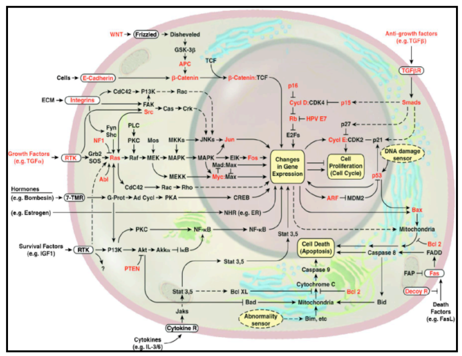
Computational Science and Engineering

Computational Cell Biology
A Brief Introduction



Yang Cao

Department of Computer Science

Virginia Tech

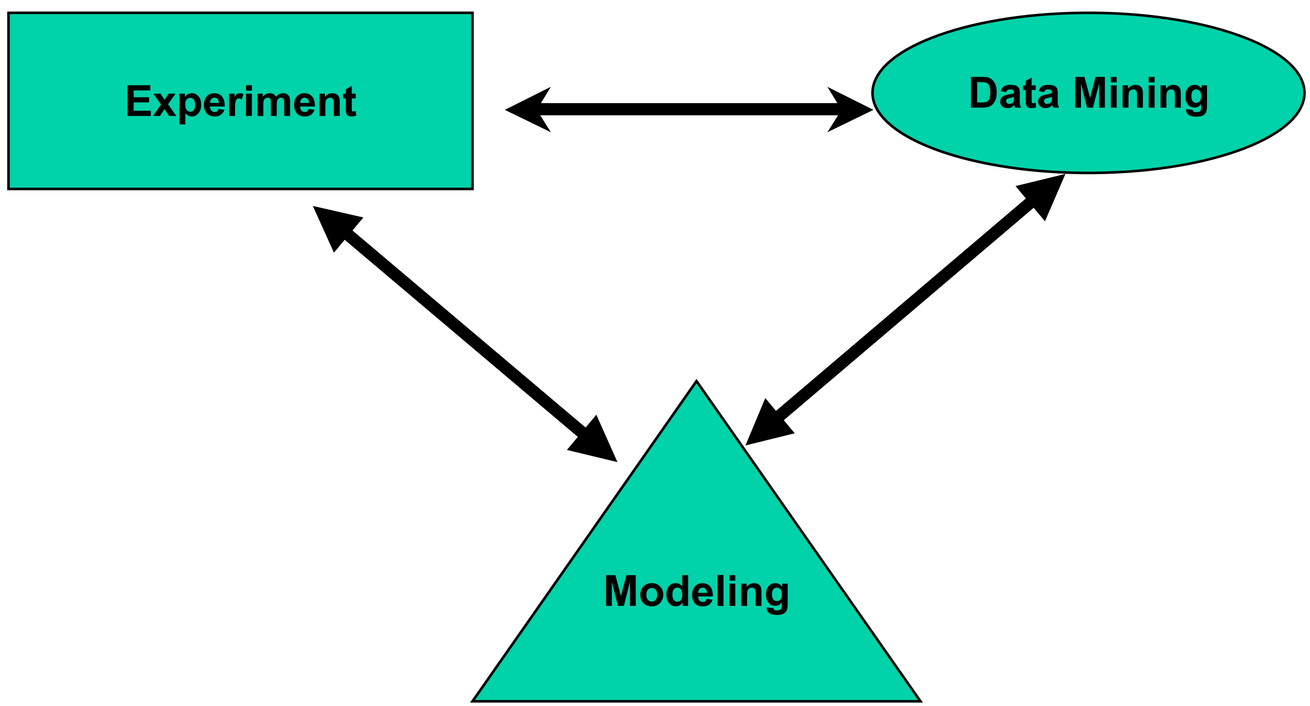
Summary

- General Introduction
- Modeling: from Simple Structures to Complex Systems
- Modeling with ODEs

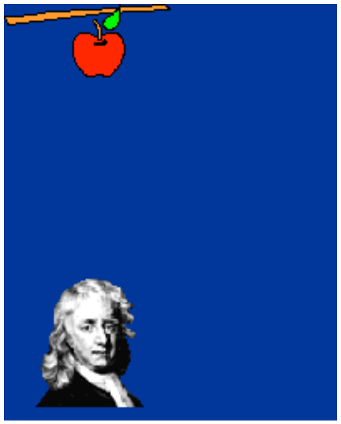
Traditional Scientific Research

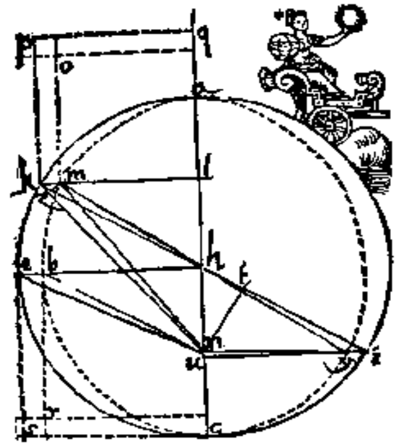
- A Common Pattern in Scientific Research

What will people usually do to study a problem?



Newton’s Apple: A Story from Astrology





Calculus

Newton’s Three Laws of Motion

Tycho Brahe

20 years of observation on the orbit of Mars

Johannes Kepler

Kepler’s Three Laws

Isaac Newton

Law of Universal Gravitation

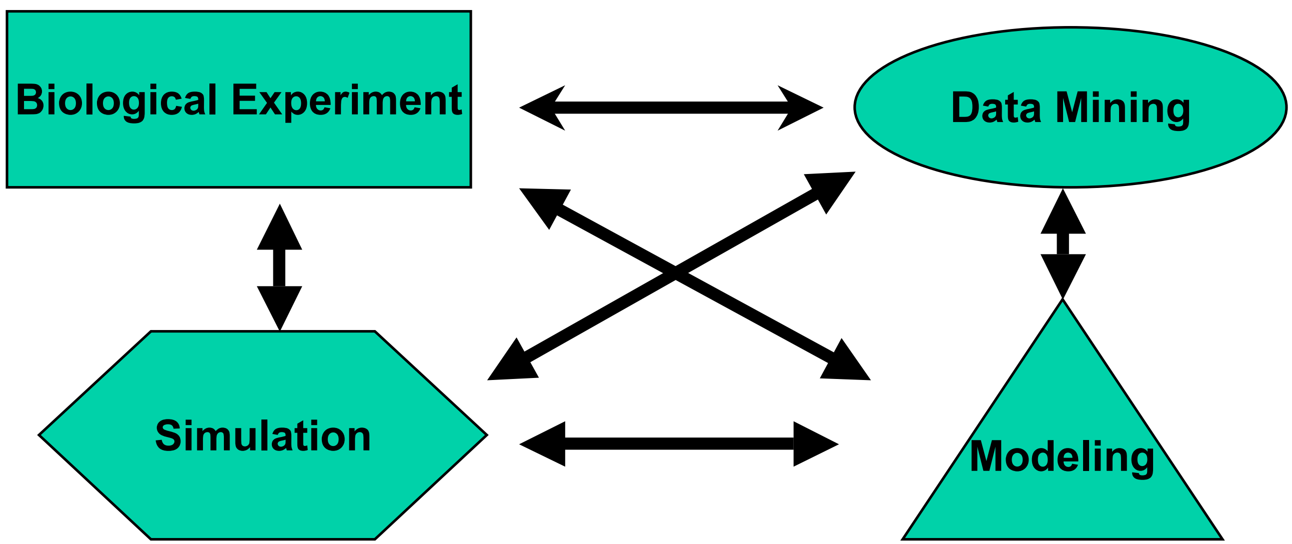
Computational Biology

- **Data Mining vs. Modeling and Simulation**

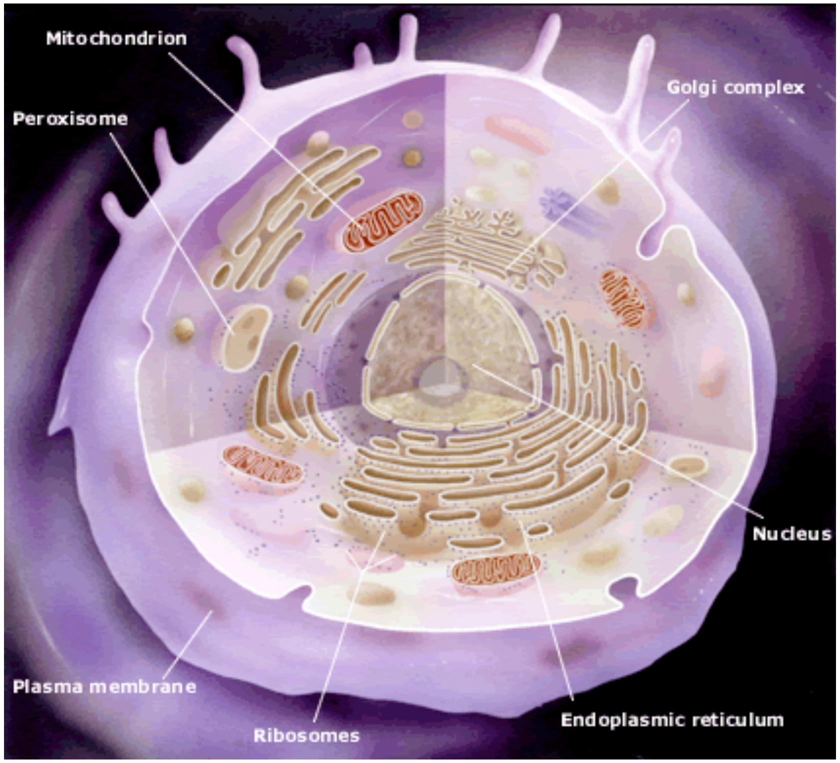
  " Computational biology has two distinct branches: knowledge discovery, or data-mining, which extracts the hidden patterns from huge quantities of experimental data, forming hypotheses as a result; and simulation-based analysis, which tests hypotheses with *in silico* experiments, providing predictions to be tested by *in vitro* and *in vivo* studies. "

  ---- H. Kitano, *Computational systems biology*, Nature, 420, 206-210, Nov. 14, 2002

- **My understanding of computational biology**

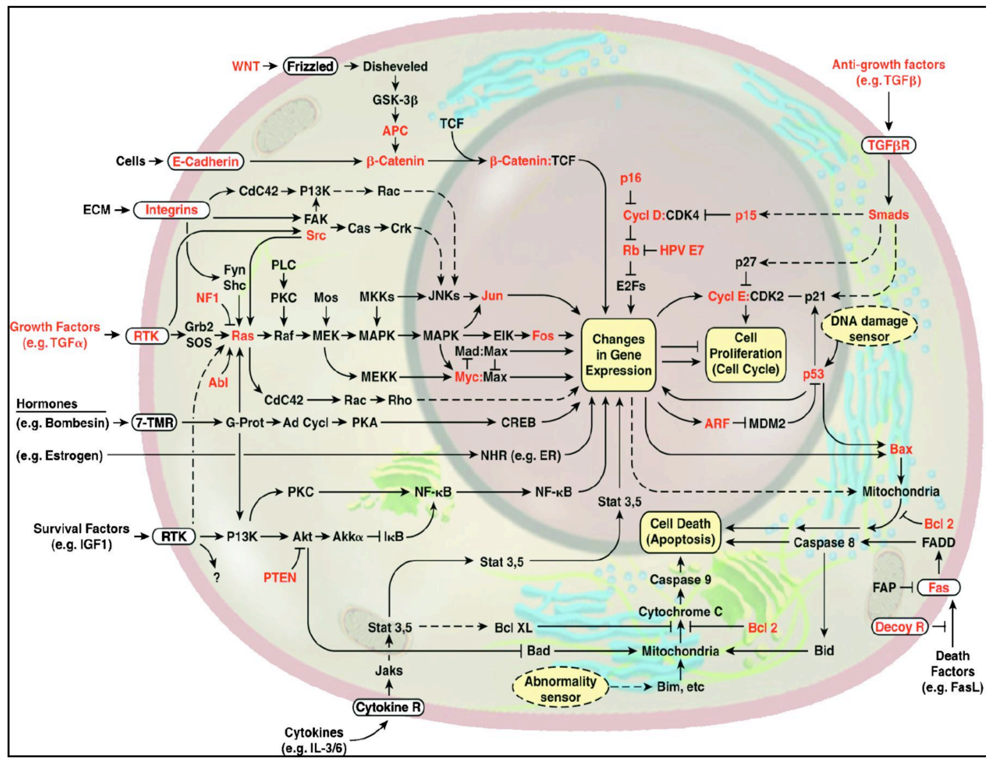


Modeling the Cell



A nice picture of Cell

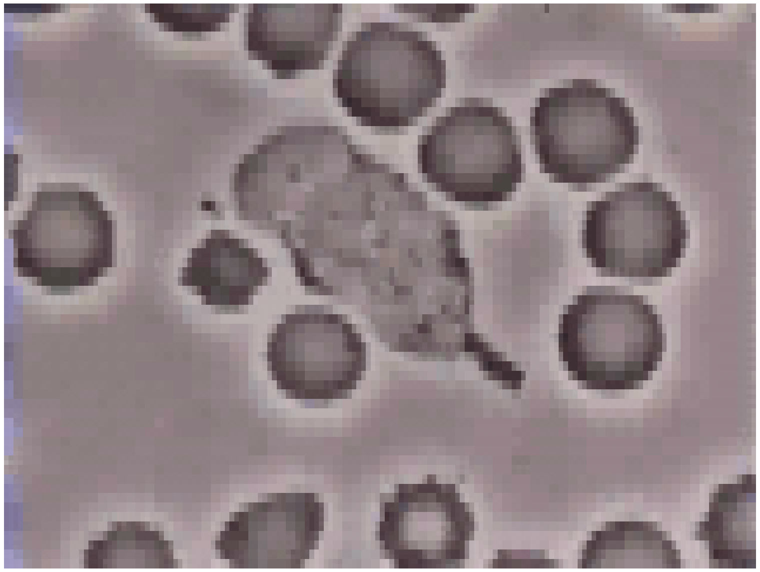
From a modeler's point of view

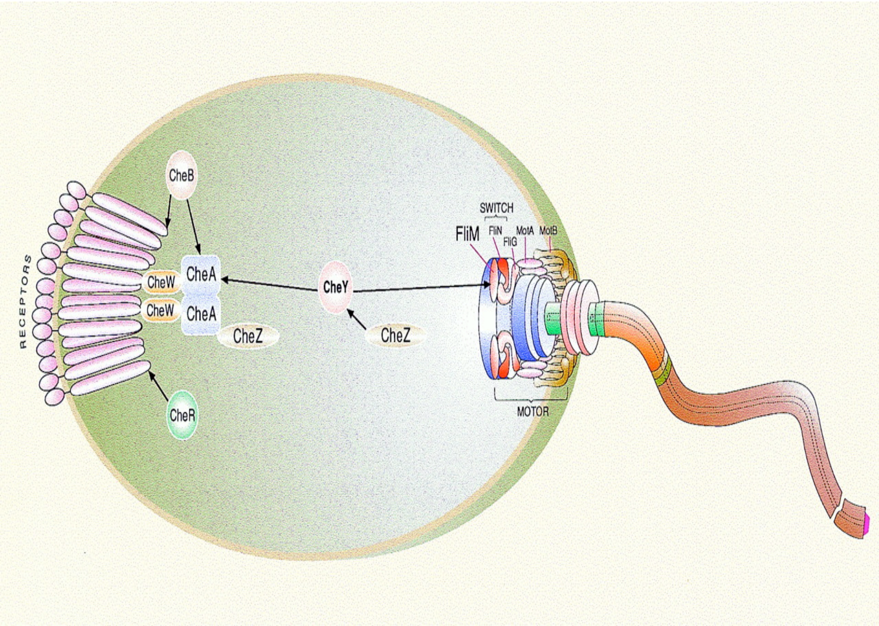


Modeling from Physics Point of View

Computational Science and Engineering

- Different Modeling Methods
  - Top down vs. Bottom up
  - Behavior vs. Mechanism
  - From Physics vs. from Chemistry
  - Deterministic vs. Stochastic





