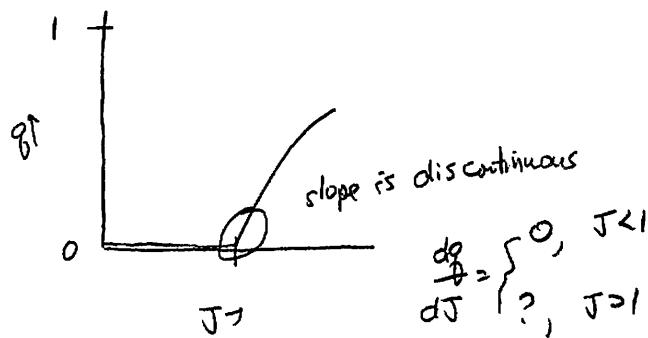
Expand to lowest non-vanishing order
of g and x

$$g = 1 - e^{-q(1+x)} = 1 - e^{-q} e^{-qx} = 1 - \left(1 - q + \frac{q^2}{2}\right) \left(1 - qx + \frac{q^2x^2}{2}\right)$$

$$g = 1 - \left[1 - q + \frac{q^2}{2} - qx + \frac{q^2x^2}{2} + \text{higher order terms}\right]$$

$$g = q - \frac{q^2}{2} + qx \Rightarrow g = 2x$$

slope $\frac{dg}{dx} \approx 2$ at $J=1$



Modules

Modules = motifs = patterns

Basic pattern: cliques

k -clique = k vertices, all edges $\rightarrow \binom{k}{2} = \frac{k(k-1)}{2}$

How many k -cliques in a random network?

<u>k</u>	<u>picture</u>	<u>n_k</u>
0	:	N
1	\bullet	N
2	---	E
3		$\binom{N}{3} f^3$ also: $\binom{N}{2} f^2$ $\xleftarrow[\text{ways to choose 3}]{}$ prob that 3 are connected $\xleftarrow[\text{ways of choosing 2 vertices}]{}$ given 2 vertices, prob (edge)
4		$\binom{N}{4} f^6 \approx \frac{N^4 f^6}{4!} = \frac{1}{N^2} \frac{J^6}{4!} \rightarrow 0 \text{ as } N \rightarrow \infty$ does not depend on $N!$
k		$\binom{N}{k} f^{\frac{k(k-1)}{2}} \approx \frac{N^k}{k!} f^{\frac{k(k-1)}{2}} \cdot N^{\frac{k(k-1)}{2}} \cdot N^{-\frac{k(k-1)}{2}} = \frac{J^k}{k!} \cdot N^{\frac{-k^2+k+2k}{2}}$ $= \frac{J^k}{k!} N^{\frac{-k(k-3)}{2}}$

Class social networks:

$$N \approx 100, J \approx 4 - 5$$

Expect:

$$n_2 = 200$$

$$n_3 = \frac{4^3}{6} \approx 10$$

$$n_4 = \left(\frac{44}{24}\right)^{\approx 10} \cdot \frac{1}{10^4} \approx \frac{1}{10^3} \quad \begin{matrix} \text{only 1 in 1000} \\ \text{will have a 4-clique!} \end{matrix}$$

How many 4-cliques do we have?

What is the p-value?

p-values for motifs.

$$\text{Have } \langle n_k \rangle = \frac{\sum^k}{k!} N^{-k(k-3)/2} = \binom{N}{k} f^{k(k-1)/2}$$

what is $\text{Var}(n_k)$?

