Problem solving by inverse methods in systems biology

Peter Schuster

Institut für Theoretische Chemie, Universität Wien, Austria
and
The Santa Fe Institute, Santa Fe, New Mexico, USA

High-performance computational and systems biology
HiBi 2010

Twente, 30.09.– 01.10.2010

Web-Page for further information:

http://www.tbi.univie.ac.at/~pks

Heinz W. Engl, Christoph Flamm, Philipp Kügler, James Lu, Stefan Müller, Peter Schuster. 2009.

Inverse problems in systems biology.

Inverse Problems 25:123014 (51pp)

Free download until December 31, 2010.

IOP PUBLISHING INVERSE PROBLEMS

Inverse Problems **25** (2009) 123014 (51pp) doi:10.1088/0266-5611/25/12/123014

TOPICAL REVIEW

Inverse problems in systems biology

Heinz W Engl[1], Christoph Flamm[2], Philipp Kügler[3], James Lu[1], Stefan Müller[1] and Peter Schuster[2]

[1] Johann Radon Institute for Computational and Applied Mathematics, Austrian Academy of Sciences, Altenbergerstraße 69, 4040 Linz, Austria
[2] Institute for Theoretical Chemistry, University of Vienna, Währingerstraße 17, 1090 Wien, Austria
[3] Industrial Mathematics Institute, Johannes Kepler University Linz, Altenbergerstraße 69, 4040 Linz, Austria

E-mail: heinz.engl@oeaw.ac.at, xtof@tbi.univie.ac.at, philipp.kuegler@jku.at, james.lu@oeaw.ac.at, stefan.mueller@oeaw.ac.at and pks@tbi.univie.ac.at

Received 7 July 2009, in final form 12 November 2009
Published 3 December 2009
Online at stacks.iop.org/IP/25/123014

Abstract

Systems biology is a new discipline built upon the premise that an understanding of how cells and organisms carry out their functions cannot be gained by looking at cellular components in isolation. Instead, consideration of the interplay between the parts of systems is indispensable for analyzing, modeling, and predicting systems' behavior. Studying biological processes under this premise, systems biology combines experimental techniques and computational methods in order to construct predictive models. Both in building and utilizing models of biological systems, inverse problems arise at several occasions, for example, (i) when experimental time series and steady state data are used to construct biochemical reaction networks, (ii) when model parameters are identified that capture underlying mechanisms or (iii) when desired qualitative behavior such as bistability or limit cycle oscillations is engineered by proper choices of parameter combinations. In this paper we review principles of the modeling process in systems biology and illustrate the ill-posedness and regularization of parameter identification problems in that context. Furthermore, we discuss the methodology of qualitative inverse problems and demonstrate how sparsity enforcing regularization allows the determination of key reaction mechanisms underlying the qualitative behavior.

(Some figures in this article are in colour only in the electronic version)

1. Introduction and motivation

Systems biology is a relatively young biological discipline that claims to consider cells and organisms as entities in a holistic way. At the same time, it focuses on the interplay

What is (computational) systems biology?

Systems biology is an attempt to understand integral systems like cells and whole organisms and their properties by means of the knowledge from molecular biology. The goal of the computational approach is prediction of changes in phenotypes from known changes in molecular structures and environmental conditions.

The current methods apply a combination of *bottom-up* techniques like using data from *in vitro* measurements on isolated molecules and *top-down* data on time series of gene acitivities and metabolite concentrations from array studies.

1. From biochemical kinetics to quantitative biology
2. Forward and inverse problems in reaction kinetics
3. Modeling biochemical reaction kinetics
4. Examples of inverse bifurcation analysis and design
5. Problems and perspectives of systems biology

1. **From biochemical kinetics to quantitative biology**
2. Forward and inverse problems in reaction kinetics
3. Modeling biochemical reaction kinetics
4. Examples of inverse bifurcation analysis and design
5. Problems and perspectives of systems biology

Biochemical kinetics

1910 – 1960 Conventional enzyme kinetics

Biochemical kinetics

1910 – 1960	Conventional enzyme kinetics
1950 – 1975	Theory of biopolymers, macroscopic properties

Biochemical kinetics

1910 – 1960	Conventional enzyme kinetics
1950 – 1975	Theory of biopolymers, macroscopic properties
1958	Gene regulation through repressor binding

Biochemical kinetics

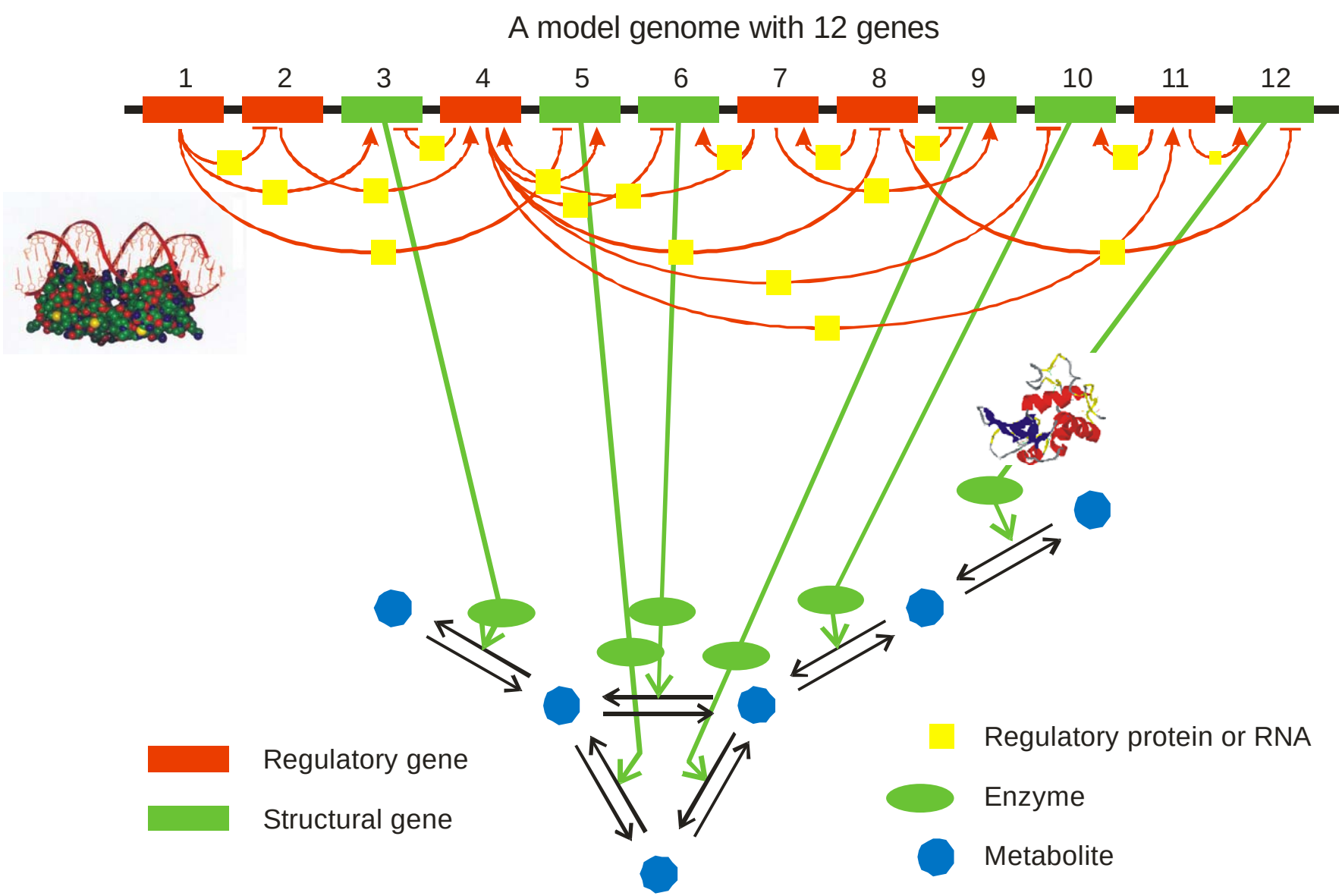
1910 – 1960	Conventional enzyme kinetics
1950 – 1975	Theory of biopolymers, macroscopic properties
1958	Gene regulation through repressor binding
1965 – 1975	Allosteric effects, cooperative transitions

Biochemical kinetics

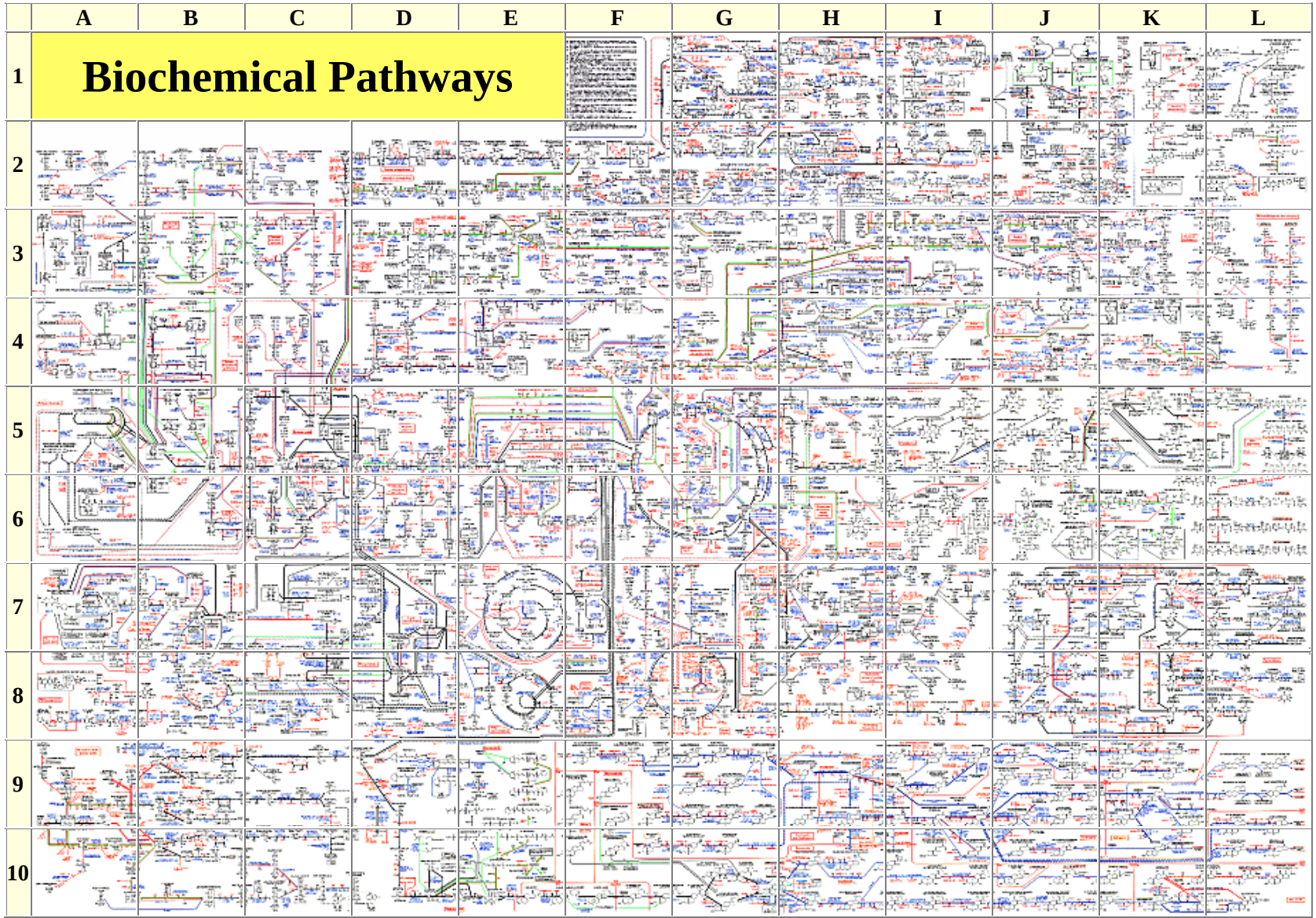
1910 – 1960	Conventional enzyme kinetics
1950 – 1975	Theory of biopolymers, macroscopic properties
1958	Gene regulation through repressor binding
1965 – 1975	Allosteric effects, cooperative transitions
1965 – 1975	Theory of cooperative binding to nucleic acids

Biochemical kinetics

1910 – 1960	Conventional enzyme kinetics
1950 – 1975	Theory of biopolymers, macroscopic properties
1958	Gene regulation through repressor binding
1965 – 1975	Allosteric effects, cooperative transitions
1965 – 1975	Theory of cooperative binding to nucleic acids
1990 -	Revival of biochemical kinetics in systems biology

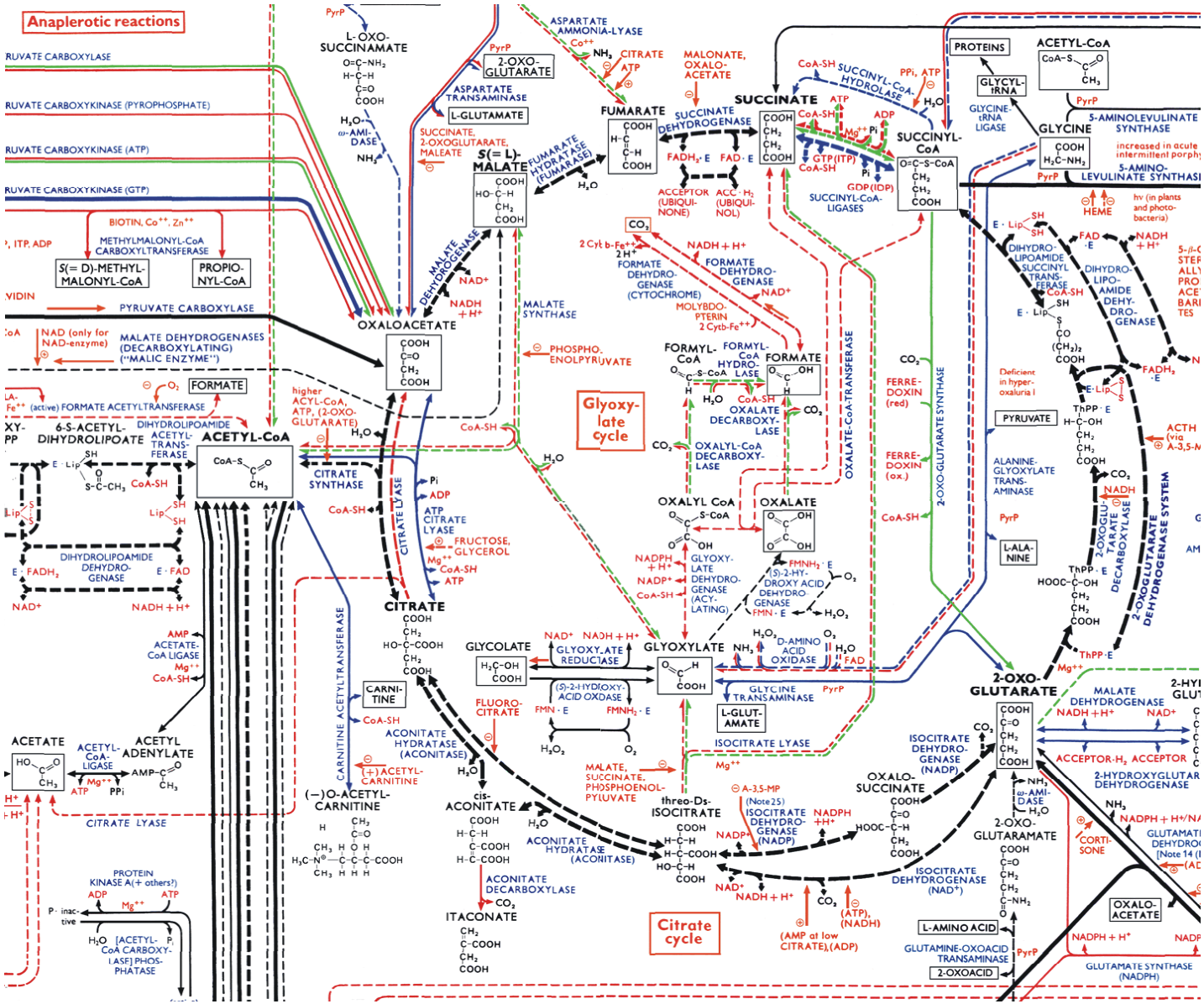


Sketch of a genetic and metabolic network



The reaction network of cellular metabolism published by Boehringer-Mannheim.

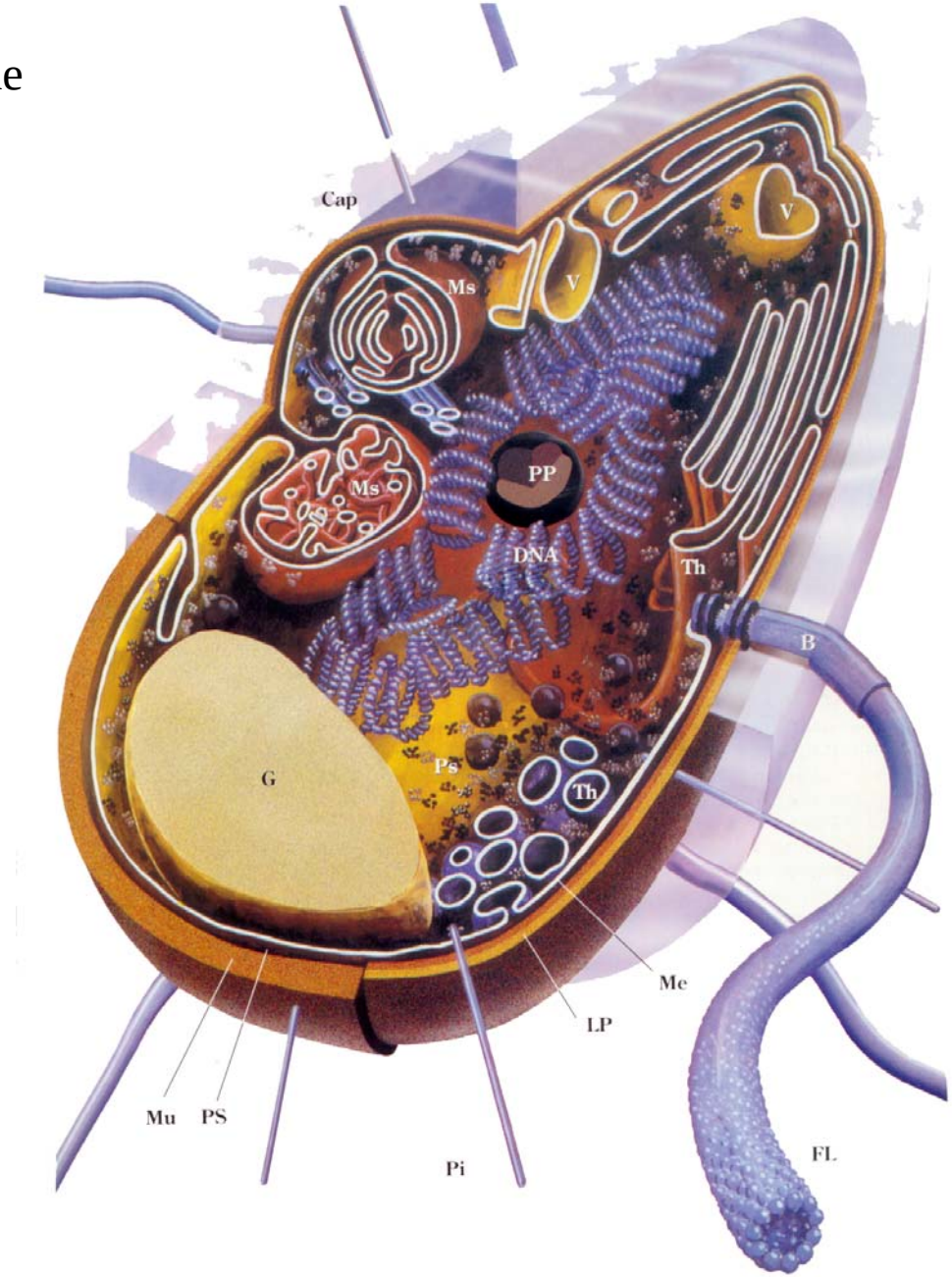
The citric acid or Krebs cycle (enlarged from previous slide).