Chemotaxis

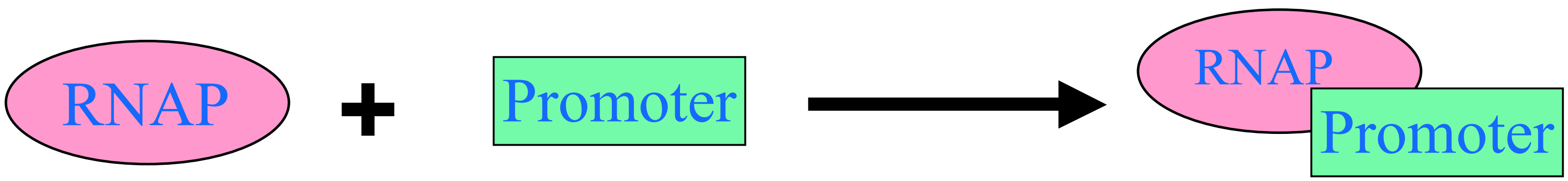
- General Introduction

- Modeling: from Simple Structures to Complex Systems

- Modeling with ODEs

A Chemically Reacting System

- Molecules of N chemical species $S_1, \ldots, S_N$
  - In a Volume Ω, at temperature T
  - Different conformation or excitation levels are considered different species if they behave differently



- M elemental reaction channels $R_1, \ldots, R_M$
  - Each R_j describes a single instantaneous physical event which changes the population of at least one species. For example,

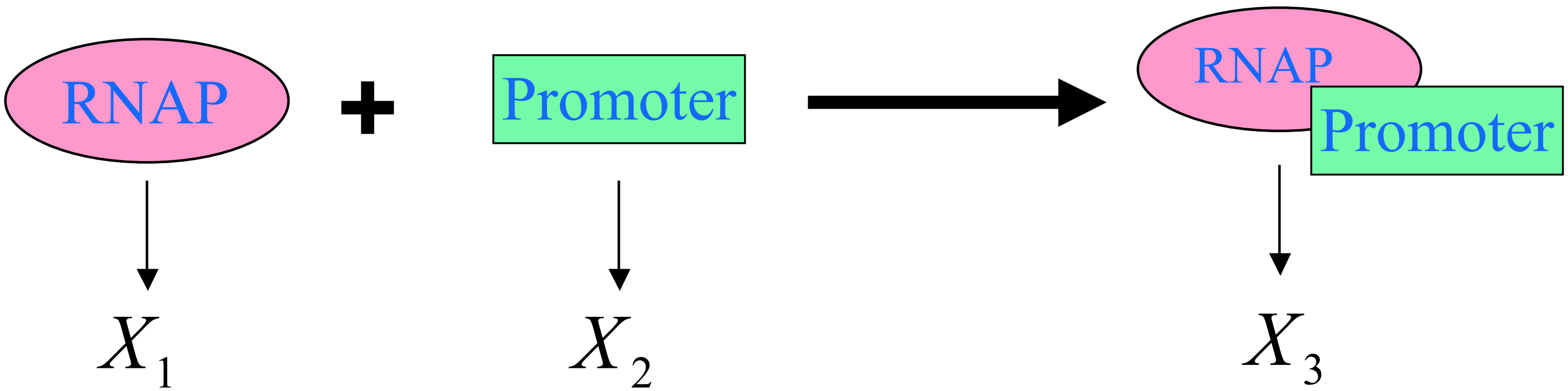
$$A \rightarrow S_i,$$

$$\text{or} \quad S_i \rightarrow \text{something else},$$

$$\text{or} \quad S_i + S_j \rightarrow \text{something else}.$$

Ordinary Differential Equation

For each species, assign a state variable, which describes its concentration or population.



Basic Deterministic Assumption:

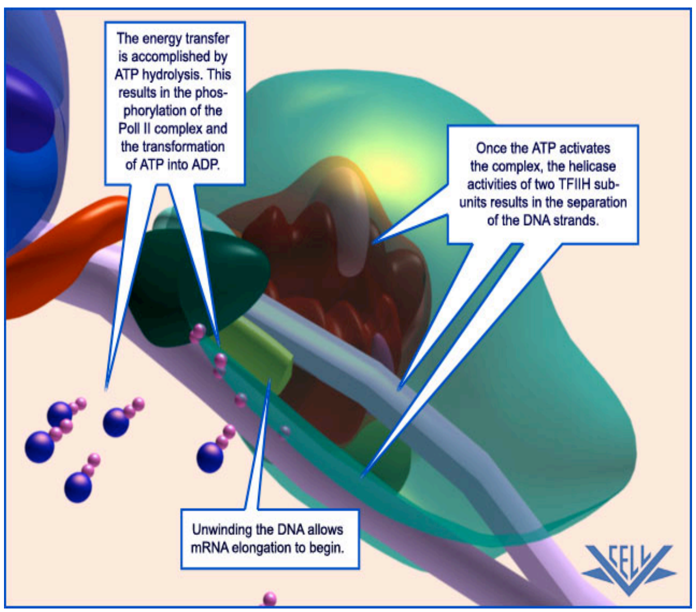
The state change is proportional to the state of the reactants and time

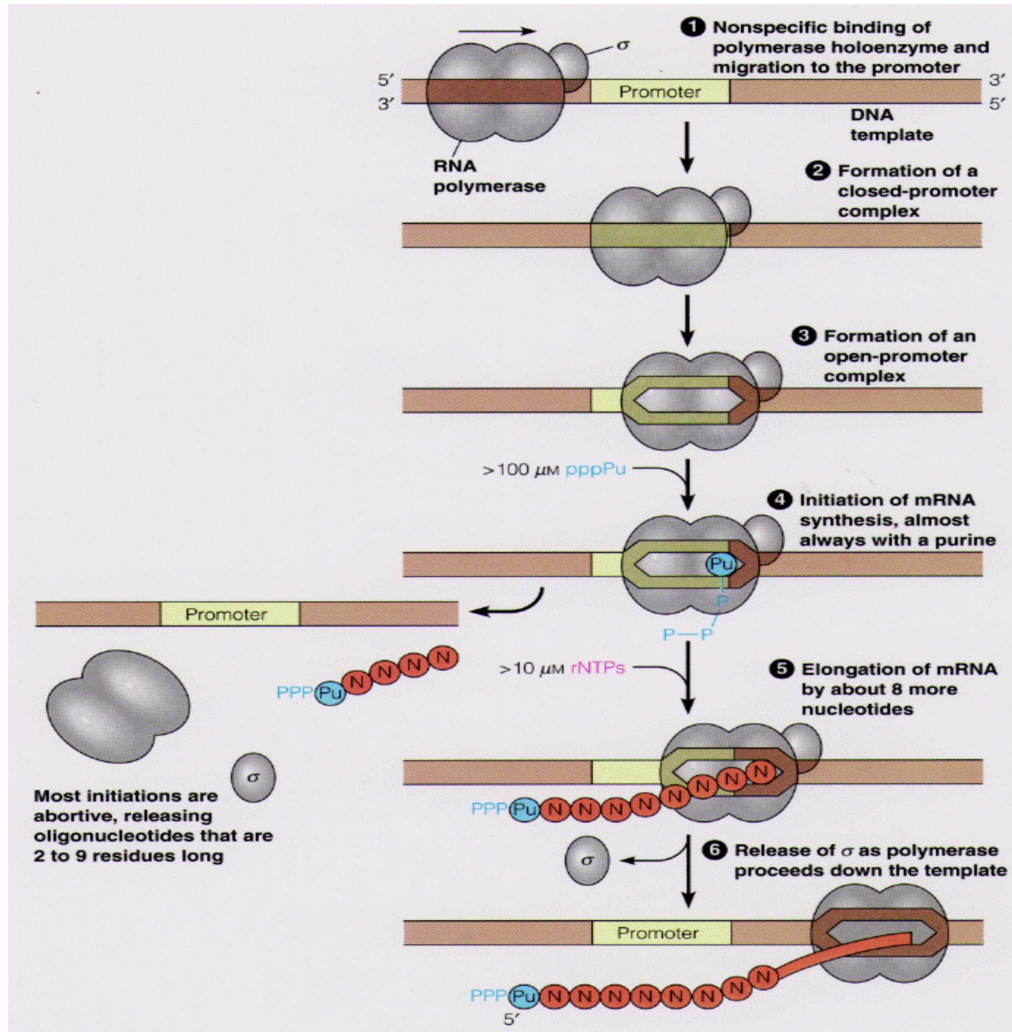
$$\Delta X_1(t) = -kX_1(t)X_2(t)\Delta t$$

$$X_1'(t) = -kX_1(t)X_2(t)$$

The Process of Transcription (in gene expression)

1. **Binding**
2. **Initiation**
3. **Elongation**
4. **Termination**





Initiation figure

A Model for Prokaryotic Gene Expression

Computational Science and Engineering

1. **Transcription Initiation (the binding and initiation)**

$$\mathrm{P+RNAP} \rightarrow \mathrm{P \bullet RNAP} \qquad k_1 = 10^8\,\mathrm{M}^{-1}s^{-1}$$

$$\mathrm{P \bullet RNAP} \rightarrow \mathrm{P+RNAP} \qquad k_2 = 10s^{-1}$$

$$\mathrm{P \bullet RNAP} \rightarrow \mathrm{TrRNAP} \qquad k_3 = 1s^{-1}$$

2. **Elongation (RBS is available before elongation terminates**

$$\mathrm{TrRNAP} \rightarrow \mathrm{RBS + P + ElRNAP} \qquad k_4 = 1s^{-1}$$

3. **Translation Initiation**

$$\mathrm{Ribosome + RBS} \rightarrow \mathrm{RibRBS} \qquad k_5 = 10^8\,\mathrm{M}^{-1}s^{-1}$$

$$\mathrm{RibRBS} \rightarrow \mathrm{Ribosome + RBS} \qquad k_6 = 2.25s^{-1}$$

$$\mathrm{RibRBS} \rightarrow \mathrm{ElRib + RBS} \qquad k_7 = 0.5s^{-1}$$

$$\mathrm{RBS} \rightarrow \mathrm{decay} \qquad k_8 = 0.3s^{-1}$$

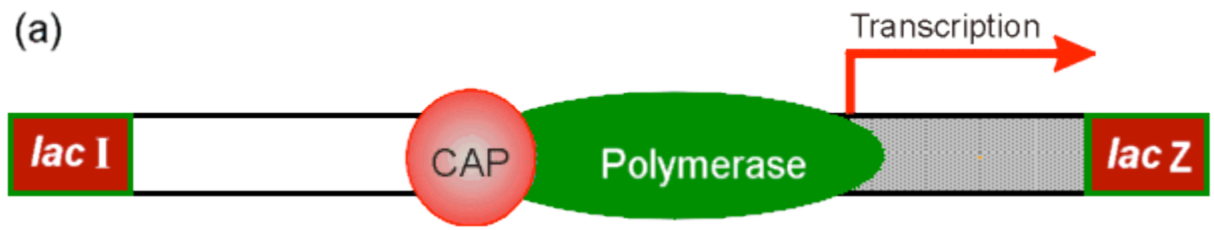
4. **Elongation**

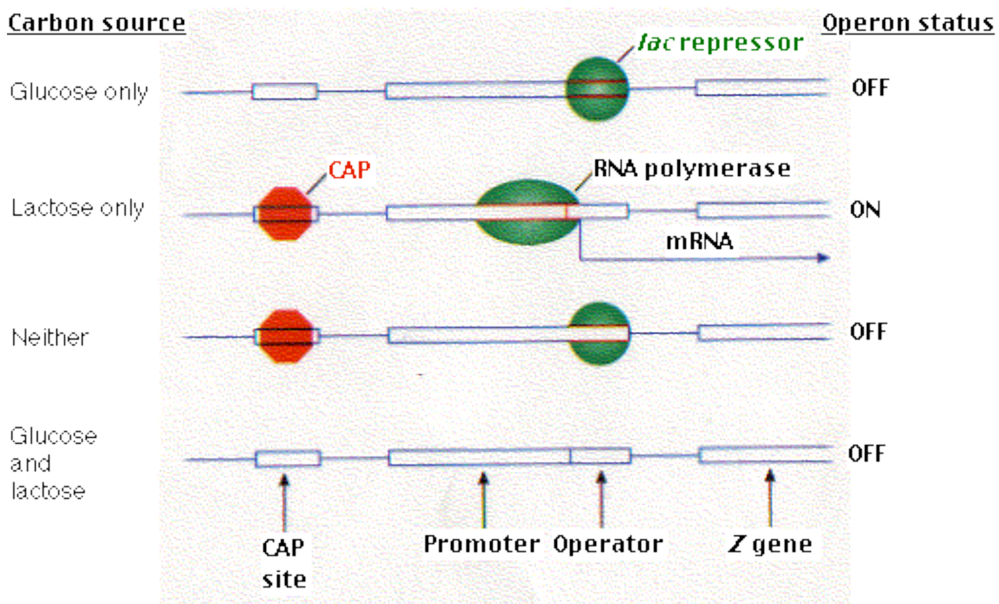
$$\mathrm{ElRib} \rightarrow \mathrm{Protein} \qquad k_9 = 0.015s^{-1}$$