Consider $n_k = \text{sum of } \binom{N}{k} \text{ trials with success prob. } f^k$

$$n_k = \sum_{i=1}^{\binom{N}{k}} x_i \quad x_i = \begin{cases} 1, & \text{prob } f^k \\ 0, & \text{prob } 1-f^k \end{cases}$$

$$\text{then } \text{Var}(n_k) = \binom{N}{k} (\underbrace{\langle x_i \rangle (1 - \langle x_i \rangle)}_{\text{binomial}}) \quad \text{but } f \ll 1 \quad 1 - f^k \approx 1$$

$$\text{Var}(n_k) \approx \binom{N}{k} f^k = \langle n_k \rangle \Rightarrow \text{Poisson!}$$

$$Pr(n_k) \approx \frac{\langle n_k \rangle^{n_k}}{n_k!} e^{-\langle n_k \rangle}$$

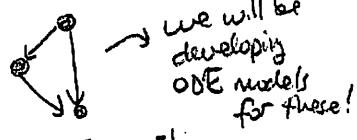
$$\text{p-value} = \sum_{n=n_k \text{ obsol}}^{\infty} Pr(n) \approx Pr(n=n_k^{\text{obsol}}) = \left[\frac{\binom{N}{k} f^k}{n_k!} \right]^{n_k} \exp[-\binom{N}{k} f^k]$$

For observing even 1 4-clique in our social network

$$\text{p-value} \approx \left[\frac{1}{1000} \right]^1 e^{-\frac{1}{1000}} \approx 10^{-3} \text{ highly significant.}$$

Biological networks:

gene regulation



feed forward

highly signif enrichment

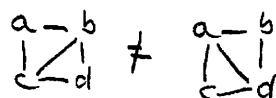
protein interactions



protein complex

careful with counting!

$a-b-c$ = 1 triangle, not 3 (or 6)



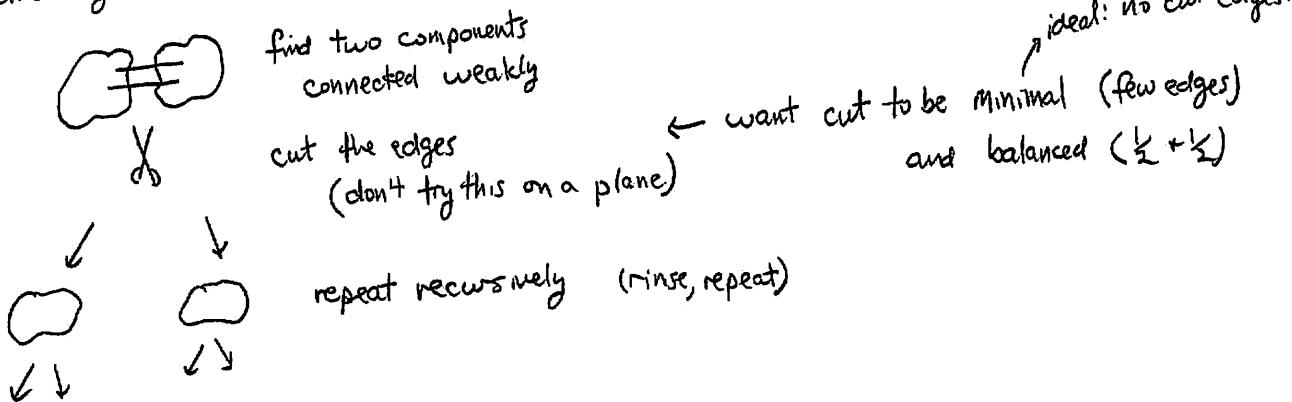
Chapter 24

Network Partitioning and Spectral Clustering

Spectral Clustering

welcome back old friend the Laplacian. $L = D - A$ what are you doing here?

Clustering idea:



Breakthrough by Horst Simon.

Name the components $+1, -1$. $x_i = \begin{cases} +\frac{1}{2} \text{ pos. component} \\ -\frac{1}{2} \text{ neg. component} \end{cases}$

Think about minimizing

$$\sum_{i,j} A_{ij} (x_i - x_j)^2 = 2 \times (\# \text{ cut edges})$$

$\downarrow 0$ for same side i, j and j, i are both in sum
 1 for opposite side

$$\text{Minimize } \sum_{i,j} A_{ij} (x_i - x_j)^2 = \sum_{i,j} A_{ij} x_i^2 + A_{ij} x_j^2 - 2 A_{ij} x_i x_j$$

$$= \sum_{i,j} x_i [D_{ij} + D_{ji} - 2 A_{ij}] x_j = 2 x^+ (D - A) x$$

Aha! Found you!