The bacterial cell as an example for the simplest form of autonomous life

The human body:

10^{14} cells = 10^{13} eukaryotic cells +
$\approx 9\times10^{13}$ bacterial (prokaryotic) cells,
and $\approx$ 200 eukaryotic cell types



The spatial structure of the bacterium *Escherichia coli*

1. From biochemical kinetics to quantitative biology
2. **Forward and inverse problems in reaction kinetics**
3. Modeling biochemical reaction kinetics
4. Examples of inverse bifurcation analysis and design
5. Problems and perspectives of systems biology

Kinetic differential equations

$$\frac{dx}{dt} = f(x;k);\ x=(x_1,\ldots,x_n);\ k=(k_1,\ldots,k_m)$$

Reaction diffusion equations

$$\frac{\partial x}{\partial t} = D\,\nabla^2 x + f(x;k)$$

Parameter set

$k_j(\mathrm{T},\mathrm{p},\mathrm{pH},\mathrm{I},\ldots);\ \mathrm{j}=1,2,\ldots,\mathrm{m}$

General conditions: T, p, pH, I, ...

Initial conditions: $x(0)$

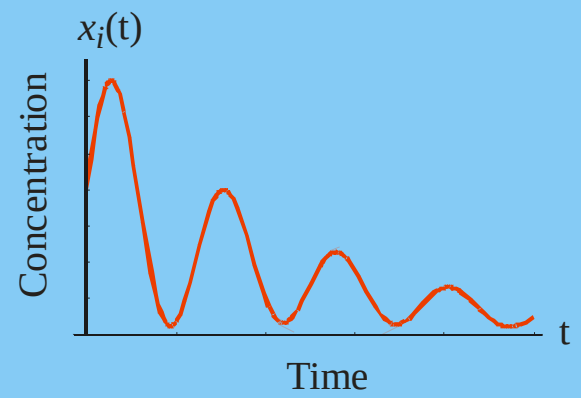
Boundary conditions:

boundary ... S, normal unit vector ... $\hat{u}$

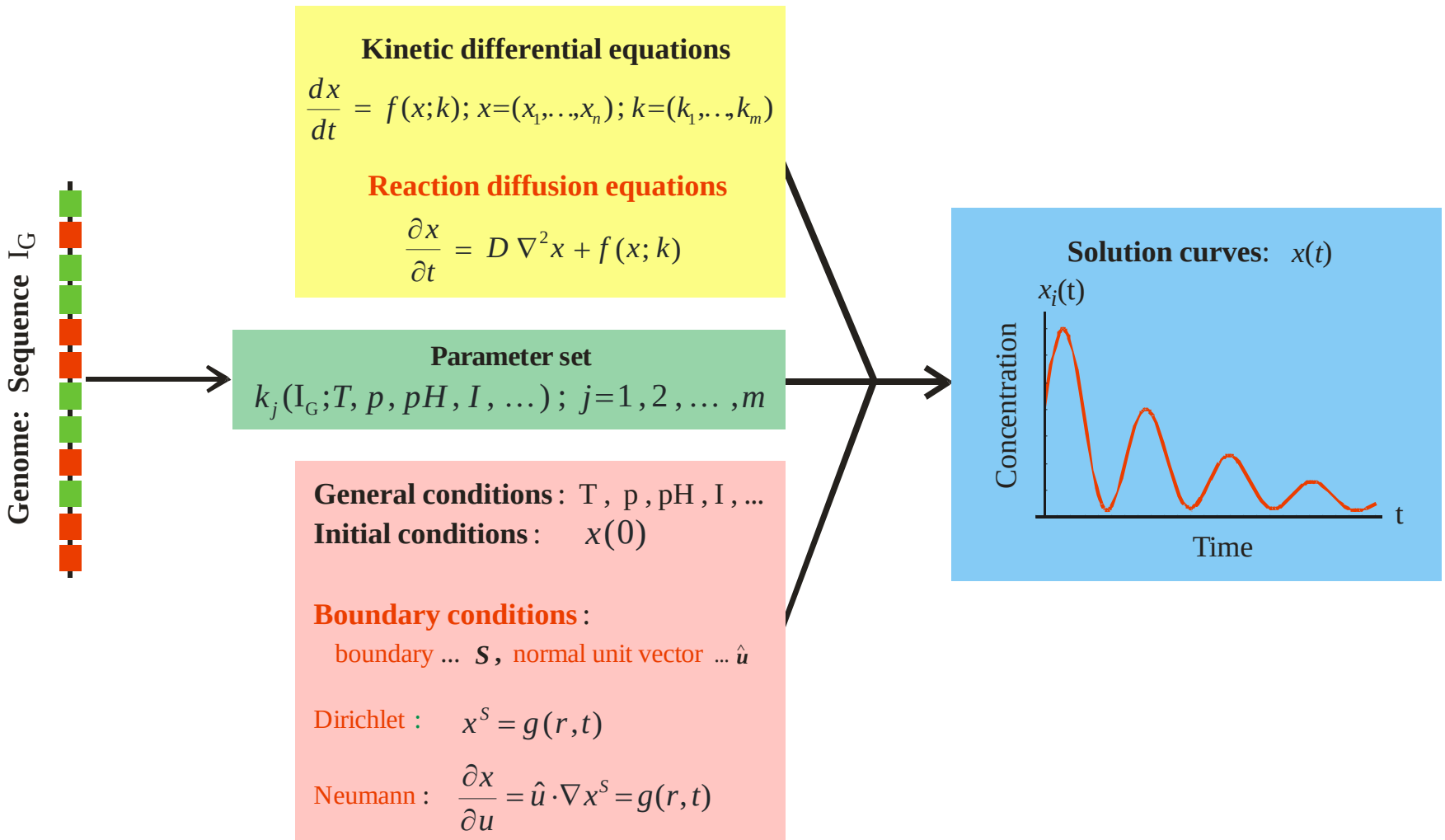
Dirichlet : $x^S = g(r,t)$

Neumann : $\frac{\partial x}{\partial u} = \hat{u}\cdot\nabla x^S = g(r,t)$

Solution curves: $x(t)$



The forward problem of chemical reaction kinetics (Level I)



The forward problem of biochemical reaction kinetics (Level I)

Kinetic differential equations

$$\frac{dx}{dt} = f(x;k);\ x=(x_1,\ldots,x_n);\ k=(k_1,\ldots,k_m)$$

Reaction diffusion equations

$$\frac{\partial x}{\partial t} = D\,\nabla^2 x + f(x;k)$$

General conditions : T , p , pH , I , ...

Initial conditions : $x(0)$

Boundary conditions :

boundary ... S, normal unit vector... $\hat{u}$

Dirichlet : $x^S = g(r,t)$

Neumann : $\frac{\partial x}{\partial u} = \hat{u}\cdot\nabla x^S = g(r,t)$

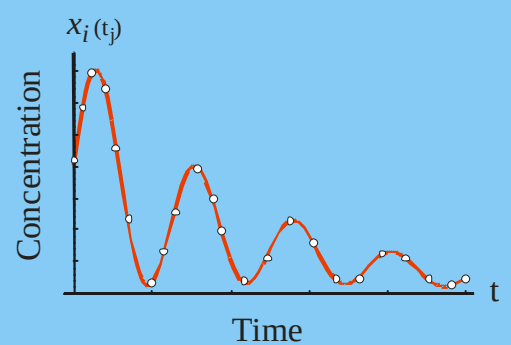
Inverse problem:
Parameter determination

Parameter set

$k_j(\mathrm{I_G};T, p, pH, I, \ldots);\ j=1,2,\ldots,m$

Data from measurements

$x(t_j);\ j = 1, 2, \ldots, N$



The inverse problem of biochemical reaction kinetics (Level I)

Inverse problem:
Parameter determination

Parameter set

$$k_j(I_G; T, p, pH, I, \ldots);\ j = 1, 2, \ldots, m$$

Kinetic differential equations

$$\frac{dx}{dt} = f(x;k);\ x=(x_1,\ldots,x_n);\ k=(k_1,\ldots,k_m)$$

Reaction diffusion equations

$$\frac{\partial x}{\partial t} = D\,\nabla^2 x + f(x;k)$$

General conditions : T , p , pH , I , ...
Initial conditions : $x(0)$

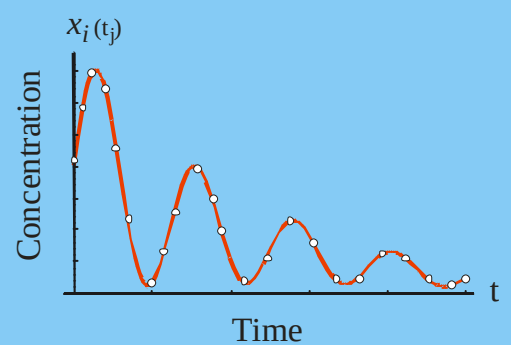
Boundary conditions :
boundary ... S, normal unit vector... $\hat{u}$

Dirichlet : $x^S = g(r,t)$

Neumann : $\frac{\partial x}{\partial u} = \hat{u} \cdot \nabla x^S = g(r,t)$

Data from measurements

$x(t_j);\ j = 1, 2, \ldots, N$



The inverse problem of biochemical reaction kinetics (Level I)

$$F(\mathbf{p}) \;=\; \mathbf{y}_\delta \;, \quad \text{in particular} \quad F(\mathbf{p}) \;=\; \mathbf{x}(t_k; k = 1, \ldots, n)$$

Optimization criterion: $\quad \|\mathbf{y}_\delta \;-\; F(\mathbf{p})\|^2_\Upsilon \qquad \rightarrow \; \min\limits_{\mathbf{p} \in \mathbf{P}}$

The parameter identification problem

$$F(\mathbf{p}) \;=\; \mathbf{y}_\delta \;, \quad \text{in particular} \quad F(\mathbf{p}) \;=\; \mathbf{x}(t_k; k = 1, \ldots, n)$$

$$\text{Optimization criterion:} \quad \|\mathbf{y}_\delta \;-\; F(\mathbf{p})\|^2_\Upsilon \qquad \rightarrow \; \min_{\mathbf{p} \in \mathbf{P}}$$

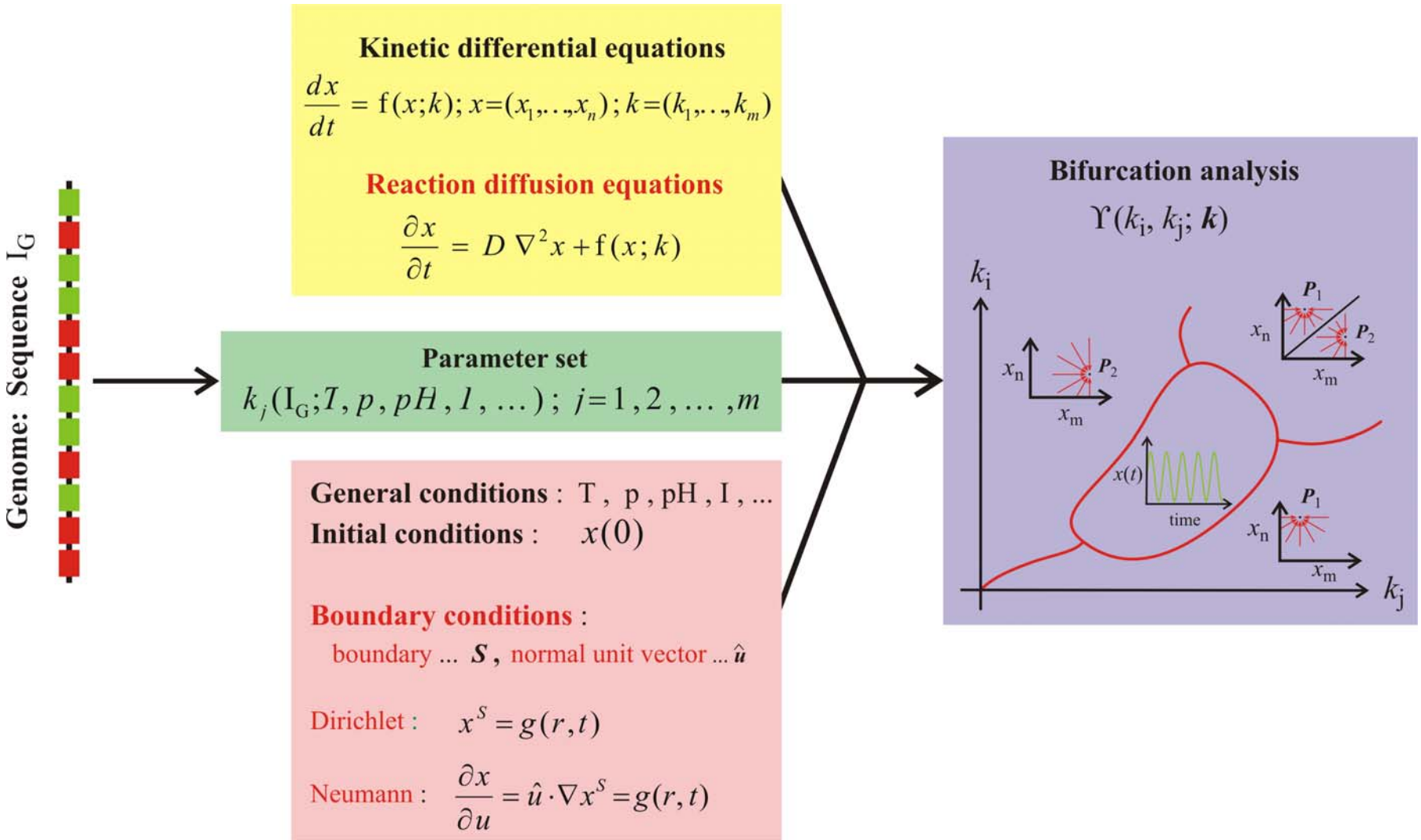
$$\text{Regularization:} \quad \|\mathbf{y}_\delta \;-\; F(\mathbf{p})\|^2_\Upsilon \;+\; \alpha\, \mathcal{R}(\mathbf{p}, \mathbf{p}_0) \qquad \rightarrow \; \min_{\mathbf{p} \in \mathbf{P}}$$

$$\text{Tikhonov regularization:} \quad \mathcal{R}(\mathbf{p}, \mathbf{p}_0) \;=\; \|\mathbf{p} \;-\; \mathbf{p}_0\|^2_{\mathbf{P}}$$

$$\text{Maximum entropy regularization:} \quad \mathcal{R}(\mathbf{p}, \mathbf{p}_0) \;=\; \int_\Omega \mathbf{p}(\omega)\, \log \frac{\mathbf{p}(\omega)}{\mathbf{p}_0(\omega)}\, d\omega$$

$$\text{Bounded variation regularization:} \quad \mathcal{R}(\mathbf{p}, \mathbf{p}_0) \;=\; \int_\Omega |\nabla \mathbf{p}(\omega)|\, d\omega$$

The parameter identification problem



The forward problem of bifurcation analysis in cellular dynamics (Level II)

Kinetic differential equations

$$\frac{dx}{dt} = f(x;k);\ x=(x_1,\ldots,x_n);\ k=(k_1,\ldots,k_m)$$

Reaction diffusion equations

$$\frac{\partial x}{\partial t} = D\,\nabla^2 x + f(x;k)$$

General conditions : T , p , pH , I , ...

Initial conditions : $x(0)$

Boundary conditions :

boundary ... $\boldsymbol{S}$, normal unit vector ... $\hat{\boldsymbol{u}}$

Dirichlet : $x^S = g(r,t)$

Neumann : $\frac{\partial x}{\partial u} = \hat{u}\cdot\nabla x^S = g(r,t)$

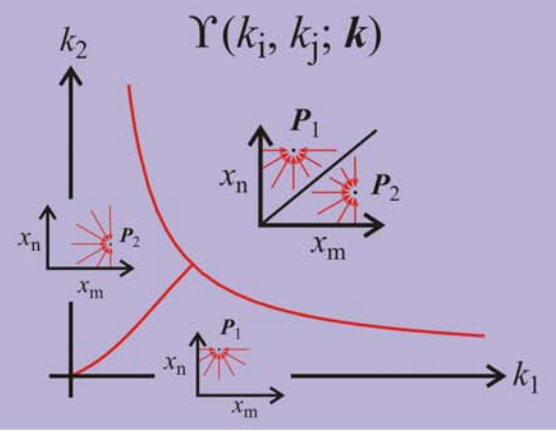
Inverse problem:
Design of bifurcation behavior

Parameter set

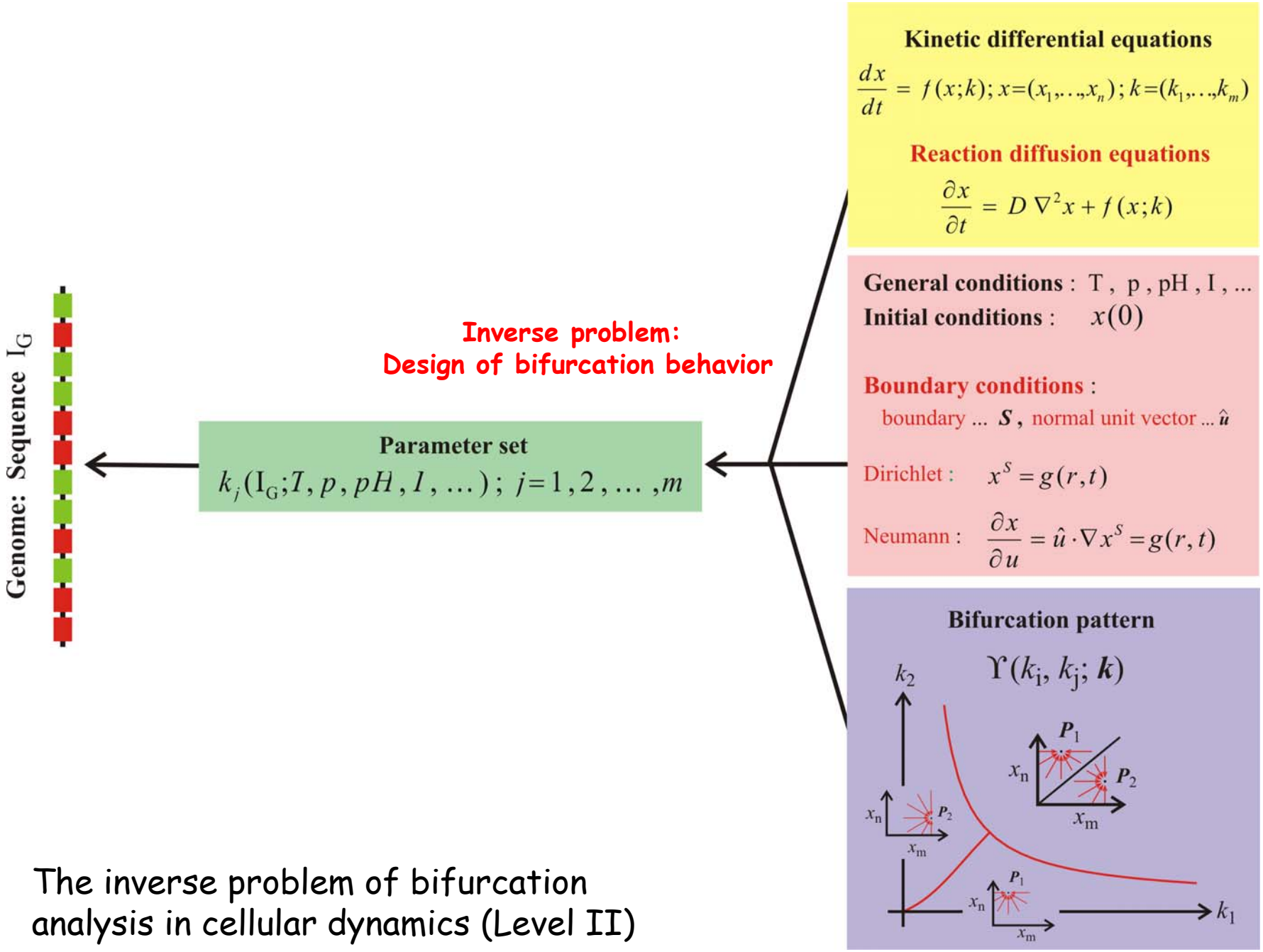
$$k_j(\mathrm{I_G};T,p,pH,I,\ldots);\ j=1,2,\ldots,m$$

Bifurcation pattern

$$\Upsilon(k_i, k_j; \boldsymbol{k})$$



The inverse problem of bifurcation analysis in cellular dynamics (Level II)



The inverse problem of bifurcation analysis in cellular dynamics (Level II)

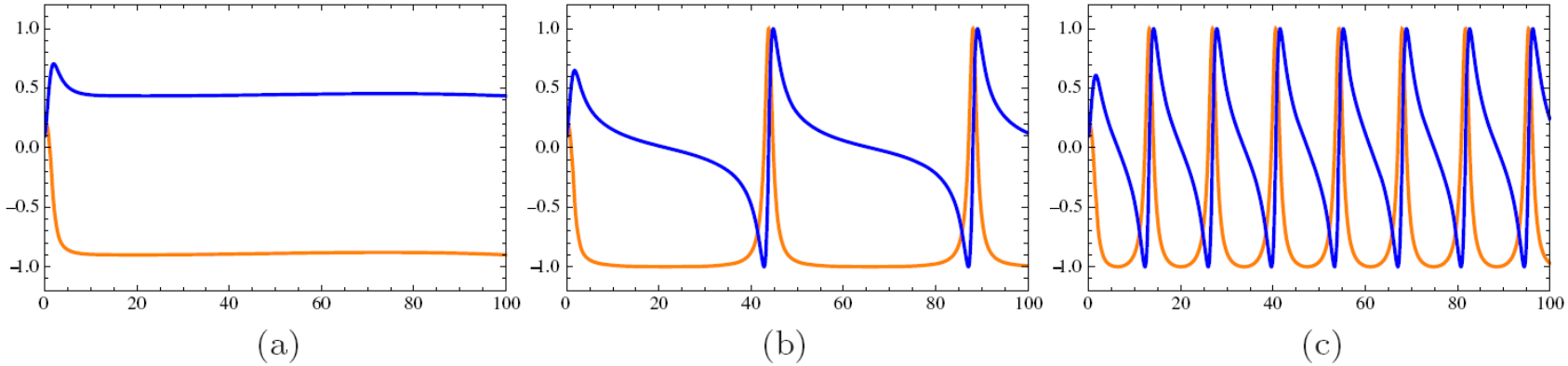


Figure 4. Transition across a Saddle-Node Invariant Cycle (SNIC) bifurcation as S varies: time courses. (a) Signal $S < S_0$, (b) signal $S = S_0+$ and (c) signal $S > S_0$.

