

Perspectives of Molecular Life Sciences on the Crossroads of Mathematics, Computation, and Biology

Peter Schuster

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



Workshop „Bridging Mathematics and Life Sciences“

CEMM-RICAM, Wien, 16.02.2012

Web-Page for further information:

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

Peter Schuster. Is there a Newton of the blade of grass? The complex relation between mathematics, physics, and biology. *Complexity* **16/6**: 5-9, 2011.

Peter Schuster. Mathematical modeling of evolution. Solved and open problems. *Theory in Biosciences* **130**:71-89, 2011

1. Prologue: Mathematics and biology
2. Modeling specific biological systems
3. Networks and evolution
4. Perspectives

- 1. Prologue: Mathematics and biology**
2. Modeling specific biological systems
3. Networks and evolution
4. Perspectives

Immanuel Kant. 1786. *Metaphysical Foundations of Science*:

... I maintain only that **in every special doctrine of nature only so much science proper can be found as there is mathematics in it.** ...



Immanuel Kant, 1724-1804

Immanuel Kant. 1786. *Metaphysische Anfangsgründe der Naturwissenschaft*:

.....Ich behaupte nur, daß in jeder besonderen Naturlehre nur so viel **eigentliche Wissenschaft** angetroffen werden könnte, als darin **Mathematik** anzutreffen ist.

Immanuel Kant. 1790. Critique of Judgment,
chapter 85:

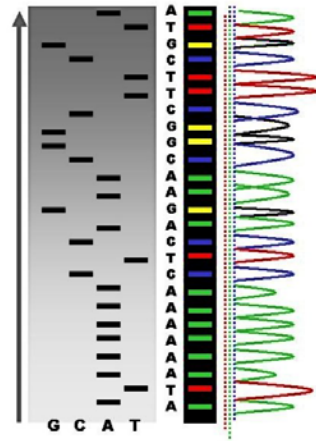
... **there will never be a Newton of the blade of grass,**
because human science will never be able to explain
how a living being can originate from inanimate
matter. ...



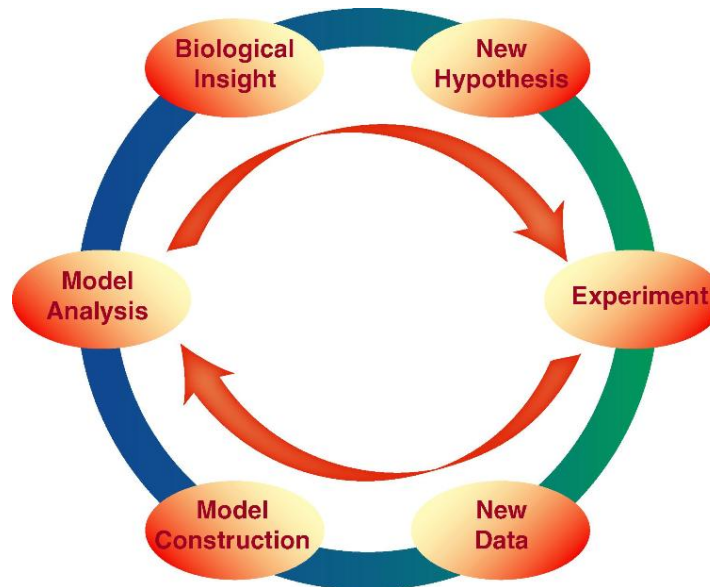
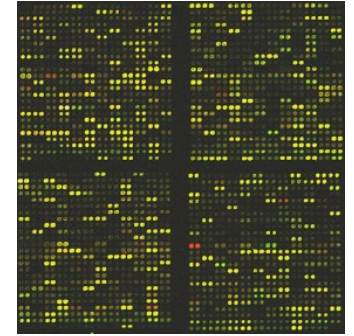
Immanuel Kant, 1724-1804

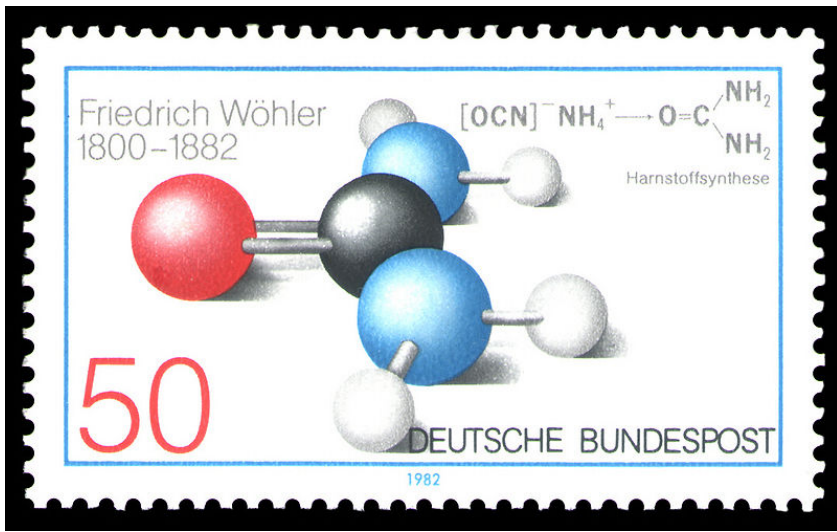
Immanuel Kant. 1790. Kritik der Urteilskraft, Kap.85:

.....und zwar so gewiß, daß man dreist sagen kann, es ist für Menschen ungereimt zu hoffen, daß noch etwa dereinst ein Newton aufstehen könnte, der auch nur die Erzeugung eines Grashalms nach Naturgesetzen, die keine Absicht geordnet hat, begreiflich machen werde, sondern man muß diese Einsicht den Menschen schlechterdings absprechen.



- 1953 molecular biology
- 1958 protein sequencing
- 1977 DNA sequencing
- 1978 bioinformatics
- 1986 genomics
- 1997 proteomics
- 1997 functional genomics
- 2000 systems biology





synthesis of urea
from ammonium
cyanate

1828

synthetic biology and biotechnology

Ian Wilmut and Keith Campbell



Dolly, the sheep
1996 - 2003

gave birth to
Bonnie,
Sally & Rosie,
Lucy & Darcy & Cotton

1996

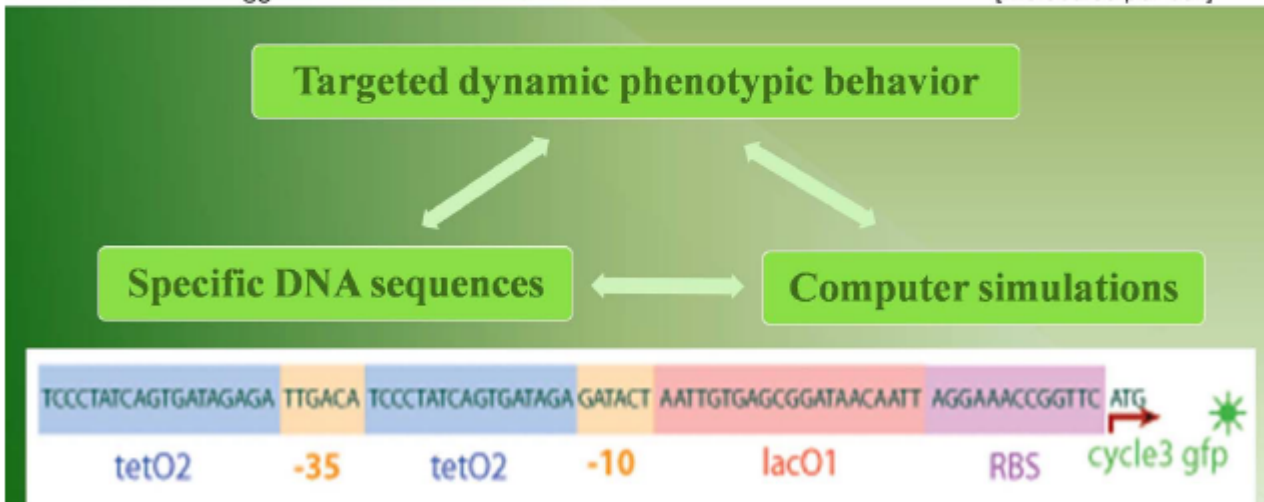
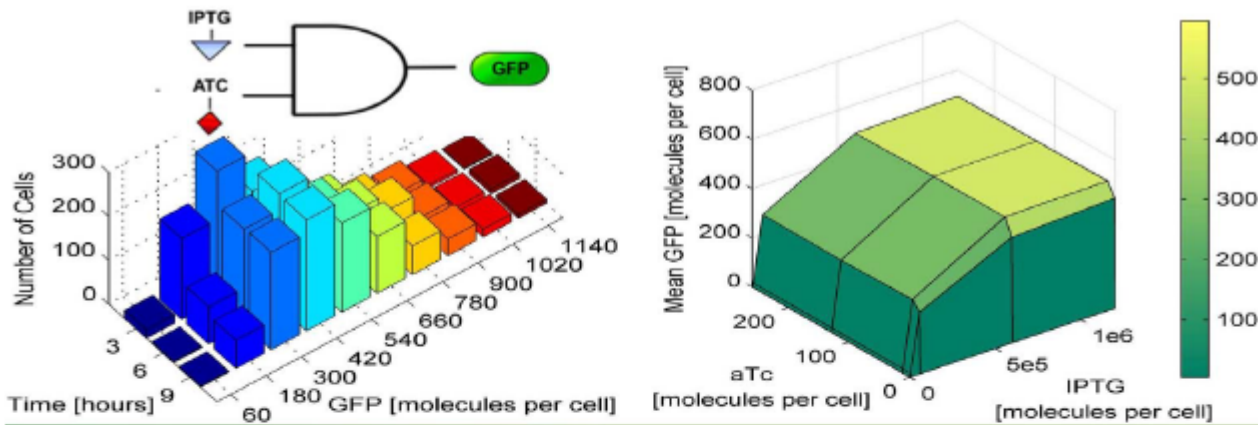


Figure 1

A major challenge in synthetic biology is to rationally select DNA sequences that result in targeted dynamic phenotypes. For example, with simulations using Hy3S [29] we are experimenting with multiple alternative promoter sequences to identify the optimal AND gate synthetic gene network, with tetracycline (atc) and IPTG as inputs and green fluorescence protein (GFP) as output.

2007

Yiannis N. Kaznessis. Models for synthetic biology.
BMC Systems Biology 1:e47, 2007.

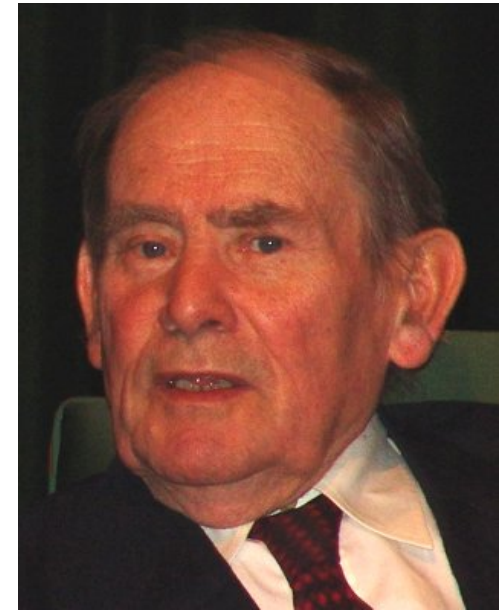


Sir Peter Brian Medawar,
1915 - 1987

..... no new principle will declare itself
from below a heap of facts.

Torbjörn Fagerström, Peter Jagers, Peter Schuster, and Eörs Szathmáry.
Biologists put on mathematical glasses. *Science* **271**:2039-240, 1996.

..... the prime intellectual task of the future lies in constructing an appropriate theoretical framework for biology theoretical biology has a bad name because of its past I have decided to forget and forgive the past and call it - *the badly required new discipline - theoretical biology*.



Sydney Brenner, 1927 -

Sydney Brenner. Theoretical biology in the third millenium.
Phil.Trans.Roy.Soc.London B **354**:1963-1965, 1999

Nothing makes sense in biology
except in the light of evolution, ...



Theodosius Dobzhansky,
1900 - 1975

Theodosius Dobzhansky. Biology, molecular and organismic.
American Zoologist 4:443-452, 1974.

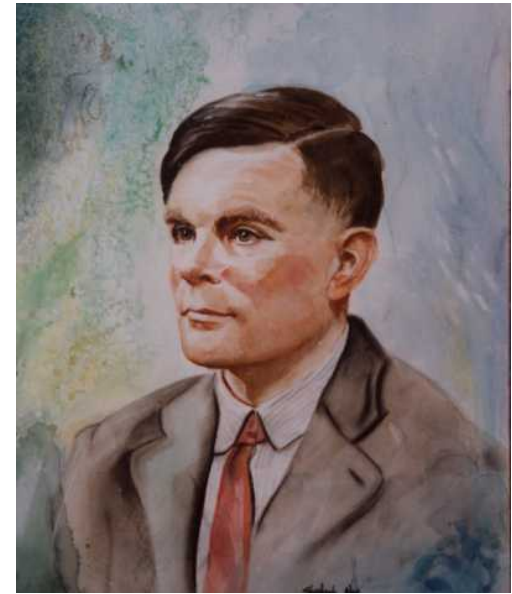
1. Prologue: Mathematics and biology
- 2. Modeling specific biological systems**
3. Networks and evolution
4. Perspectives



Philip Maini, 1959 -

Turing patterns in embryological morphogenesis:

„..... although reaction-diffusion theory provides a very elegant mechanism for segmentation, nature seems to have chosen a much less elegant way of doing it.“



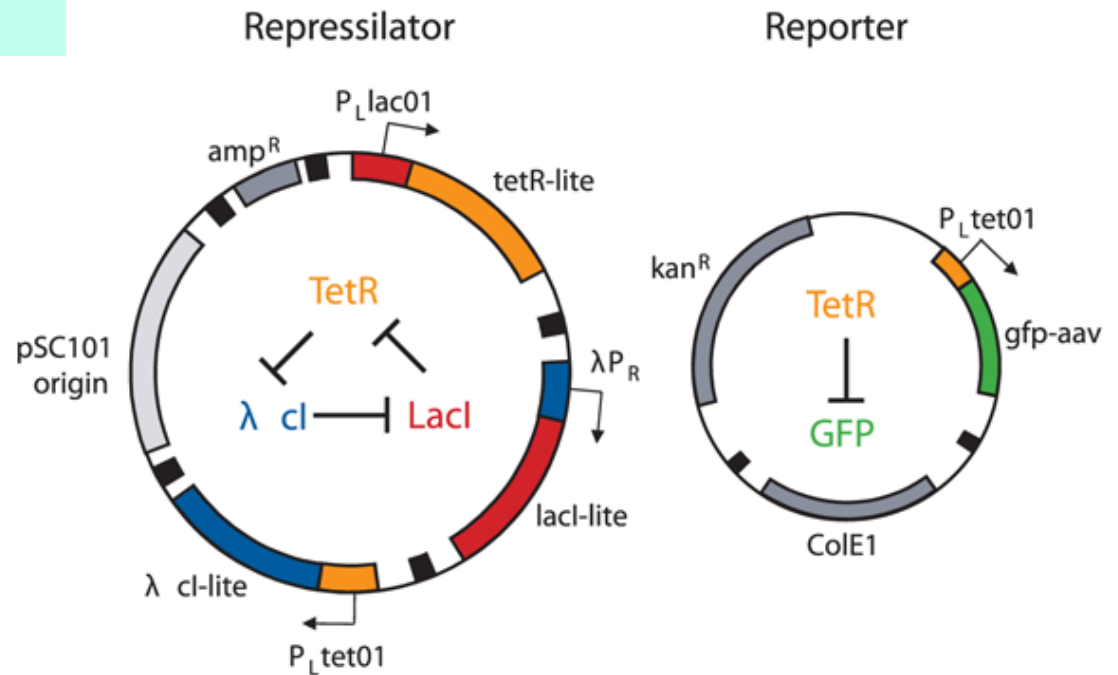
Alan M.Turing, 1912 - 1954

Philip K. Maini, Kevin J. Painter, and Helene Nguyen Phong Chau.
Spatial pattern formation in chemical and biological systems.
J.Chem.Soc., Fraday Trans. **93**:3602-3610, 1997.