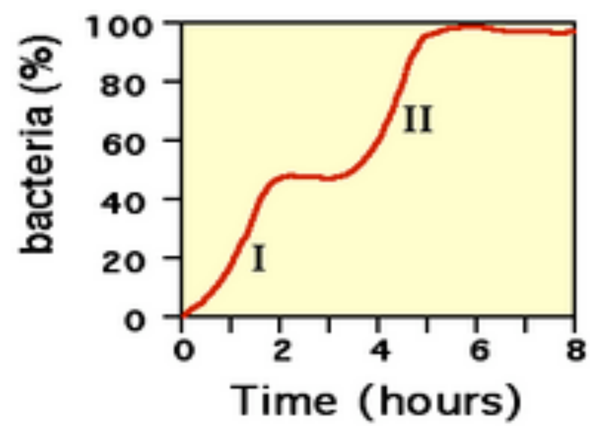
$$\mathrm{Protein} \rightarrow \mathrm{decay} \qquad k_{10} = 6.42 \times 10^{-5} s^{-1}$$

Gene Regulation

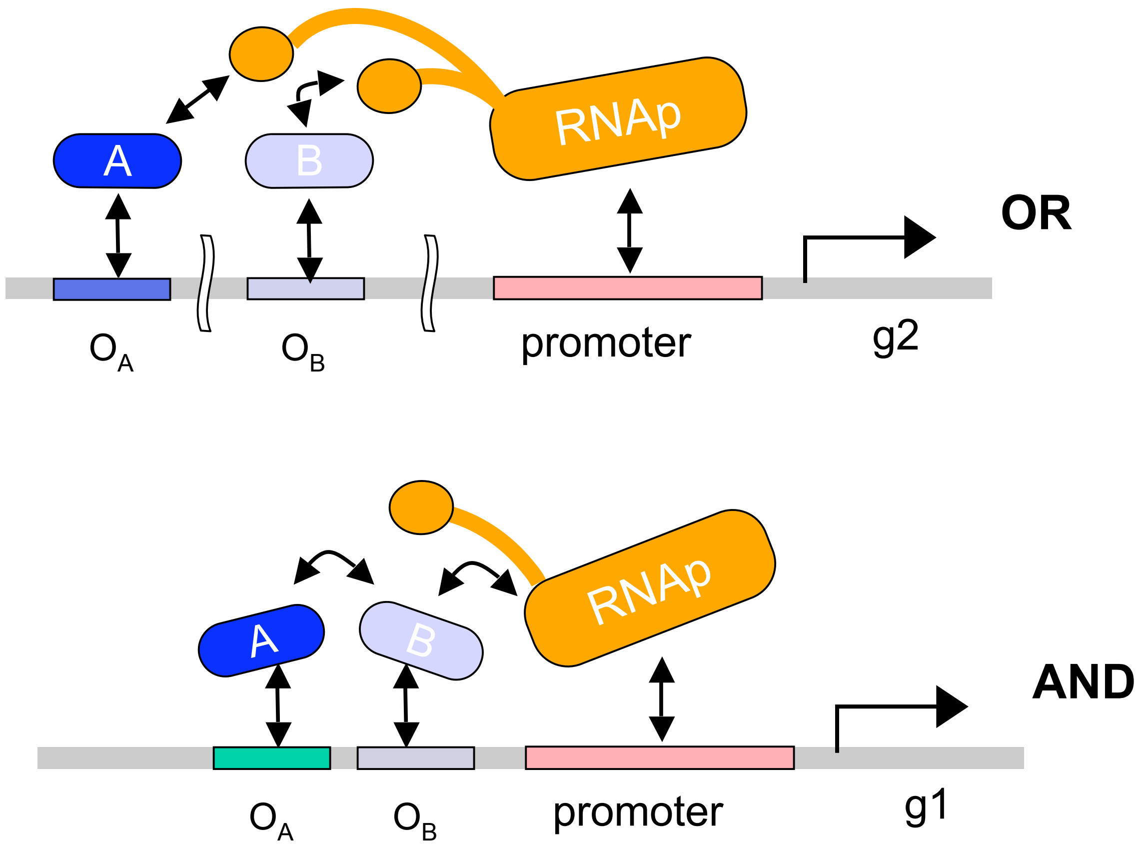
- **Activator**



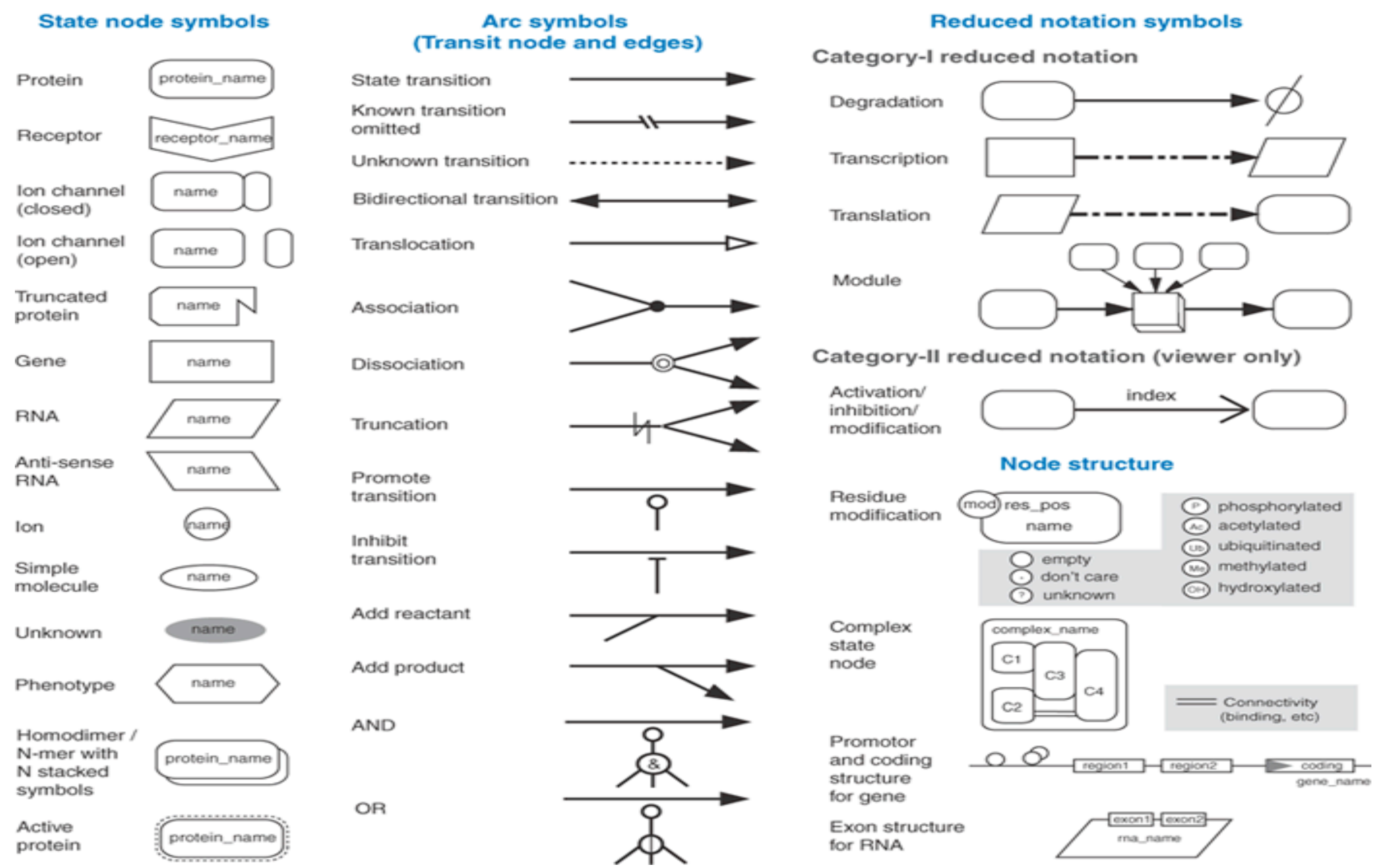




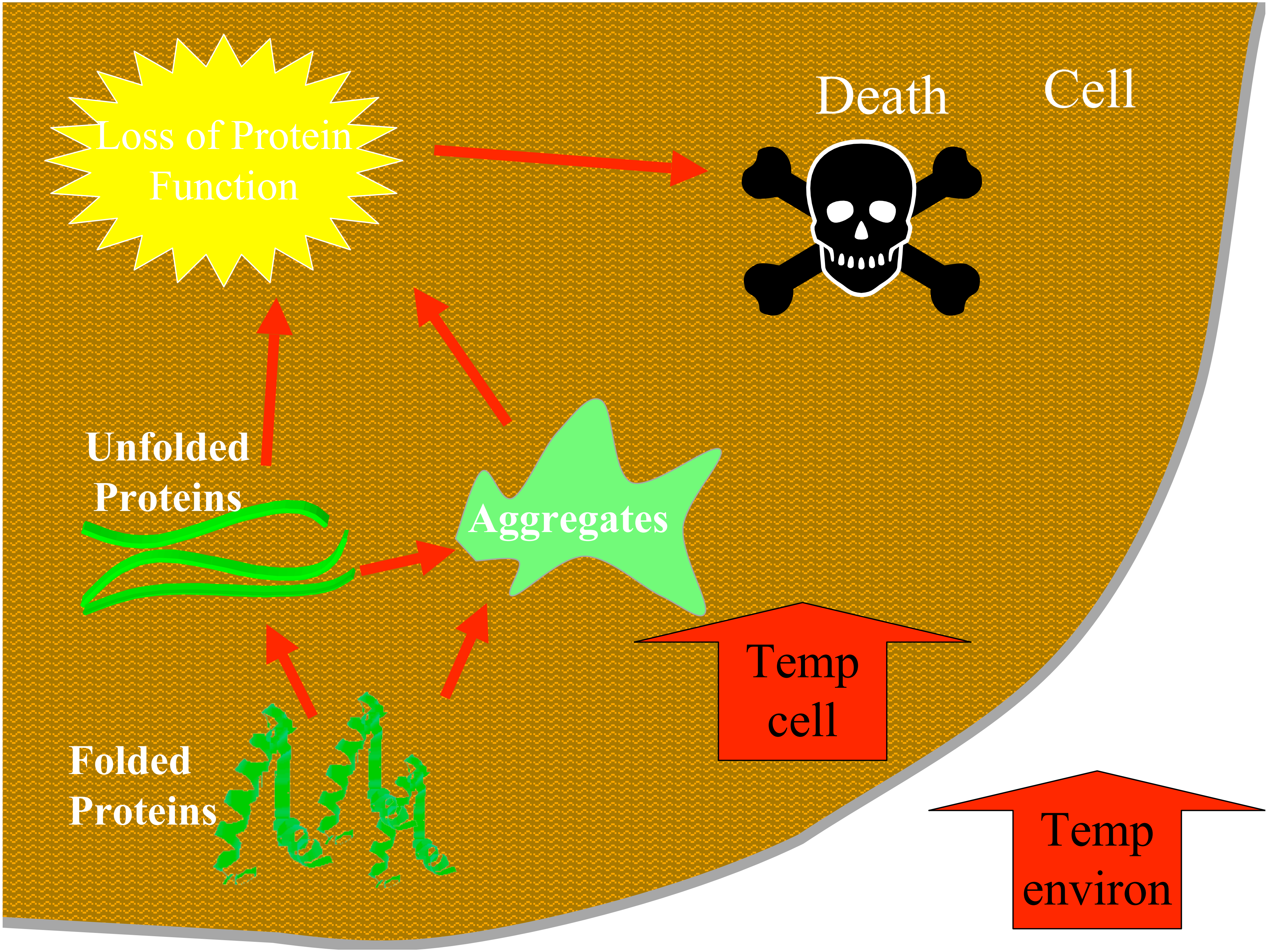
Simple Regulation in Biology – Circuits?