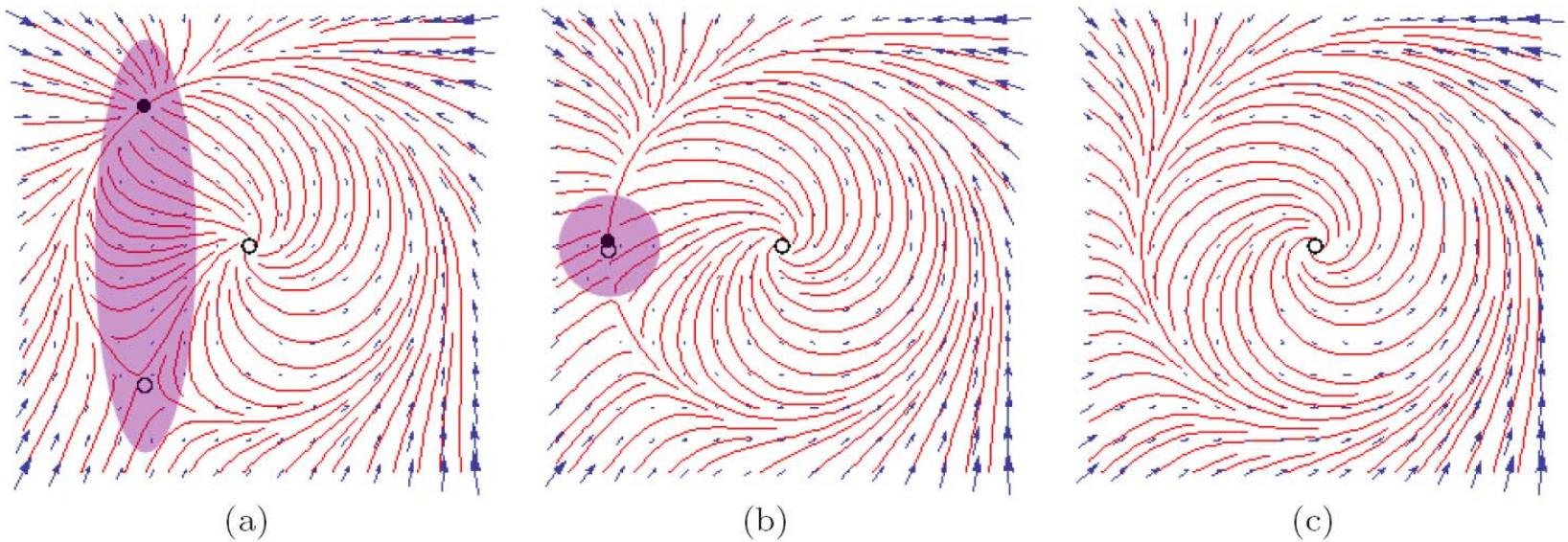


Figure 5. Transition across a Saddle-Node Invariant Cycle (SNIC) bifurcation as S varies: phase portraits. (a) Signal $S < S_0$, (b) signal $S = S_0-$ and (c) signal $S > S_0$.

A dynamical system with an oscillatory regime between a saddle node - invariant cycle (SNIC) bifuraction and a Hopf bifurcation.

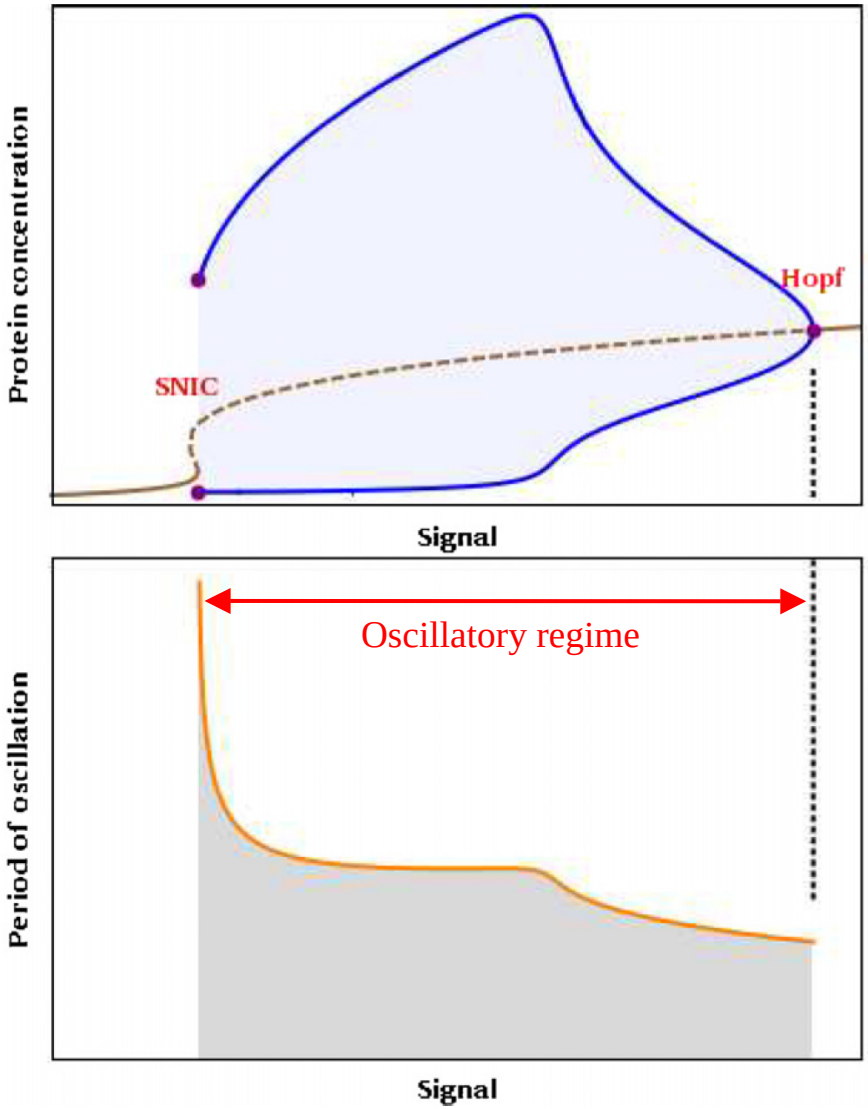


Figure 6. Bifurcation diagram of a system undergoing Hopf and SNIC bifurcations. Top: at low values of the signal, there exists a stable steady state (solid line) which subsequently loses stability via the SNIC bifurcation (the curve of unstable equilibrium is shown as a dashed line); further along the curve, the equilibrium undergoes a Hopf bifurcation, from which a stable limit cycle solution comes off (the curve of minima and maxima of limit cycle is denoted by the solid blue lines). Bottom: the period of oscillation for the limit cycle solution shows a blow-up $\sim \frac{1}{\sqrt{S-S_0}}$ near the SNIC bifurcation.

1. From biochemical kinetics to quantitative biology
2. Forward and inverse problems in reaction kinetics
3. **Modeling biochemical reaction kinetics**
4. Examples of inverse bifurcation analysis and design
5. Problems and perspectives of systems biology

$$A + B \xrightarrow{E} C$$

$$* \longrightarrow A$$

$$* \longrightarrow B$$

$$E + A \rightleftharpoons EA$$

$$E + B \rightleftharpoons EB$$

$$EA + B \rightleftharpoons EAB$$

$$EB + A \rightleftharpoons EAB$$

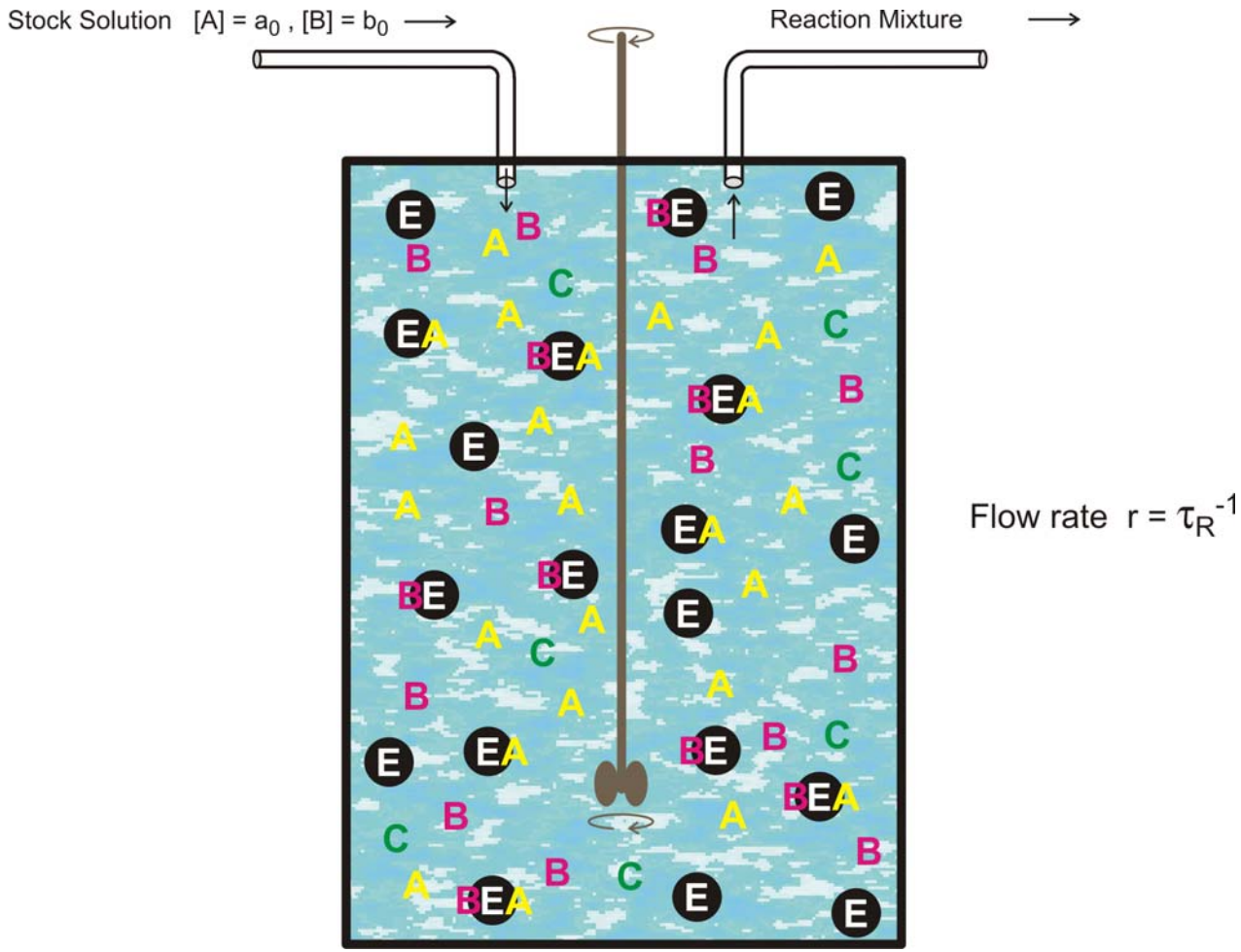
$$EAB \longrightarrow E + C$$

$$A \longrightarrow \varnothing$$

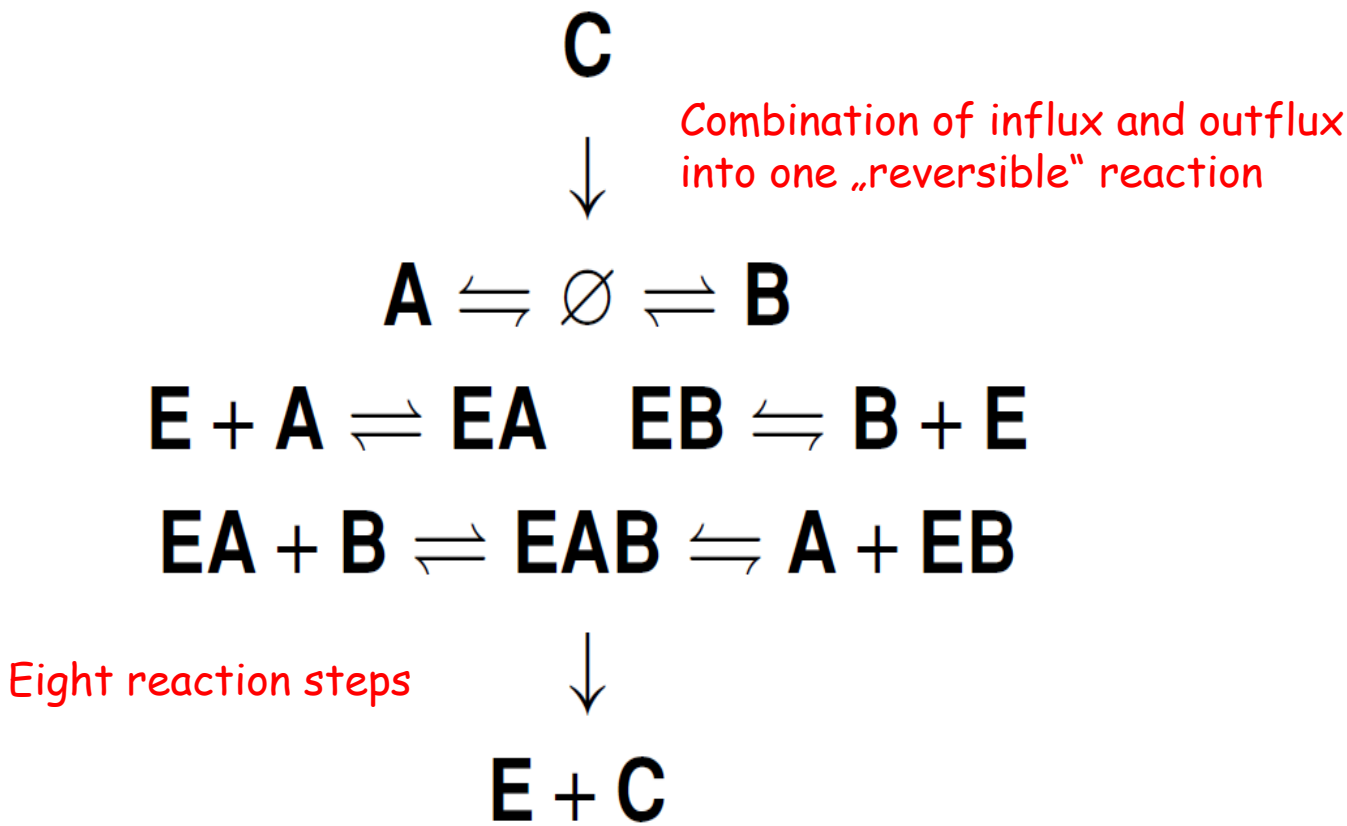
$$B \longrightarrow \varnothing$$

$$C \longrightarrow \varnothing$$

Ten reaction steps



An enzyme catalyzed addition reaction,
A + B → C in the flow reactor



A model reaction network with 12 complexes:

$\mathcal{C}$ = { ø , A , B , C , EA , EB , EAB , E+A , E+B , EA+B , EB+A , E+C }

$$
\begin{aligned}
R_1 &: \quad \varnothing \quad \rightleftharpoons \mathsf{A} \\
R_2 &: \quad \varnothing \quad \rightleftharpoons \mathsf{B} \\
R_3 &: \quad \varnothing \quad \leftarrow \mathsf{C} \\
R_4 &: \quad \mathsf{E}+\mathsf{A} \quad \rightleftharpoons \mathsf{EA} \\
R_5 &: \quad \mathsf{E}+\mathsf{B} \quad \rightleftharpoons \mathsf{EB} \\
R_6 &: \quad \mathsf{EA}+\mathsf{B} \quad \rightleftharpoons \mathsf{EAB} \\
R_7 &: \quad \mathsf{EB}+\mathsf{A} \quad \rightleftharpoons \mathsf{EAB} \\
R_8 &: \quad \mathsf{EAB} \quad \rightarrow \mathsf{E}+\mathsf{C}
\end{aligned}
$$

$$
S = \begin{array}{c} \\ \mathsf{A} \\ \mathsf{B} \\ \mathsf{C} \\ \mathsf{E} \\ \mathsf{EA} \\ \mathsf{EB} \\ \mathsf{EAB} \end{array}
\begin{array}{c}
\begin{array}{cccccccc} R_1 & R_2 & R_3 & R_4 & R_5 & R_6 & R_7 & R_8 \end{array} \\
\left[\begin{array}{cccccccc}
+1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 \\
0 & +1 & 0 & 0 & -1 & -1 & 0 & 0 \\
0 & 0 & -1 & 0 & 0 & 0 & 0 & +1 \\
0 & 0 & 0 & -1 & -1 & 0 & 0 & +1 \\
0 & 0 & 0 & +1 & 0 & -1 & 0 & 0 \\
0 & 0 & 0 & 0 & +1 & 0 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 & +1 & +1 & -1
\end{array}\right]
\end{array}
$$

Figure 1. Representation of a chemical reaction network as chemical equations (left) and as a stoichiometric matrix (right). The equations describe the mechanism for an enzyme catalyzed reaction with unordered substrate binding which is found in the cycline-dependent kinase-catalyzed phosphorylation reaction at the heart of the cell cycle.