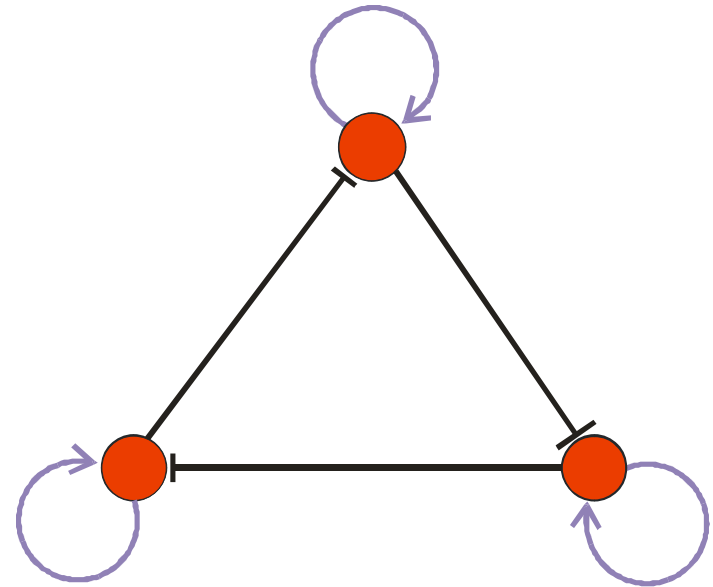
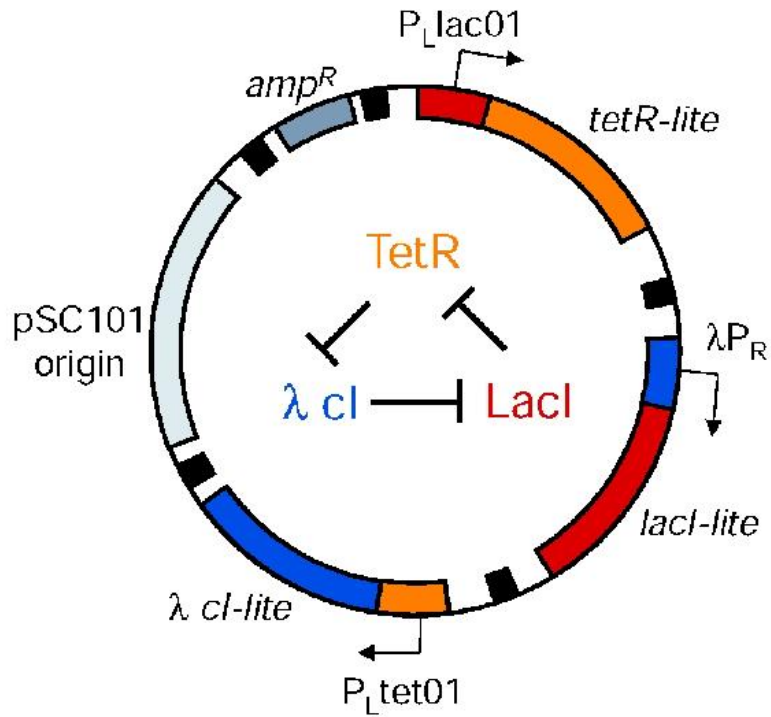
Synthetic biology

Analysis of synthetic genetic network

The repressilator

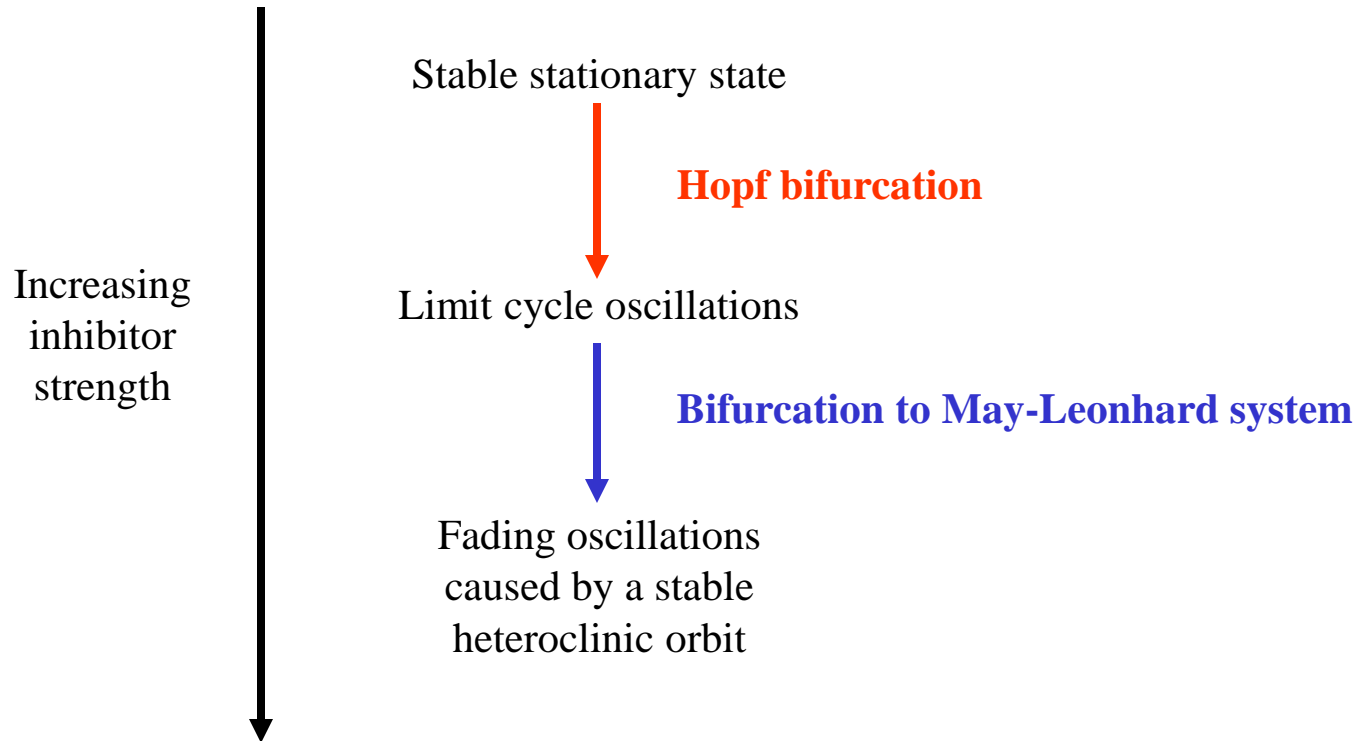


Michael B. Elowitz and Stanislas Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature* **403**:335-338, 2000.

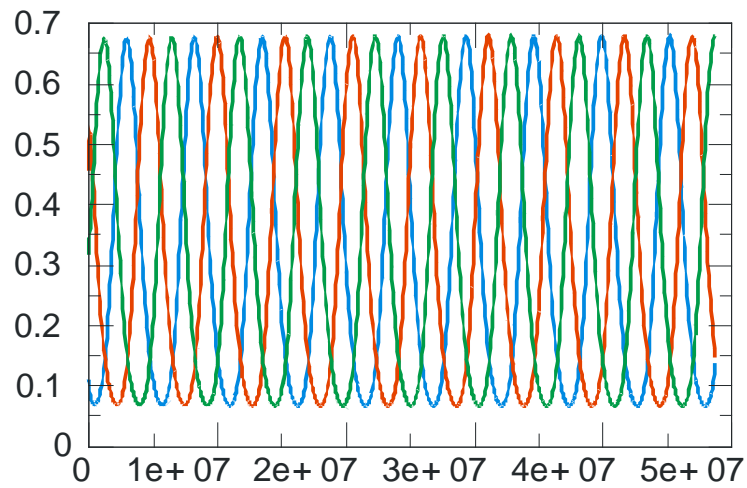


An example analyzed and simulated by MiniCellSim

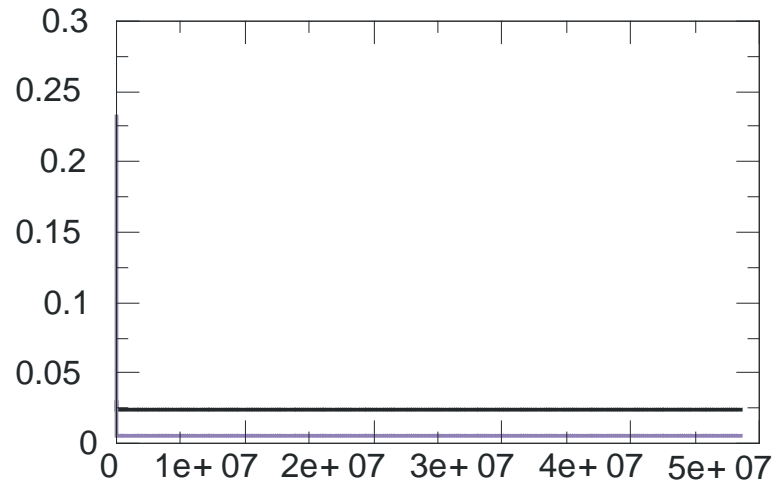
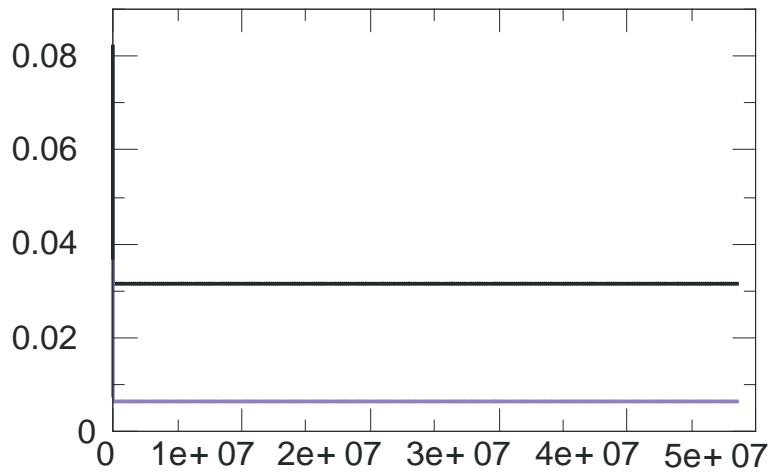
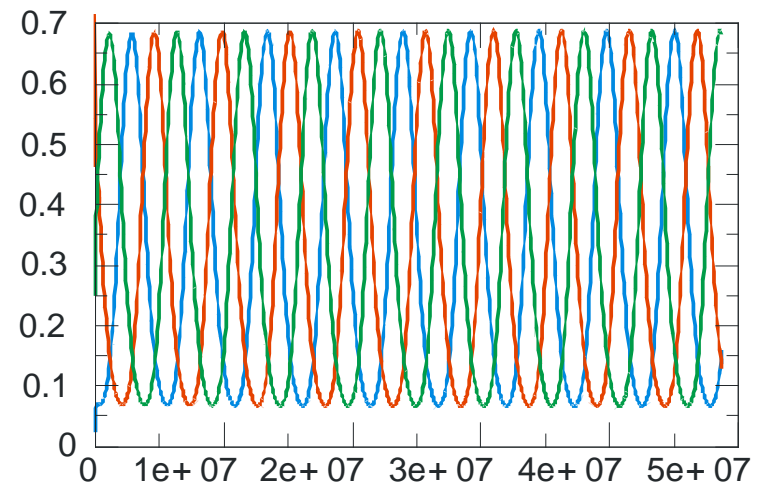
Stefan Müller, Josef Hofbauer, Lukas Endler, Christoph Flamm, Stefanie Widder and Peter Schuster.
 A generalized model of the repressilator. *J. Math. Biol.* **53**:335-338, 2006.



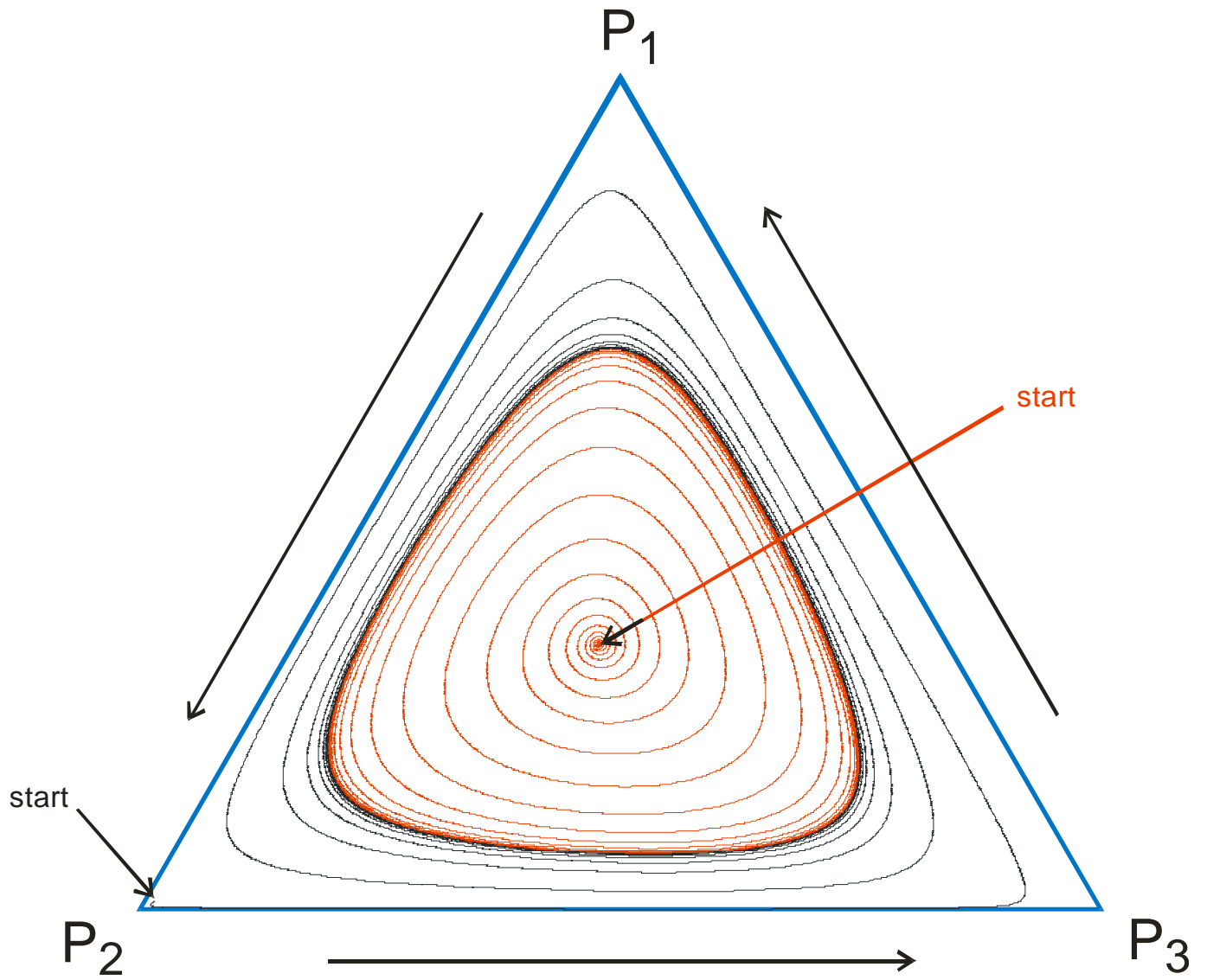
Proteins



mRNAs

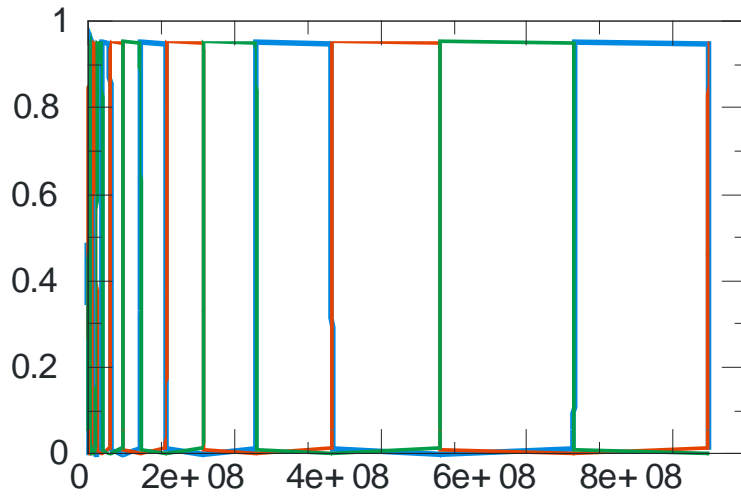


The repressilator limit cycle

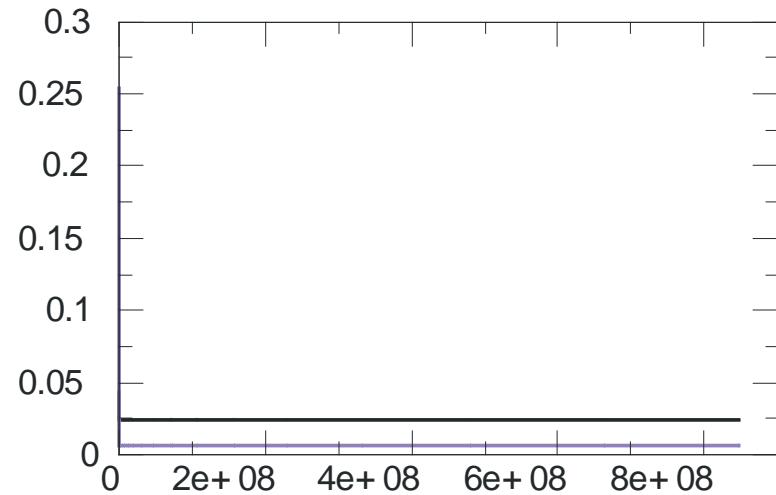
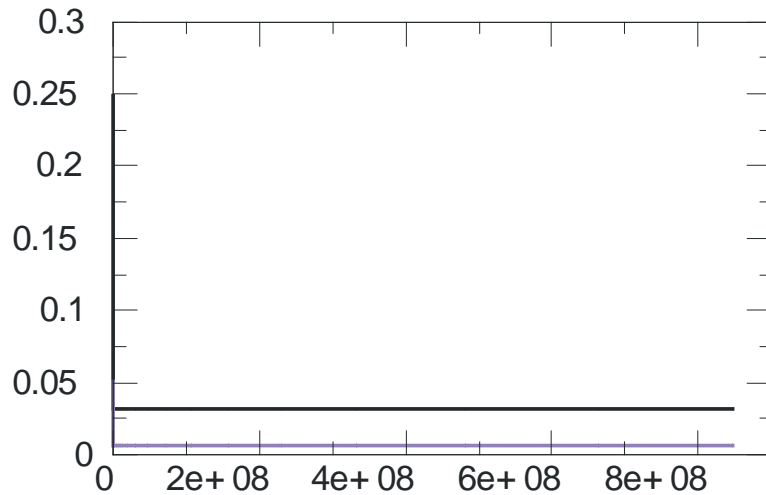
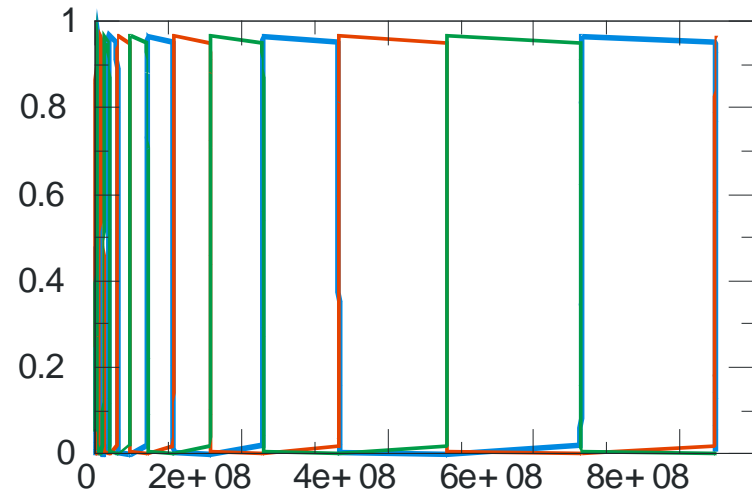