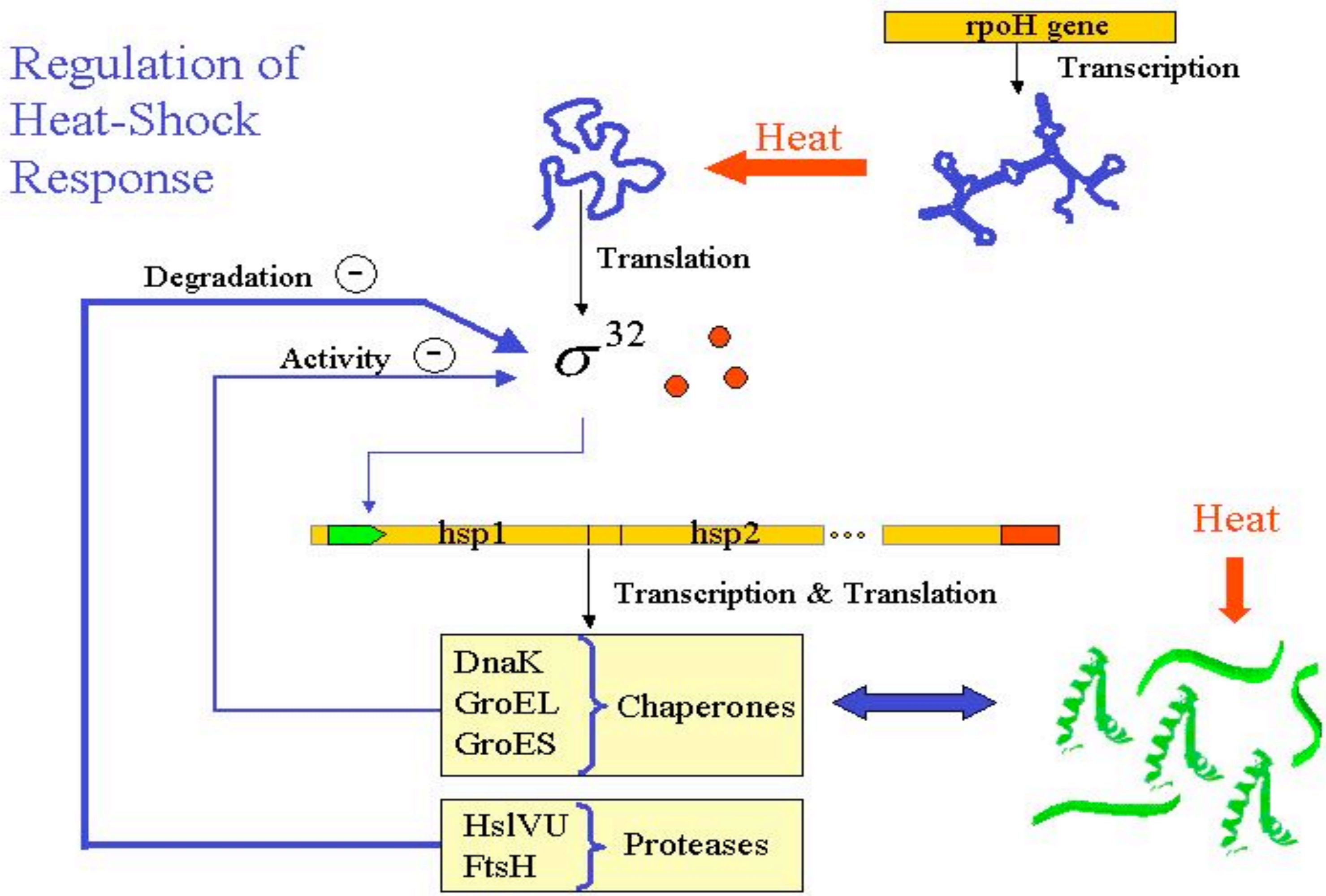
Yes! Circuits!



Kitano H, Funahashi A, Matsuoka Y, et al., Using process diagrams for the **graphical representation of biological networks, NATURE BIOTECHNOLOGY 23 (8): 961-966 AUG 2005**


Loss of Protein Function
Death
Cell
Unfolded Proteins
Aggregates
Temp cell
Folded Proteins
Temp environ

A Complex Model



Involved Subject

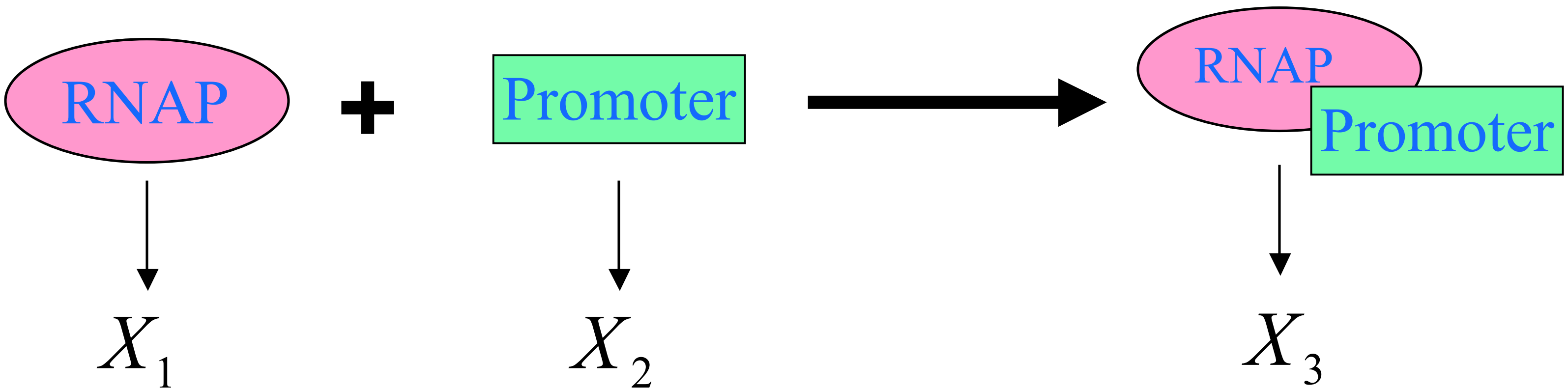
- **Calculus, ODE (Mathematics)**
- **Probability (Statistics)**
- **Programming language (Computer Science)**
- **Systems Biology**

- Come on! Is this possible?

- **General Introduction**
- **Modeling: from Simple Structures to Complex Systems**
- **Modeling with ODEs**

Ordinary Differential Equation

For each species, assign a state variable, which describes its concentration or population.



Basic Deterministic Assumption:

The state change is proportional to the state of the reactants and time

$$\Delta X_1(t) = -kX_1(t)X_2(t)\Delta t$$

$$X_1'(t) = -kX_1(t)X_2(t)$$

Ordinary Differential Equations

Many scientific applications result in the following system of equations

$$\frac{dy}{dx} = f(x, y)$$

which is called ordinary differential equations (ODEs).

Example: Newton's Motion Law $F = m\ddot{x}$

Thomas Malthus

An Essay on the Principle of Population

An Essay on the Principle of Population, as it Affects the Future Improvement of Society with Remarks on the Speculations of Mr. Godwin, M. Condorcet, and Other Writers.

LONDON, PRINTED FOR J. JOHNSON, IN ST. PAUL'S CHURCH-YARD, 1798.

"I SAID that population, when unchecked, increased in a geometrical ratio, and subsistence for man in an arithmetical ratio. "

---- Thomas Malthus

Malthus Model:

Assumption: the reproduction rate is proportional to the size of the population

$$\frac{dP}{dt} = kP, \quad k = \text{growth rate per capita}$$

Solution: $P(t) = P(0)e^{kt}$

$k > 0$: exponential growth, $k < 0$: exponential decay

Malthus Model

The reproduction rate is proportional to the population

$$P(t+\Delta t) = P(t) + kP(t)\Delta t$$

Solve it we have

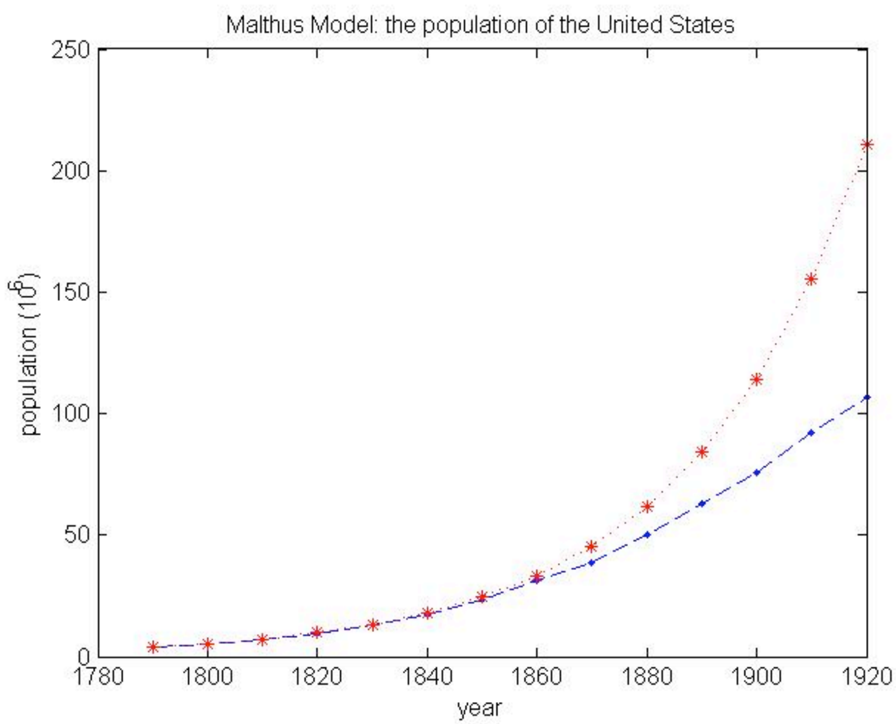
$$P(t) = P_0 e^{k(t-t_0)}$$

The population in the United States in year 1790 is 3.9×10^6 **.**

The corresponding population in year 1800 is 5.3×10^6 **.**

With a data fitting, we obtain:

$$P(t) = 3.9 \times 10^6 e^{0.0307(t-1790)}$$



Logistic Population Model

- Developed by Belgian mathematician Pierre Verhulst (1838) in 1838
- The rate of population increase may be limited, i.e., it may depend on population density

$$P(t+\Delta t) = P(t) + k(P(t))\Delta t$$

where

$$k(P(t)) = k_0\left(1 - \frac{P(t)}{P_m}\right)P(t)$$

The solution is

$$P(t) = P_0 e^{k_0(t-t_0)} \frac{P_m}{P_0 e^{k_0(t-t_0)} + (P_m - P_0)} = \frac{P_m}{1 + (\frac{P_m}{P_0} - 1)e^{-k_0(t-t_0)}}$$

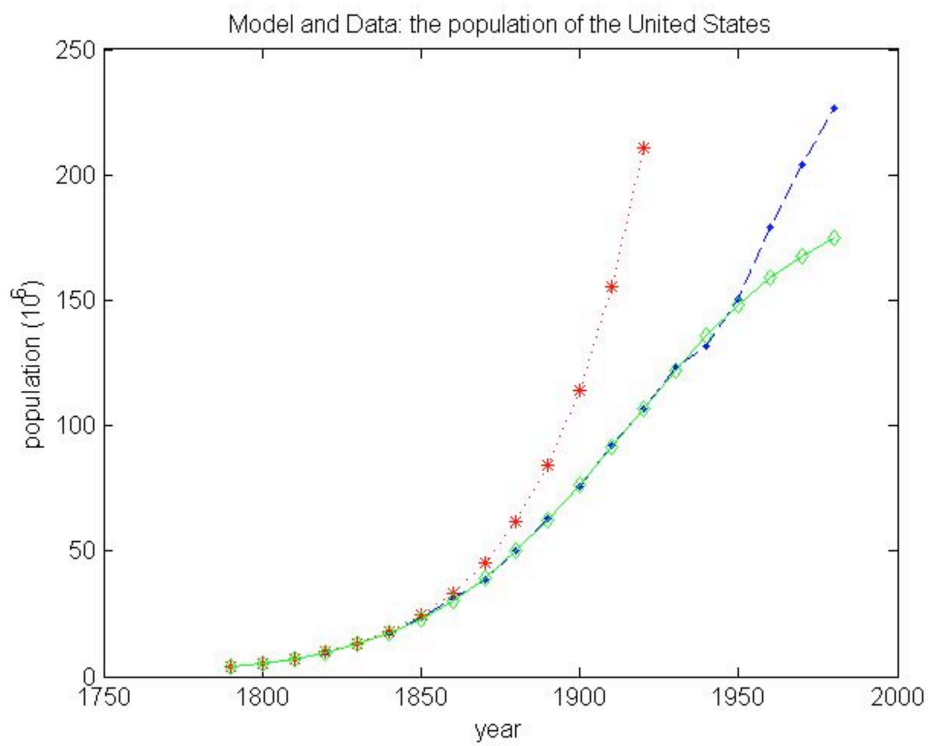
Logistic Population Model

The solution of the Logistic model

$$P(t) = P_0 e^{k_0(t-t_0)} \frac{P_m}{P_0 e^{k_0(t-t_0)} + (P_m - P_0)} = \frac{P_m}{1 + (\frac{P_m}{P_0} - 1)e^{-k_0(t-t_0)}}$$

With a data fitting

$$P_m = 197 \times 10^6, \qquad k_0 = 0.03134$$



Model of two species (Competition)

Let the population of two species be $x(t)$ and $y(t)$, and they compete in the same environment. If there is no competition, the population of X will satisfy

$$\dot{x}(t) = r_1 x(t)(1 - \frac{x}{N_1})$$

With the competition,

$$\dot{x}(t) = r_1 x(t)(1 - \frac{x + \alpha y}{N_1})$$

For another species, there is a similar equation

$$\dot{y}(t) = r_2 y(t)(1 - \frac{y + \beta x}{N_2})$$

The physical meaning of α and β can be understood as:

$$\alpha = \frac{\text{the resource each X species consume}}{\text{the resource each Y species consume}}.$$

Thus we have

$$\alpha\beta = 1$$

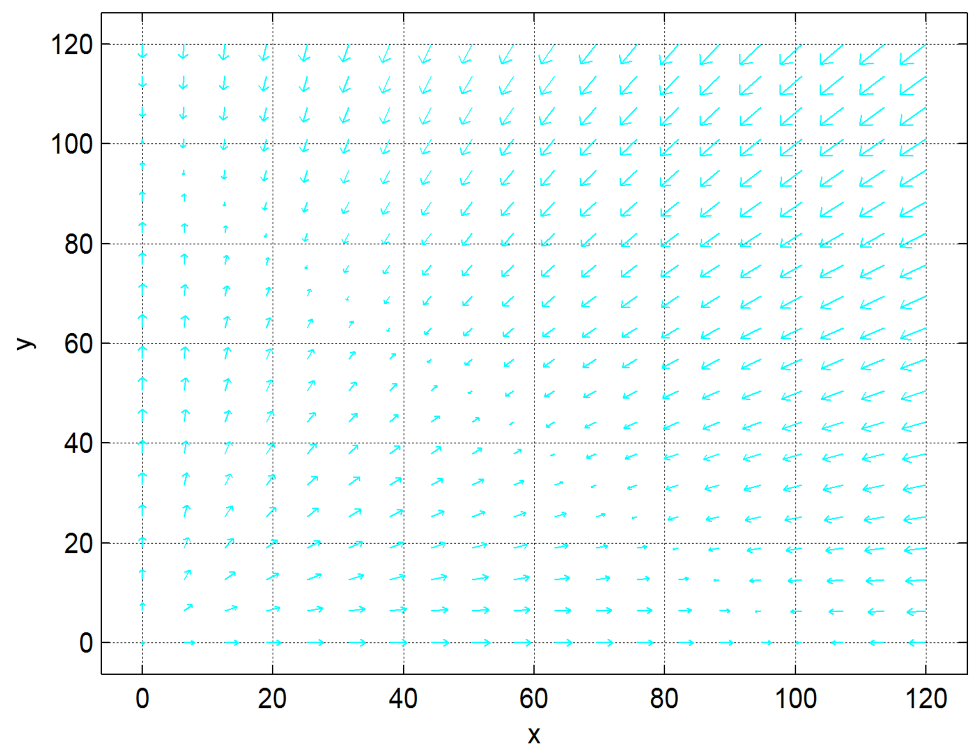
State Dynamics Plot vs Phase Plot

State Dynamics Plot: state vs time,

Phase Plot: the state space, use arrow to represent the tangent vector

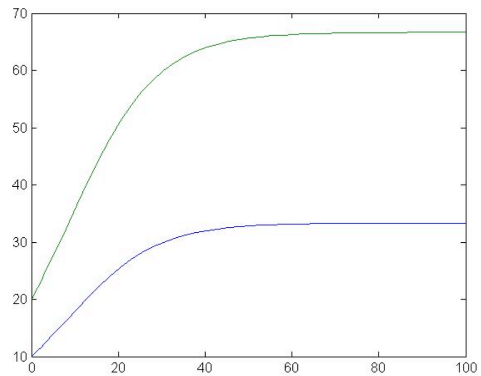
The phase plot reveals the geometric property of a dynamic system represented by a pair of ODEs.