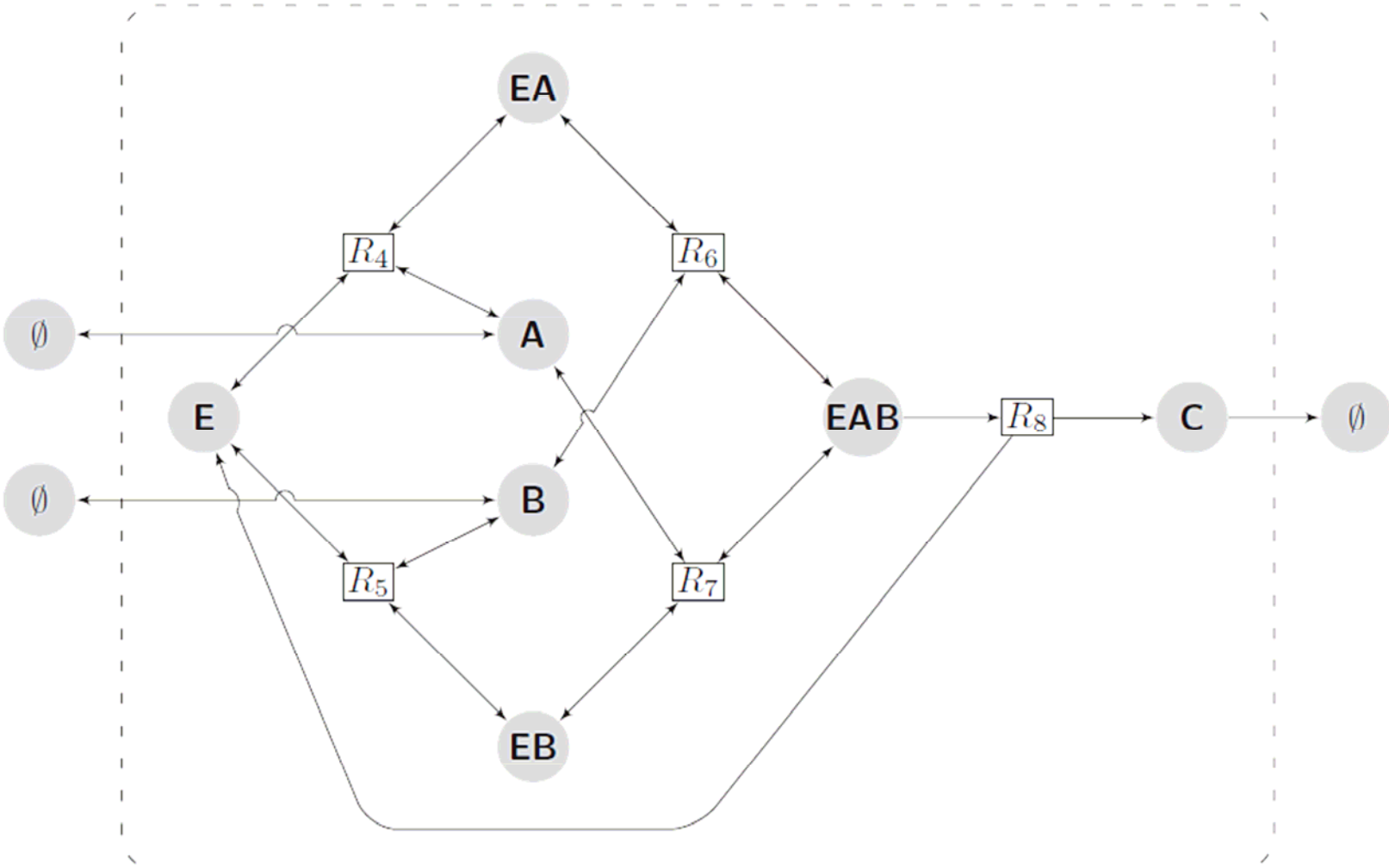


Figure 2. Graph representation of the chemical reaction network from figure 1. Species nodes are drawn as gray circles and reaction nodes as squares (for clarity the in- and outflow reaction nodes (R_1–R_3) have been omitted).

$$\frac{d}{dt}x = S \cdot v = \frac{d}{dt}\begin{pmatrix} x_A \\ x_B \\ x_C \\ x_E \\ x_{EA} \\ x_{EB} \\ x_{EAB} \end{pmatrix} = \begin{pmatrix} +1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 \\ 0 & +1 & 0 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & +1 \\ 0 & 0 & 0 & -1 & -1 & 0 & 0 & +1 \\ 0 & 0 & 0 & +1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & +1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & +1 & +1 & -1 \end{pmatrix} \cdot \begin{pmatrix} v_1(x_A, q) \\ v_2(x_B, q) \\ v_3(x_C, q) \\ v_4(x_A, x_E, x_{EA}, q) \\ v_5(x_B, x_E, x_{EB}, q) \\ v_6(x_B, x_{EA}, x_{EAB}, q) \\ v_7(x_A, x_{EB}, x_{EAB}, q) \\ v_8(x_{EAB}, q) \end{pmatrix}$$

Figure 3. The stoichiometric matrix S as linear transformation between the reaction rate vector v and the vector of changes in species concentrations x. In many cases, entries in the reaction rate vector v cannot be measured directly and hence must be inferred indirectly from biological data.

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} x_A \\ x_B \\ x_C \\ x_E \\ x_{EA} \\ x_{EB} \\ x_{EAB} \end{pmatrix} = \begin{pmatrix} +1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 \\ 0 & +1 & 0 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & +1 \\ 0 & 0 & 0 & -1 & -1 & 0 & 0 & +1 \\ 0 & 0 & 0 & +1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & +1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & +1 & +1 & -1 \end{pmatrix} \begin{pmatrix} k_1 - k_{-1} \cdot x_A \\ k_2 - k_{-2} \cdot x_B \\ k_{-3} \cdot x_C \\ k_4 \cdot x_E \cdot x_A - k_{-4} \cdot x_{EA} \\ k_5 \cdot x_E \cdot x_B - k_{-5} \cdot x_{EB} \\ k_6 \cdot x_{EA} \cdot x_B - k_{-6} \cdot x_{EAB} \\ k_7 \cdot x_{EB} \cdot x_A - k_{-7} \cdot x_{EAB} \\ k_8 \cdot x_{EAB} \end{pmatrix}$$

Kinetic differential equation of the reaction network with mass action kinetics

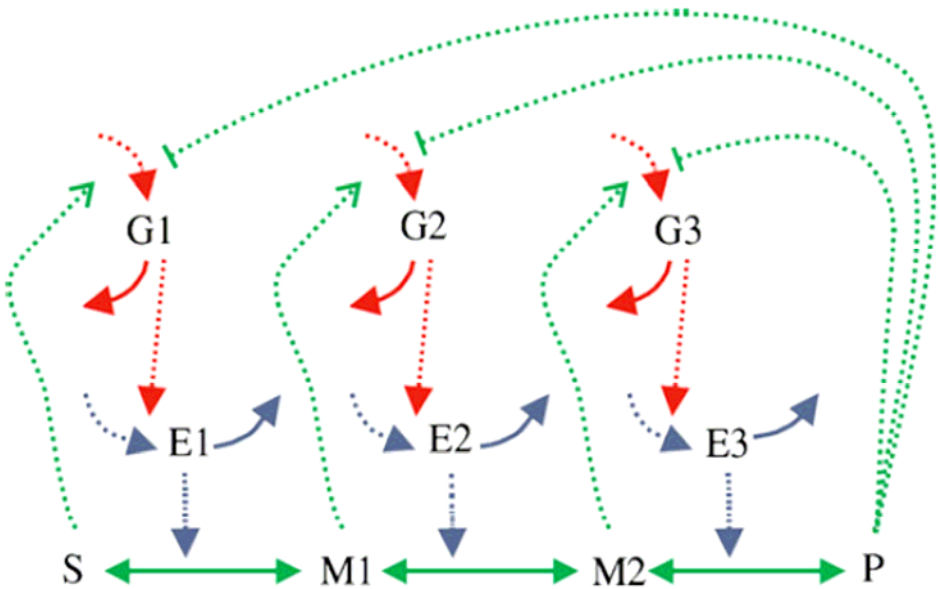


Figure 32. Three-step metabolic pathway (reproduced from [113]). Solid arrows represent mass flow, and dashed arrows represent regulation, where $\rightarrow$ denotes activation and $\dashv$ denotes inhibition. Three genes are producing mRNAs G1, G2, G3 and enzymes E1, E2, E3 to regulate the transformation of substrate S into product P via the intermediate metabolites M1, M2.

$$A + B \rightarrow X$$
$$2\,X \rightarrow Y$$
$$Y + X \rightarrow D$$

Stoichiometric equations

↓

SBML – systems biology markup language

↓

Kinetic differential equations

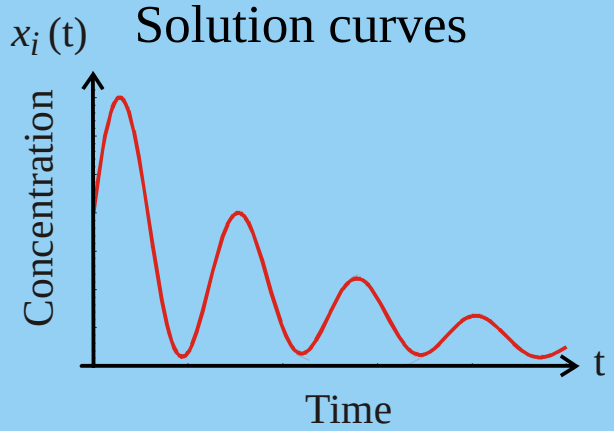
$$\frac{da}{dt} = \frac{db}{dt} = -k_1 a b$$
$$\frac{dx}{dt} = k_1 a b - k_2 x^2 - k_3 x y$$
$$\frac{dy}{dt} = k_2 x^2 - k_3 x y$$
$$\frac{dd}{dt} = k_3 x y$$

↓

ODE Integration by means of *CVODE*

↓

Solution curves



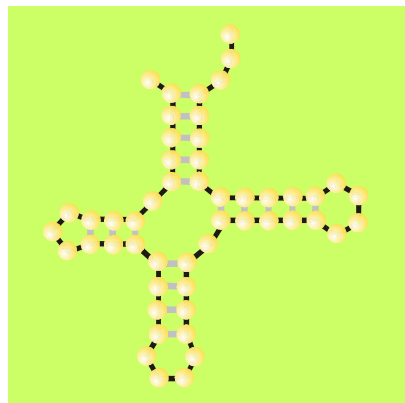
Sequences

↓

Vienna RNA Package

↓

Structures and kinetic parameters



The elements of the simulation tool MiniCellSim

SBML: *Bioinformatics* **19**:524-531, 2003; *CVODE*: *Computers in Physics* **10**:138-143, 1996

Peter Schuster. 2006.

Prediction of RNA secondary structures: From theory to models and real molecules.

Rep.Prog.Phys. 69:1419-1477

INSTITUTE OF PHYSICS PUBLISHING REPORTS ON PROGRESS IN PHYSICS

Rep. Prog. Phys. **69** (2006) 1419–1477 doi:10.1088/0034-4885/69/5/R04

Prediction of RNA secondary structures: from theory to models and real molecules

Peter Schuster[1,2]

[1]Institut für Theoretische Chemie der Universität Wien, Währingerstraße 17, A-1090 Vienna, Austria
[2]The Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA

E-mail: pks@tbi.univie.ac.at

Received 8 January 2006, accepted for publication 22 March 2006
Published 18 April 2006
Online at stacks.iop.org/RoPP/69/1419

Abstract

RNA secondary structures are derived from RNA sequences, which are strings built form the natural four letter nucleotide alphabet, {**AUGC**}. These coarse-grained structures, in turn, are tantamount to constrained strings over a three letter alphabet. Hence, the secondary structures are discrete objects and the number of sequences always exceeds the number of structures. The sequences built from two letter alphabets form perfect structures when the nucleotides can form a base pair, as is the case with {**GC**} or {**AU**}, but the relation between the sequences and structures differs strongly from the four letter alphabet. A comprehensive theory of RNA structure is presented, which is based on the concepts of *sequence space* and *shape space*, being a space of structures. It sets the stage for modelling processes in ensembles of RNA molecules like evolutionary optimization or kinetic folding as dynamical phenomena guided by mappings between the two spaces.

The number of minimum free energy (mfe) structures is always smaller than the number of sequences, even for two letter alphabets. Folding of RNA molecules into mfe energy structures constitutes a non-invertible mapping from sequence space onto shape space. The preimage of a structure in sequence space is defined as its *neutral network*. Similarly the set of *suboptimal structures* is the preimage of a sequence in shape space. This set represents the *conformation space* of a given sequence. The evolutionary optimization of structures in populations is a process taking place in sequence space, whereas kinetic folding occurs in molecular ensembles that optimize free energy in conformation space. Efficient folding algorithms based on dynamic programming are available for the prediction of secondary structures for given sequences. The inverse problem, the computation of sequences for predefined structures, is an important tool for the design of RNA molecules with tailored properties. Simultaneous folding or *cofolding* of two or more RNA molecules can be modelled readily at the secondary structure level

This article was invited by Professor T J Newman.

1. From biochemical kinetics to quantitative biology
2. Forward and inverse problems in reaction kinetics
3. Modeling biochemical reaction kinetics
4. **Examples of inverse bifurcation analysis and design**
5. Problems and perspectives of systems biology

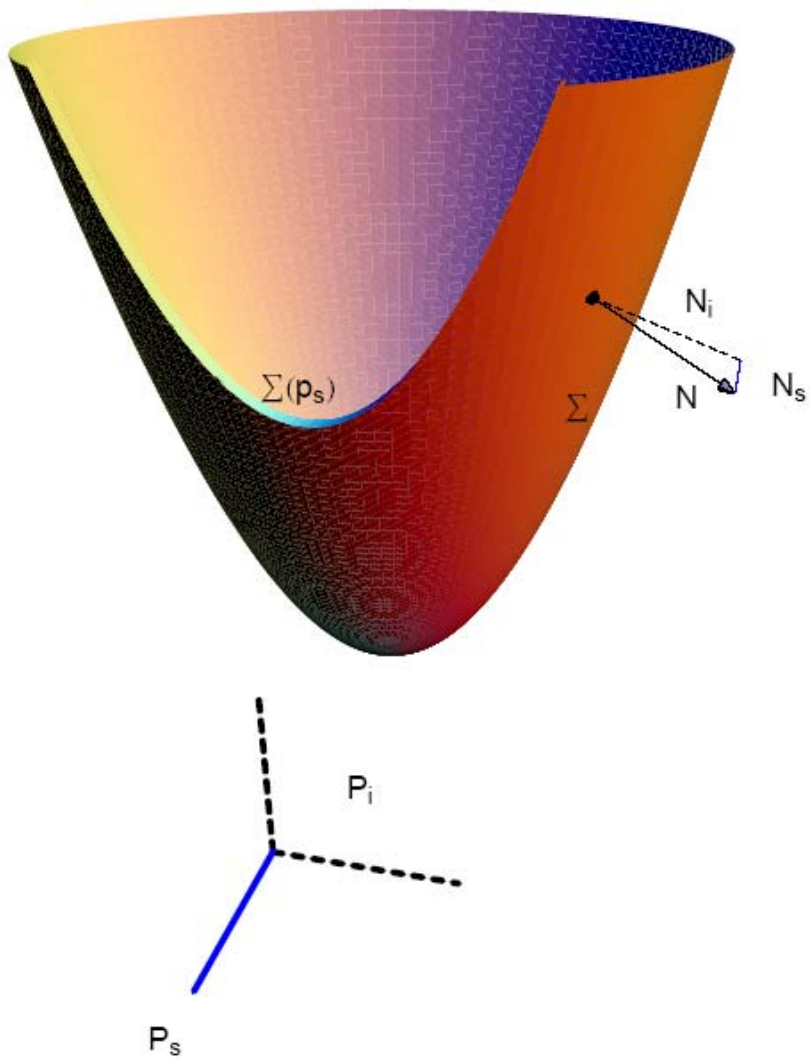
$$\dot{x} = f(x\,;\,p); \quad x=(x_1,\ldots,x_n); \quad p=(p_1,\ldots,p_m);\; p\in \mathrm{P}\subset \mathbf{R}^m$$

Σ … bifurcation manifold

$$p=(p_i\,,\,p_s)\in \mathrm{P}_i\times\mathrm{P}_s; \quad \mathrm{P}=\mathrm{P}_i\oplus\mathrm{P}_s\,;\,\Sigma(p_s)\equiv\ \Sigma\cap\{p_s\}$$

Inverse bifurcation analysis

J. Lu, H.W. Engl, P. Schuster. Inverse bifurcation analysis: Application to simple gene systems. *AMB Algorithms for Molecular Biology* **1**, no.11, 2006.



The bifurcation manifold

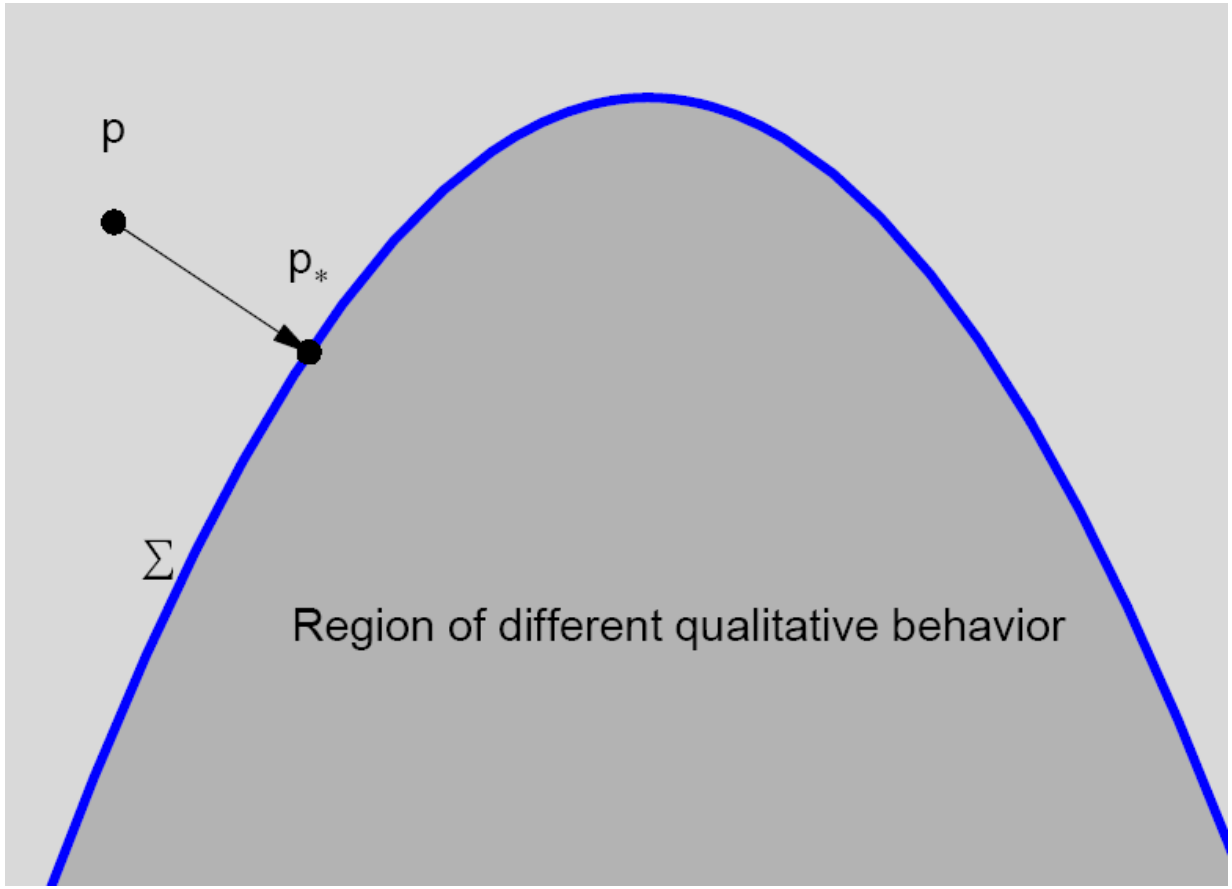
$$\dot{x} = f(x\,;\,p); \quad x=(x_1,\ldots,x_n); \quad p=(p_1,\ldots,p_m);\; p\in \mathrm{P}\subset \mathbf{R}^m$$

Σ … bifurcation manifold

$$p=(p_i\,,\,p_s)\in \mathrm{P}_i\times\mathrm{P}_s\,; \quad \mathrm{P}=\mathrm{P}_i\oplus\mathrm{P}_s\,;\,\Sigma(p_s)\equiv\ \Sigma\cap\{p_s\}$$

$$F(p)\equiv(F(p)_i\,,F(p_s))=\left(\pi_{\perp\Sigma(p_s)}\,p_i\ ,\ p_s\right)\ \ldots\ \text{forward operator}$$

J. Lu, H.W. Engl, P. Schuster. Inverse bifurcation analysis: Application to simple gene systems. *AMB Algorithms for Molecular Biology* **1**, no.11, 2006.