The repressilator limit cycle

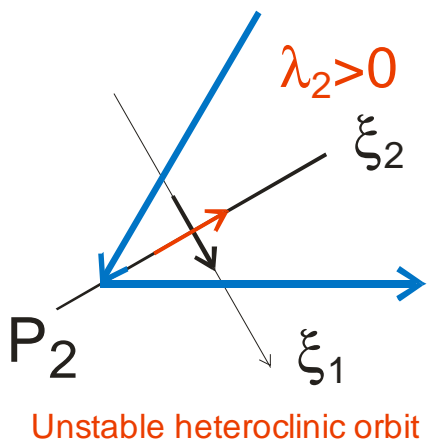
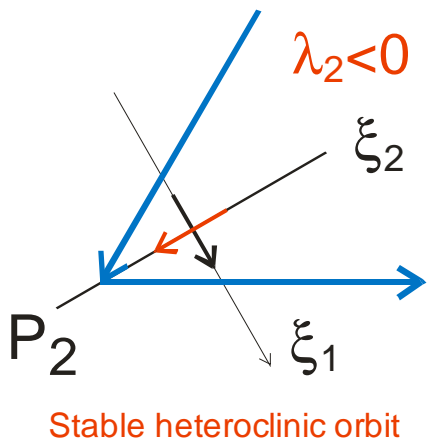
Proteins



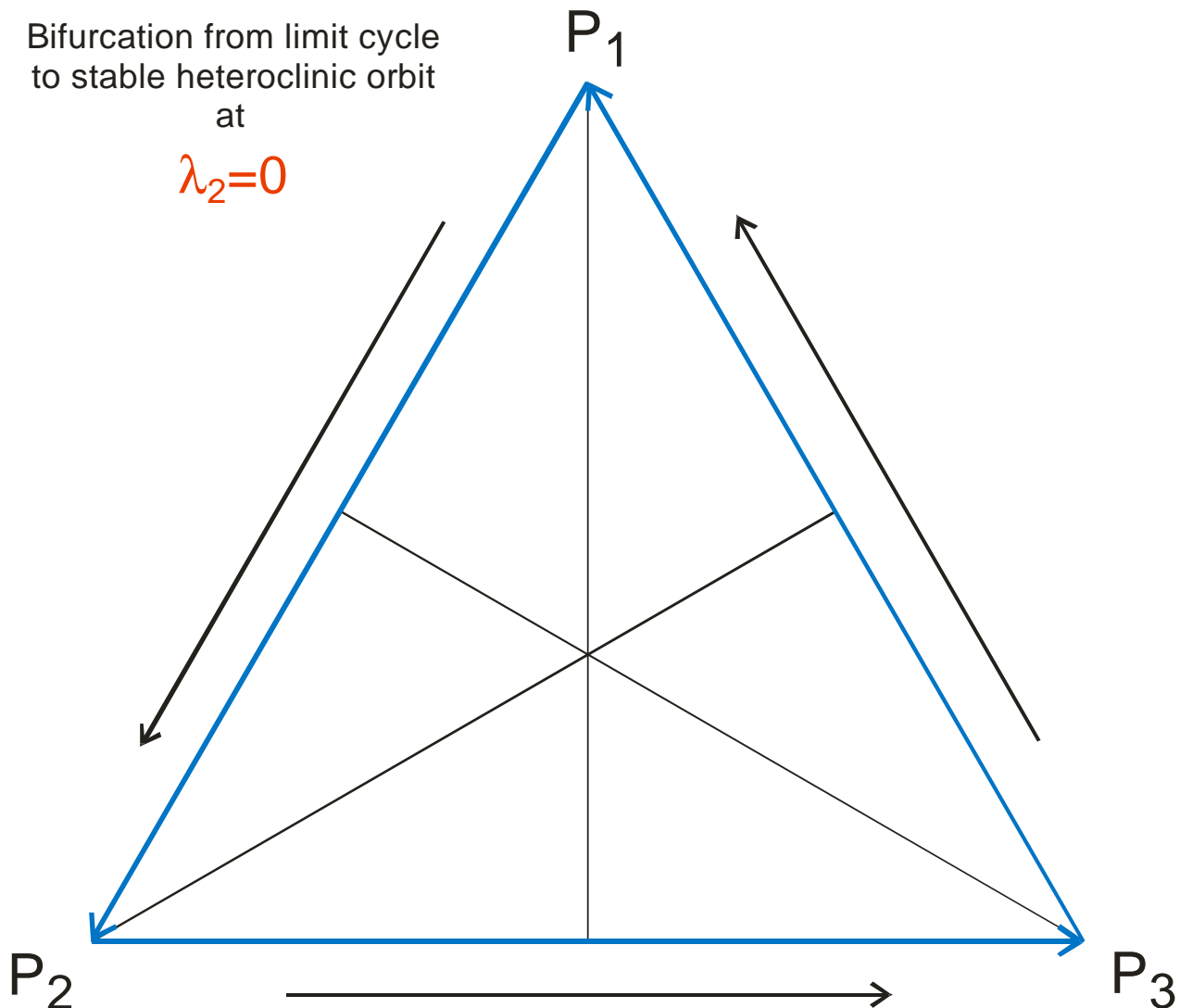
mRNAs



The repressilator heteroclinic orbit



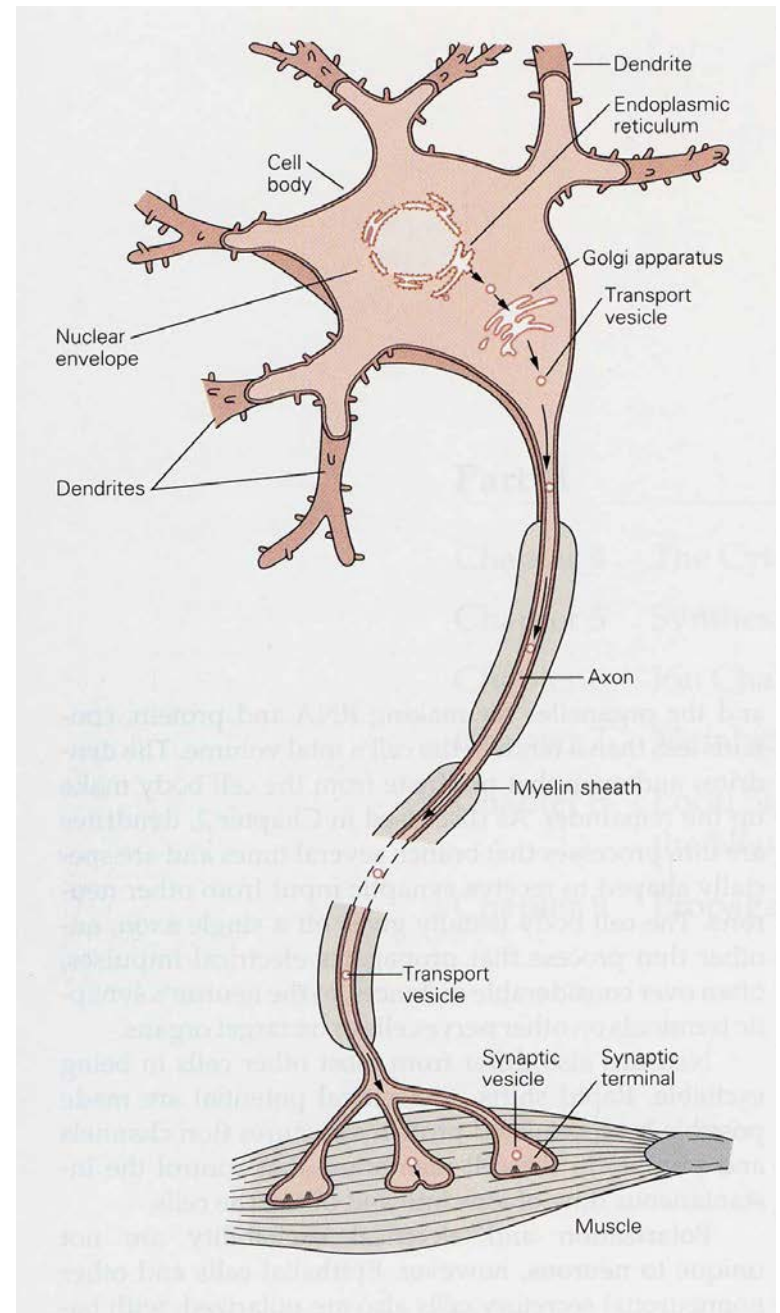
Bifurcation from limit cycle
to stable heteroclinic orbit
at
 $\lambda_2 = 0$



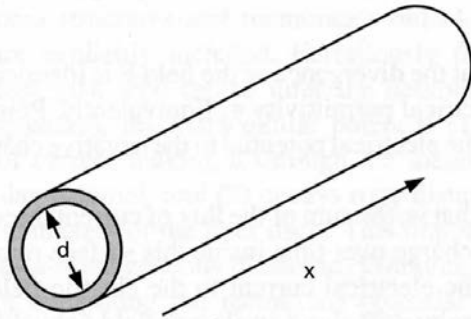
The repressilator heteroclinic orbit

Neurobiology

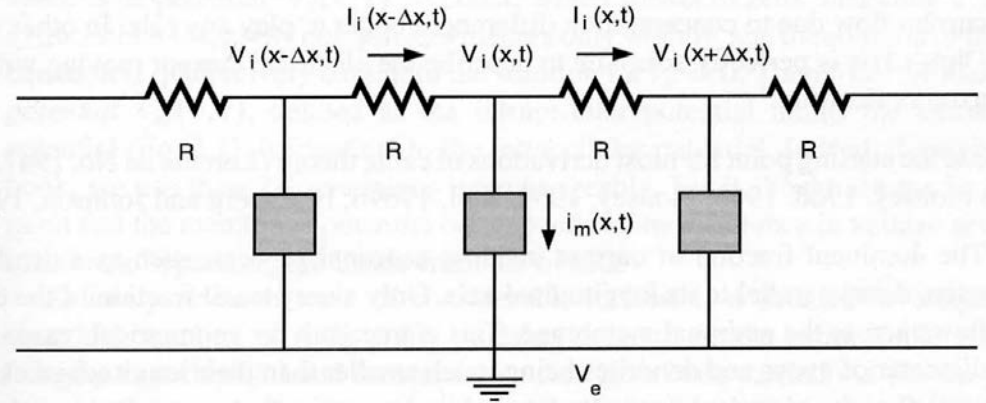
Propagation of a nerve pulse along the axon



A single neuron signaling to a muscle fiber



A



B

Fig. 2.2 ELECTRICAL STRUCTURE OF A CABLE (A) Idealized cylindrical axon or dendrite at the heart of one-dimensional cable theory. Almost all of the current inside the cylinder is longitudinal due to geometrical (the radius is much smaller than the length of the cable) and electrical factors (the membrane covering the axon or dendrite possesses a very high resistivity compared to the intracellular cytoplasm). As a consequence, the radial and angular components of the current can be neglected, and the problem of determining the potential in these structures can be reduced from three spatial dimensions to a single one. On the basis of the bidomain approximation, gradients in the extracellular potentials are neglected and the cable problem is expressed in terms of the transmembrane potential $V_m(x, t) = V_i(x, t) - V_e$. (B) Equivalent electrical structure of an arbitrary neuronal process. The intracellular cytoplasm is modeled by the purely ohmic resistance R . This tacitly assumes that movement of carriers is exclusively due to drift along the voltage gradient and not to diffusion. Here and in the following the extracellular resistance is assumed to be negligible and V_e is set to zero. The current per unit length across the membrane, whether it is passive or contains voltage-dependent elements, is described by i_m and the system is characterized by the second-order differential equation, Eq. 2.5.

$$\frac{dV}{dt} = \frac{1}{C_M} \left[I - \bar{g}_{Na} m^3 h (V - V_{Na}) - \bar{g}_K n^4 (V - V_K) - \bar{g}_l (V - V_l) \right]$$

$$\frac{dm}{dt} = \alpha_m (1 - m) - \beta_m m$$

$$\frac{dh}{dt} = \alpha_h (1 - h) - \beta_h h$$

$$\frac{dn}{dt} = \alpha_n (1 - n) - \beta_n n$$

Hodgkin-Huxley ODEs

Simulation of space independent Hodgkin-Huxley equations:
Voltage clamp and constant current

Alan L. Hodgkin, Andrew F. Huxley. A Quantitative Description of Membrane Current and its Application to Conduction and Excitation in Nerve.
Journal of Physiology **117**:500–544, 1952.

$$\alpha_m = \frac{x}{e^x - 1}, \quad x = \frac{25 - V}{10}; \quad \beta_m = 4 \exp\left[-\frac{V}{18}\right]$$

$$\alpha_h = 0.07 \exp\left[-\frac{V}{20}\right]; \quad \beta_h = \frac{1}{e^x - 1}, \quad x = \frac{30 - V}{10}$$

$$\alpha_n = \frac{x}{10(e^x - 1)}, \quad x = \frac{10 - V}{10}; \quad \beta_n = 0.125 \exp\left[-\frac{V}{80}\right]$$

Gating functions of the Hodgkin-Huxley equations

$$\frac{\partial m}{\partial t} = \Theta(T) [\alpha_m(1 - m) - \beta_m m]$$

$$\frac{\partial h}{\partial t} = \Theta(T) [\alpha_h(1 - h) - \beta_h h]$$

$$\frac{\partial n}{\partial t} = \Theta(T) [\alpha_n(1 - n) - \beta_n n] ,$$

$$\text{where } \Theta(T) = 3^{(T-6.3)/10}$$

Temperature dependence of the Hodgkin-Huxley equations

$$\frac{1}{R} \frac{\partial^2 V}{\partial x^2} = C \frac{\partial V}{\partial t} + \left[g_{Na} m^3 h (V - V_{Na}) + g_K n^4 (V - V_K) + g_l (V - V_l) \right] 2\pi r L$$

$$\frac{\partial m}{\partial t} = \alpha_m (1 - m) - \beta_m m$$

$$\frac{\partial h}{\partial t} = \alpha_h (1 - h) - \beta_h h$$

$$\frac{\partial n}{\partial t} = \alpha_n (1 - n) - \beta_n n$$

Hodgkin-Huxley PDEquations

Travelling pulse solution: $V(x, t) = V(\xi)$ with
 $\xi = x + \theta t$

Hodgkin-Huxley equations describing pulse propagation along nerve fibers

$$\frac{1}{R} \frac{d^2 V}{d \xi^2} = C_M \theta \frac{d V}{d \xi} + \left[\bar{g}_{Na} m^3 h (V - V_{Na}) + \bar{g}_K n^4 (V - V_K) + \bar{g}_l (V - V_l) \right] 2\pi r L$$

$$\theta \frac{d m}{d \xi} = \alpha_m (1 - m) - \beta_m m$$

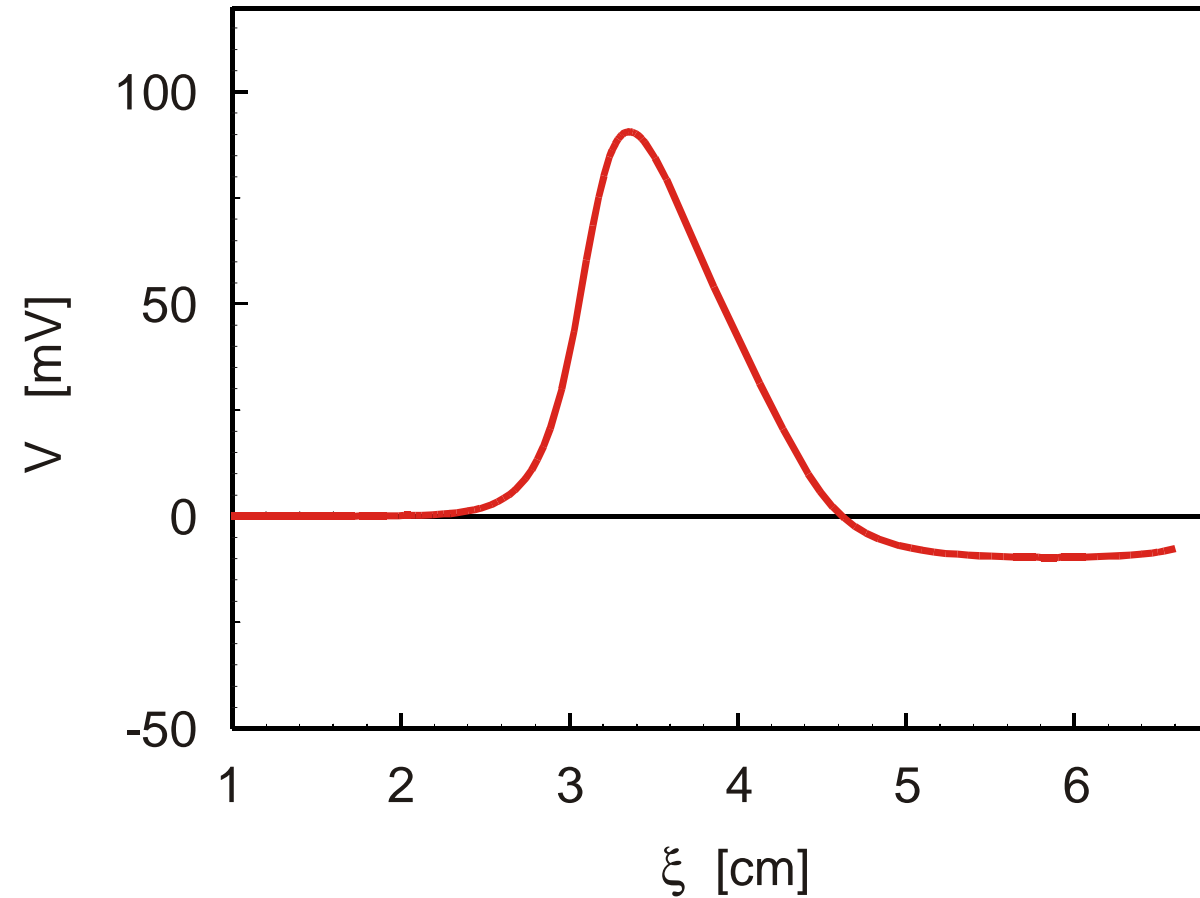
$$\theta \frac{d h}{d \xi} = \alpha_h (1 - h) - \beta_h h$$