Trivial minimum: no cut, all $x = \pm \frac{1}{2}$

$$(D - A) \left(\frac{k}{2} \right) = \frac{1}{2} \sum_j (D_{ij} - A_{ij}) = \frac{1}{2} (D_{ii} - D_{ii}) = 0$$

What about next eigenvalue?

Since $x^+ (D - A) x = \sum_{i,j} A_{ij} (x_i - x_j)^2 \geq 0$, $D - A$ is positive semidefinite with positive eigenvalues ≥ 0 semi

Idea: Suppose x is an eigenvector of $(D-A)$

$$(D-A)x = \lambda x$$

Choosing x with smallest eigenvalue minimizes $x^T(D-A)x$!

↳ subject to $x^T x = 1$ of course

Does this solve our problem?

Almost $\rightarrow x_i$ will be real, not $\pm i$.

Reasonable ideas: split on $x_i > 0$ vs $x_i < 0$

split on largest gap

choose split to minimize $x^T(D-A)x$

:

what does this actually minimize?

Imagine 2 components with size N_1, N_2 .

$$L = \begin{pmatrix} L_1 & 0 \\ 0 & L_2 \end{pmatrix} \quad \text{block diagonal.} \quad x = \begin{pmatrix} a_1 \\ \dots \\ a_2 \end{pmatrix}$$

Trivial solution: $a_1 = a_2$ since $L \cdot \underset{\text{vector}}{\text{const}} = 0$.

Next eigenvector (2nd eigenvector, Fiedler vector)

$$\textcircled{1} \text{ orthogonal to } \underline{1} \Rightarrow 1^T \begin{pmatrix} a_1 \\ \dots \\ a_2 \end{pmatrix} = 0 = N_1 a_1 + N_2 a_2 \Rightarrow a_2 = -\frac{N_1}{N_2} a_1$$

\textcircled{2} normalized

$$N_1 a_1^2 + N_2 a_2^2 = 1 \Rightarrow N_1 a_1^2 + N_2 \cdot \frac{N_1^2}{N_2^2} a_1^2 = 1 = a_1^2 \left[N_1 + \frac{N_1^2}{N_2} \right]$$

$$a_1^2 \frac{N_1(N_1+N_2)}{N_2} = 1$$

$$a_1 = \frac{\sqrt{N_2/N_1}}{\sqrt{N}} \quad a_2 = \frac{\sqrt{N_1/N_2}}{\sqrt{N}}$$

(3)

Now... what are we actually minimizing when $x_1 = \frac{\sqrt{N_2}M_1}{\sqrt{N}}$ $x_2 = -\frac{\sqrt{M_1}N_2}{\sqrt{N}}$?

$$\frac{1}{2} \sum_{i,j} A_{ij} (x_i - x_j)^2 = \frac{1}{2} \cdot 2E_{12} \cdot \left[\frac{\sqrt{N_2}M_1}{\sqrt{N}} + \frac{\sqrt{M_1}N_2}{\sqrt{N}} \right]^2$$

only contribution
is from cuts
have E_{cut} of these
or E_{12}

$$= \frac{E_{12}}{N} \left[\underbrace{\frac{N_2}{M_1} + \frac{M_1}{N_2} + 2}_{(N_1^2 + N_2^2 + 2N_1N_2) / N_1N_2} \right]$$

$$\frac{1}{2} x^T (D - A) x = E_{12} \cdot \frac{N}{N_1N_2} = E_{12} \cdot \frac{N}{N_1N_2} = E_{12} \left[\frac{1}{M_1} + \frac{1}{N_2} \right]$$

↑
favors balanced cuts

Improvements:

Multi-way: look at several eigenvectors

Modularity: Instead of E_{12} , look at $E_{12} - \langle E_{12} \rangle$ for a random network