Defininition of the forward operator $F(p)$

$$\dot{x} = f(x\,;\,p); \quad x=(x_1,\ldots,x_n); \quad p=(p_1,\ldots,p_m);\; p\in \mathrm{P}\subset \mathbf{R}^m$$

Σ … bifurcation manifold

$$p=(p_i\,,\,p_s)\in \mathrm{P}_i\times\mathrm{P}_s; \quad \mathrm{P}=\mathrm{P}_i\oplus\mathrm{P}_s;\; \Sigma(p_s)\equiv\ \Sigma\cap\{p_s\}$$

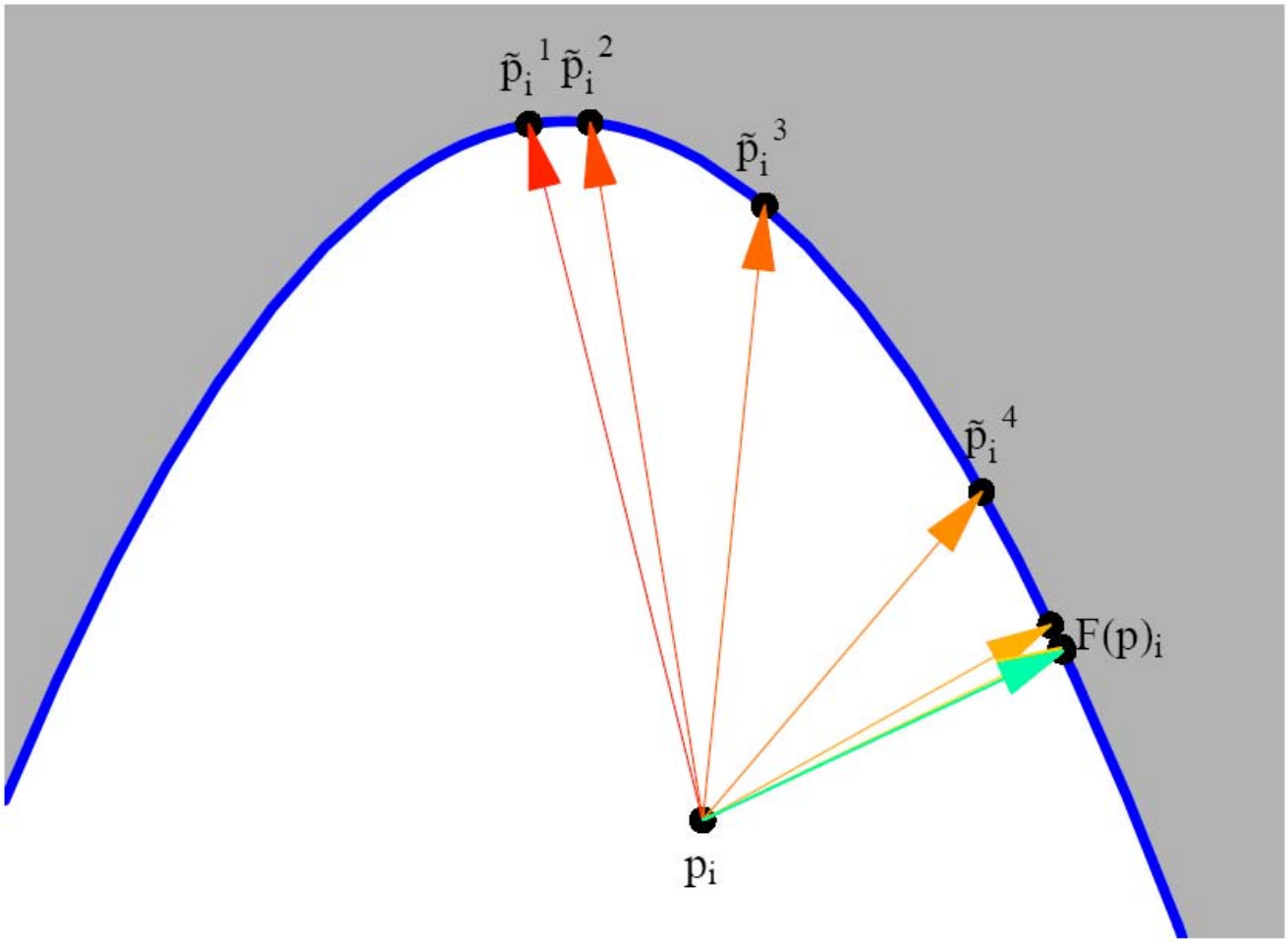
$$F(p)\equiv(F(p)_i\,,F(p_s))=\left(\pi_{\perp\Sigma(p_s)}\,p_i\ ,\ p_s\right)\ \ldots\ \text{forward operator}$$

$$\min_{p_s} J(p)=\min_{p_s}\left\|F(p)_i-p_i\right\| \quad \text{... formulation of the inverse problem}$$

subject to $p_{\text{low}}\le p\le p_{\text{upp}}$

and $0\le c(F(p)_i)$

J. Lu, H.W. Engl, P. Schuster. Inverse bifurcation analysis: Application to simple gene systems. *AMB Algorithms for Molecular Biology* **1**, no.11, 2006.



Iterative solution for $\min J(p)$

Examples of inverse bifurcation analysis and design

- **Oscillatory behavior in *Escherichia coli***
- The repressilator
- The mitotic cell cycle
- Time scales and oscillatory bursts
- Circadian rhythms.

$$\frac{dx_{2k-1}}{dt} = \beta_{2k-1}(f_{2k-1} - x_{2k-1})$$
$$\frac{dx_{2k}}{dt} = \beta_{2k}(x_{2k-1} - x_{2k}),$$

$k = 1, 2, 3$

$$f_1 = \begin{cases} B & \text{for } x_2^{g12} x_4^{g14} \le B \\ x_2^{g12} x_4^{g14} & \text{for } B < x_2^{g12} x_4^{g14} < M \\ M & \text{for } x_2^{g12} x_4^{g14} \ge M \end{cases}$$

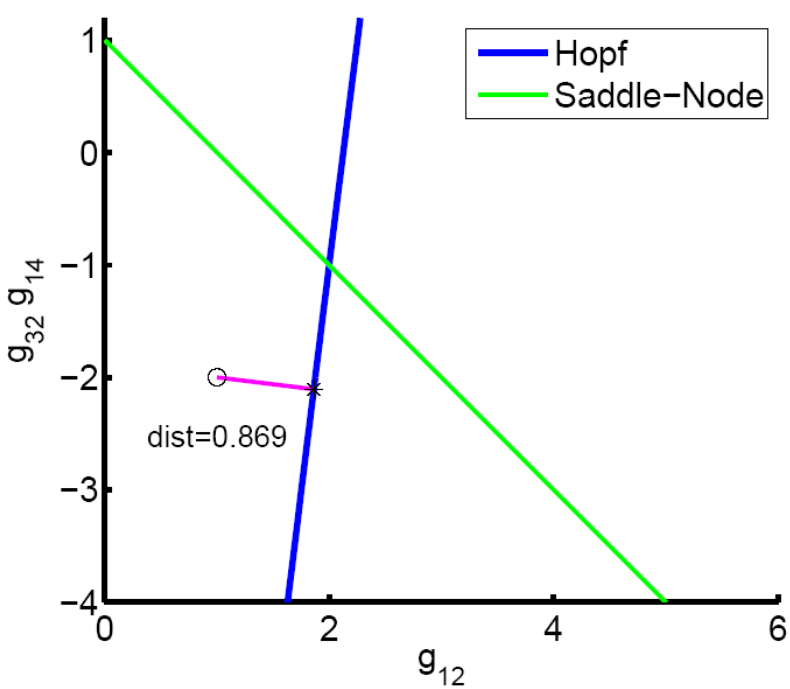
$$f_3 = \begin{cases} B & \text{for } x_2^{g32} \le B \\ x_2^{g32} & \text{for } B < x_2^{g32} < M \\ M & \text{for } x_2^{g32} \ge M \end{cases}$$

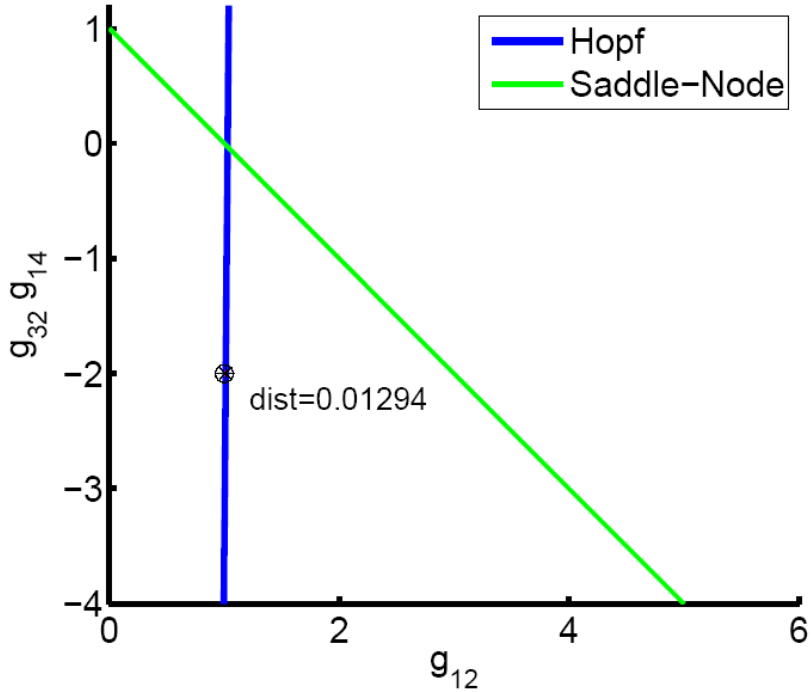
$$f_5 = \begin{cases} 1/M & \text{for } x_4^{g54} \le 1/M \\ x_4^{g54} & \text{for } 1/M < x_4^{g54} < 1/B \\ 1/B & \text{for } x_4^{g54} \ge 1/B \end{cases},$$

Switch or oscillatory behavior in Escherichia coli

T.S. Gardner, C.R. Cantor, J.J. Collins. Construction of a genetic toggle switch in Escherichia coli. *Nature* **403**:339-342, 2000.

M.R. Atkinson, M.A. Savageau, T.J. Myers, A.J. Ninfa. Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in Escherichia coli. *Cell* **113**:597-607, 2003.



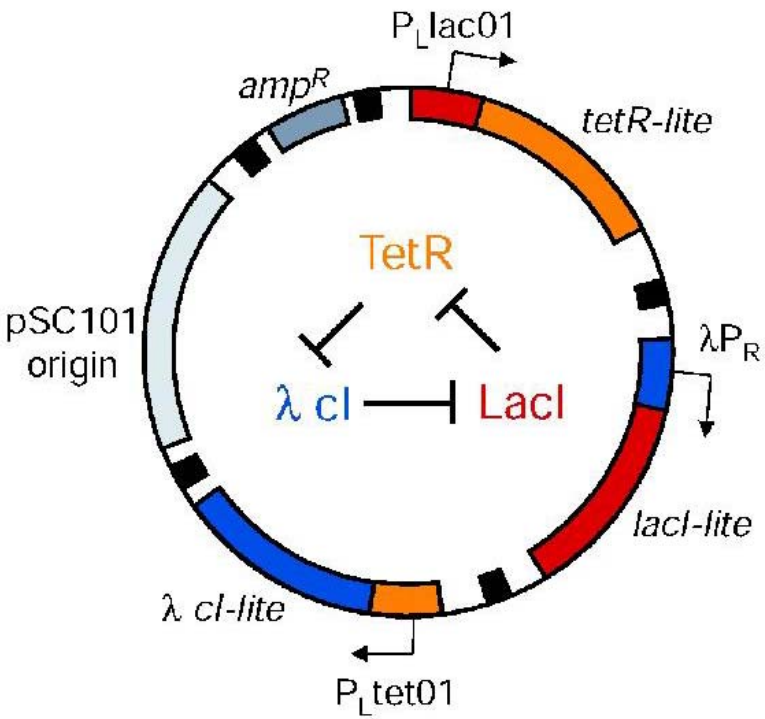


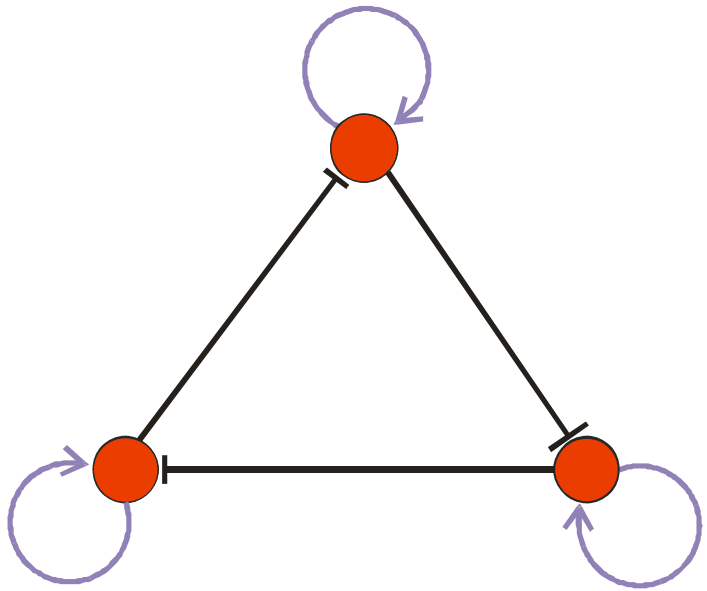
Inverse bifurcation analysis of switch or oscillatory behavior in *Escherichia coli*

J. Lu, H.W. Engl, P. Schuster. Inverse bifurcation analysis: Application to simple gene systems. *AMB Algorithms for Molecular Biology* **1**:11, 2006.

Examples of inverse bifurcation analysis and design

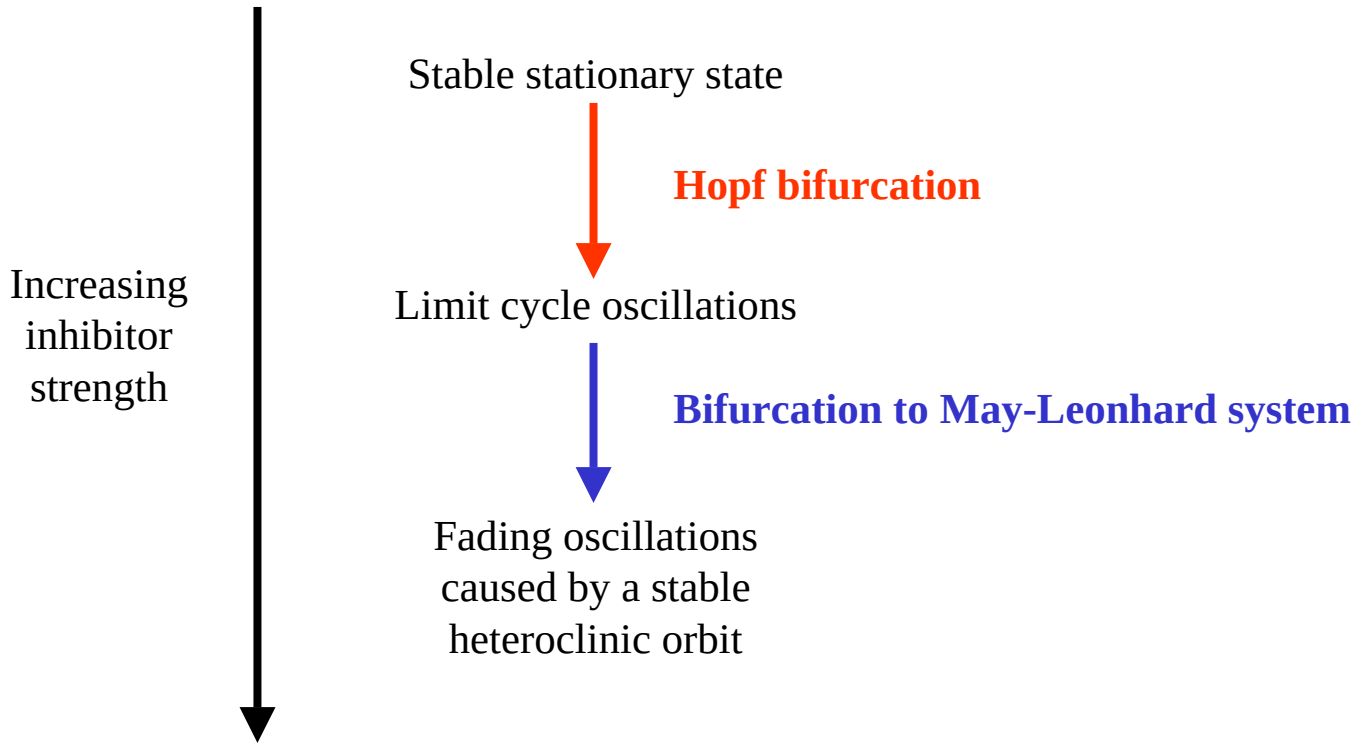
- Oscillatory behavior in *Escherichia coli*
- **The repressilator**
- The mitotic cell cycle
- Time scales and oscillatory bursts
- Circadian rhythms.