$$\theta \frac{d n}{d \xi} = \alpha_n (1 - n) - \beta_n n$$

Hodgkin-Huxley PDEquations

Travelling pulse solution: $V(x, t) = V(\xi)$ with
 $\xi = x + \theta t$

Hodgkin-Huxley equations describing pulse propagation along nerve fibers



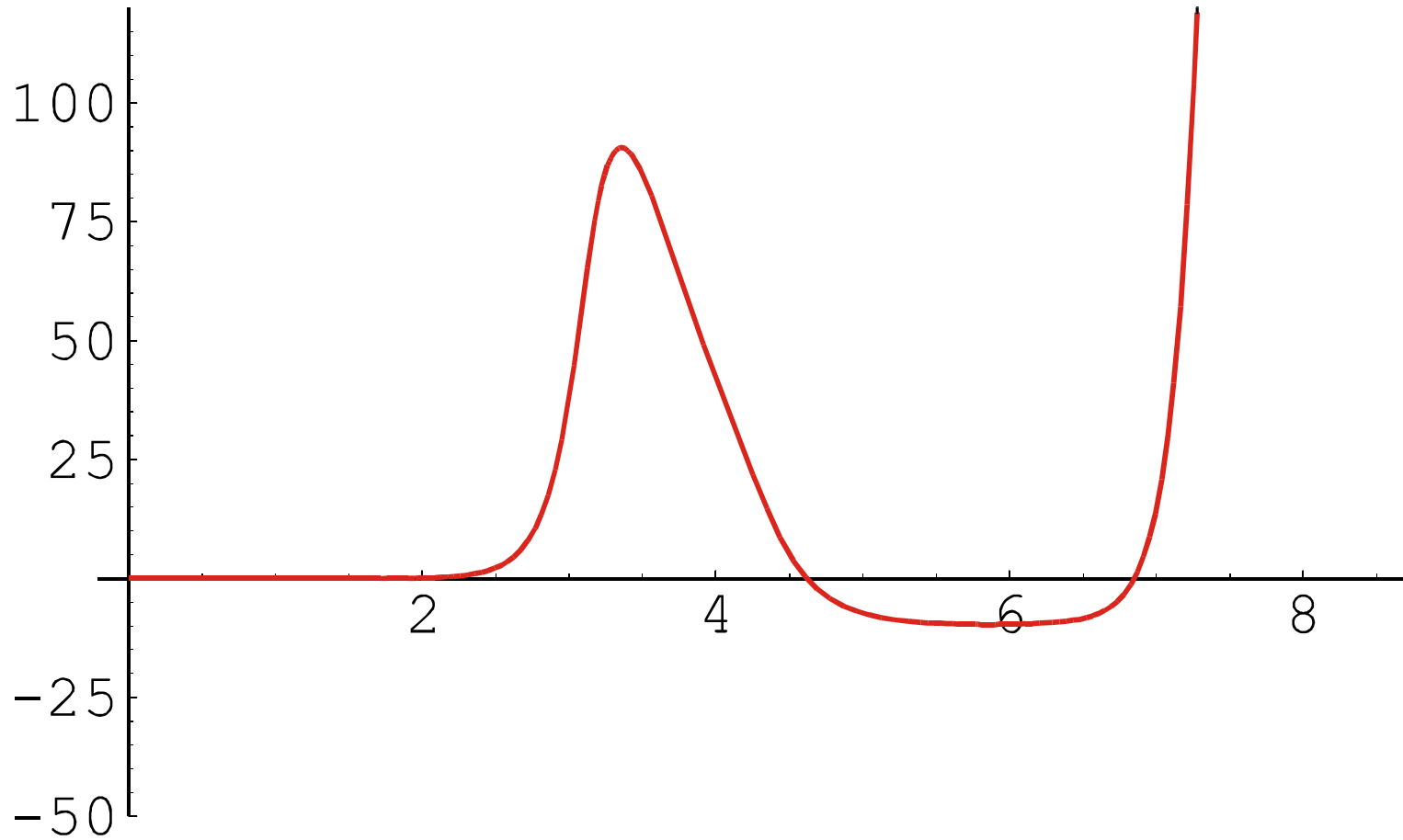
$T = 18.5$ C; $\theta = 1873.33$ cm / sec

Paul E. Phillipson and Peter Schuster. Analytical dynamics of neuron pulse propagation.

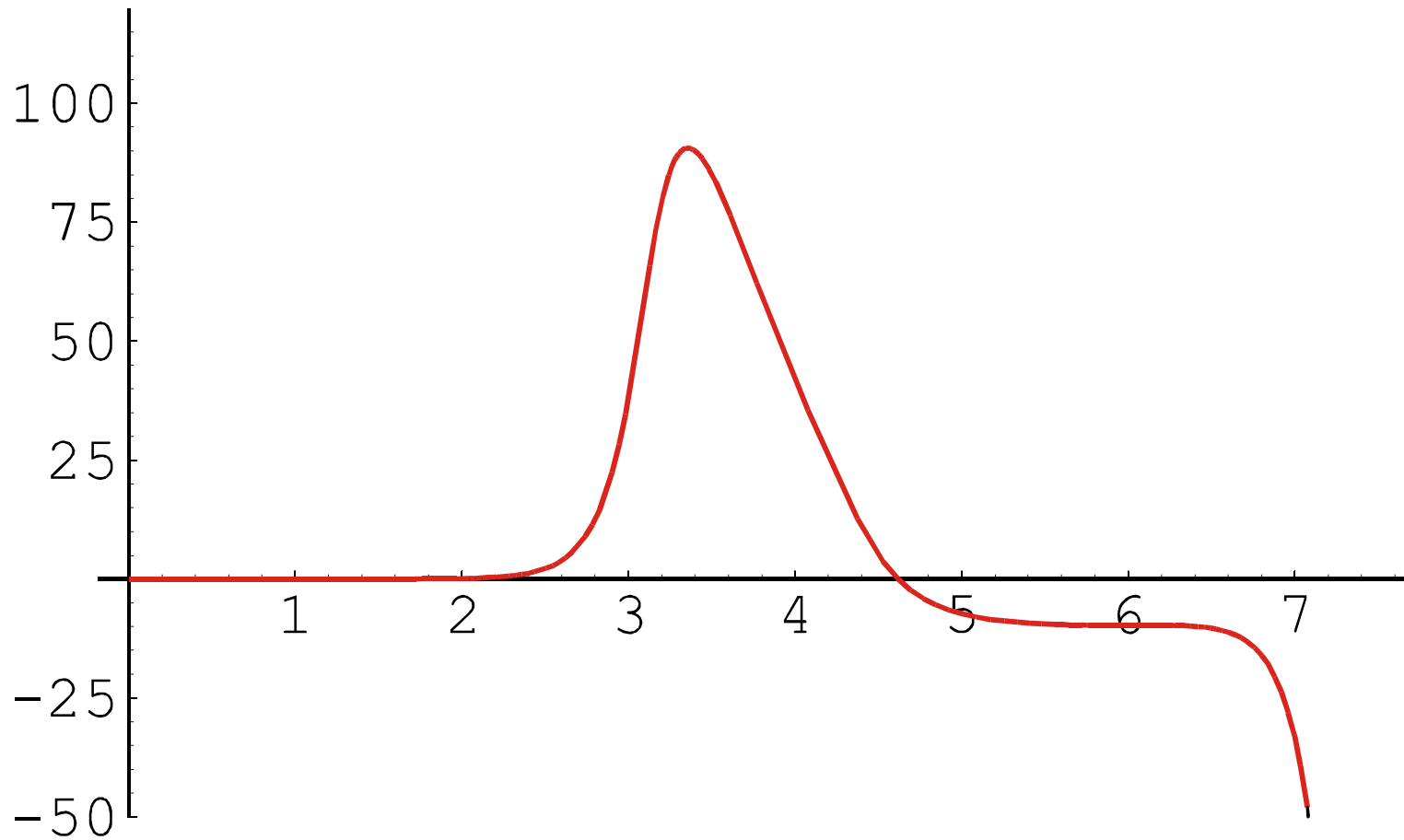
International Journal of Bifurcation and Chaos **16**:3605-3616, 2006.

- " - , - " - . A comparative study of the Hodgkin-Huxley and the Fitzhugh-Nagumo models of neuron pulse propagation.

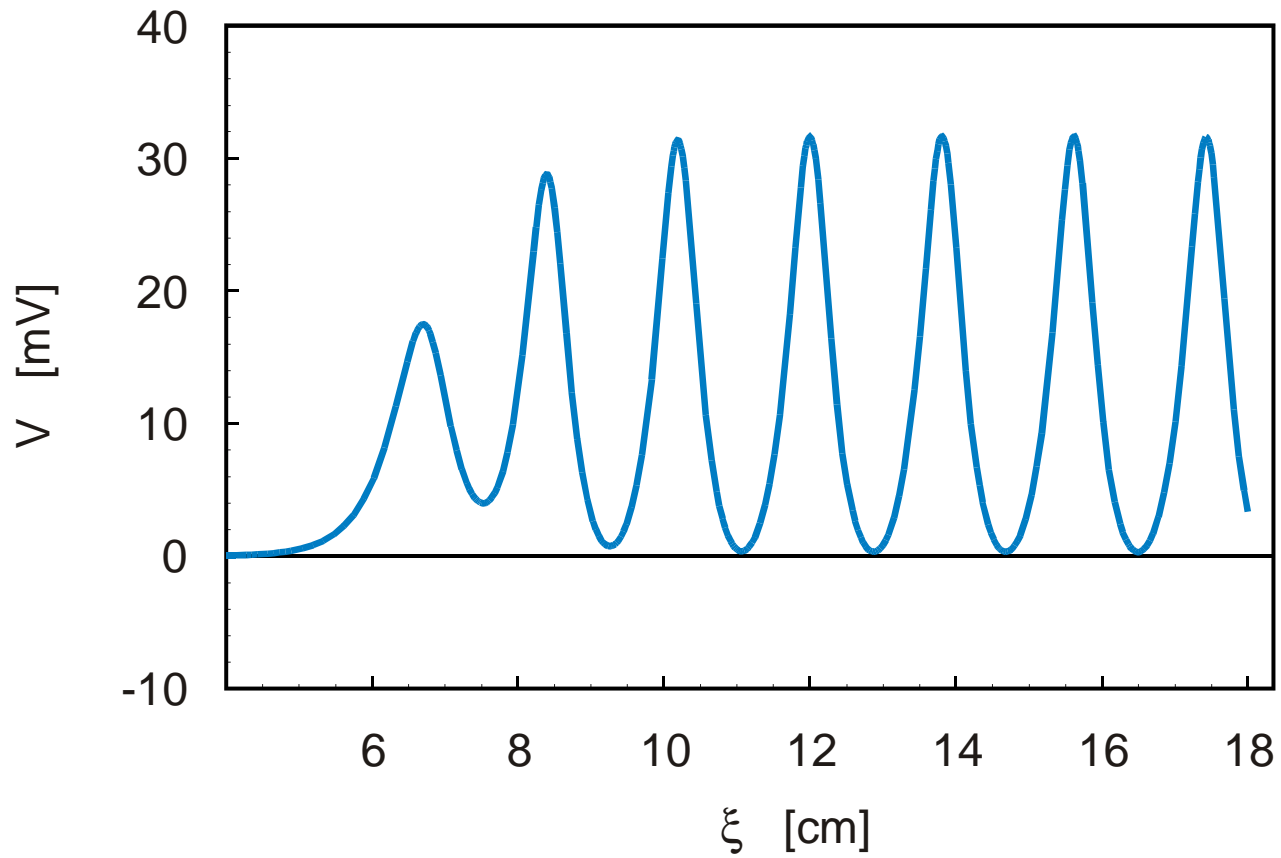
International Journal of Bifurcation and Chaos **15**:3851-3866, 2005.



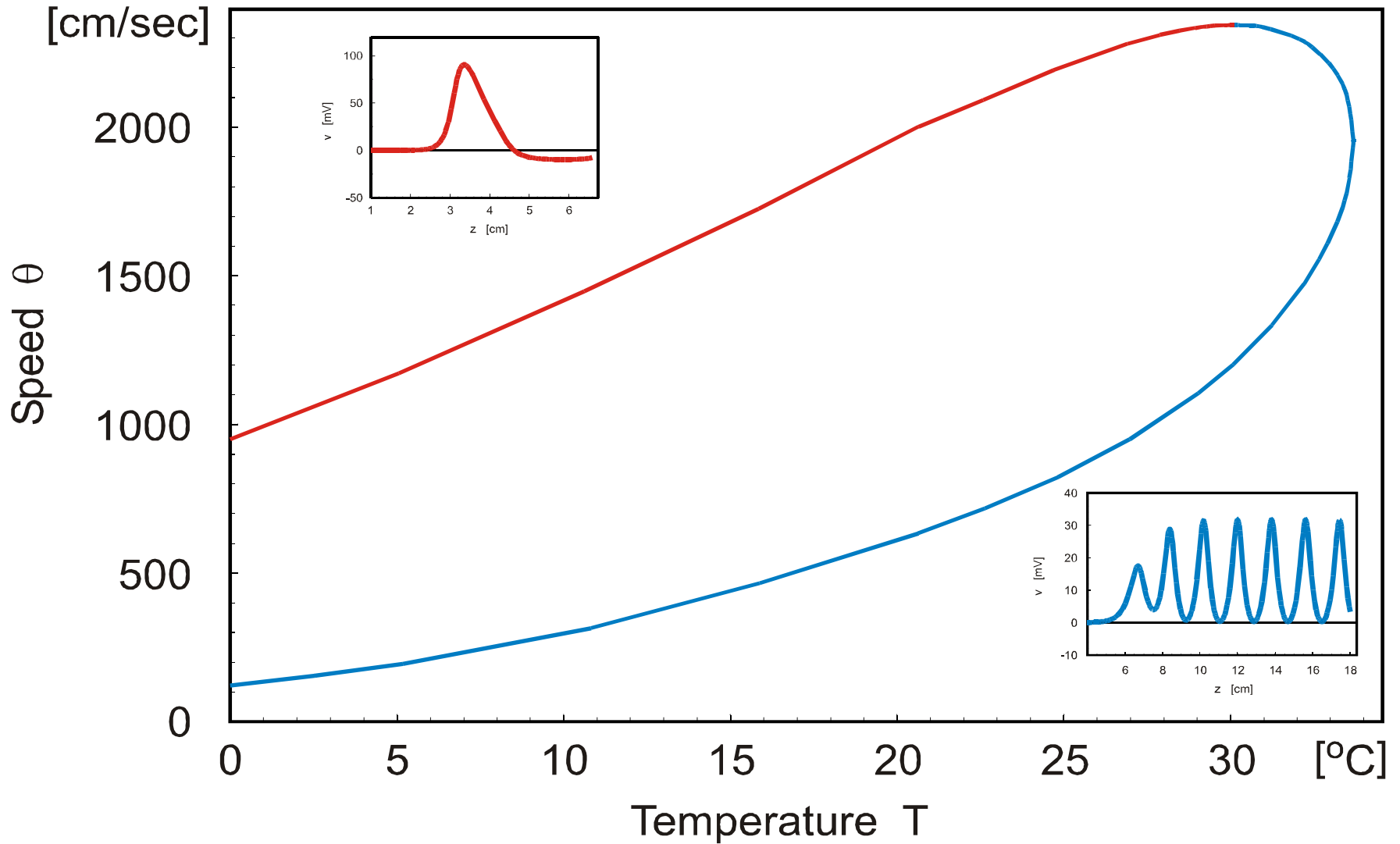
$T = 18.5 \text{ C}; \theta = 1873.3324514717698 \text{ cm / sec}$