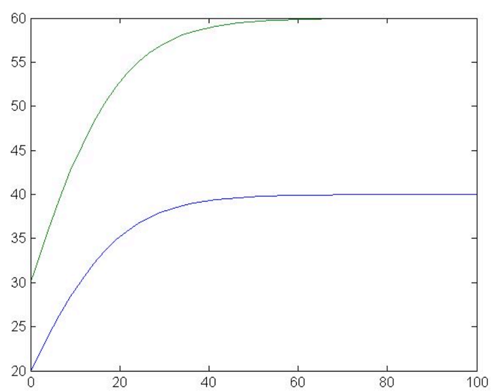
$$\begin{cases} \dot{x}(t) = 0.1x(1 - \dfrac{x+y}{100}) \\ \dot{y}(t) = 0.1y(1 - \dfrac{x+y}{100}) \end{cases}$$

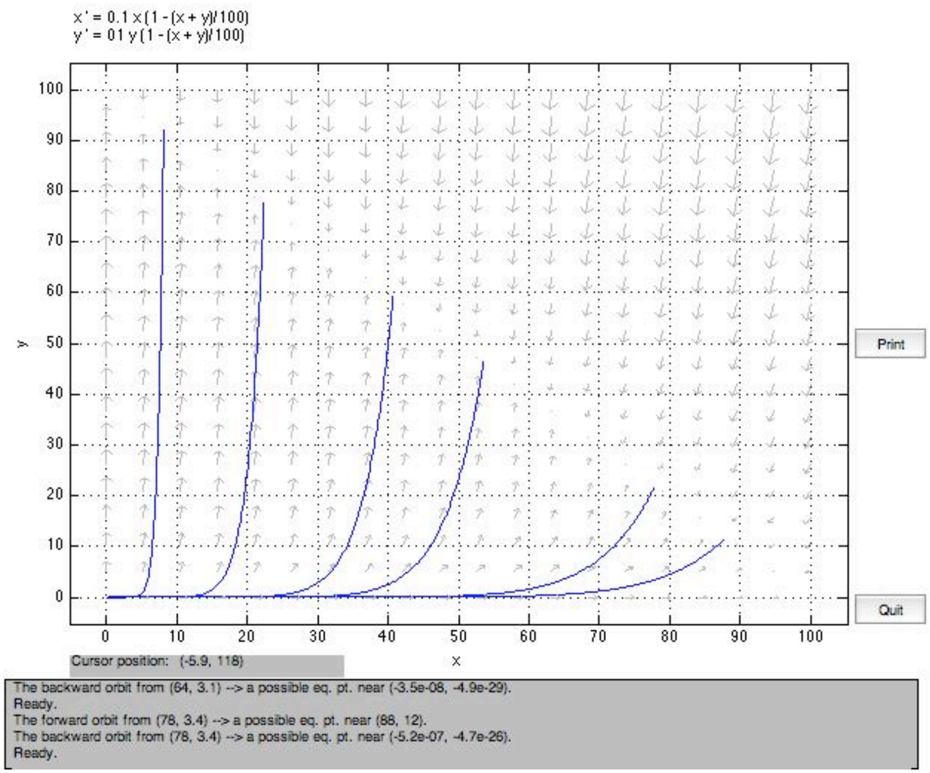


State Dynamics Plot vs Phase Plot

Example: from different initial value, the trajectory follow the direction of the arrows and reaches to its equilibrium state





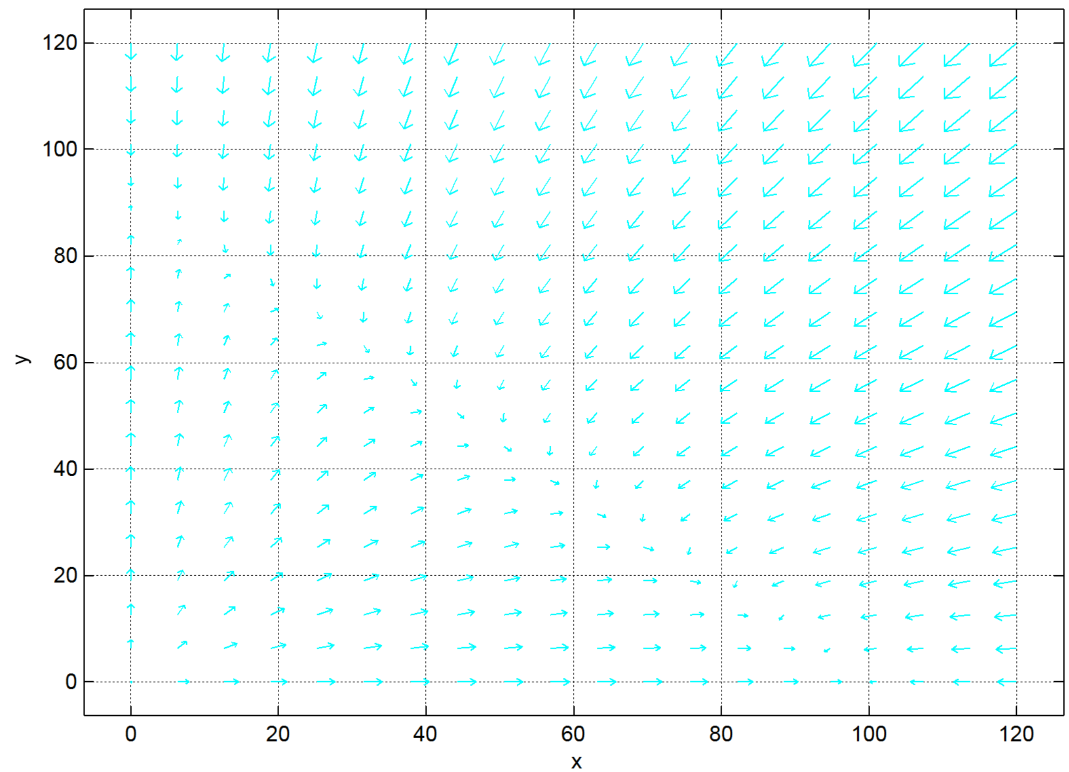


State Dynamics Plot vs Phase Plot

Computational Science and Engineering

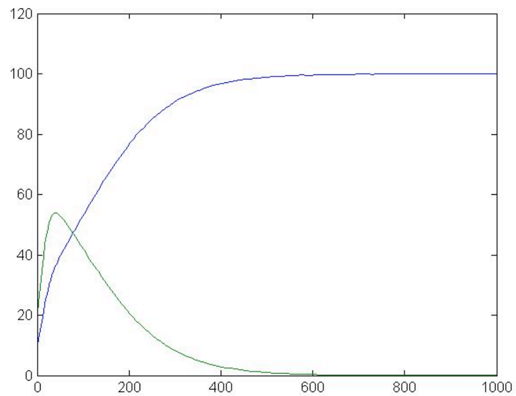
However, a slight change of parameters make a big difference in phase plot and lead to a different conclusion

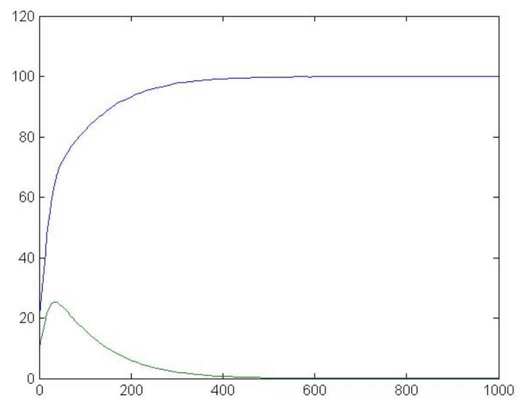
$$\begin{cases} \dot{x}(t) = 0.1x(1-\dfrac{x+y}{100}) \\ \dot{y}(t) = 0.1y(1-\dfrac{x+y}{90}) \end{cases}$$

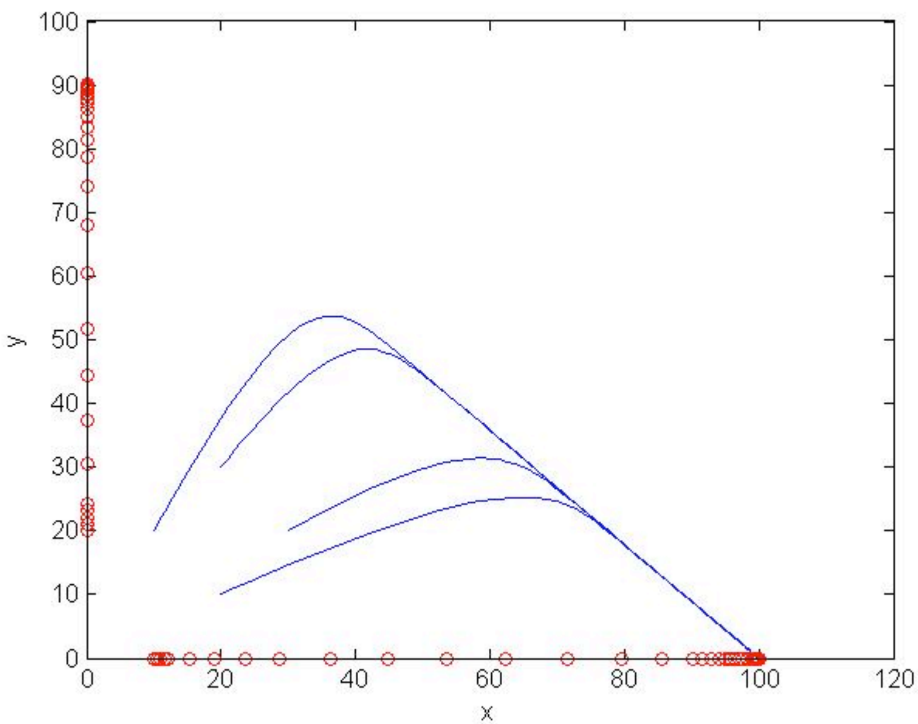


State Dynamics Plot vs Phase Plot

$$\begin{cases} \dot{x}(t) = 0.1x(1 - \frac{x+y}{100}) \\ \dot{y}(t) = 0.1y(1 - \frac{x+y}{90}) \end{cases}$$







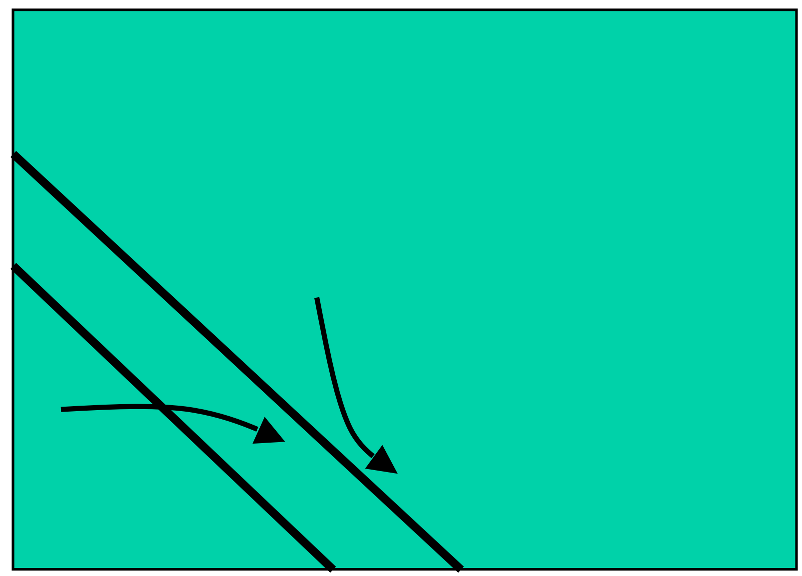
State Dynamics Plot vs Phase Plot

A direct analysis through the phase plot

$$\begin{cases} \dot{x}(t) = r_1 x(1 - \dfrac{x + \alpha y}{N_1}) \\ \dot{y}(t) = r_2 y(1 - \dfrac{\beta x + y}{N_2}) \end{cases}$$

The sign of the derivatives are decided by two values

$N_1 - (x + \alpha y)$ **and** $\alpha N_2 - (x + \alpha y)$

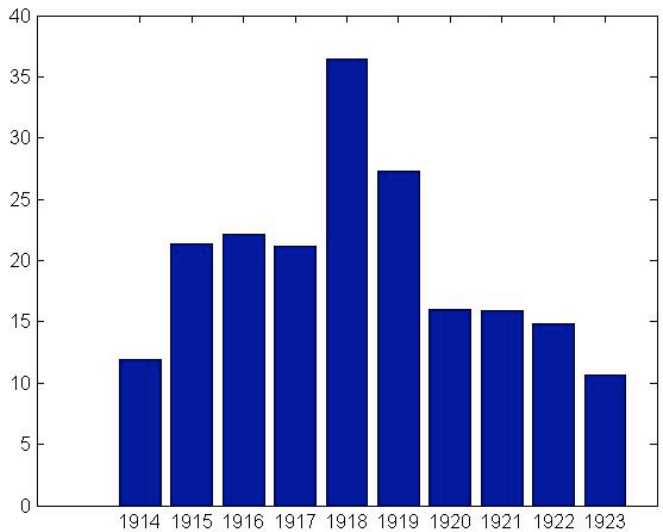


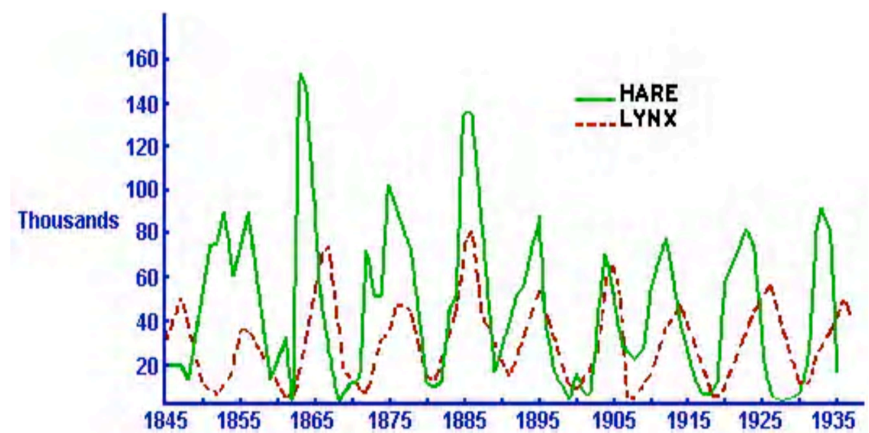
If $N_1 > \alpha N_2$ **, X species will win.**