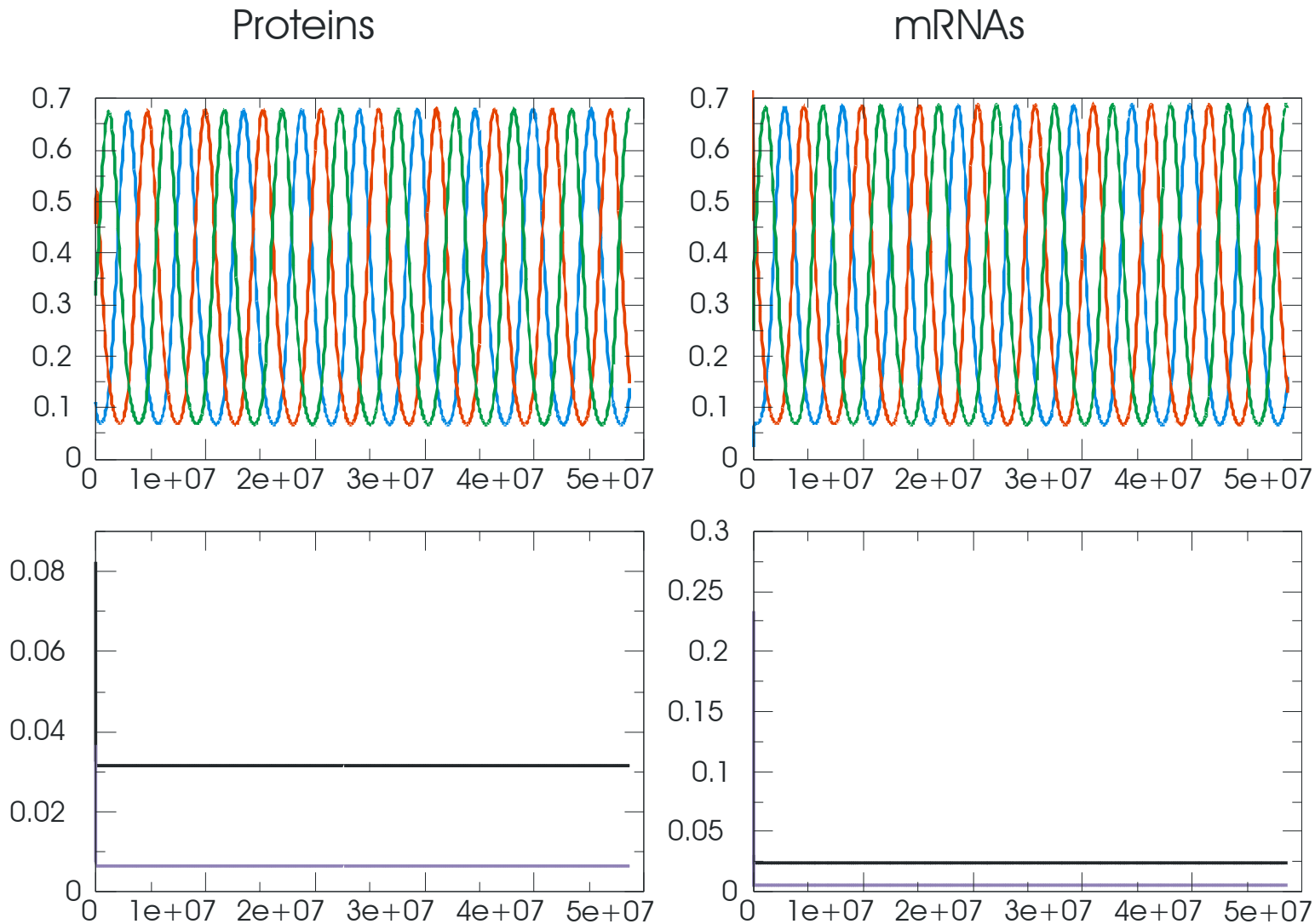




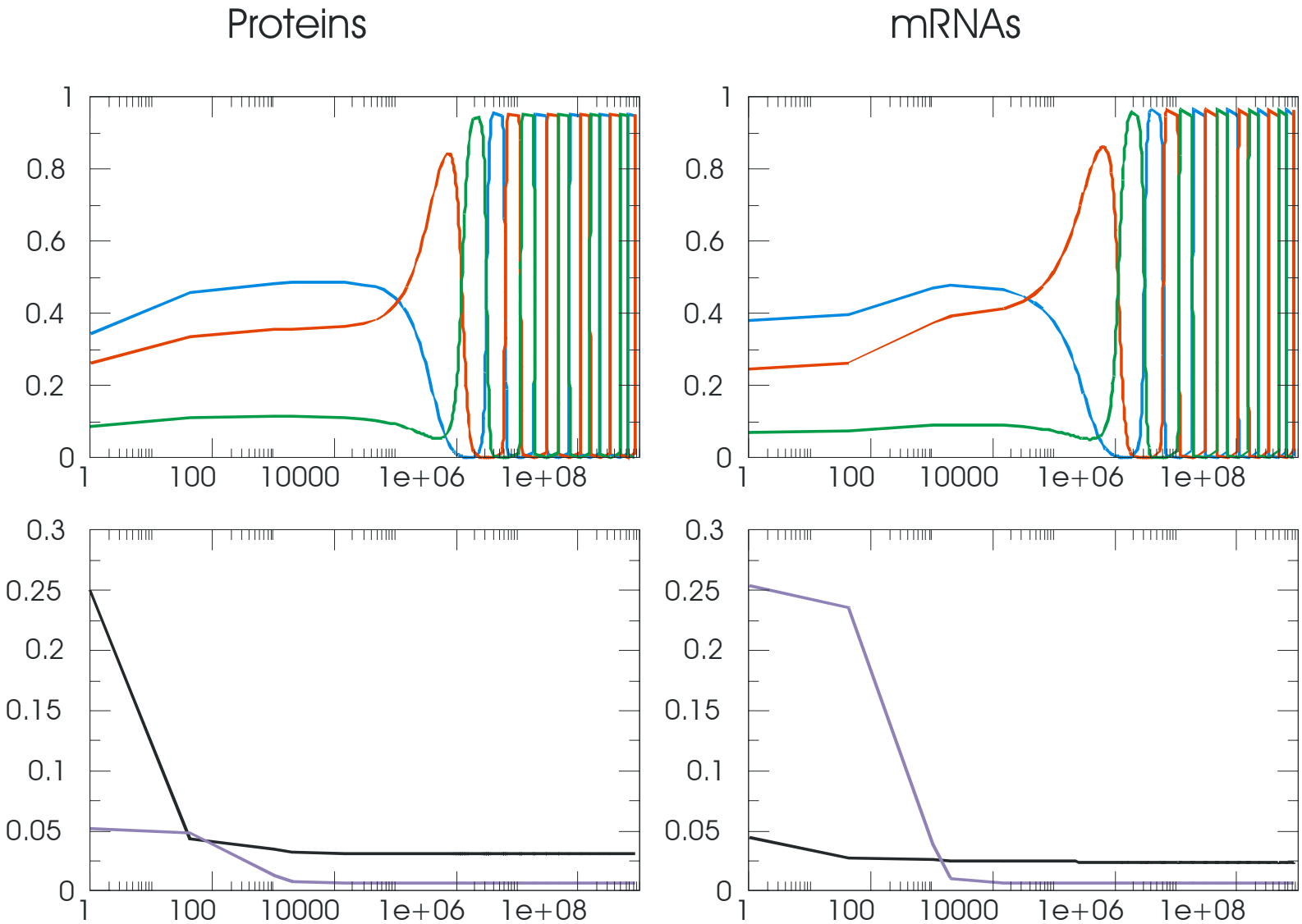
An example analyzed and simulated by MiniCellSim

The repressilator: M.B. Ellowitz, S. Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature* **403**:335-338, 2002

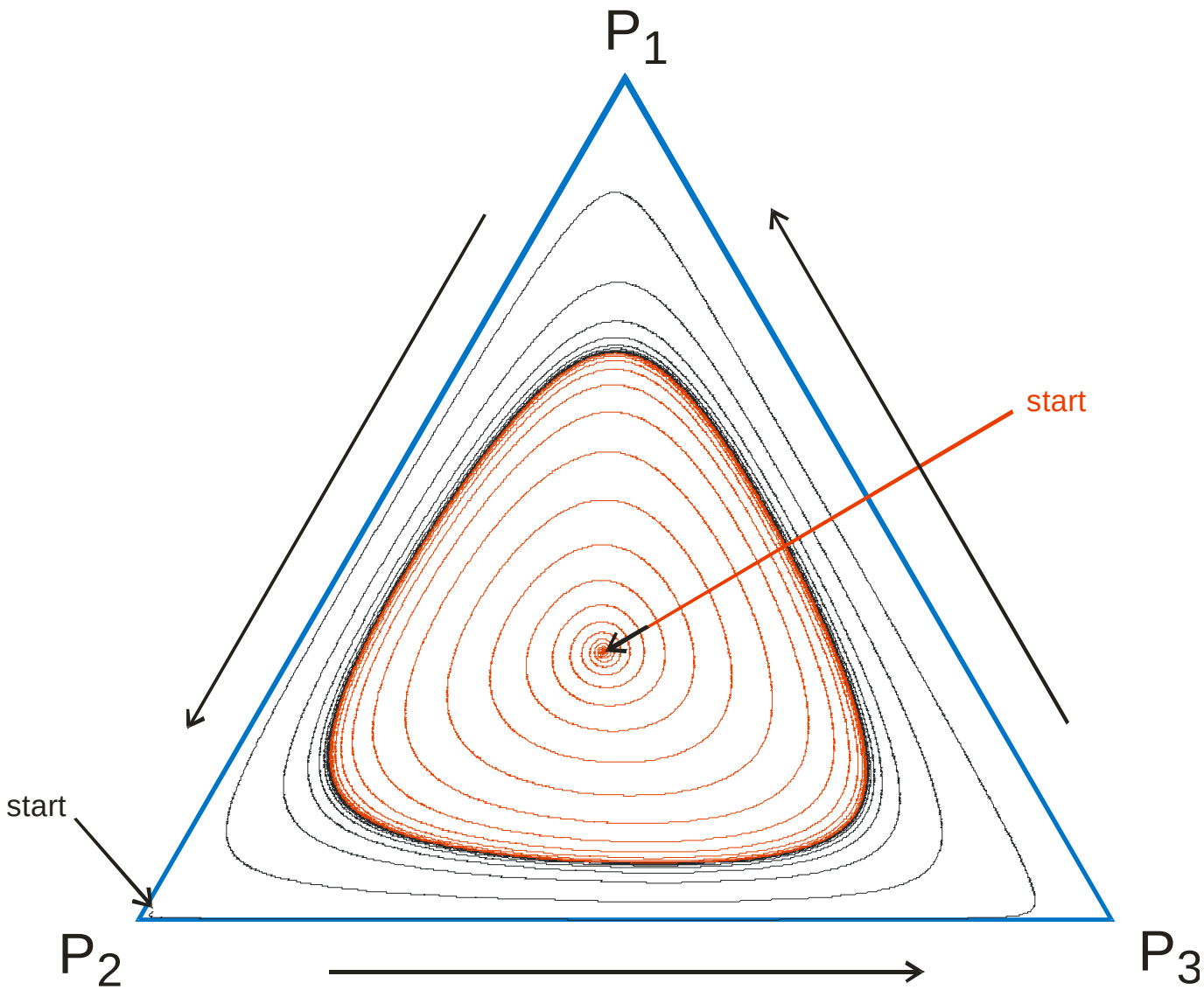

Increasing inhibitor strength
Stable stationary state
Hopf bifurcation
Limit cycle oscillations
Bifurcation to May-Leonhard system
Fading oscillations caused by a stable heteroclinic orbit



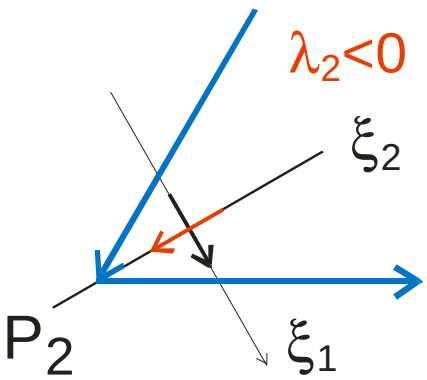
The repressilator limit cycle



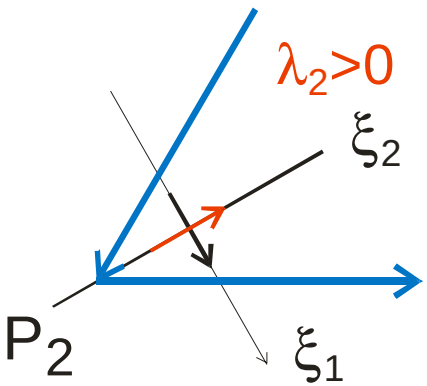
The repressilator heteroclinic orbit (logarithmic time scale)



The repressilator limit cycle

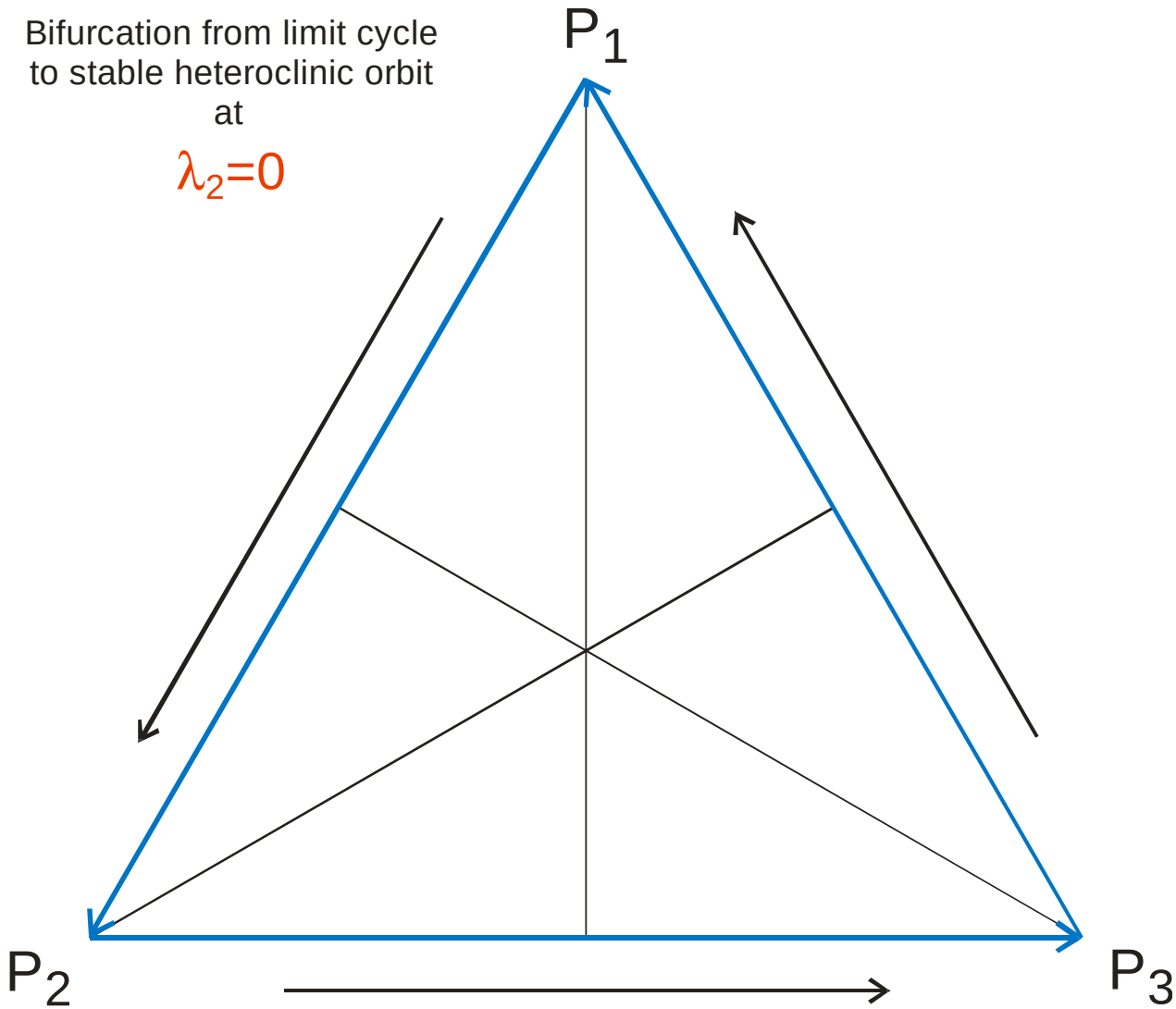


Stable heteroclinic orbit

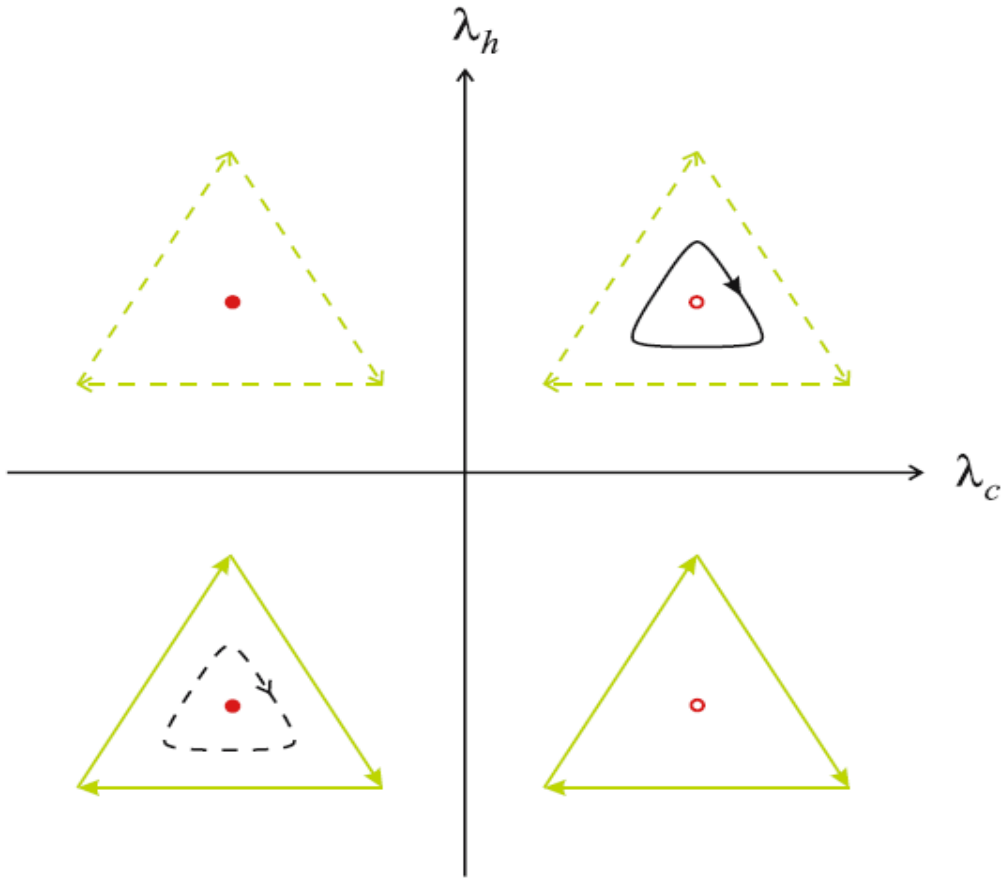


Unstable heteroclinic orbit

Bifurcation from limit cycle to stable heteroclinic orbit at $\lambda_2=0$



The repressilator heteroclinic orbit



All possible scenarios of repressilator dynamics

$$\begin{aligned}\dot{x}_i &= \beta_i(y_i - x_i) \\ \dot{y}_i &= \alpha_i \left(\frac{1-\delta_i}{1 + x_{i-1 \mod n}^{h_i}} + \delta_i \right) - y_i, \; i = 0, \cdots, n-1\end{aligned}$$

$$\alpha_i = \alpha, \beta_i = \beta, h_i = h, \delta_i = \delta$$

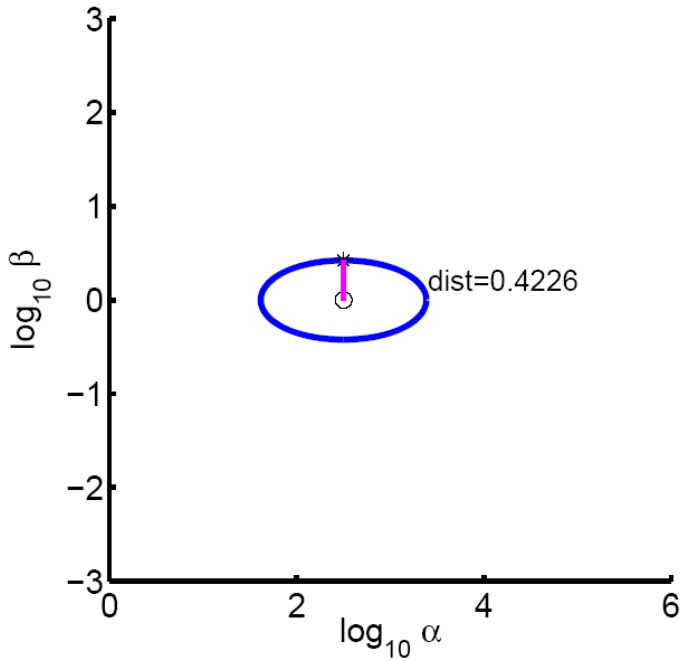
$$p_i = (\alpha, \beta)$$

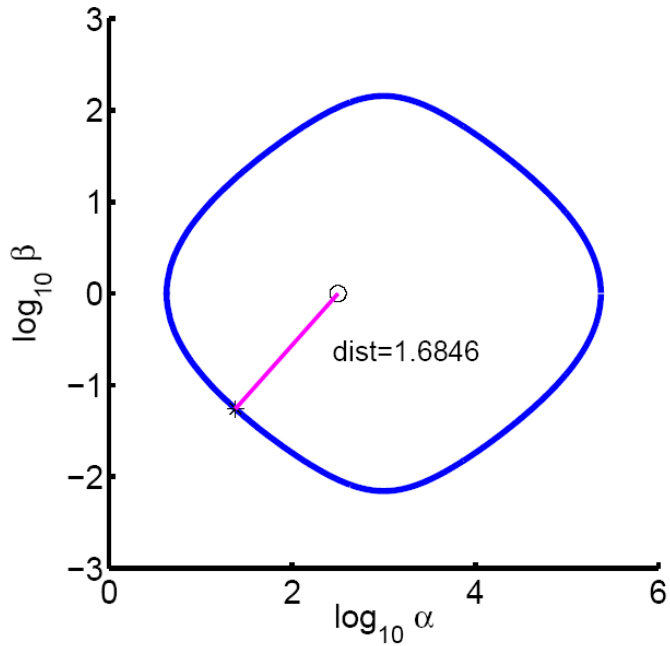
$$p_s = (\delta, h)$$

$$(10^{-4}, 0) \leq (\delta, h) \leq (10^{-1}, 2)$$

Inverse bifurcation analysis of the repressilator model

S. Müller, J. Hofbauer, L. Endler, C. Flamm, S. Widder, P. Schuster. A generalized model of the repressilator. *J. Math. Biol.* **53**:905-937, 2006.