$T = 18.5 \text{ C}; \theta = 1873.3324514717697 \text{ cm / sec}$

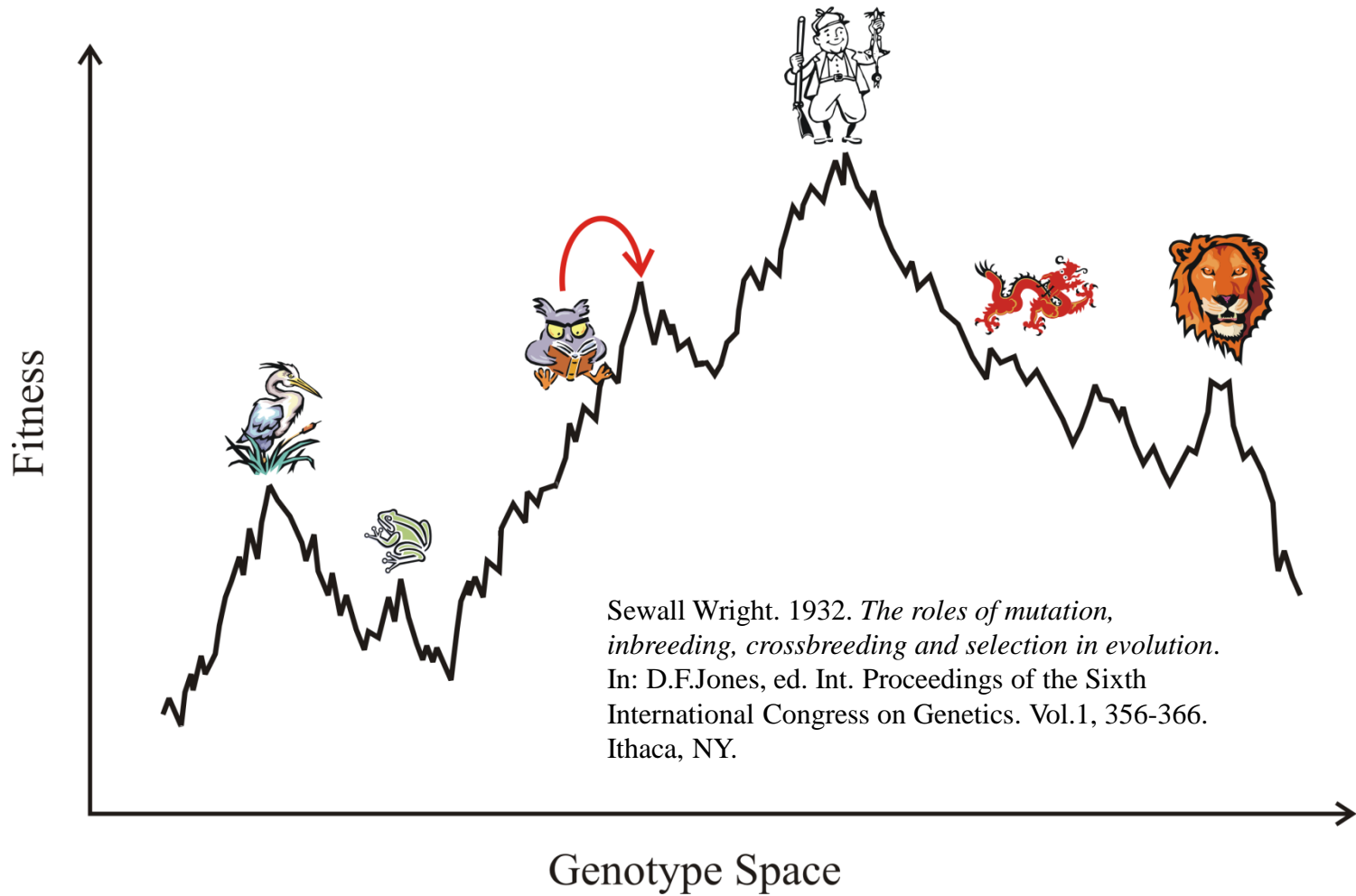


$T = 18.5$ C; $\theta = 544.070$ cm / sec

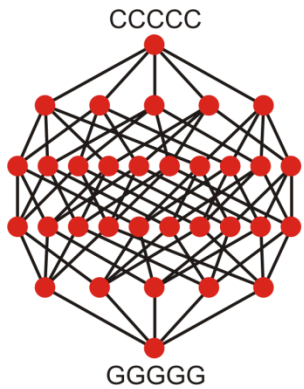


Propagating wave solutions of the Hodgkin-Huxley equations

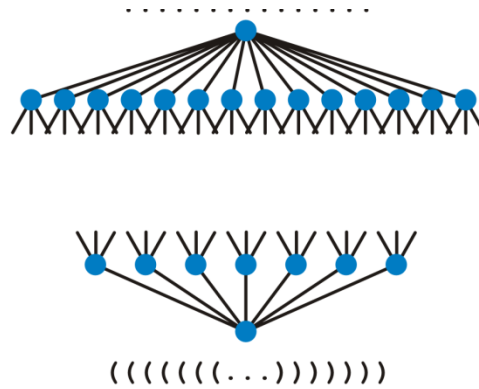
1. Prologue: Mathematics and biology
2. Modeling specific biological systems
- 3. Networks and evolution**
4. Perspectives



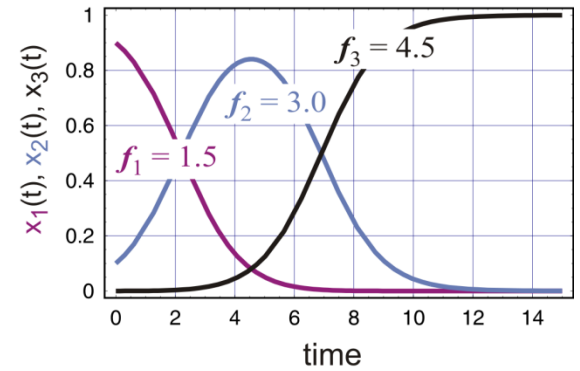
Sewall Wright's fitness landscape as metaphor for Darwinian evolution



sequence space



shape space



parameter space

$$\Phi: (\mathcal{Q}, d_H) \Rightarrow (\mathcal{Y}, d_Y)$$

$$\Psi: (\mathcal{Y}, d_Y) \Rightarrow \mathbf{R}^1$$

S



Y = Φ(S)



f = Ψ(Y)

sequence

structure

function

The paradigm of structural biology

RNA sequence

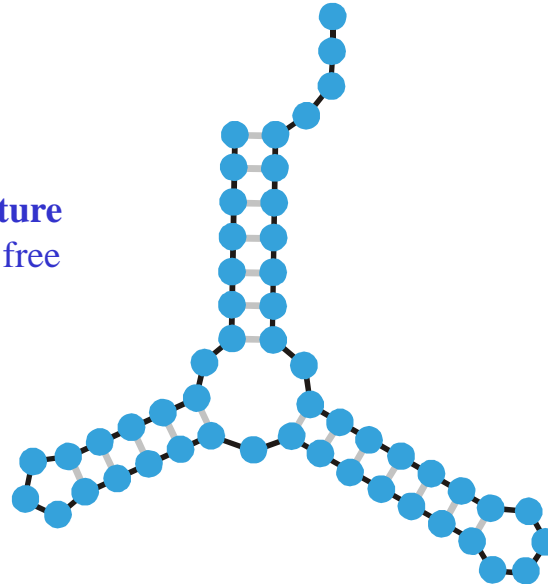
GUAUCGAAAUACGUAGCGUAUGGGGAUGCUGGACGGUCCCAUCGGUACUCCA

RNA folding:
Structural biology,
spectroscopy of
biomolecules,
understanding
molecular function

Biophysical chemistry:
thermodynamics and
kinetics

Empirical parameters

RNA structure
of minimal free
energy



Sequence, structure, and design

RNA sequence

GUAUCGAAAUACGUAGCGUAUGGGGAUGCUGGACGGUCCCAUCGGUACUCCA

RNA folding:
Structural biology,
spectroscopy of
biomolecules,
understanding
molecular function

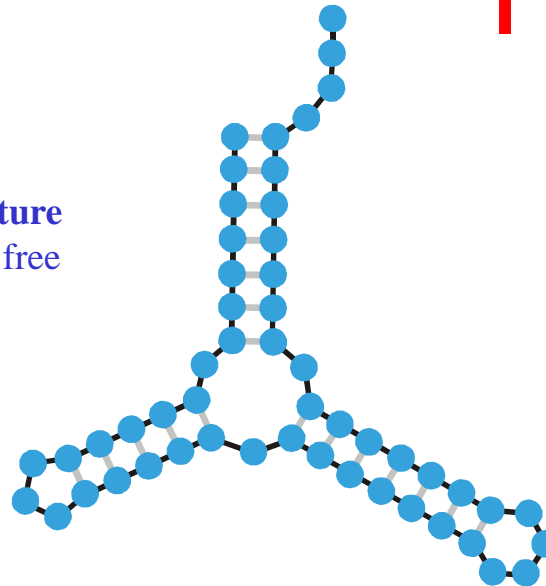
Iterative determination
of a sequence for the
given secondary
structure

**Inverse Folding
Algorithm**

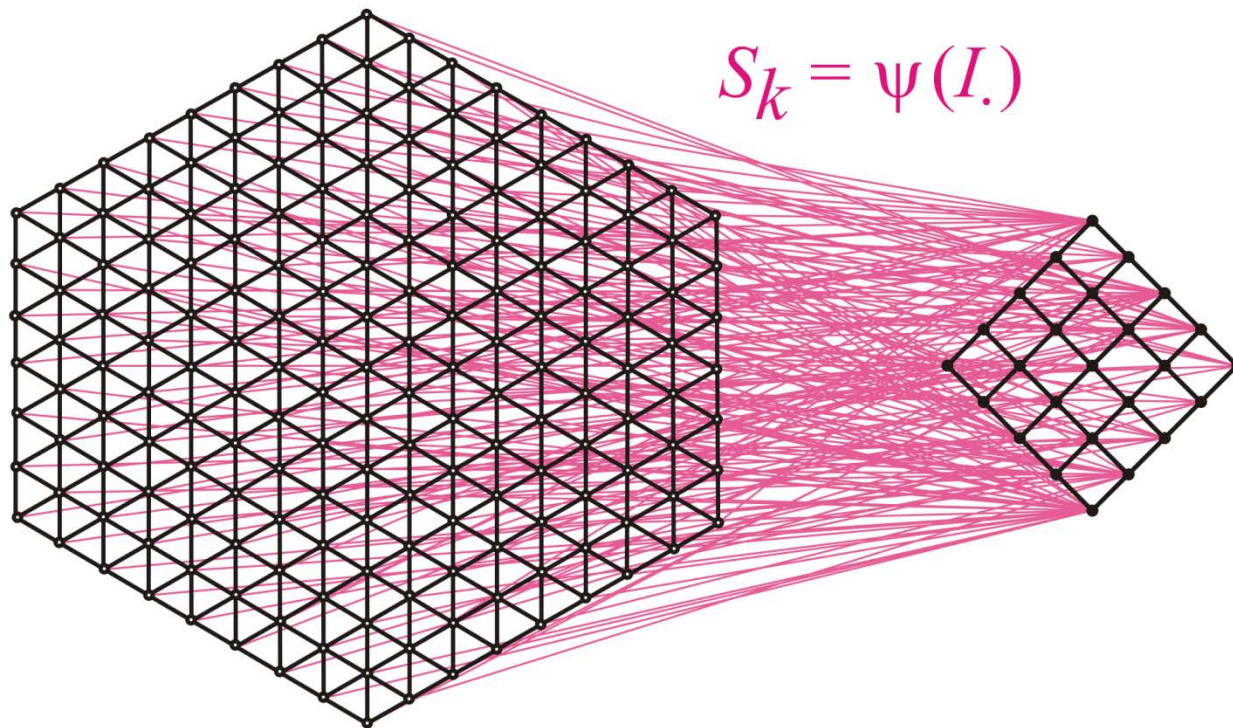
Inverse folding of RNA:

Biotechnology,
design of biomolecules
with predefined
structures and functions

RNA structure
of minimal free
energy

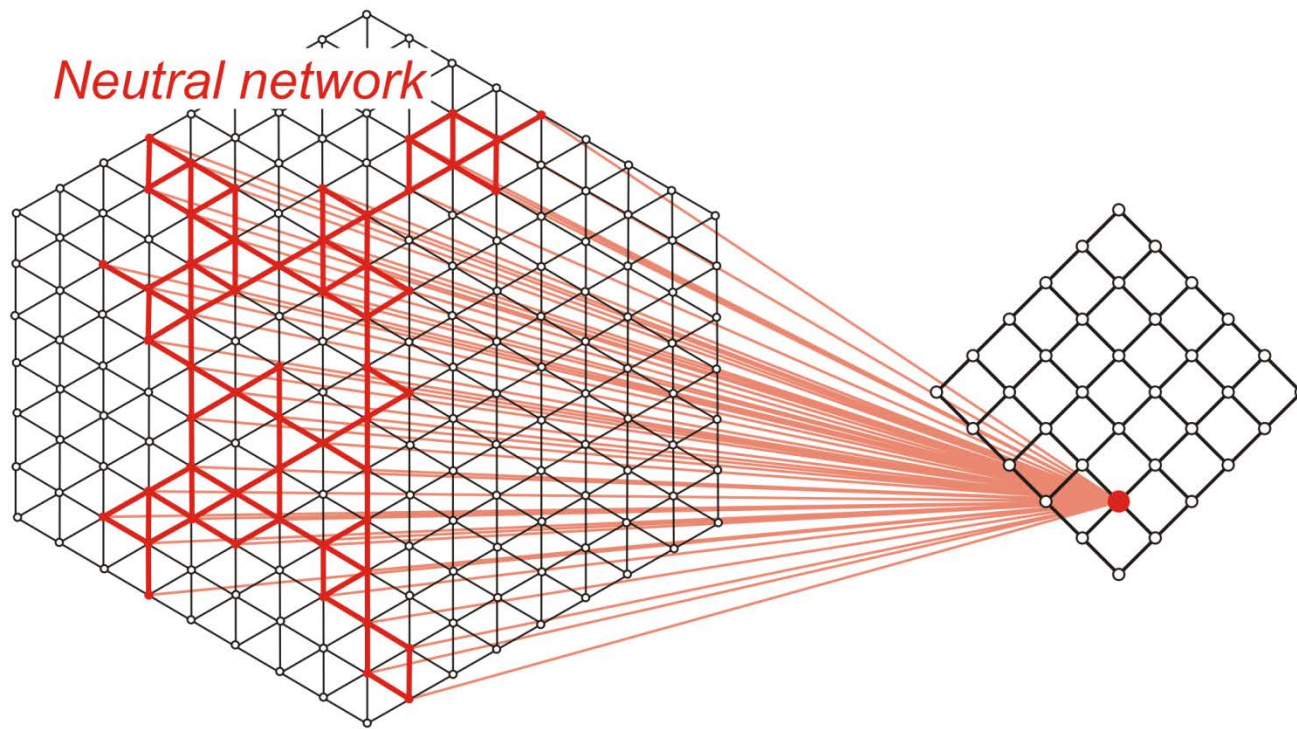


Sequence, structure, and design



Sequence space

Structure space



Neutral network

Sequence space

Structure space

Space of sequences: $I = \{I_1, I_2, I_3, I_4, \dots, I_N\}$; Hamming metric

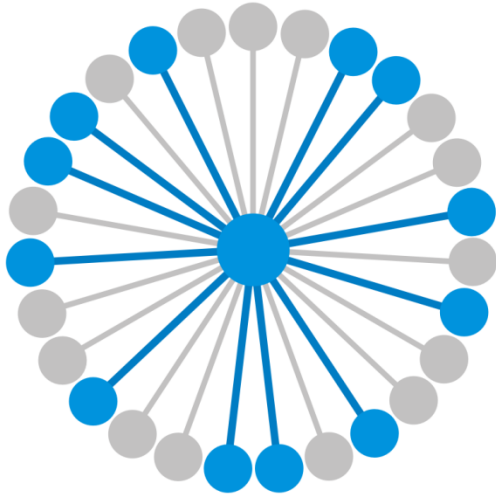
Space of structures: $S = \{S_1, S_2, S_3, S_4, \dots, S_M\}$; metric (not required)

$$N \gg M$$

$$\psi(I_j) = S_k$$

Neutral network: $G_k = \psi^{-1}(S_k) \equiv \{ I_j \mid \psi(I_j) = S_k \}$

A mapping ψ and its inversion



$$\lambda_j = 12 / 27 = 0.444$$

$$\mathbf{G}_k = \psi^{-1}(\mathbf{S}_k) \doteq \{ I_j \mid \psi(I_j) = \mathbf{S}_k \}$$

$$\bar{\lambda}_k = \frac{\sum_{j \in |\mathbf{G}_k|} \lambda_j(k)}{|\mathbf{G}_k|}$$

Alphabet size κ :

κ	λ_{cr}	
2	0.5	AU,GC,DU
3	0.423	AUG , UGC
4	0.370	AUGC

$\bar{\lambda}_k > \lambda_{cr}$ network \mathbf{G}_k is connected

$\bar{\lambda}_k < \lambda_{cr}$ network \mathbf{G}_k is **not** connected

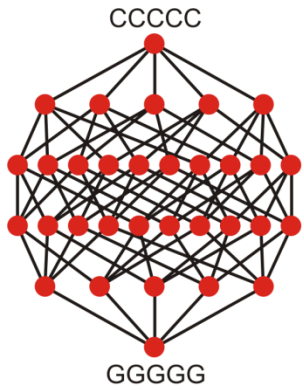
Connectivity threshold: $\lambda_{cr} = 1 - \kappa^{-1/(\kappa-1)}$

Degree of neutrality of neutral networks and the connectivity threshold

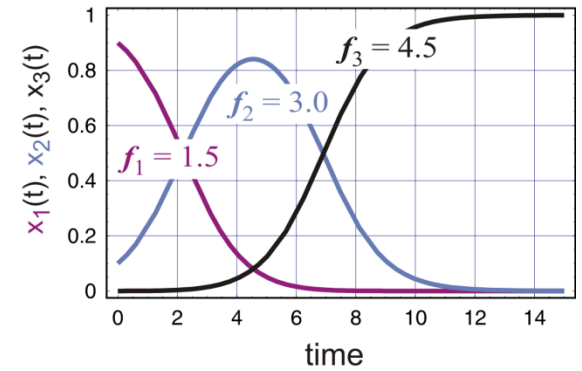
Conclusions concerning realistic fitness landscapes:

1. Adaptations take place in high-dimensional spaces,
2. fitness landscapes are rugged, and
3. fitness landscapes show a substantial degree of neutrality.