If $N_1 < \alpha N_2$ **, Y species will win.**

Model of two species (Predator and Prey)

- **Lotka-Volterra Model**
- **The simplest model of predator-prey interactions developed independently by Lotka (1925) and Volterra (1926)**
- **Ancona's observation on Shark's population during world war I.**





Model of two species (Predator and Prey)

Assumption:

- **The predator species is totally dependent on a single prey species as its only food supply,**
- **The prey species has an unlimited food supply, and there is no threat to the prey other than the specific predator.**

Let X represent the prey and Y represent the predator, without the predator, the Malthus model can be applied

$$\dot{x} = ax$$

However, because of the predator, r has to be modified

$$\dot{x} = (a - by)x$$

For the predator, the situation is just the opposite.

$$\dot{y} = (-c + dx)y$$

Thus we get the ODEs for this model

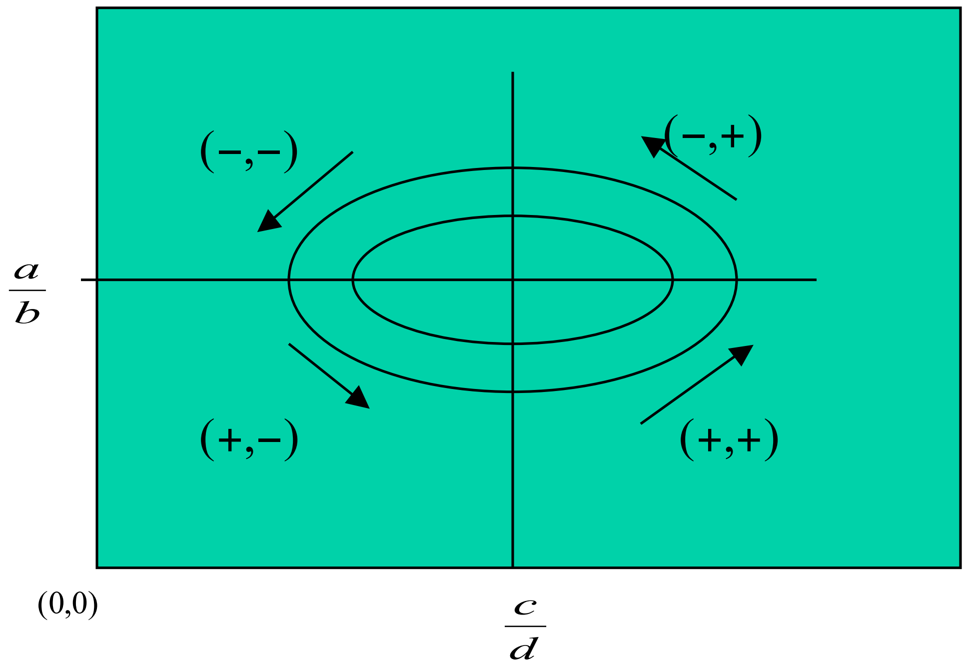
$$\begin{cases} \dot{x} = (a - by)x \\ \dot{y} = (-c + dx)y \end{cases}$$

Phase Plot Analysis

$$\begin{cases} \dot{x} = (a - by)x \\ \dot{y} = (-c + dx)y \end{cases}$$

There are two corresponding equilibrium points:

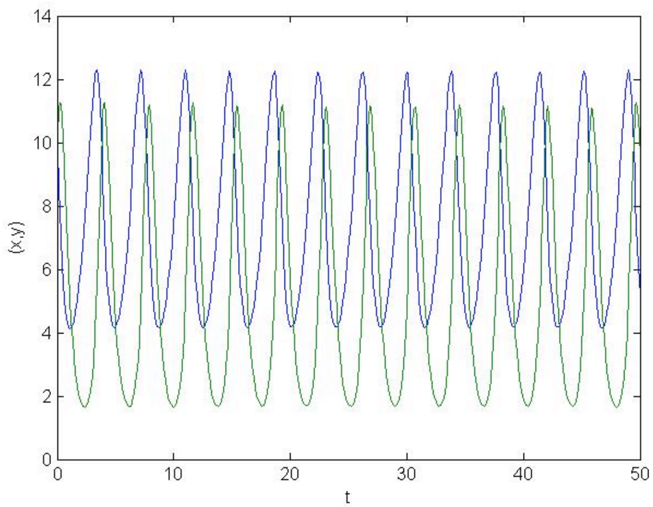
$(0,0)$ **or** $(\frac{c}{d}, \frac{a}{b})$

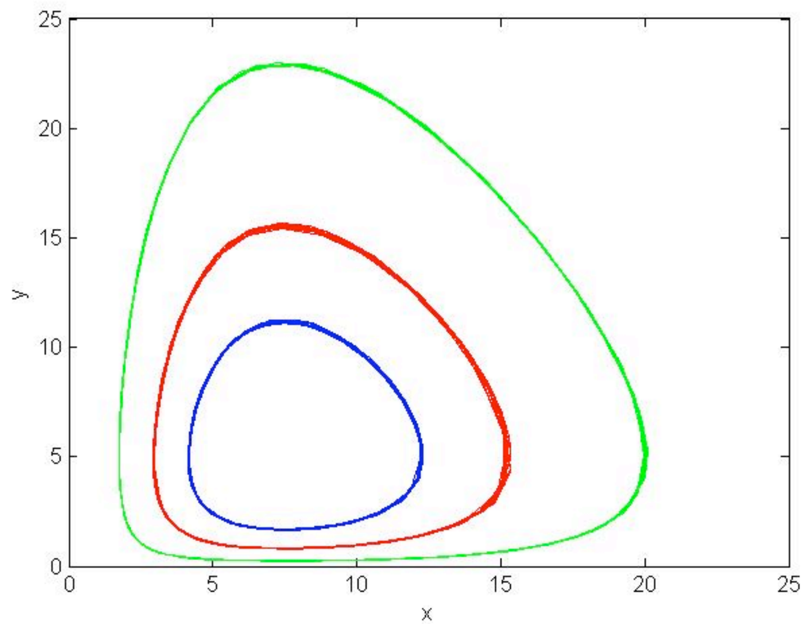


Matlab Simulation Result

Based on example:

$$\begin{cases} \dot{x} = (1 - 0.2y)x \\ \dot{y} = (-3 + 0.4x)y \end{cases}$$





SIR Model (Kermak – McKendrick Model)

S: Susceptible

I: Infected

R: Recovered/Removed

$$\frac{dS}{dt} = -\beta S(t) I(t)$$

$$\frac{dI}{dt} = \beta S(t) I(t) - \gamma I(t)$$

$$\frac{dR}{dt} = -\gamma I(t)$$

$$S(t) + I(t) + R(t) = \mathrm{N} = \text{the total population}$$

SIR Model

From the original model, we have

$$\frac{dS}{dt} = -\beta S(t)I(t)$$

$$\frac{dI}{dt} = \beta S(t)I(t) - \gamma I(t) = [\beta S(t) - \gamma]I(t)$$

Note that this type of equations is similar to what we've seen before.

The Simplest Chemical Reaction

- Decaying Process

This process can be modeled as

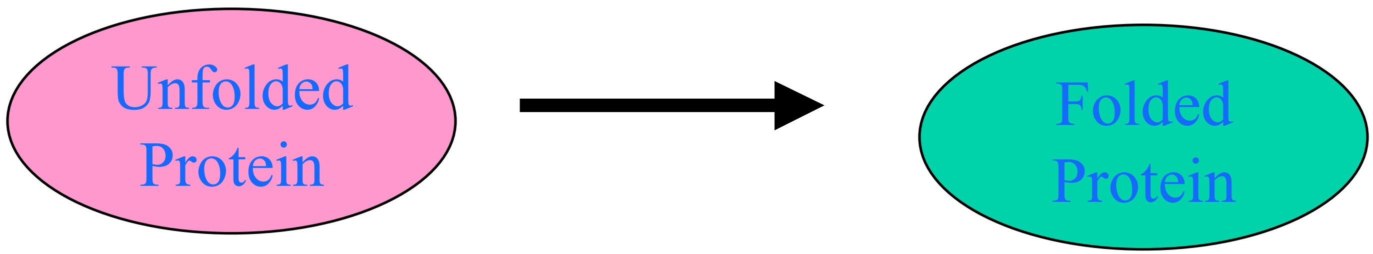
$$S \xrightarrow{k} \varnothing$$

which can be further formulated into reaction rate equations (RREs)

$$\frac{dx}{dt} = -kx$$

Simple Chemical Reaction

- Isomerization



This process can be modeled as

$$S_1 \xrightarrow{k} S_2$$

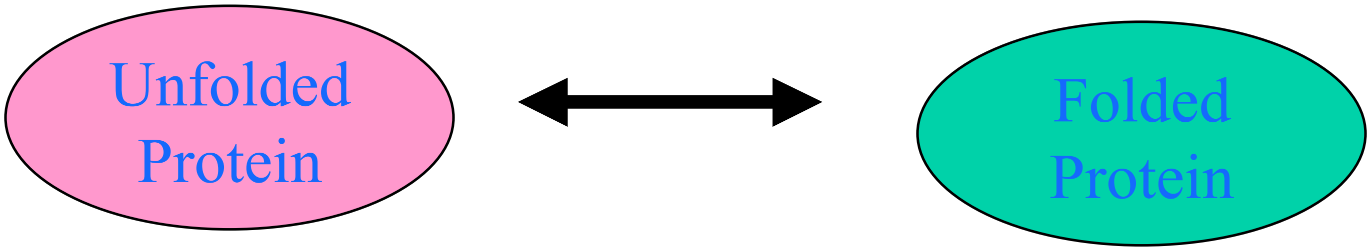
which can be further formulated into RREs

$$\frac{dx_1}{dt} = -kx_1$$

$$\frac{dx_2}{dt} = kx_1$$

Simple Chemical Reaction

- Reversible Isomerization



This process can be modeled as

$$S_1 \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} S_2$$

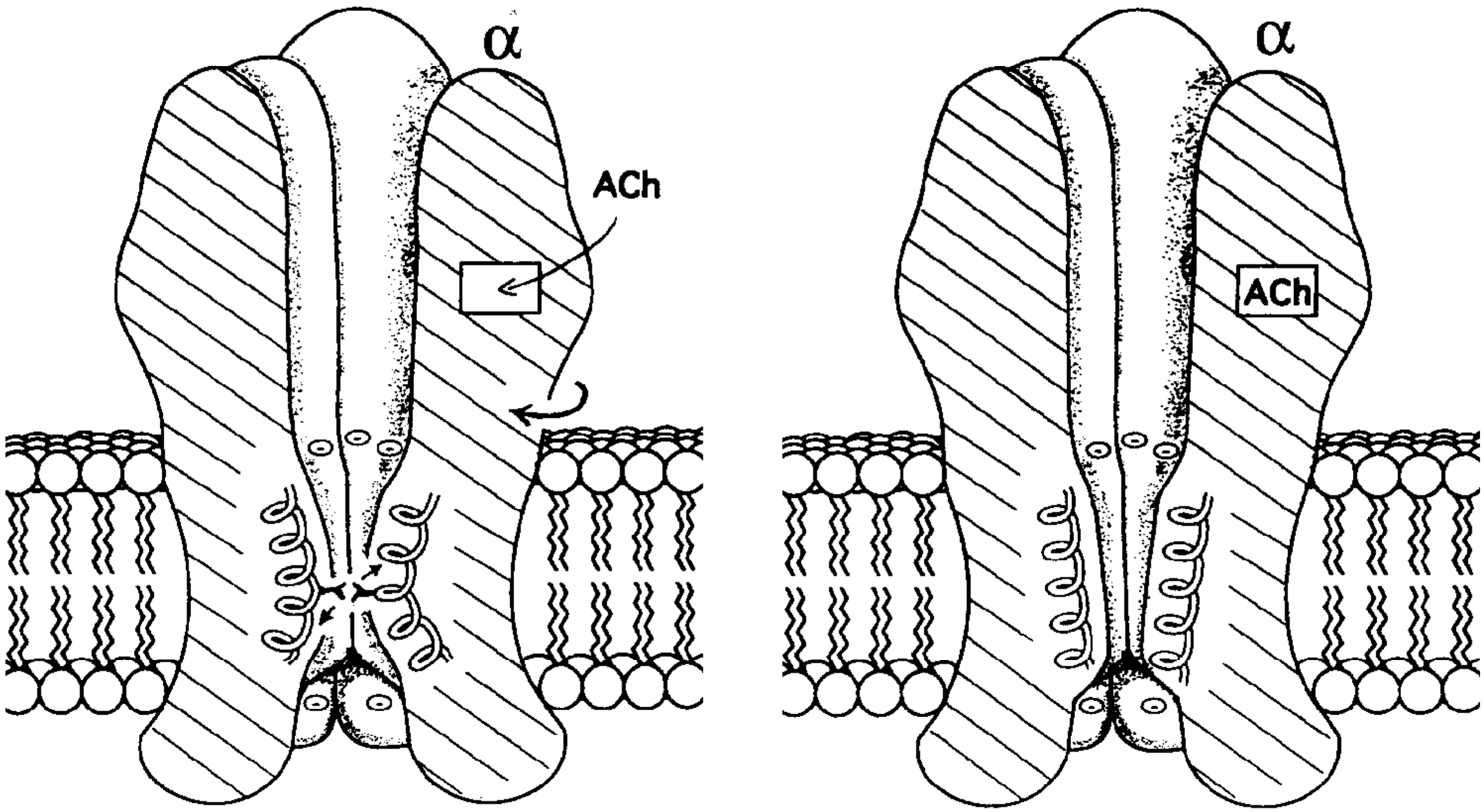
which can be further formulated into RREs

$$\frac{dx_1}{dt} = -k_1 x_1 + k_{-1} x_2$$

$$\frac{dx_2}{dt} = k_1 x_1 - k_{-1} x_2$$

Channel Gating Mechanisms

AChR: Proposed gating mechanism (Unwin, 1995)