Inverse bifurcation analysis of the repressilator model

J. Lu, H.W. Engl, P. Schuster. Inverse bifurcation analysis: Application to simple gene systems. *AMB Algorithms for Molecular Biology* **1**:11, 2006.

Examples of inverse bifurcation analysis and design

- Oscillatory behavior in *Escherichia coli*
- The repressilator
- **The mitotic cell cycle**
- Time scales and oscillatory bursts
- Circadian rhythms.

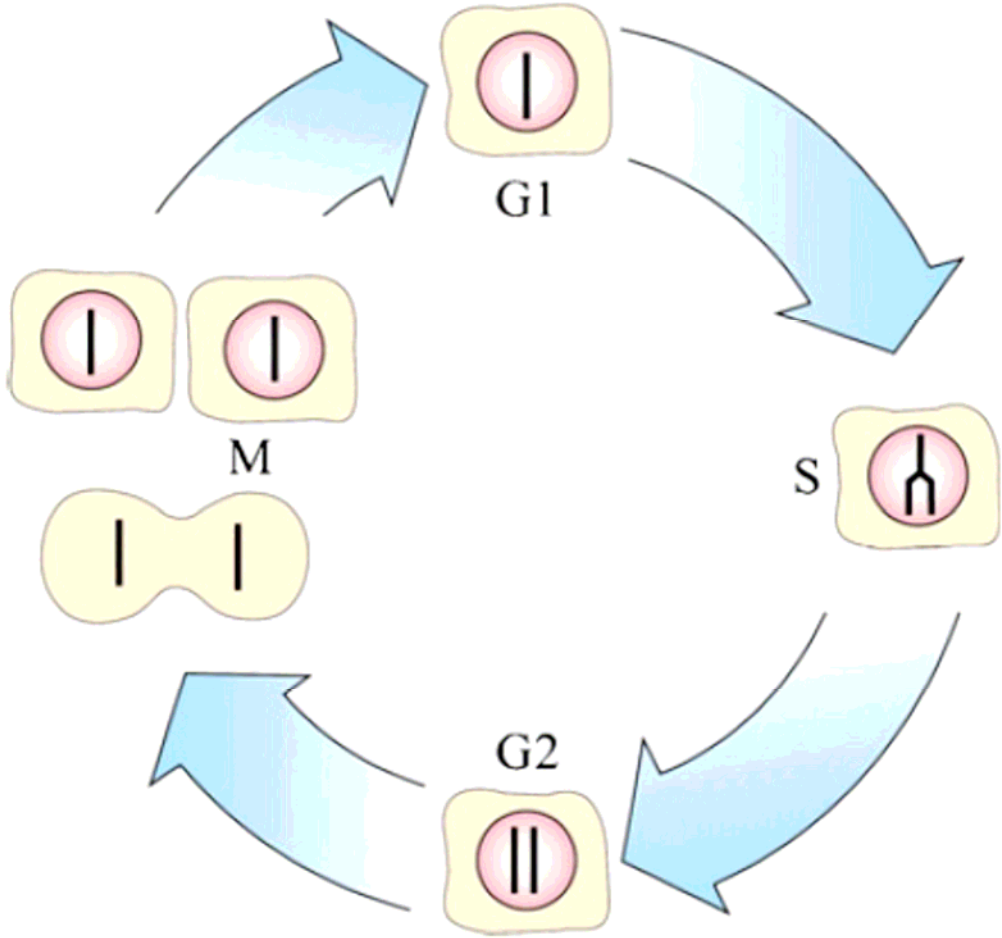


Figure 7. The cell division cycle: in Gap 1 (G_1) phase cells grow in size, preparing for the synthesis (S) phase during which DNA duplication occurs; in Gap 2 (G_2) phase cells continue to grow until the mitosis (M) phase is reached and the cells divide.

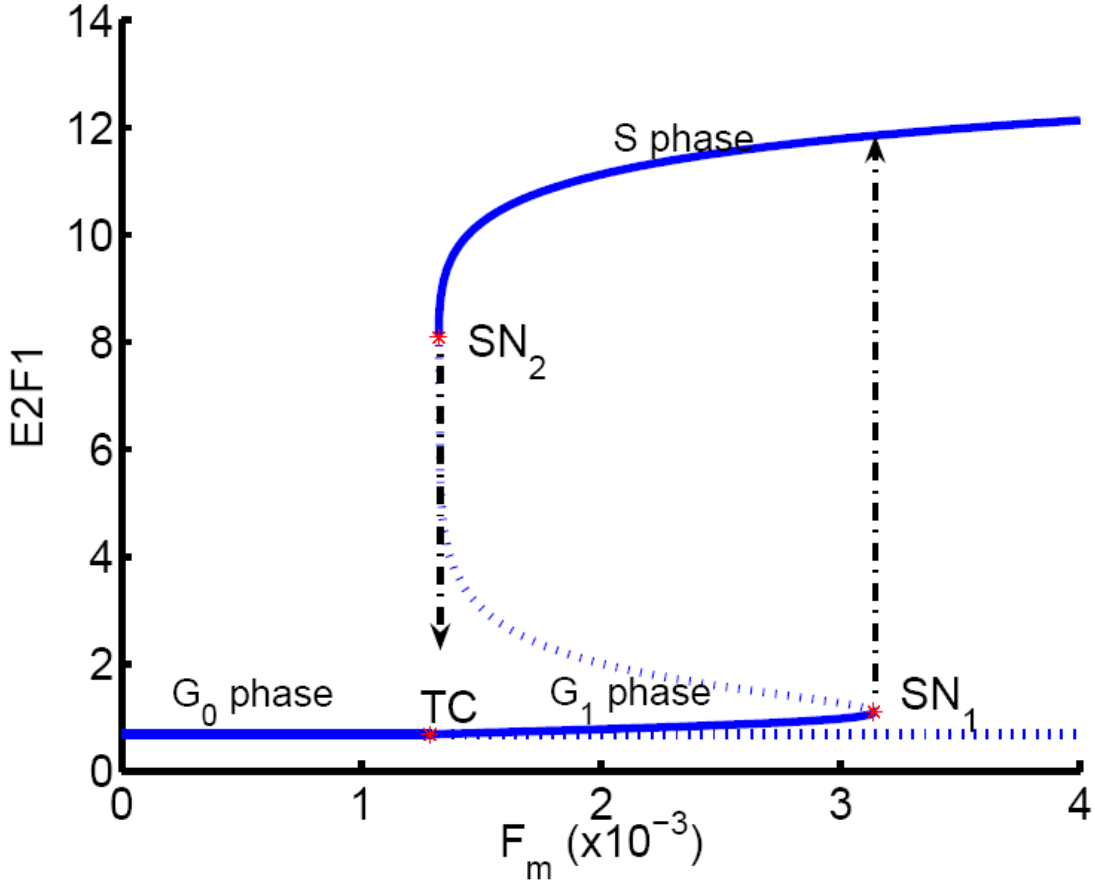
$$\frac{d}{dt}[\text{pRB}] = k_1 \frac{[\text{E2F1}]}{K_{m1} + [\text{E2F1}]} \frac{J_{11}}{J_{11} + [\text{pRB}]} - \phi_{\text{pRB}}[\text{pRB}]$$

$$\frac{d}{dt}[\text{E2F1}] = k_P + k_1 \frac{a^2 + [\text{E2F1}]^2}{K_{m2}^2 + [\text{E2F1}]^2} \frac{J_{12}}{J_{12} + [\text{pRB}]} - \phi_{\text{E2F1}}[\text{E2F1}]$$

$$\frac{d}{dt}[\text{AP1}] = F_m + k_{25}[\text{E2F1}] \frac{J_{15}}{J_{15} + [\text{p}RB]} \frac{J_{65}}{J_{11} + [\text{pRB}']} - \phi_{\text{AP1}}[\text{AP1}]$$

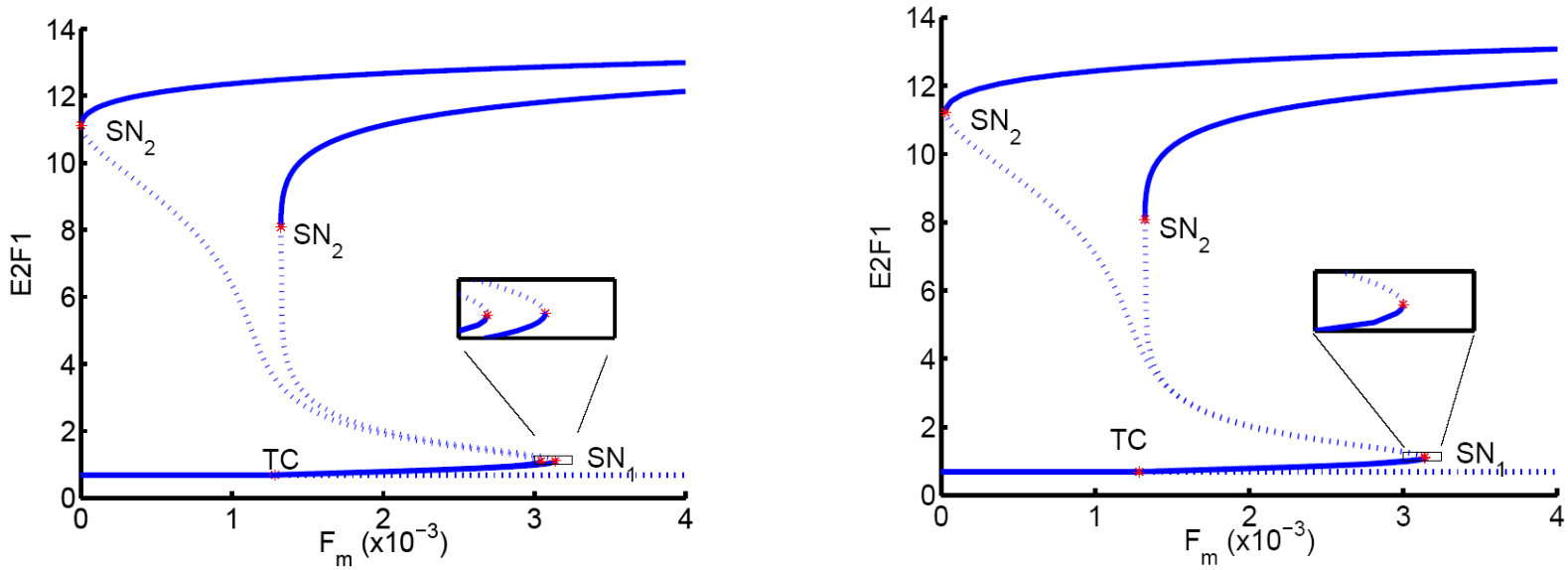
A simple dynamical cell cycle model

J.J. Tyson, A. Csikasz-Nagy, B. Novak. The dynamics of cell cycle regulation. *Bioessays* **24**:1095-1109, 2002



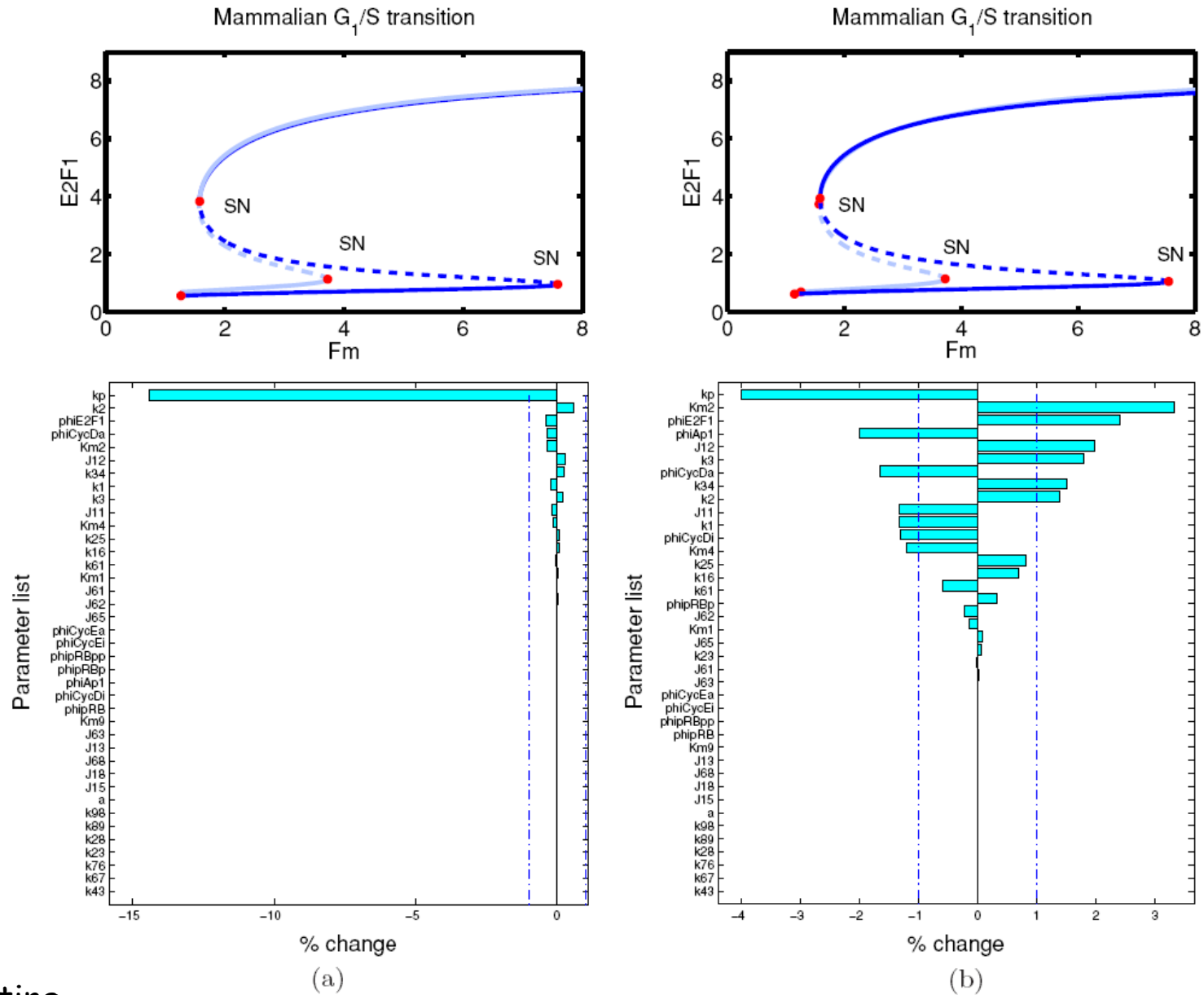
A simple dynamical cell cycle model

J.J. Tyson, A. Csikasz-Nagy, B. Novak. The dynamics of cell cycle regulation. *Bioessays* **24**:1095-1109, 2002



Inverse bifurcation analysis of a dynamical cell cycle model

J. Lu, H.W. Engl, P. Schuster. Inverse bifurcation analysis: Application to simple gene systems. *AMB Algorithms for Molecular Biology* **1**:11, 2006.



Sparsity promoting functional

Figure 23. Identified parameters leading to elongating saddle-nose using l_p and l_2 penalty functions. (a) Sparsity-promoting $l_{p,\epsilon}$ functional. (b) Standard l_2 functional.

Examples of inverse bifurcation analysis and design

- Oscillatory behavior in *Escherichia coli*
- The repressilator
- The mitotic cell cycle
- **Time scales and oscillatory bursts**
- Circadian rhythms.

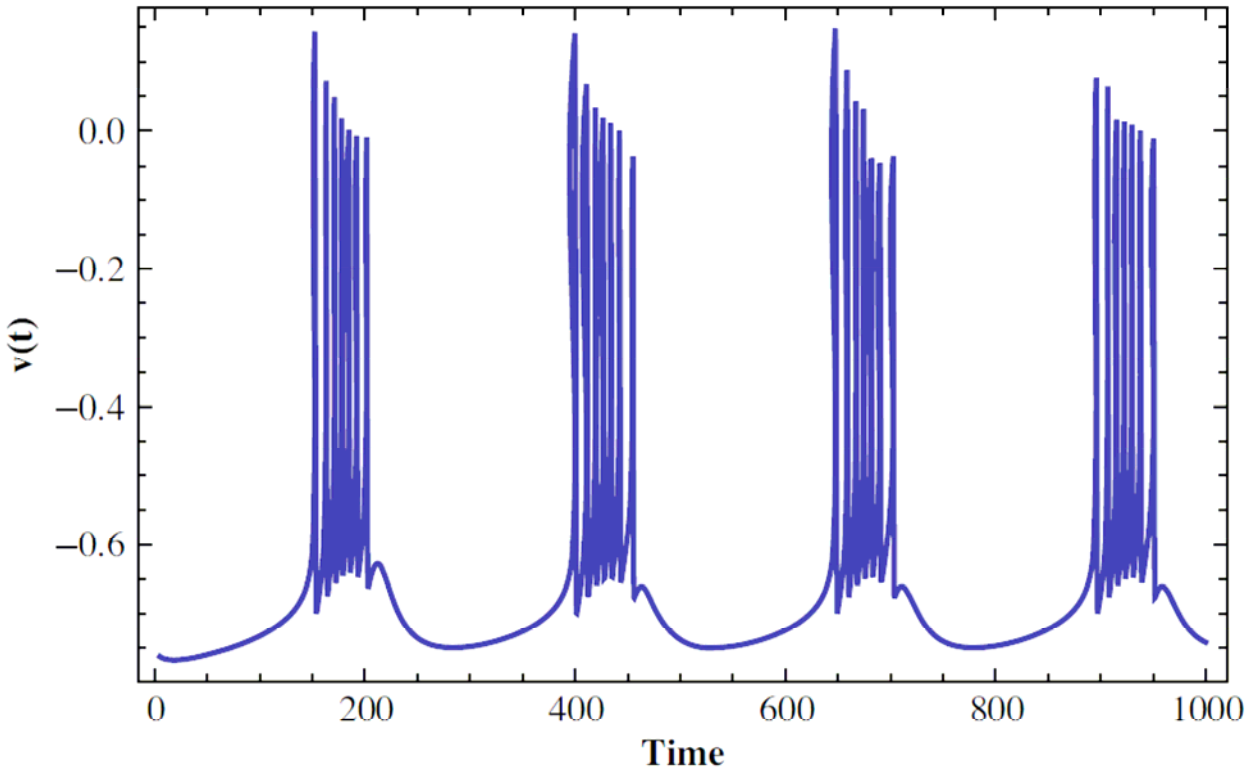


Figure 11. Bursting dynamics: an interplay of fast and slow time scales.