Roger D. Kouyos, Gabriel E. Leventhal, Trevor Hinkley, Mojgdan Haddad, Jeannette M. Whitcomb, Christos J. Petropoulos, and Sebastian Bonhoeffer. Exploring the complexity of the HIV-1 fitness landscape. *PLoS Genetics*, 2012, in press.



sequence space



parameter space

S
sequence



f = Ψ (**Y**)
function

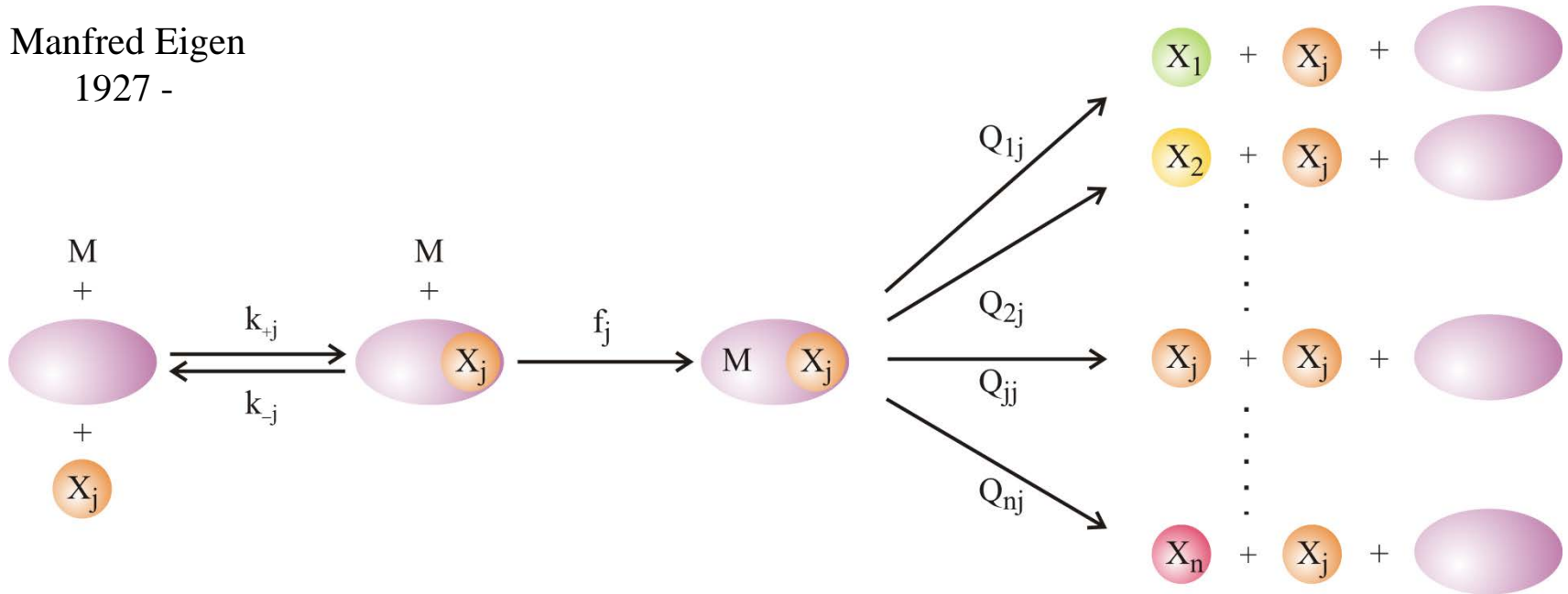
The simplified mapping from genotypes into function



$$\frac{dx_j}{dt} = \sum_{i=1}^n W_{ji} x_i - x_j \Phi ; j = 1, 2, \dots, n$$

$$\Phi = \sum_{i=1}^n f_i x_i / \sum_{i=1}^n x_i$$

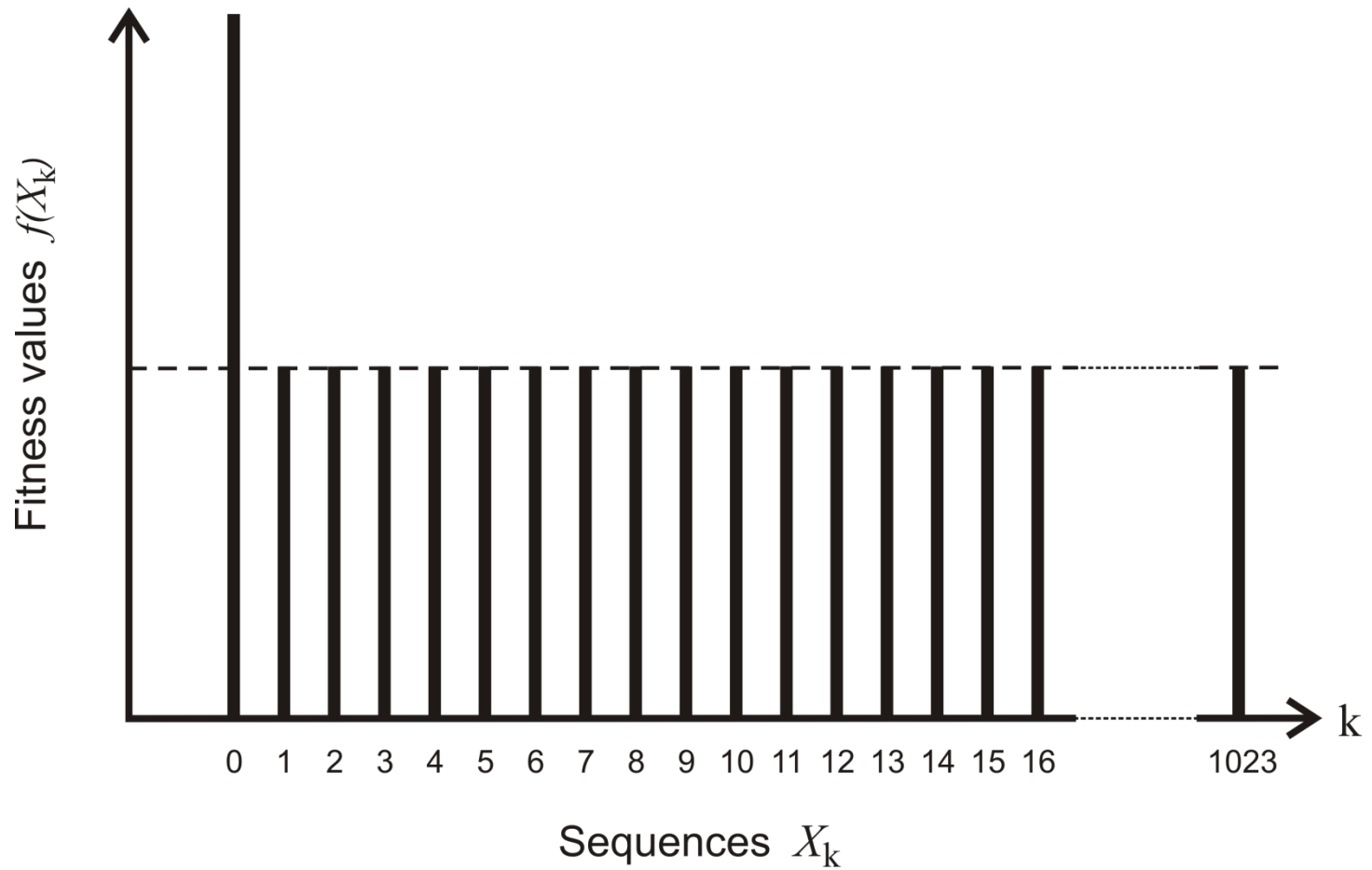
Manfred Eigen
1927 -



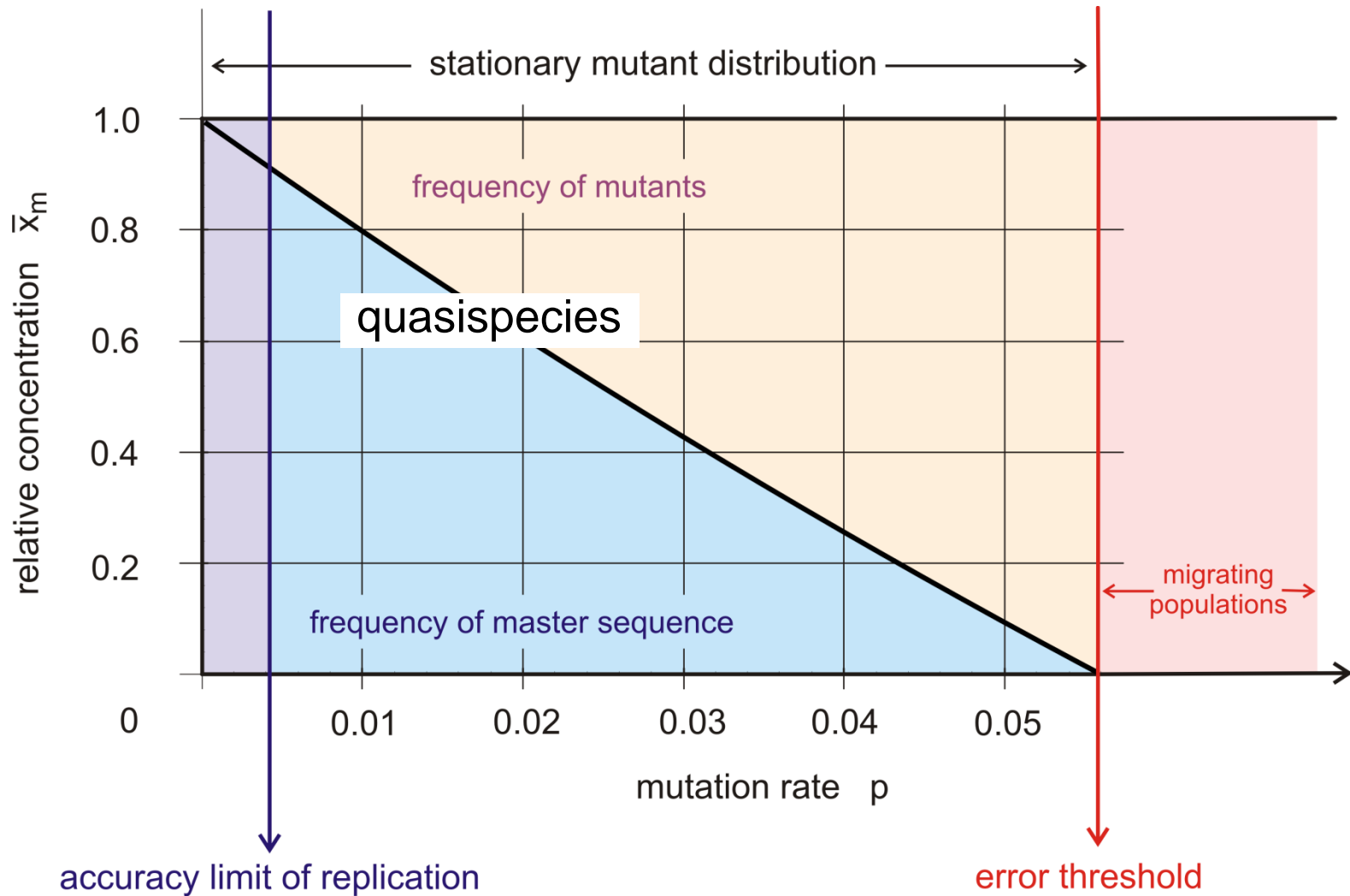
Mutation and (correct) replication as parallel chemical reactions

M. Eigen. 1971. *Naturwissenschaften* 58:465,

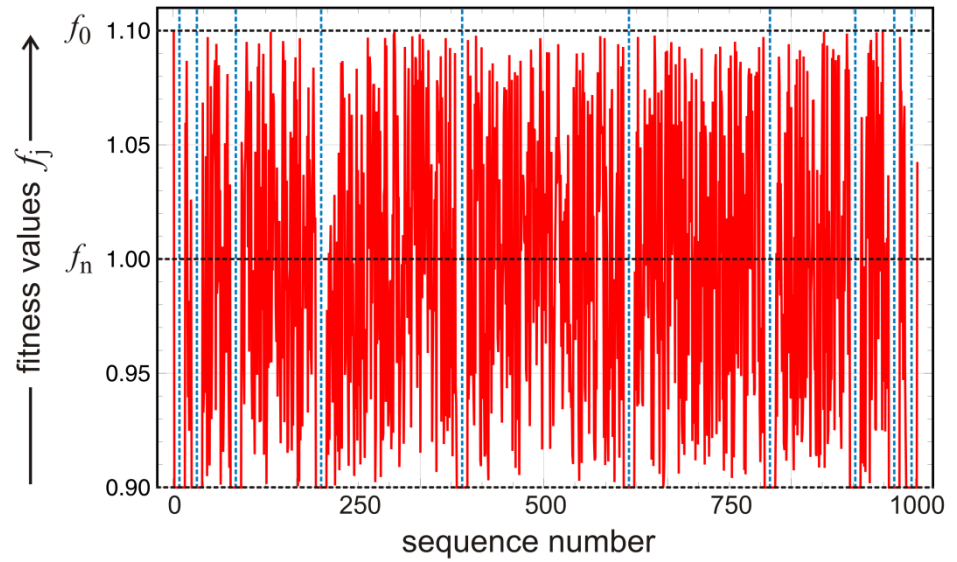
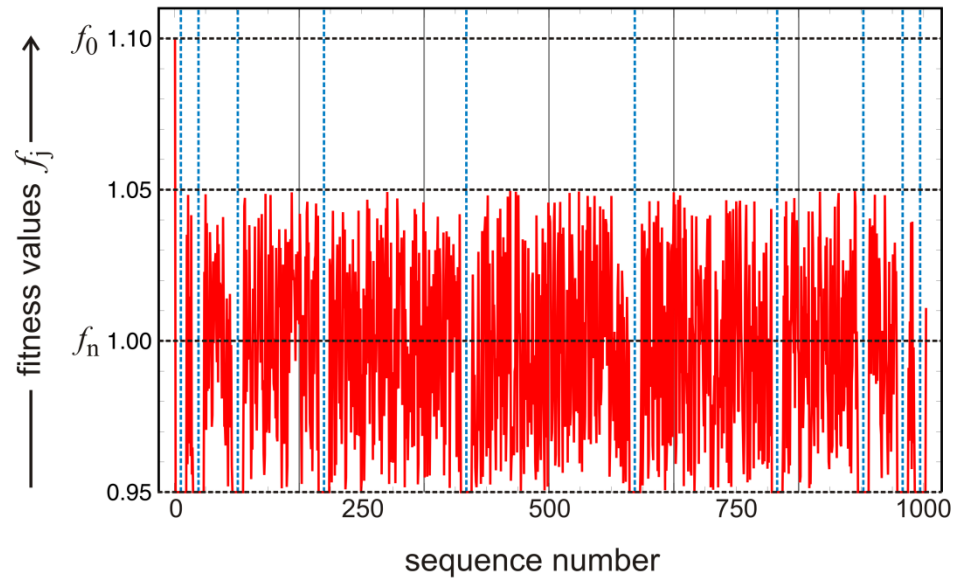
M. Eigen & P. Schuster. 1977. *Naturwissenschaften* 64:541, 65:7 und 65:341