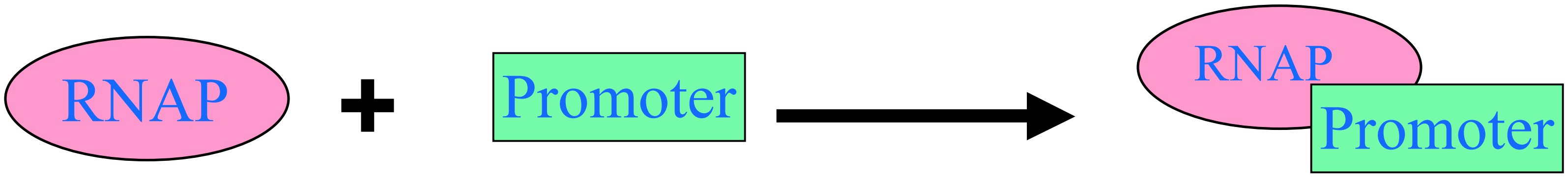


Closed

Open

Simple Chemical Reaction

- Dimerization (Bi-molecular Reaction)



This process can be modeled as

$$S_1 + S_2 \xrightarrow{k} S_3$$

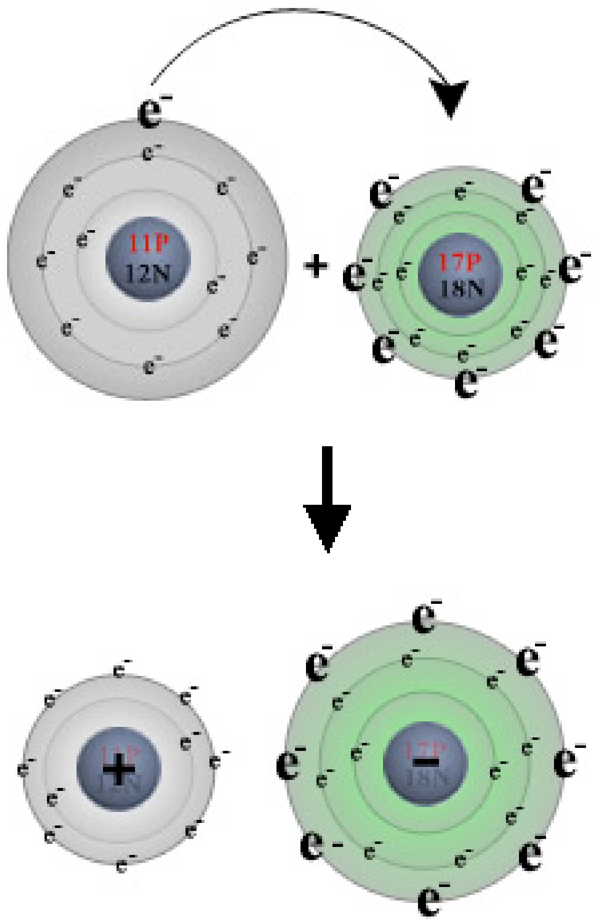
which can be further formulated into RREs

$$\frac{dx_1}{dt} = -kx_1x_2$$

$$\frac{dx_2}{dt} = -kx_1x_2$$

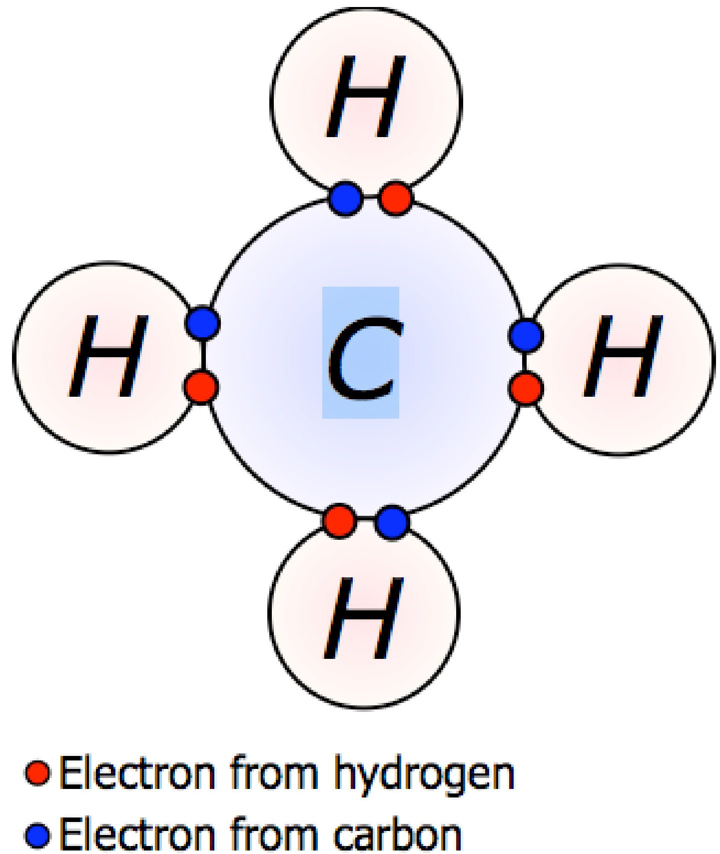
$$\frac{dx_3}{dt} = kx_1x_2$$

Bi-molecular Reaction



Na + Cl = Na Cl

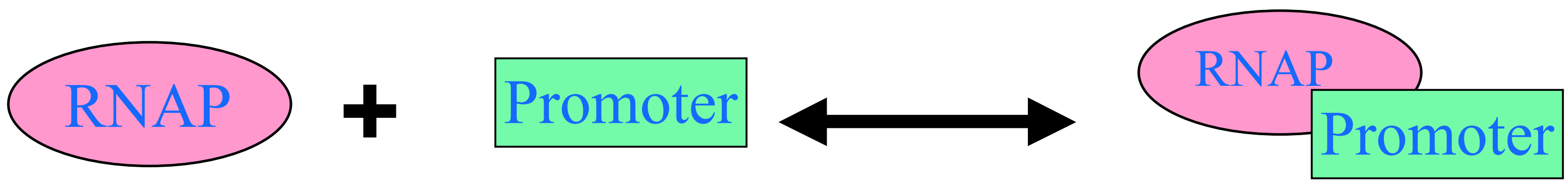
Ionic bond



Covalent bond

Simple Chemical Reaction

- Reversible Dimerization



This process can be modeled as

$$S_1 + S_2 \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} S_3$$

which can be further formulated into RRE

$$\frac{dx_1}{dt} = -k_1 x_1 x_2 + k_{-1} x_3$$
$$\frac{dx_2}{dt} = -k_1 x_1 x_2 + k_{-1} x_3$$
$$\frac{dx_3}{dt} = k_1 x_1 x_2 - k_{-1} x_3$$

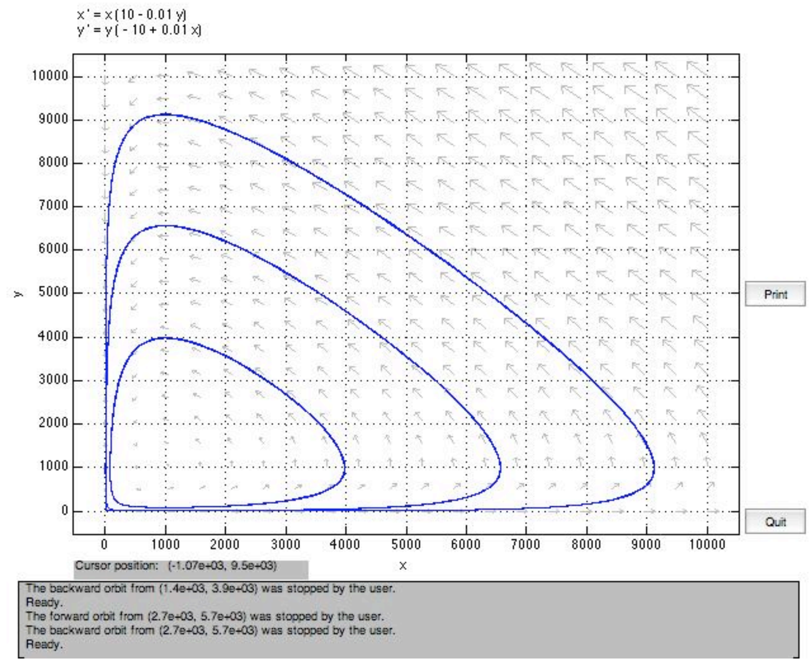
Chemically Reacting Network

Lotka reactions:

$$A + X \xrightarrow{c_1} 2X$$

$$X + Y \xrightarrow{c_2} 2Y$$

$$Y \xrightarrow{c_3} Z$$



Lead to ODEs

$$\begin{cases} \dot{x} = (c_1 A - c_2 y)x \\ \dot{y} = (-c_3 + c_2 x)y \end{cases}$$

We can use the following values to simulate this system.

$$c_1 A = 10,$$

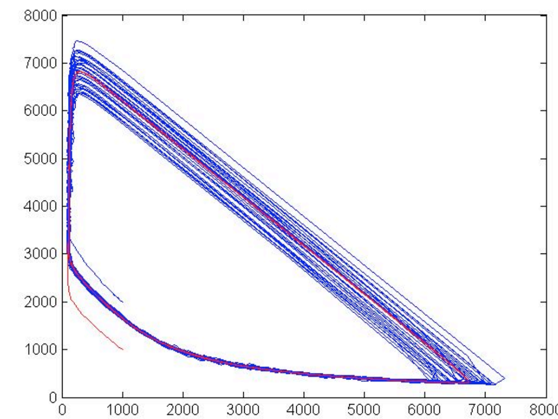
$$c_2 = 0.01,$$

$$c_3 = 10$$

Brusselator

$$A \xrightarrow{c_1} X$$
$$B + X \xrightarrow{c_2} Y + C$$
$$2X + Y \xrightarrow{c_3} 3X$$
$$X \xrightarrow{c_4} D$$

$c_1 A = 5000,$
$c_2 B = 50,$
$c_3 = 0.00005,$
$c_4 = 5.$



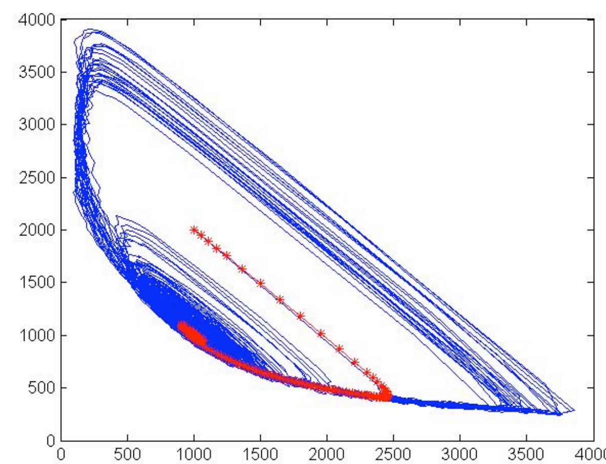
Lead to ODEs

$$\begin{cases} \dot{x} = c_1 A - c_2 B x + \frac{c_3}{2} x^2 y - c_4 x \\ \dot{y} = c_2 B x - \frac{c_3}{2} x^2 y \end{cases}$$

Bifurcation happens around the condition:

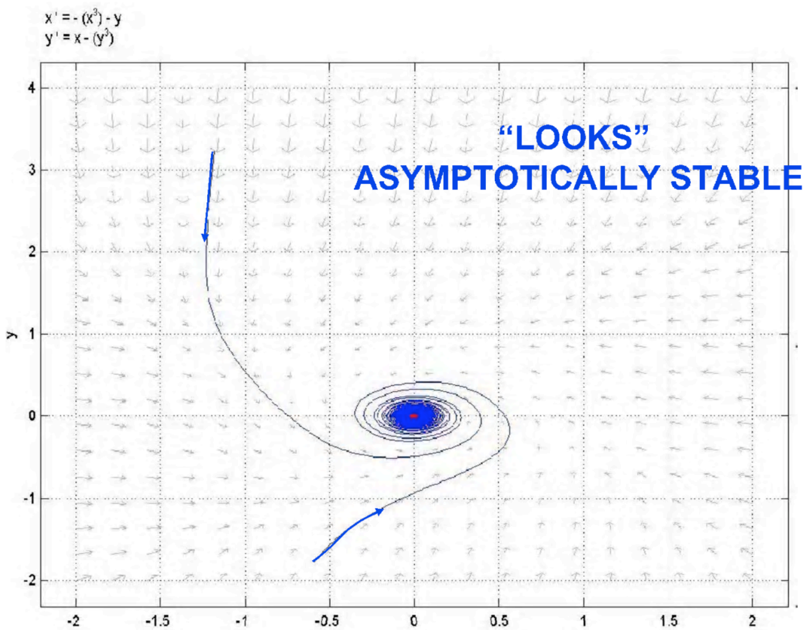
$$\frac{2 c_2 B}{c_3} = \frac{(c_1 A)^2}{c_4^2} + \frac{2 c_4}{c_3}$$