Neurons (nerve cells) and endocrine signalling

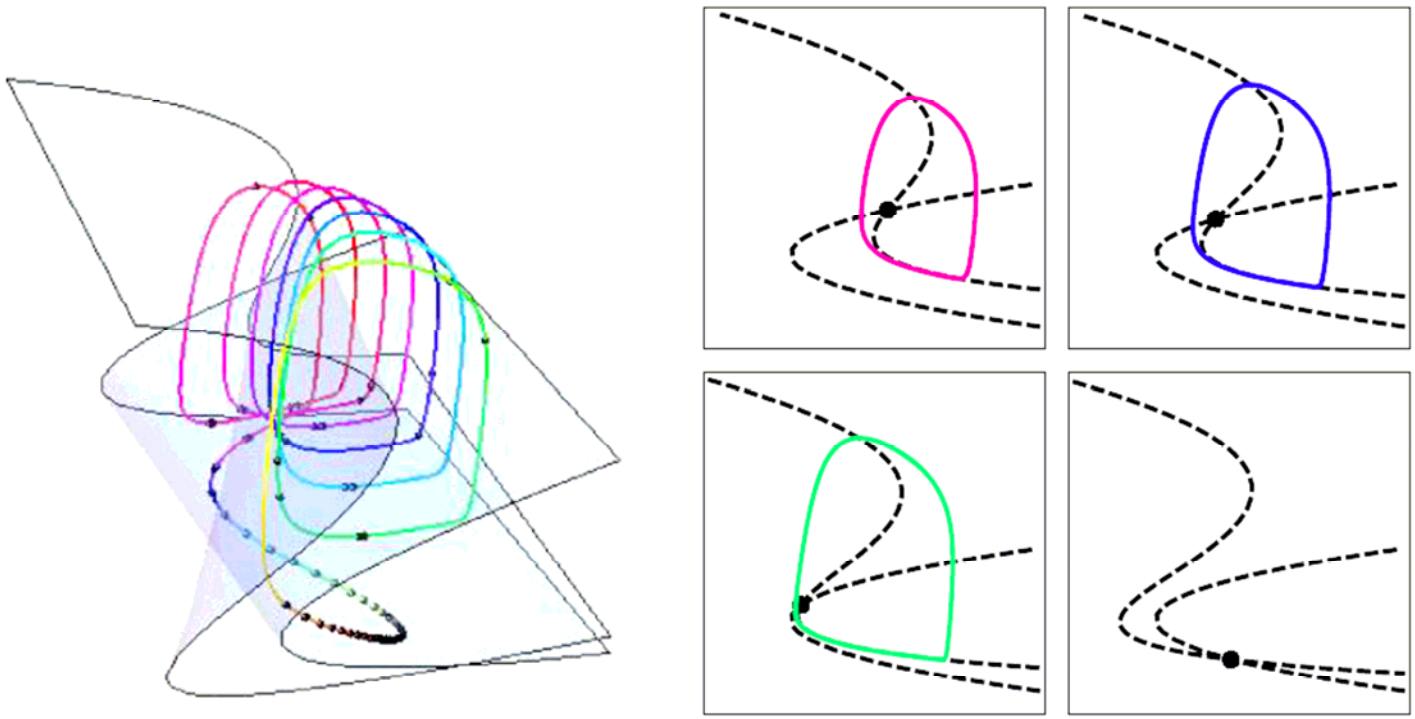


Figure 12. Trajectory of system and phase portraits. (a) State trajectory of bursting dynamics superimposed on the nullclines for the fast subsystem. (b) Trajectory of the fast subsystem, sliced at various values of the slow variable.

Bifurcation analysis of neuronal burst dynamics

Examples of inverse bifurcation analysis and design

- Oscillatory behavior in *Escherichia coli*
- The repressilator
- The mitotic cell cycle
- Time scales and oscillatory bursts
- **Circadian rhythms.**

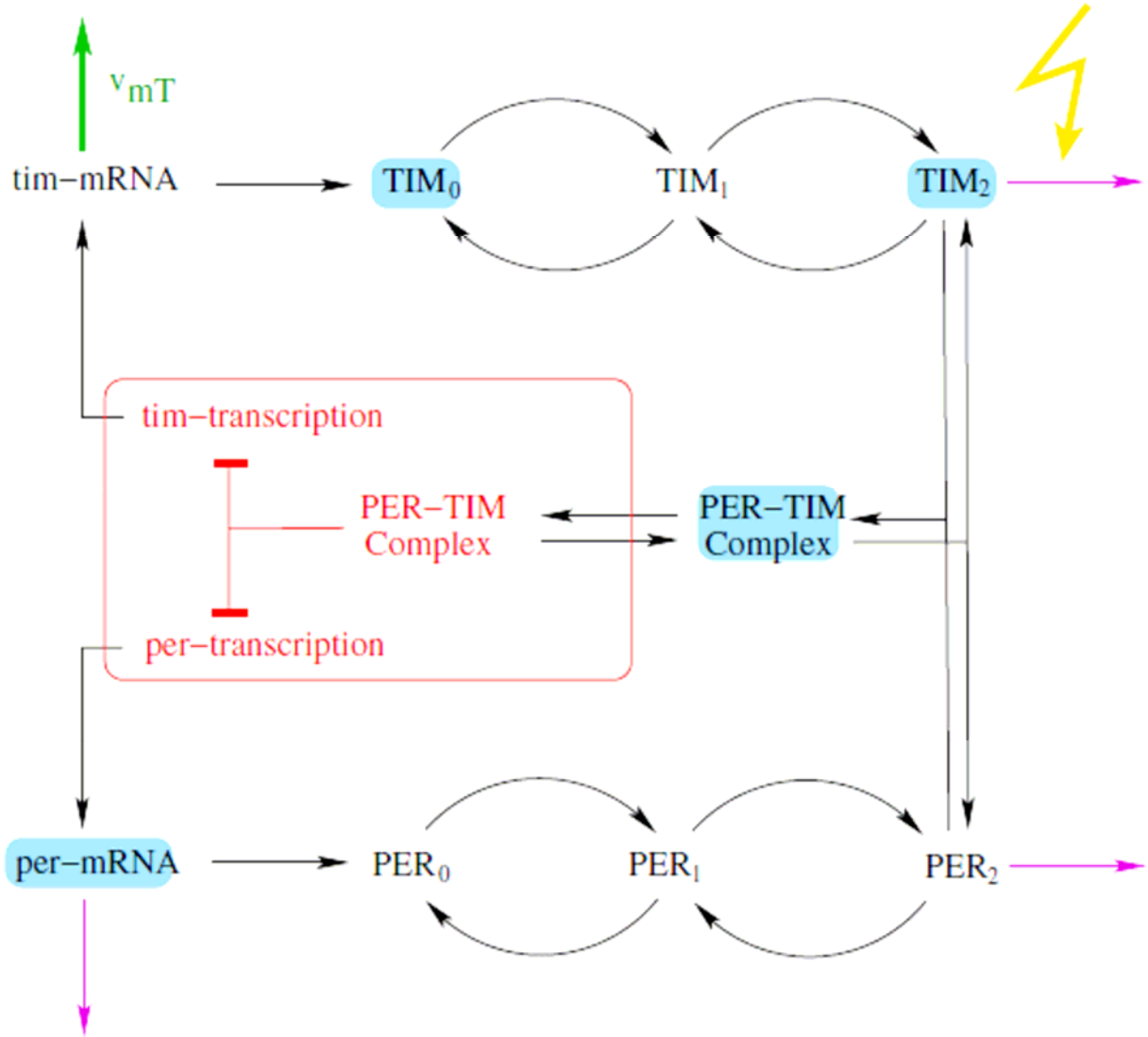


Figure 30. Species involved in the identified mechanisms from the circadian rhythm example.

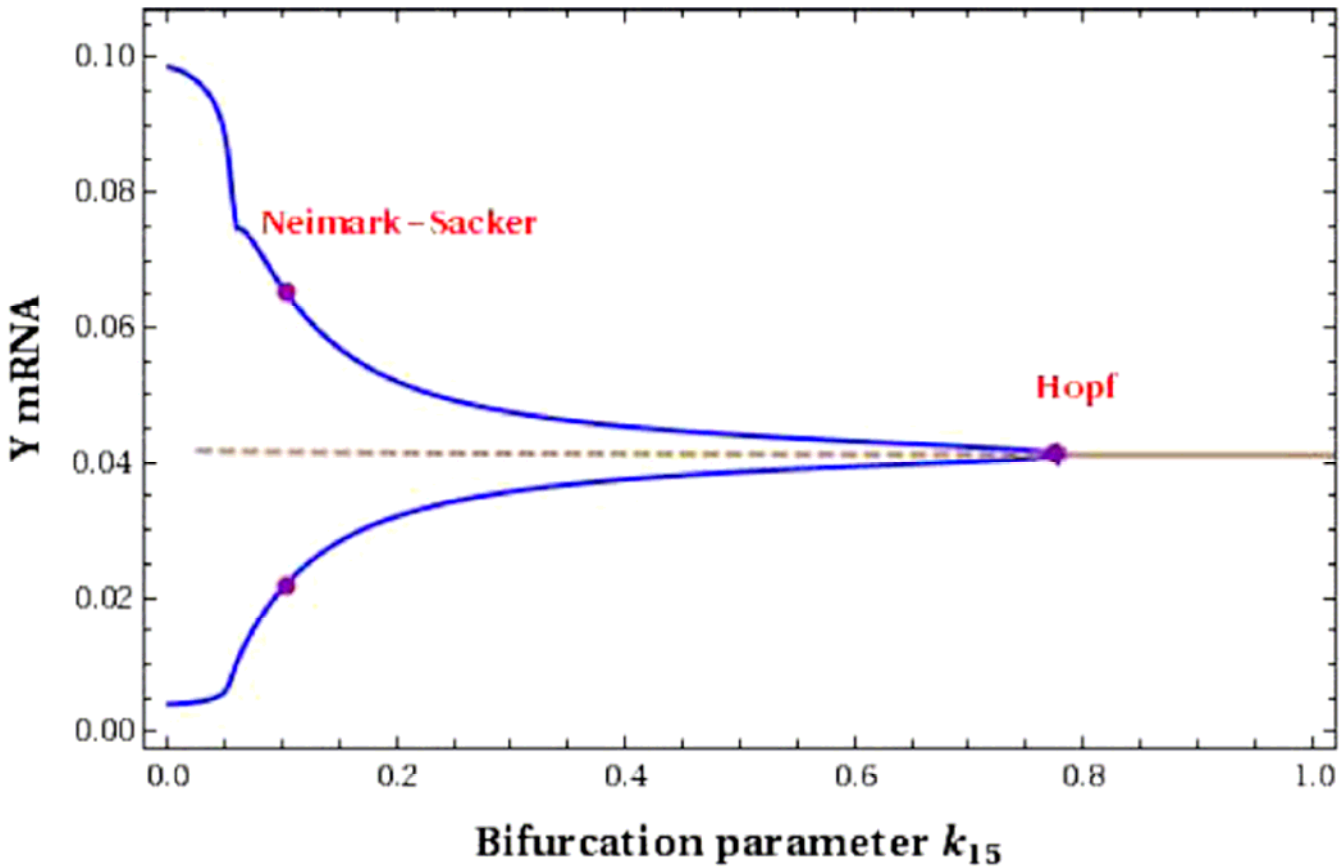


Figure 14. Bifurcation diagram of the 3-loop circadian rhythm system proposed in [102].

Bifurcation analysis of circadian rhymths

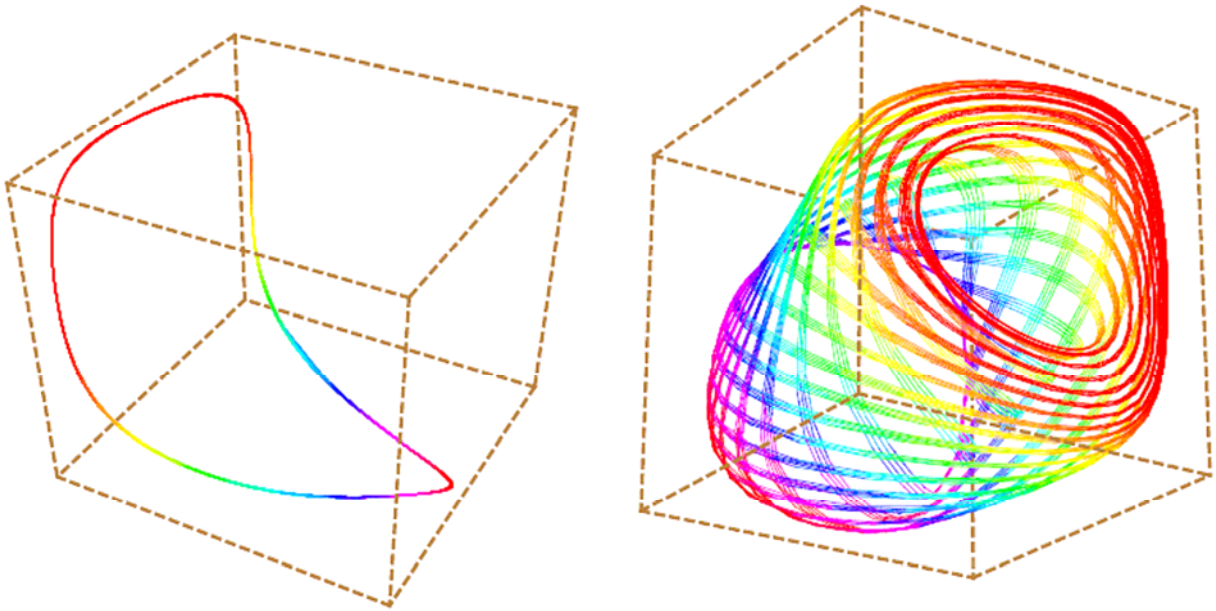


Figure 15. State trajectories of the 3-loop circadian rhythm system proposed in [102]. (a) Loop-like structure for limit cycle solution. (b) Torus-like structure after undergoing the Neimark–Sacker bifurcation.

Bifurcation analysis of circadian rhymths

1. From biochemical kinetics to quantitative biology
2. Forward and inverse problems in reaction kinetics
3. Modeling biochemical reaction kinetics
4. Examples of inverse bifurcation analysis and design
5. **Problems and perspectives of systems biology**

Challenges of quantitative biology

1. Validation of data from different sources
2. Low particle numbers and stochasticity
3. Conformational heterogeneity of biomolecules
4. Spatial heterogeneity of cells and cell organelles
5. High dimensionality of molecular dynamical systems

Challenges of quantitative biology

1. **Validation of data from different sources**
2. Low particle numbers and stochasticity
3. Conformational heterogeneity of biomolecules
4. Spatial heterogeneity of cells and cell organelles
5. High dimensionality of molecular dynamical systems

Challenges of quantitative biology

1. Validation of data from different sources
2. **Low particle numbers and stochasticity**
3. Conformational heterogeneity of biomolecules
4. Spatial heterogeneity of cells and cell organelles
5. High dimensionality of molecular dynamical systems

Challenges of quantitative biology

1. Validation of data from different sources
2. Low particle numbers and stochasticity
3. **Conformational heterogeneity of biomolecules**
4. Spatial heterogeneity of cells and cell organelles
5. High dimensionality of molecular dynamical systems

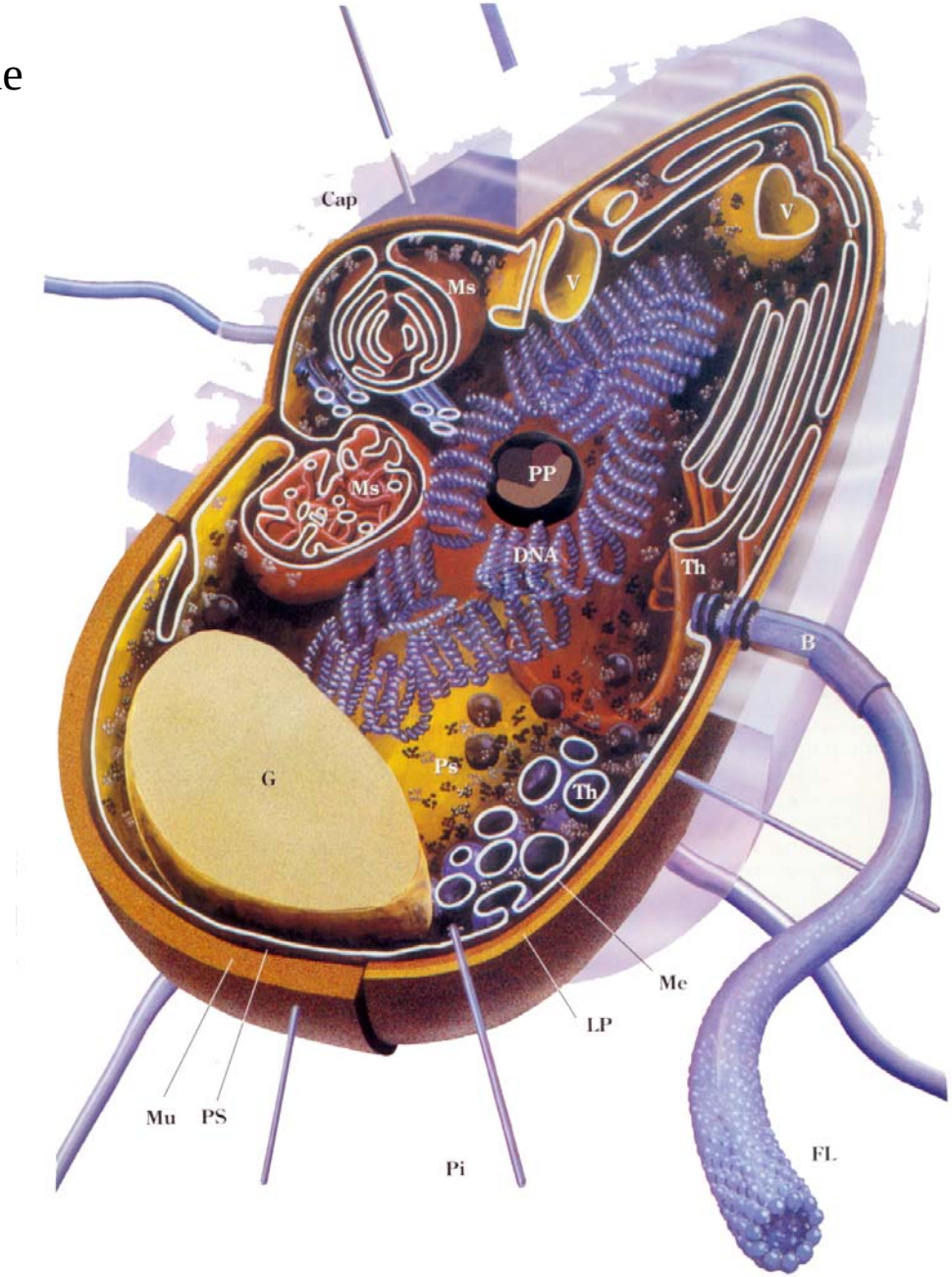
Challenges of quantitative biology

1. Validation of data from different sources
2. Low particle numbers and stochasticity
3. Conformational heterogeneity of biomolecules
4. **Spatial heterogeneity of cells and cell organelles**
5. High dimensionality of molecular dynamical systems

The bacterial cell as an example for the simplest form of autonomous life

The human body:

10^{14} cells = 10^{13} eukaryotic cells +
$\approx 9\times10^{13}$ bacterial (prokaryotic) cells,
and $\approx$ 200 eukaryotic cell types



The spatial structure of the bacterium *Escherichia coli*

Challenges of quantitative biology

1. Validation of data from different sources
2. Low particle numbers and stochasticity
3. Conformational heterogeneity of biomolecules
4. Spatial heterogeneity of cells and cell organelles
5. **High dimensionality of molecular dynamical systems**

Suitable systems for upscaling

1. Linear systems via large eigenvalue problems
2. Cascades
3. Cyclic systems
4. Sufficiently simple networks and flux analysis

Universität Wien

Coworkers

Heinz Engl, Philipp Kügler, **James Lu**, **Stefan Müller**, RICAM Linz, AT

Walter Fontana, Harvard Medical School, MA

Martin Nowak, Harvard University, MA

Christian Reidys, Nankai University, Tien Tsin, China

Christian Forst, Los Alamos National Laboratory, NM

Ivo L.Hofacker, **Christoph Flamm**, Universität Wien, AT

Stefanie Widder, Lukas Endler, Rainer Machne,
Universität Wien, AT

Universität Wien

Acknowledgement of support

Wiener Wissenschafts-, Forschungs- und Technologiefonds
(WWTF)
Project No. Mat05

Universität Wien and the Santa Fe Institute

Thank you for your attention!

Web-Page for further information:

http://www.tbi.univie.ac.at/~pks