Statistical: look at $\frac{(E_{12} - \langle E_{12} \rangle)^2}{\text{Var}(E_{12})} \approx \chi^2$ statistic \downarrow gives a stopping criterion

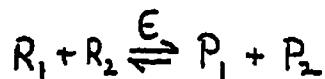
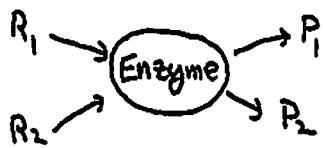
Application:

Social network clustering

Chapter 25

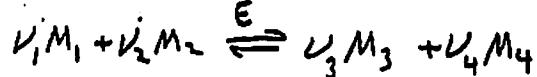
Metabolic Networks and Flux Balance Analysis

Metabolic networks: moving stuff



↓ more general stoichiometry,
call them all metabolites

Kirkhoff's Law = flux balance



$$\begin{pmatrix} \dot{M}_1 \\ \dot{M}_2 \\ \dot{M}_3 \\ \dot{M}_4 \end{pmatrix} = \begin{pmatrix} -v_1 \\ -v_2 \\ +v_3 \\ +v_4 \end{pmatrix} \quad (\text{flux through } E)$$

↑
stoichiometry
vector

multiple enzymes = multiple reaction channels

$$\begin{matrix} \text{roughly} \\ 1000 \text{ metabolites} \end{matrix} \begin{pmatrix} \dot{M}_1 \\ \dot{M}_2 \\ \vdots \\ \dot{M}_{1000} \end{pmatrix} = \begin{pmatrix} \xleftarrow{\text{rxns \in enzymes}} & & & \\ \uparrow & \text{metabs} & & \\ & \downarrow & \text{Stoch.} & \\ & & \text{matrix } S & \end{pmatrix} \begin{pmatrix} F_1 \\ F_2 \\ F_3 \\ \vdots \\ F_{1000} \end{pmatrix} \quad \xrightarrow{\text{flux vector}} F$$

$$\text{cell at steady state: } \dot{M} = 0 = S \cdot F$$

trivial solution: $F = 0$ cell is dead

other solutions: null space of S

dimension of null space: (# variables) - (# constraints)

variables = # of fluxes to choose ≈ 600

constraints = # of $\dot{M}_i = 0$ conditions ≈ 500
(1 per metabolite)

dimension of null space $\approx (\# \text{ rxns}) - (\# \text{ metabolites}) \approx 100$

100-dimension space $\rightarrow 2^{100}$ possible solutions for (on/off) switch

think about solution that optimizes cell growth: rate of metabolite production

Efficient optimization: linear programming

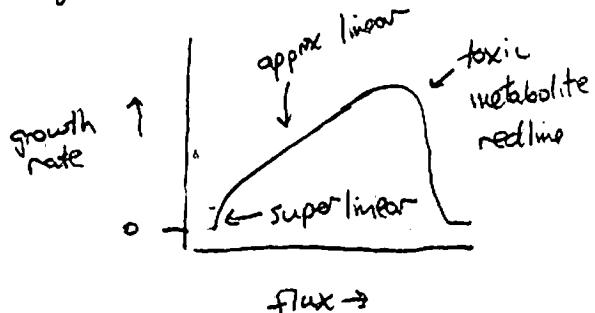
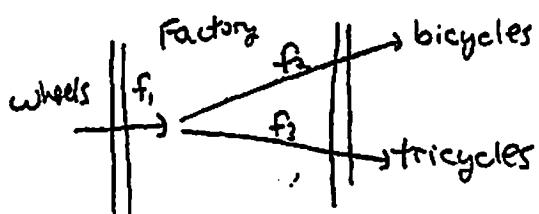
Linear constraints: $S \cdot F = 0$

Linear objective function not really known!

$$\text{biomass production rate} = \sum_{\text{amino acids}} w_{aa} \cdot M_{aa} = w \cdot F$$

\uparrow linearity is a strong assumption!

Efficient solution using linear programming



$$\begin{pmatrix} \dot{w} \\ \dot{g} \\ \dot{f} \end{pmatrix} = \begin{pmatrix} 1 & -2 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix}$$

Phase 1: Collected wheels

Phase 2: ?