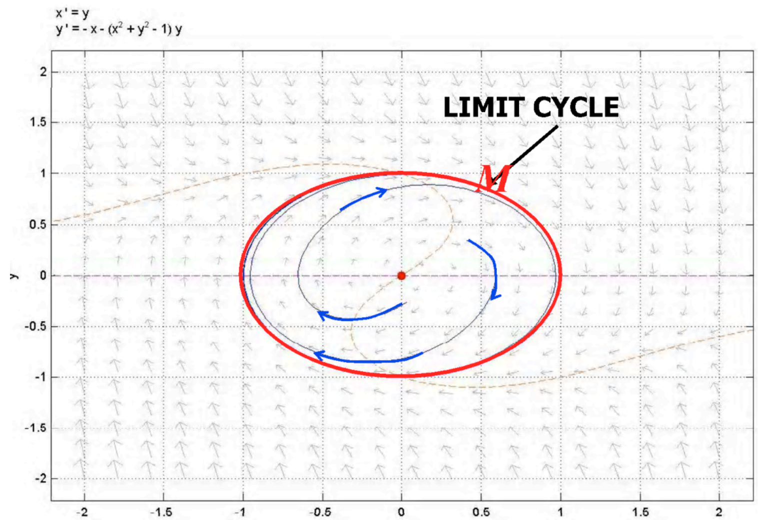
$c_1 A = 5000,$
$c_2 B = 50,$
$c_3 = 0.0001,$
$c_4 = 5.$



J. Tyson's 1973, 1974 paper

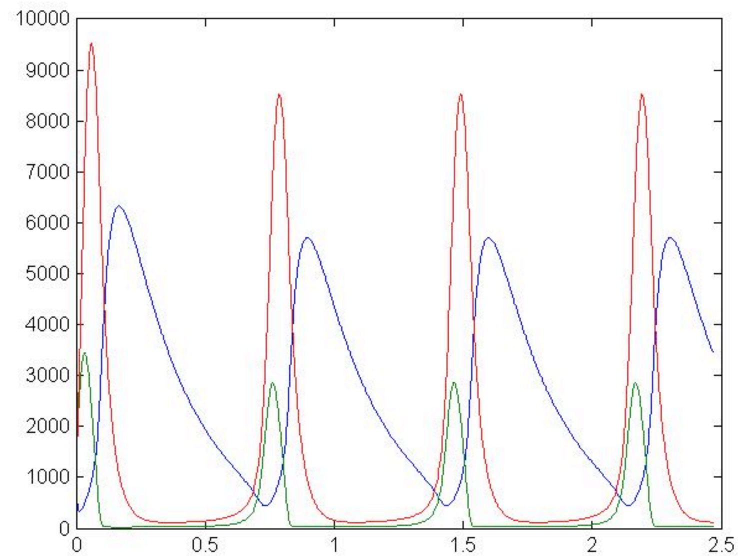
Different Dynamic Behavior





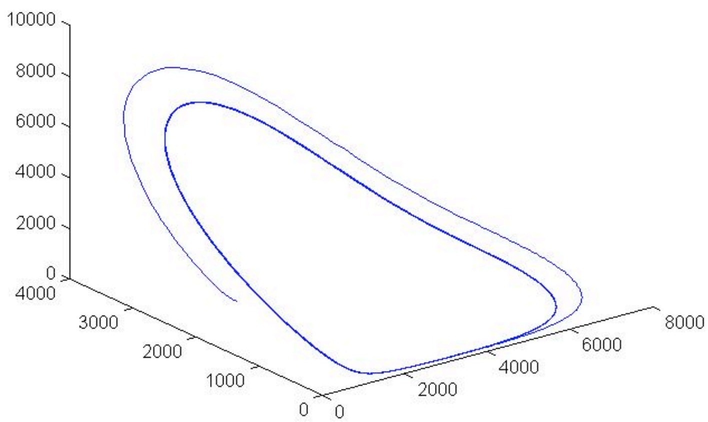
Oregonator

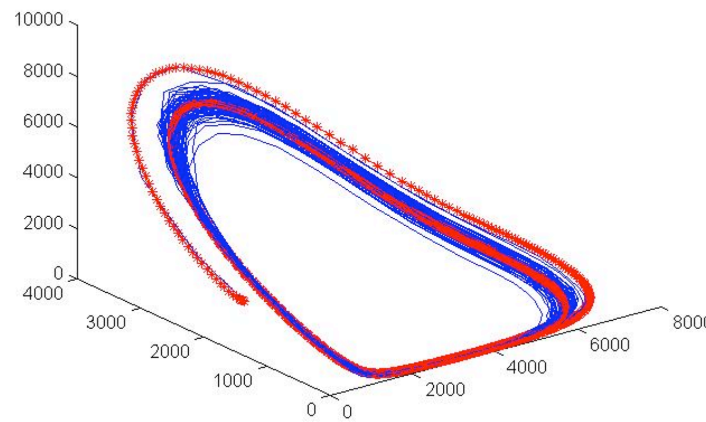
$$A + X \xrightarrow{c_1} Y$$
$$X + Y \xrightarrow{c_2} B$$
$$C + Y \xrightarrow{c_3} 2Y + Z$$
$$2Y \xrightarrow{c_4} D$$
$$E + Z \xrightarrow{c_5} X$$



$$\begin{cases} \dot{x} = -c_1 Ax - c_2 xy + c_5 Ez \\ \dot{y} = c_1 Ax - c_2 xy + c_3 Cy - c_4 y^2 \\ \dot{z} = c_3 Cy - c_5 Ez \end{cases}$$

$c_1 A$

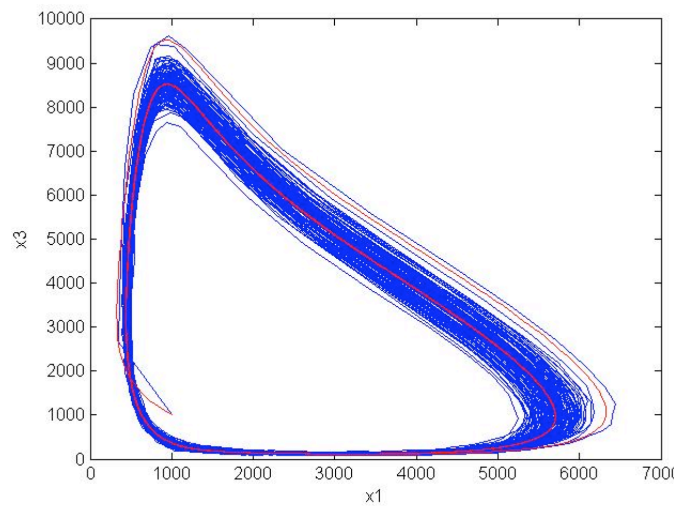


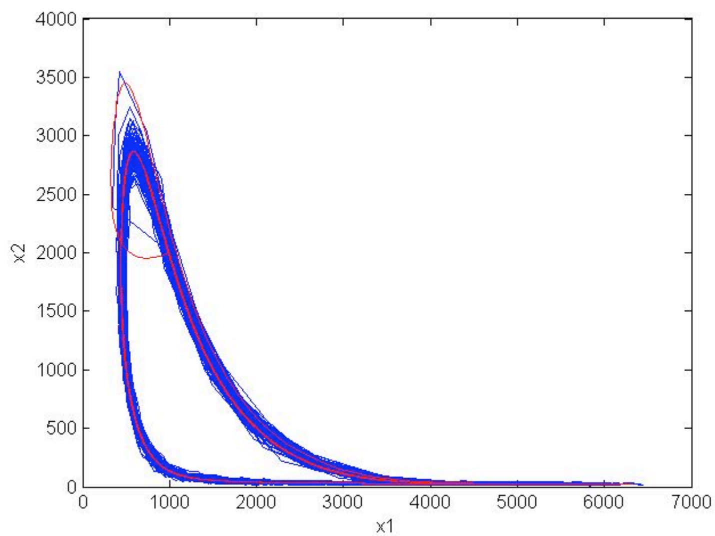


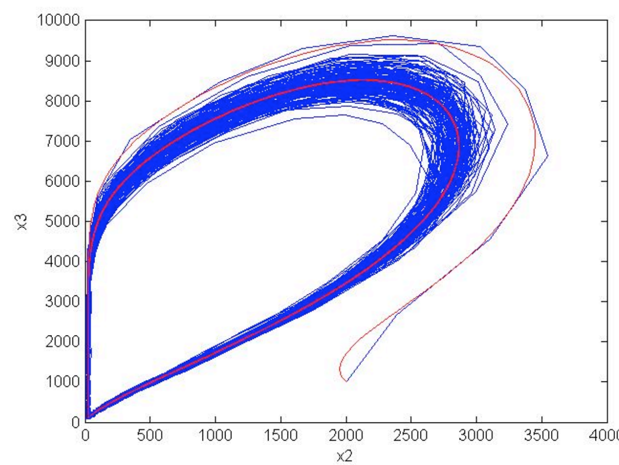
Oregonator

$$A + X \xrightarrow{c_1} Y$$
$$X + Y \xrightarrow{c_2} B$$
$$C + Y \xrightarrow{c_3} 2Y + Z$$
$$2Y \xrightarrow{c_4} D$$
$$E + Z \xrightarrow{c_5} X$$

$$c_1 A = 2, \mathrm{c}_2 = 0.1, \mathrm{c}_3 C = 104, \mathrm{c}_4 = 0.016, \mathrm{c}_5 E = 26$$

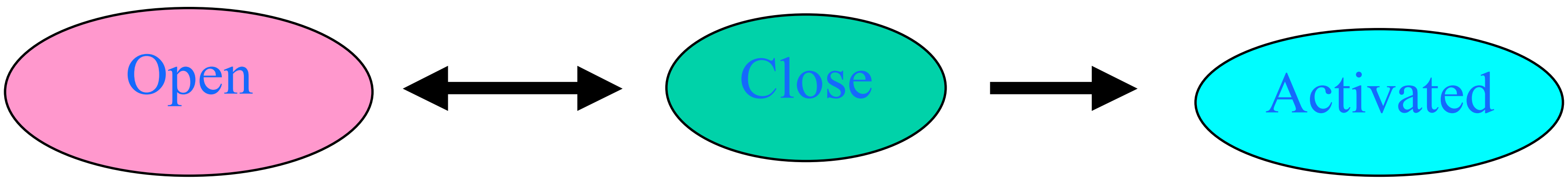






Fast and Slow Scales

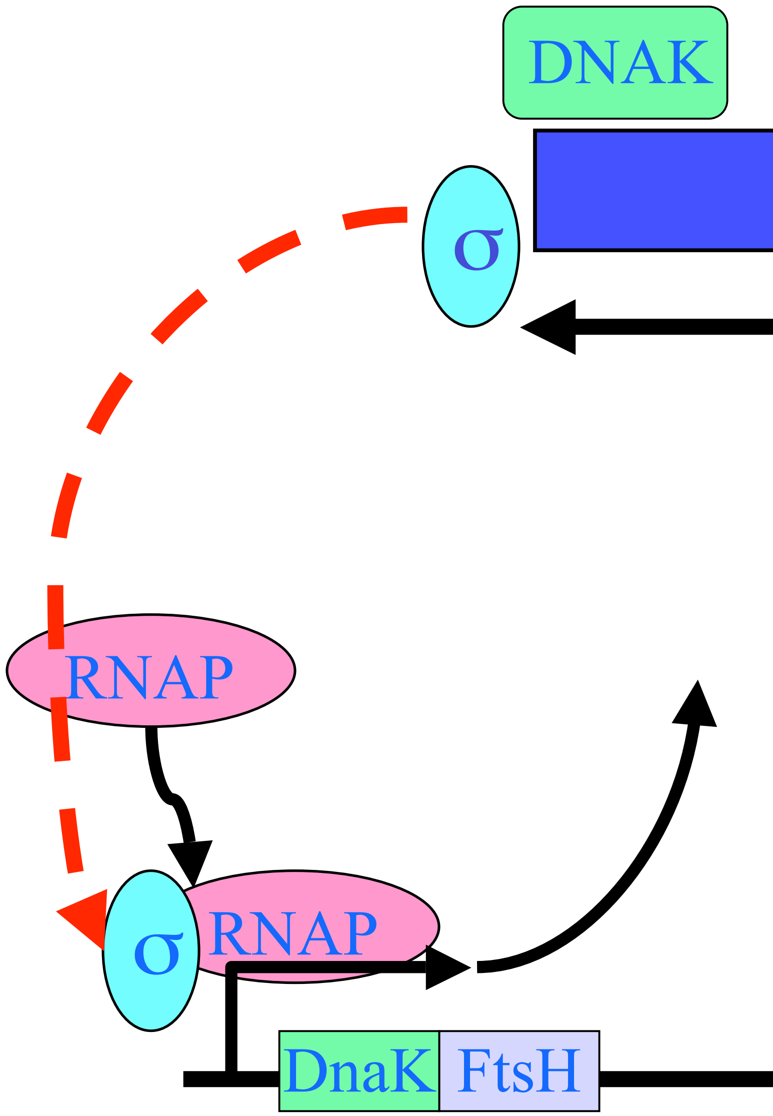
• Many practical chemically reacting systems show different time scales in different reactions. The simplest example is given by the following:



$$S_1 \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} S_2 \xrightarrow{k_2} S_3$$

Where the first two reactions are much faster than the third.

Stiffness in the Heat Shock Response (HSR) model



The total "concentration" of σ_{32} is 30-100 per cell

But the "concentration" of free σ_{32} is .01-.05 per cell

- **The multiscale behavior can be modeled in the following simple model:**

$$S_1 + S_2 \longleftrightarrow S_3$$

$$S_1 + S_4 \longrightarrow S_5$$

or a simpler model

$$S_1 \longleftrightarrow S_2$$

$$S_1 \longrightarrow S_3$$

- **Features**
  - Fast and slow reactions
  - Fast reactions usually "less important" than slow ones

MM Equation in Enzyme Kinetics

Consider the following enzyme-substrate system

$$E + S \longleftrightarrow ES \longrightarrow E + P$$

- **Partial Equilibrium Assumption**
- **Quasi-Steady State Assumption**
- **Total Quasi-Steady State Assumption**

Exercise

Exercise

1. Write down the ordinary differential equations for the following chemical reactions:

$$S_1 + S_2 \xrightarrow{k_1} S_1 + S_1$$
$$S_1 \xrightarrow{k_2} \varnothing$$
$$S_2 \xrightarrow{k_3} S_2 + S_2$$

where $k_1 = 0.01, \quad k_2 = k_3 = 10.$

2. Choose one of the following two exercises:
   a. Simulate this ordinary differential equation in Matlab with the initial condition $S_1 = S_2 = 1000$.. Plot the trajectory for these two variables.
   b. In Matlab, use pplane7 to study its dynamic behavior. Will this system oscillate, or tend to a limit point?