The single peak model landscape for all sequences
with chain lengths $n = 10$

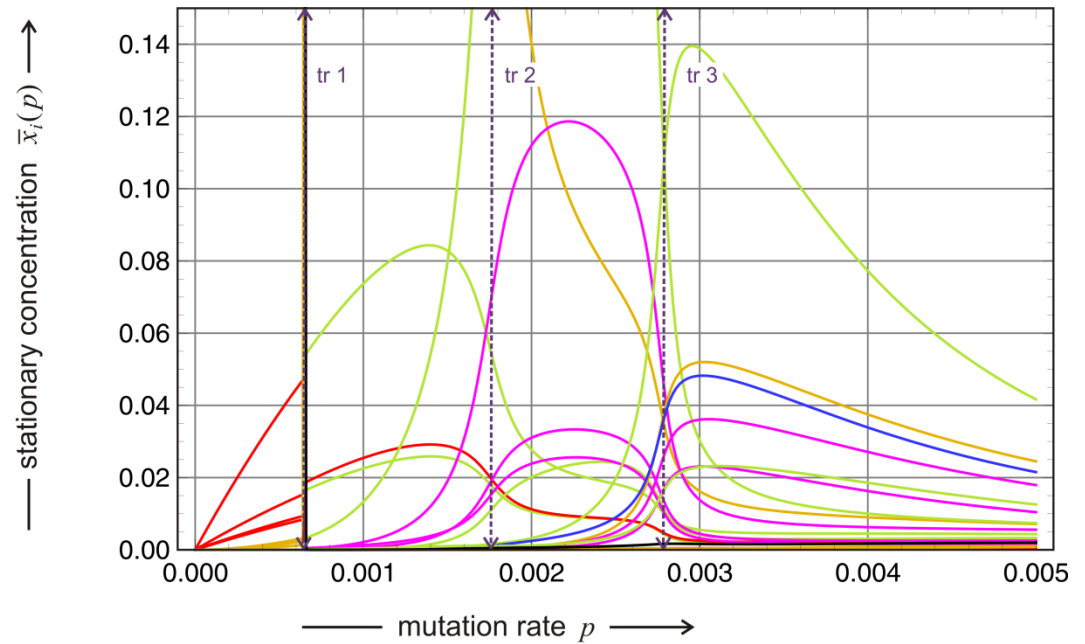
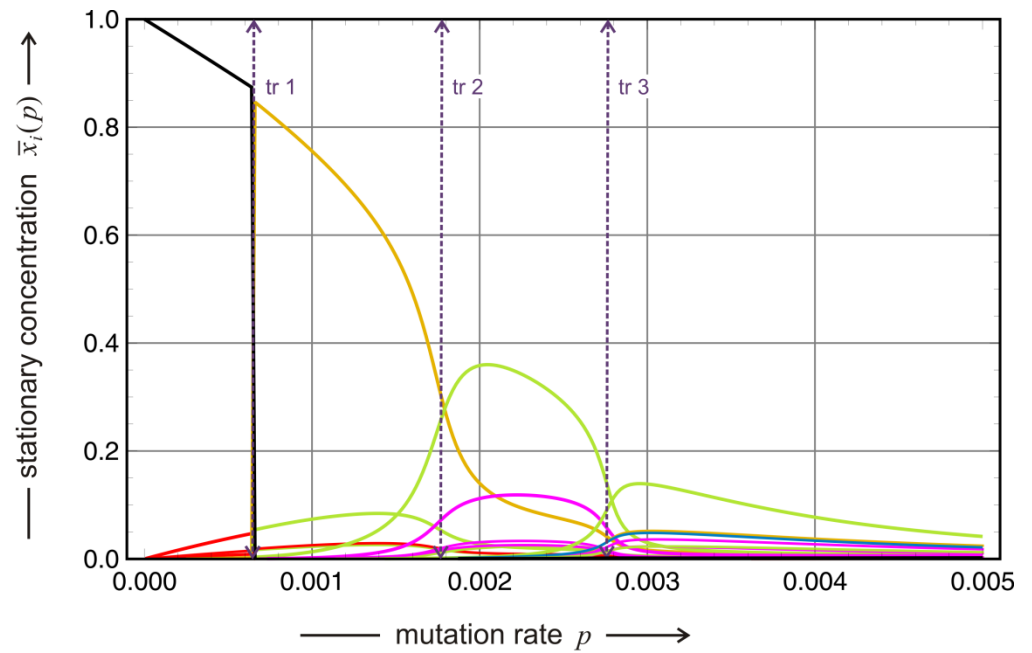


The error threshold in replication and mutation

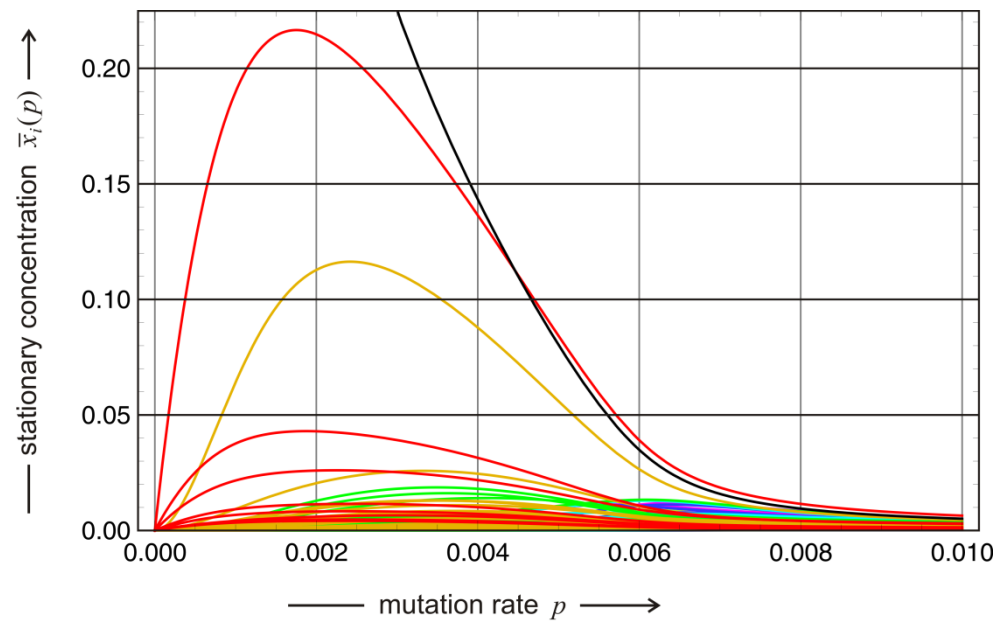
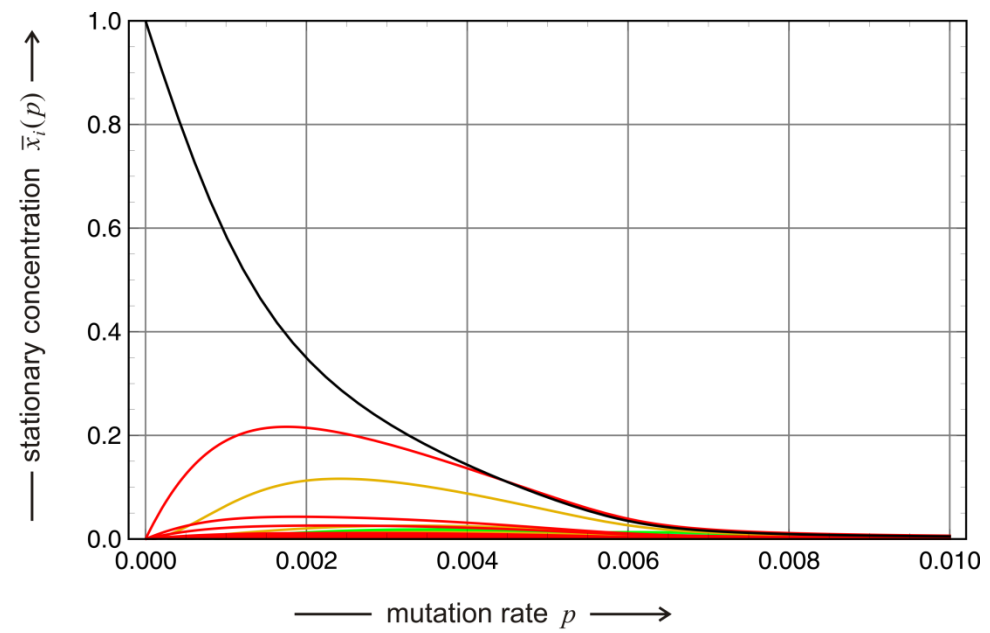


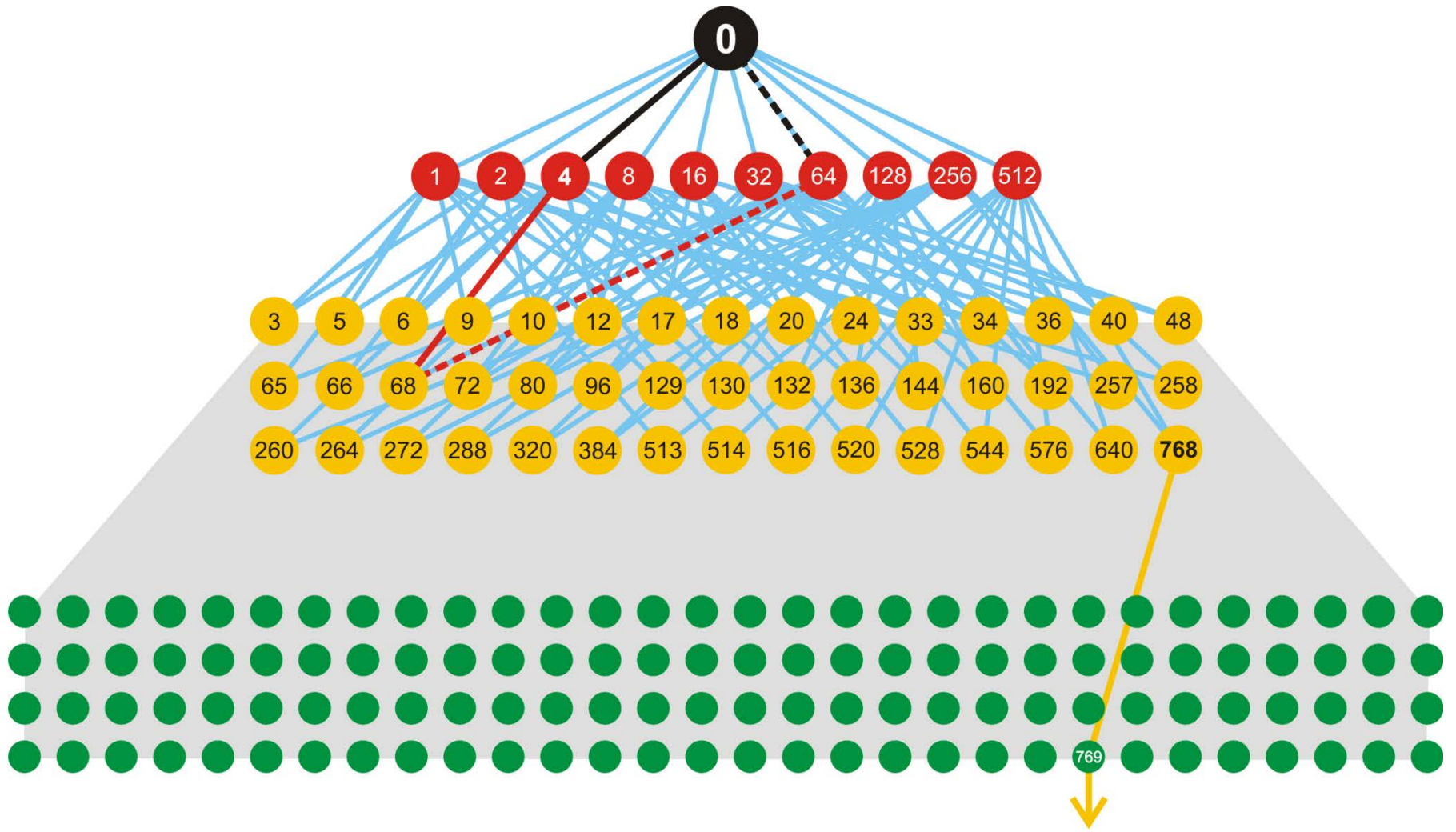
„Realistic“ fitness landscapes
with scattered fitness values

Quasispecies with phase transitions

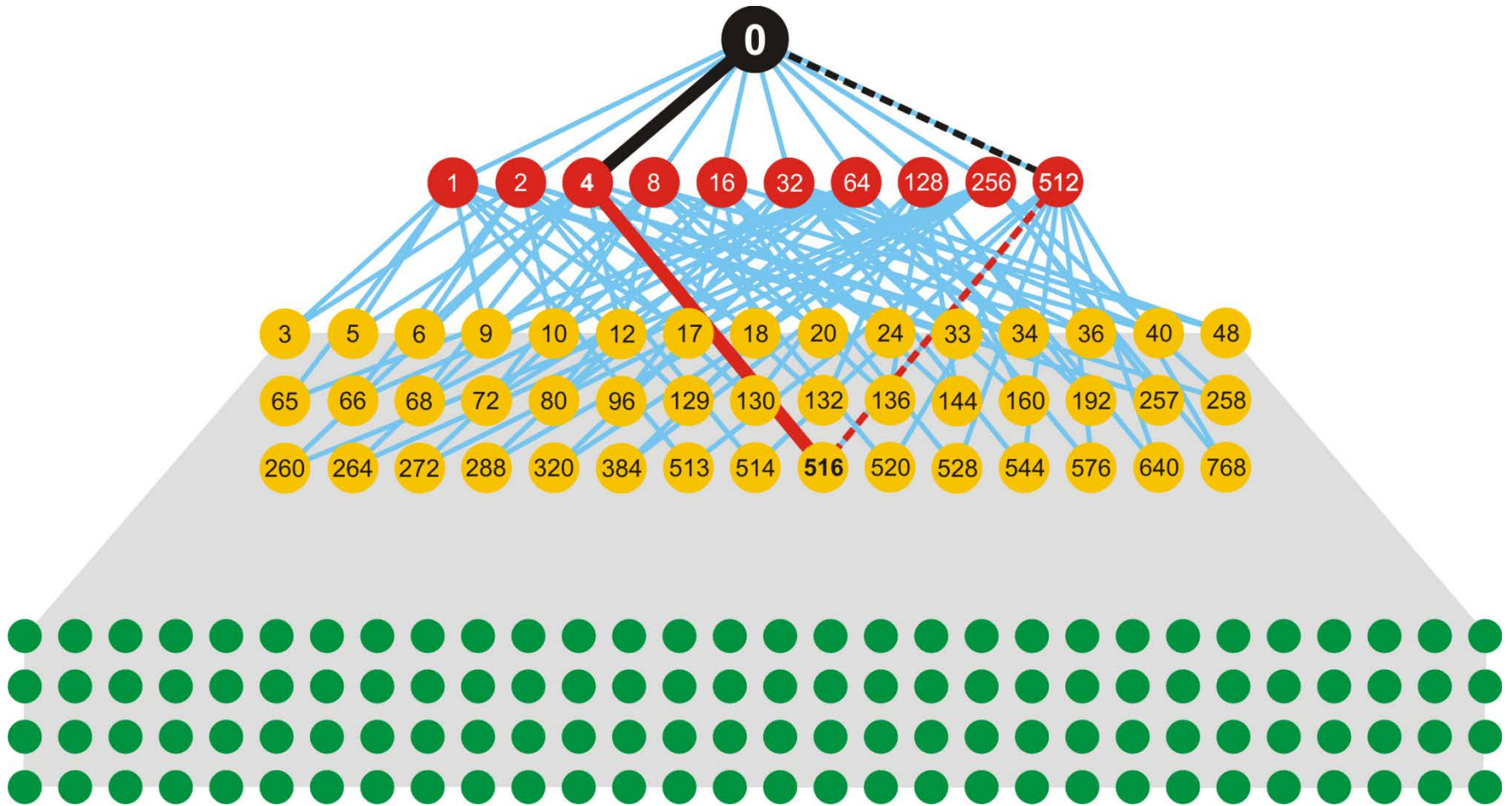


Strong quasispecies





Most probable fitness distribution in sequence space



Condition for the occurrence of a strong quasispecies



Motoo Kimuras population genetics of neutral evolution.

Evolutionary rate at the molecular level.
Nature **217**: 624-626, 1955.

The Neutral Theory of Molecular Evolution.
Cambridge University Press. Cambridge,
UK, 1983.

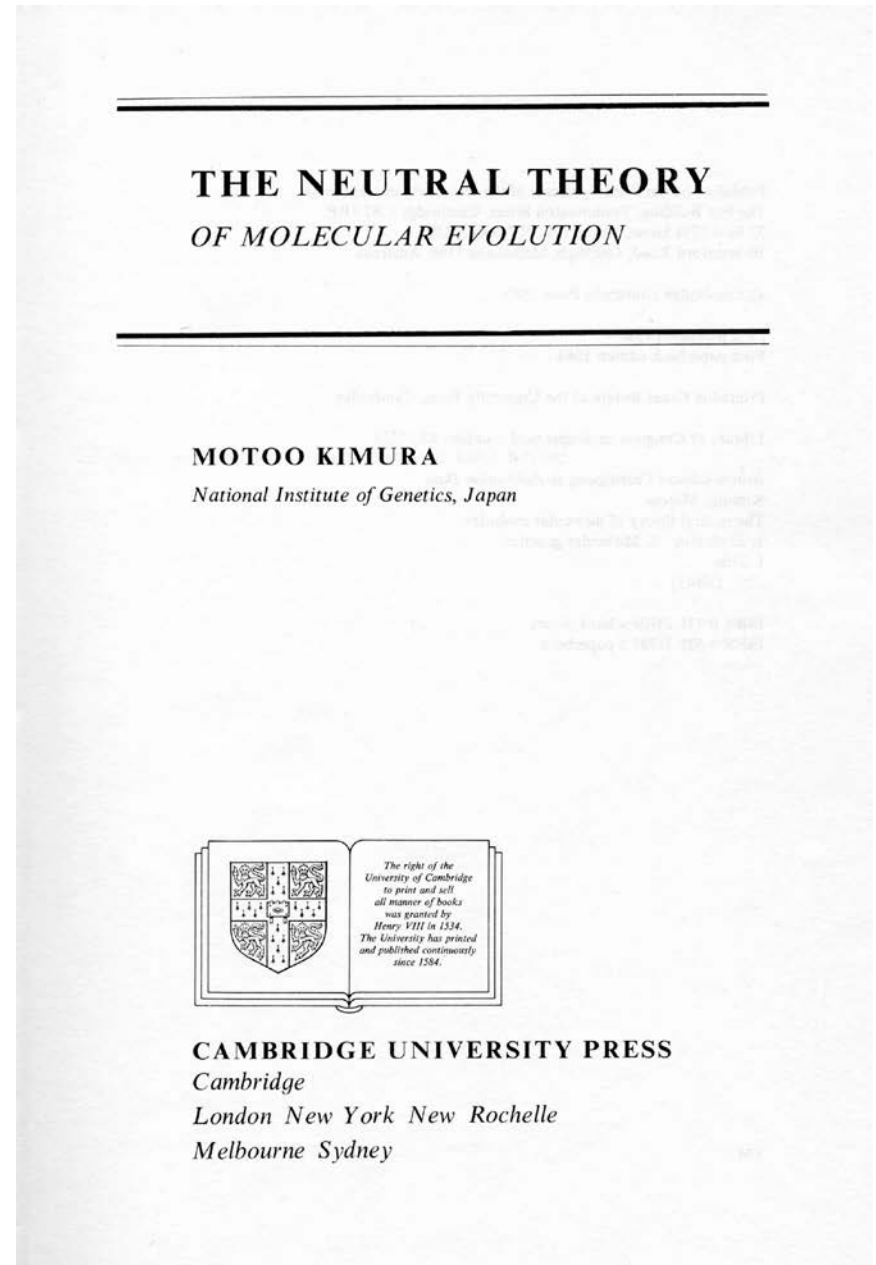
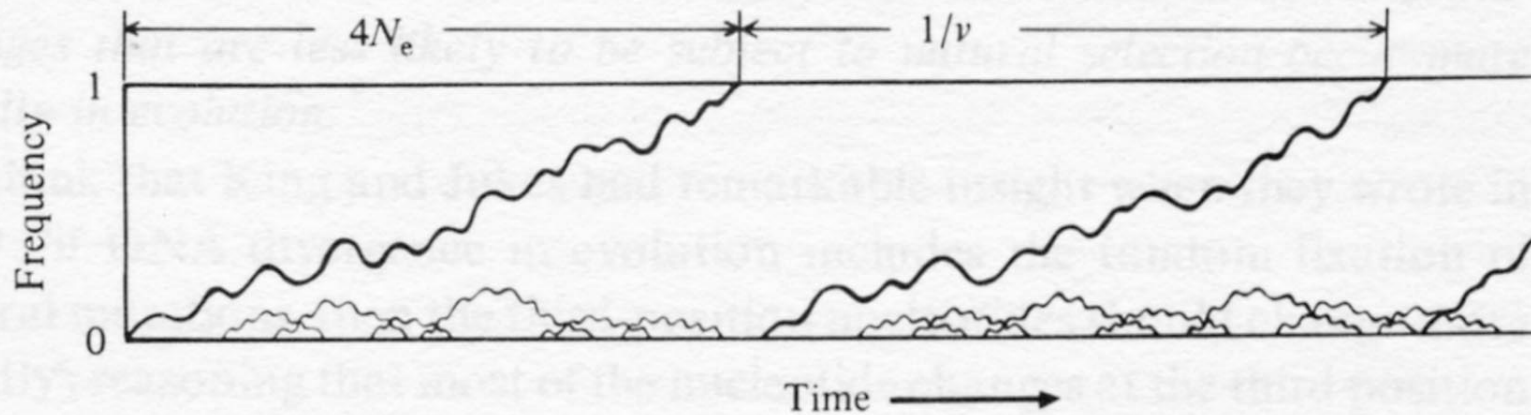