Phase 3: Profit!

Additional constraints:

$$f_1 < W_{\max} \quad \text{wheel supplier is limited}$$

$$\left. \begin{array}{l} f_2 < B_{\max} \\ f_3 < T_{\max} \end{array} \right\} \text{assembly lines are limited}$$

optimization fn:

$$\begin{aligned} \text{net profit} &= P_2 f_2 + P_3 f_3 - C_1 f_1 = \underbrace{(P_2 - 2C_1)}_{\text{net profit}} f_2 + \underbrace{(P_3 - 3C_1)}_{\text{net profit}} f_3 \\ &\uparrow \qquad \uparrow \qquad \uparrow \\ \text{price of a bike} &\qquad \text{price of a trike} \qquad \text{cost per wheel} \end{aligned}$$

$$\begin{pmatrix} \omega \\ B \\ T \end{pmatrix} = \begin{pmatrix} 1 & -2 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} f_W \\ f_B \\ f_T \end{pmatrix}$$

constraints:

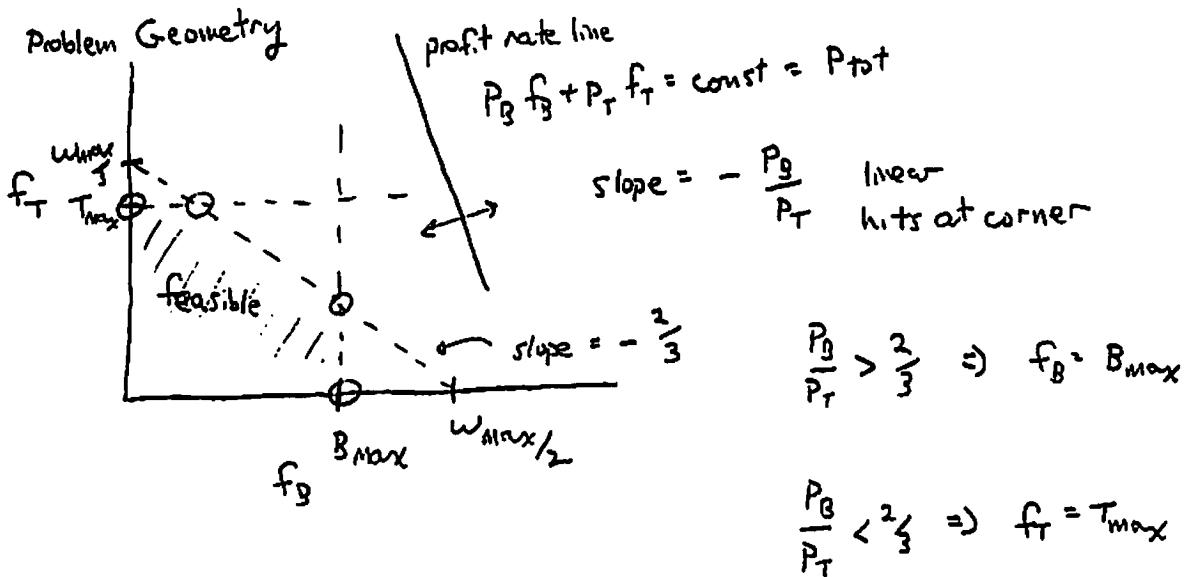
$$f_W < \omega_{\max}$$

$$f_B < B_{\max}$$

$$f_T < T_{\max}$$

optimization function:

$$\underset{\text{rate}}{\text{profit}} = (C_B - 2C_\omega) f_B + (C_T - 2C_\omega) f_T = P_B f_B + P_T f_T$$



feasible space:

$$f_B < B_{\max}$$

$$f_T < T_{\max}$$

$$\underbrace{f_W}_{\sim} < \omega_{\max}$$

$$f_W - 2f_B - 3f_T = 0$$

$$2f_B + 3f_T < \omega_{\max}$$