Fig. 3.1. Behavior of mutant genes following their appearance in a finite population. Courses of change in the frequencies of mutants destined to fixation are depicted by thick paths. N_e stands for the effective population size and v is the mutation rate.



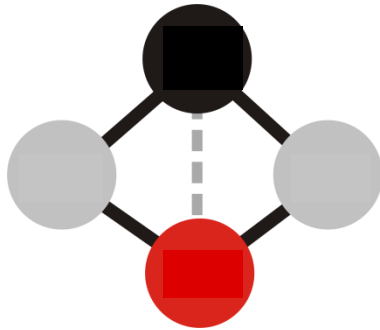
Motoo Kimura

Is the Kimura scenario correct for frequent mutations?



$$d_H = 1$$

$$\lim_{p \rightarrow 0} x_1(p) = x_2(p) = 0.5$$



$$d_H = 2$$

$$\lim_{p \rightarrow 0} x_1(p) = \alpha / (1 + \alpha)$$

$$\lim_{p \rightarrow 0} x_2(p) = 1 / (1 + \alpha)$$

$$d_H \geq 3$$

$$\lim_{p \rightarrow 0} x_1(p) = 1, \lim_{p \rightarrow 0} x_2(p) = 0 \text{ or}$$

$$\lim_{p \rightarrow 0} x_1(p) = 0, \lim_{p \rightarrow 0} x_2(p) = 1$$

Pairs of neutral sequences in replication networks

Random fixation in the sense of Motoo Kimura

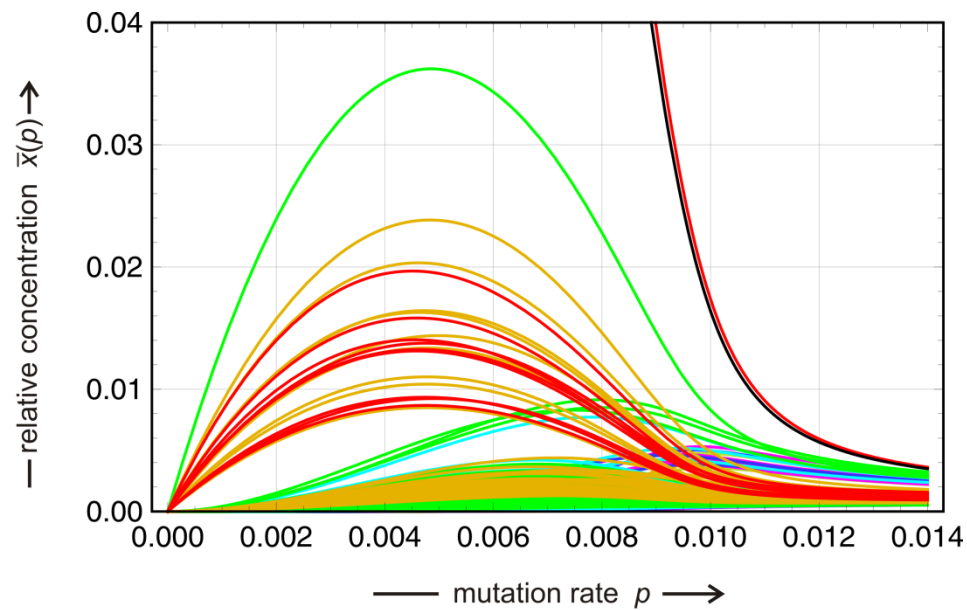
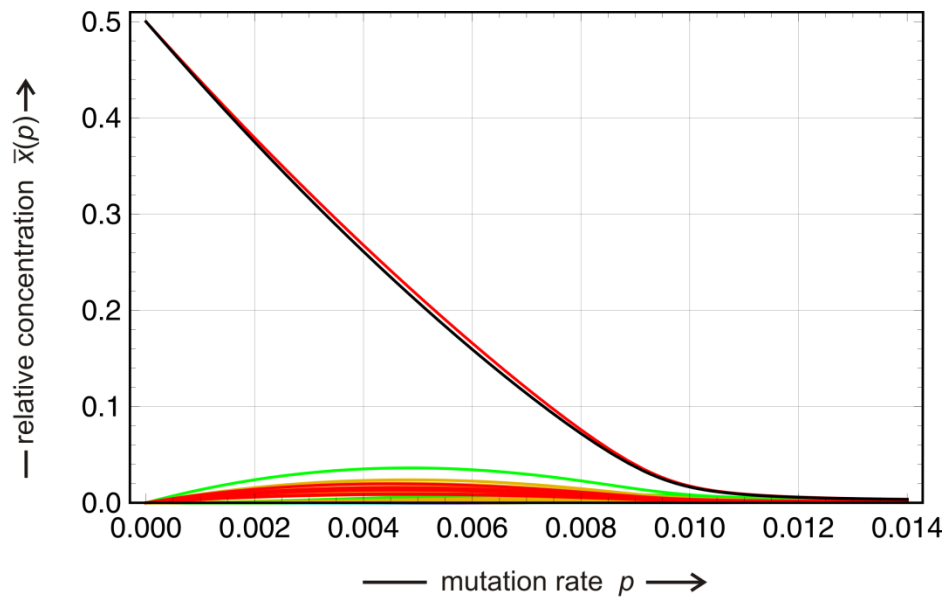


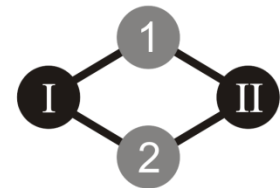
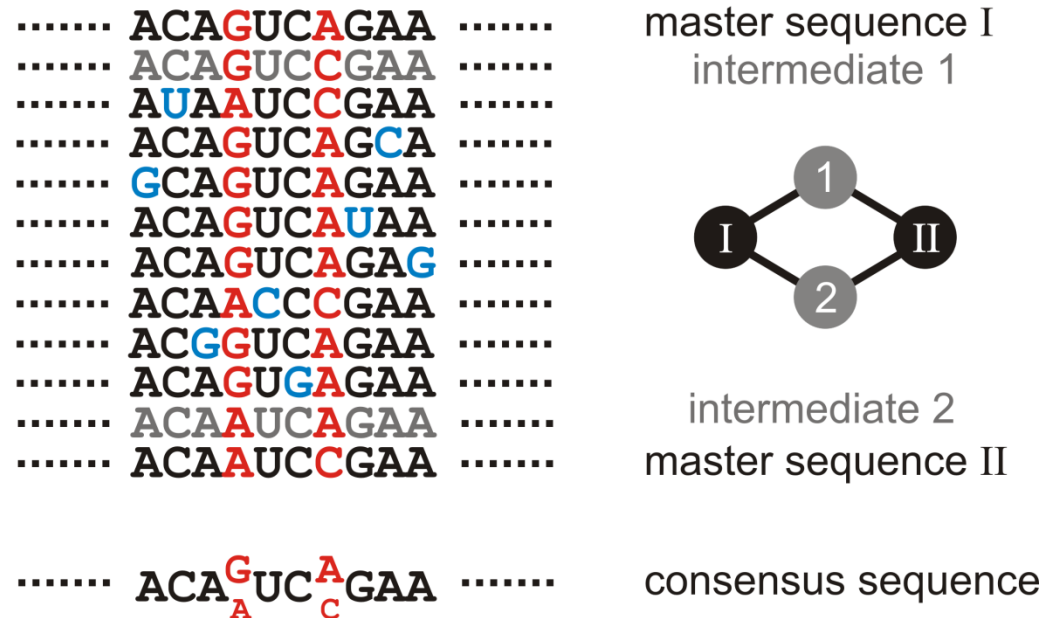
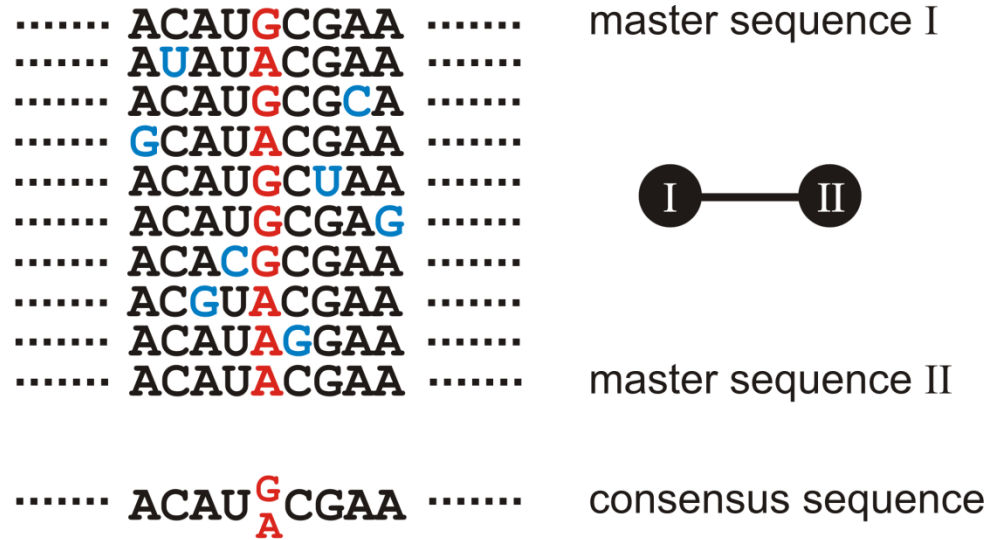
Neutral network

$\lambda = 0.01$, $s = 367$

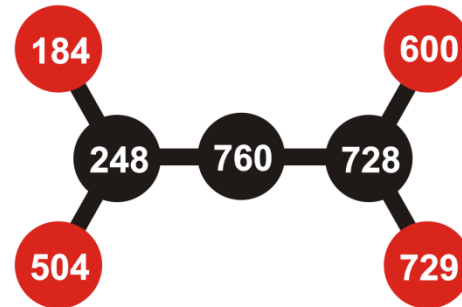
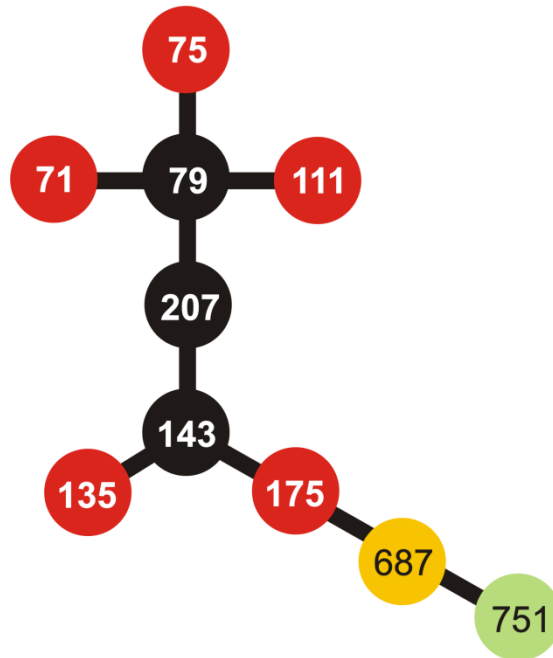
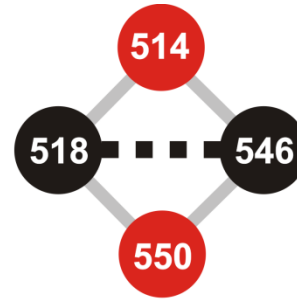
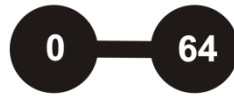
Neutral network: Individual sequences

$n = 10$, $\sigma = 1.1$, $d = 0.5$





Two neutral sequences with Hamming distances $d_H=1$ and $d_H=2$ can be detected in the consensus sequence of the population.



1. Prologue: Mathematics and biology
2. Modeling specific biological systems
3. Networks and evolution
- 4. Perspectives**



Manfred Eigen, 1927 -

..... theory cannot remove complexity,
but it shows what kind of „regular“
behavior can be expected and what
experiments have to be done to get a
grasp on the irregularities.

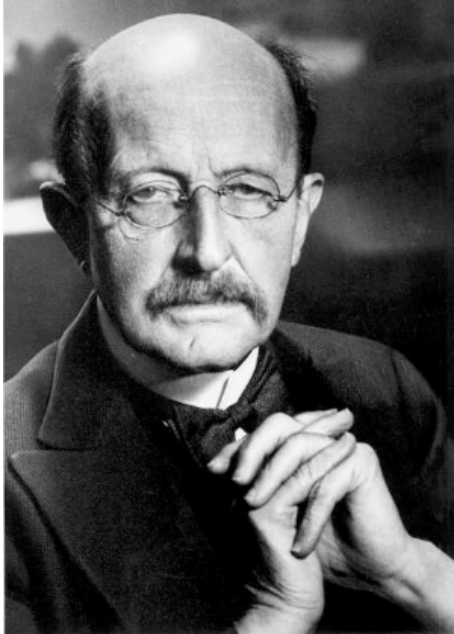
Manfred Eigen. Preface to E. Domingo, C.R. Parrish, J.J.Holland, eds.
Origin and Evolution of Viruses. Academic Press 2008

The molecular view has a clear advantage over phenomenology.

Analysis by methods of biochemical kinetics is indispensable for an understanding of cellular dynamics. Automation of kinetic analysis will come in the near future and this will change the situation completely.

It sounds commonplace but progress in theoretical biology will be very limited unless new approaches for handling networks and dealing with complexity will be developed.

Complexity in biology has one origin among others: Evolution does neither care for elegance nor for simplicity nor for intelligibility, the only thing that counts is efficiency.



Max Planck, 1859 - 1947

„Application without knowledge is dangerous“

„Anwendung ohne Wissen ist gefährlich.“

Coworkers

Peter Stadler, Bärbel M. Stadler, Universität Leipzig, GE

Paul E. Phillipson, University of Colorado at Boulder, CO

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

Jord Nagel, Kees Pleij, Universiteit Leiden, NL

Walter Fontana, Harvard Medical School, MA

Martin Nowak, Harvard University, MA

Christian Reidys, University of Southern Denmark, Odense, Denmark

Christian Forst, Los Alamos National Laboratory, NM

Thomas Wiehe, Ulrike Göbel, Walter Grüner, Stefan Kopp, Jaqueline Weber,
Institut für Molekulare Biotechnologie, Jena, GE

Ivo L.Hofacker, Christoph Flamm, Andreas Svrček-Seiler, Universität Wien, AT

**Kurt Grünberger, Michael Kospach, Andreas Wernitznig, Stefanie Widder,
Stefan Wuchty, Jan Cupal, Stefan Bernhart, Lukas Ender, Ulrike Langhammer,
Rainer Machne, Ulrike Mückstein, Erich Bornberg-Bauer**,
Universität Wien, AT



Universität Wien

Acknowledgement of support

Fonds zur Förderung der wissenschaftlichen Forschung (FWF)
Projects No. 09942, 10578, 11065, 13093
13887, and 14898

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

Jubiläumsfonds der Österreichischen Nationalbank
Project No. Nat-7813

European Commission: Contracts No. 98-0189, 12835 (NEST)

Austrian Genome Research Program – GEN-AU: Bioinformatics
Network (BIN)

Österreichische Akademie der Wissenschaften

Siemens AG, Austria

Universität Wien and the Santa Fe Institute



Universität Wien

Thank you for your attention!

Web-Page for further information:

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

