lit: Eigen/Engel/Fürstel: Physik der Evolutionsprozesse, Akademie Verlag, Berlin 190.
Eigen, M; 1971, Naturwiss. 58: 465

# 4 Competition and Selection in Biological Systems

## 4.1 Self reproducing systems

An essential prerequisite of all known life forms is the possibility of macromolecules to *self reproduce*, i. e., to produce exact copies of themselves. Self reproduction is not exclusively a biological phenomenon, it is also found in many other areas such as in chemical reactions (e. g. the number of protons in the saponification of esters) and physics (e. g. the occupation number of atomic states in lasers).

The process of self reproduction can be described by the elementary step

$$A + X \rightarrow 2X, \qquad \text{A abundant} \tag{1}$$

where $X$ describes the number of macromolecules (protons, occupied states, etc.) and $A$ some raw material needed for the reproduction.

If we assume that a certain type of macromolecule (a species $X$) reproduces with a rate $k$ and decays with a rate $k'$ to some final product $F$

$$A + X \xrightarrow{k} 2X, \quad X \xrightarrow{k'} F, \tag{2}$$

we can describe the temporal change of the concentration $n$ of that species by the differential equation

$$\frac{dn}{dt} = kn - k'n = (k - k')n. \tag{3}$$

Under the condition that the parameters $k$ and $k'$ remain constant, the solution of the differential equation (3) is simply given by

$$n(t) = n_0 e^{(k-k')t} \tag{4}$$

with $n_0$ denoting the initial concentration at time $t = 0$.

Usually there exist many different species $X_i$ of macromolecules, each being described by a reproduction constant $k_i$ and a degradation constant $k_i'$:

$$\begin{aligned} A + X_1 &\xrightarrow{k_1} 2X_1, & X_1 &\xrightarrow{k_1'} F, \\ A + X_2 &\xrightarrow{k_2} 2X_2, & X_2 &\xrightarrow{k_2'} F, \\ &\vdots & &\vdots \\ A + X_n &\xrightarrow{k_n} 2X_n, & X_n &\xrightarrow{k_n'} F. \end{aligned} \tag{5}$$

The differential equation systems reads

$$\frac{dn_i}{dt} = k_i n - k_i' n_i = (k_i - k_i')n_i. \tag{6}$$

In this simple case without interaction, the equations are uncoupled and the solutions can directly be taken from Eq. (4):

$$n_i(t) = n_{i,0} e^{(k_i - k_i')t} \tag{7}$$

with $n_{i,0}$ denoting the initial concentrations of species $X_i$ at time $t = 0$. These equations describe an exponential growth (for $k_i - k_i' > 0$) or an exponential decay (for $k_i - k_i' < 0$).

A + cX → (c+1)X
n(1) = 0
n(2) = c-1/k! stationär
n(2) n(1) instabil

## 4.2 Selection under constant total concentration

A selection process can only occur if there exists a *selection pressure* resulting in a competition between the different species. The simplest conditions to imply such a competition are given by

(a) a limited supply of raw material required by all species for growth

(b) a limited total number of macromolecules (or other particles, individuals, etc.).

We will consider the second case here as the mathematical description is simpler and analytical solutions can be found. A constant total concentration can be ensured by a permanent *dilution*, e. g. by diffusion processes into the neighboured space. Assuming that the dilution is proportional to the concentrations $n_i$, the dilution process changes the differential equation system (6) to

$$\frac{dn_i}{dt} = (k_i - k_i')n_i - \varphi \cdot n_i. \tag{8}$$

The value of $\varphi$ is determined by the condition that the total concentration remains constant, i. e.

$$0 = \sum_i \frac{dn_i}{dt} = \sum_i (k_i - k_i')n_i - \varphi \sum_i n_i. \tag{9}$$

This yields

$$\varphi = \frac{\sum_i (k_i - k_i')n_i}{\sum_i n_i} = \frac{\sum_i E_i n_i}{\sum_i n_i} = \bar{E}, \tag{10}$$

where $E_i = k_i - k_i'$ denotes the excess productivity of species $X_i$ and $\varphi = \bar{E}$ can be interpreted as the mean excess productivity. Inserting Eq. (10) into Eq. (8) yields the *selection equations*

$$\frac{dn_i}{dt} = \left( k_i - k_i' - \frac{\sum_j (k_j - k_j')n_j}{\sum_j n_j} \right) n_i = (E_i - \bar{E})n_i. \tag{11}$$

Since $\bar{E} = \bar{E}(n_i(t))$ is a function of the concentrations $n_i$ this is a system of *coupled, non-linear* differential equations.

### 4.2.1 Qualitative discussion

The selection equations Eq. (11) determine the temporal development of the concentrations $n_i = n_i(t)$:

(a) For all species $X_i$ with $E_i > \bar{E}$ the relation $dn_i/dt > 0$ holds: The concentration increases.

(b) For all species $X_i$ with $E_i < \bar{E}$ the relation $dn_i/dt < 0$ holds: The concentration decreases.

The quantity $\bar{E}$ can therefore be understood as a *selection threshold* and $E_i$ can be interpreted as the *selective value* of species $X_i$. As a consequence of these considerations, those species with a higher selective value will contribute more to the mean value given by Eq. (10) than those species with a lower selective value. Therefore, $\bar{E}$ will increase in time ($d\bar{E}/dt > 0$). This means that the selective values of more and more species will drop below the selection threshold leading to the extinction of the corresponding species. As a result of the selection process only one species, the so called *Master Species*, remains.

### 4.2.2 Explicit solution of the selection equations

Inserting the ansatz

$$n_i(t) = z_i(t)e^{-\int_0^t \bar{E}(\tau)d\tau} \qquad \text{Koordinaten Transf.} \tag{12}$$

into the equation system Eq. (11) leads to

$$\frac{dz_i}{dt}e^{-\int_0^t \bar{E}(\tau)d\tau} - z_i\bar{E}e^{-\int_0^t \bar{E}(\tau)d\tau} = E_i z_i e^{-\int_0^t \bar{E}(\tau)d\tau} - \bar{E}z_i e^{-\int_0^t \bar{E}(\tau)d\tau} \tag{13}$$

yielding

$$\frac{dz_i}{dt} = E_i z_i, \tag{14}$$

an uncoupled system of differential equations.
Because of Eq. (12) we have $z_i(0) = n_i(0) = n_{i,0}$ and the solution of Eq. (14) reads

$$z_i(t) = n_{i,0}e^{E_i t}. \tag{15}$$

From the condition that the overall concentration remains constant,

$$\sum_i n_i(t) = e^{-\int_0^t \bar{E}(\tau)d\tau}\sum_i z_i(t) = C = const., \tag{16}$$

it follows that

$$e^{-\int_0^t \bar{E}(\tau)d\tau} = \frac{C}{\sum_i z_i(t)}. \tag{17}$$

The solution of the system of differential equations (11) therefore reads

$$n_i(t) = \frac{C \cdot z_i(t)}{\sum_j z_j(t)} = \frac{C \cdot n_{i,0}e^{E_i t}}{\sum_j n_{j,0}e^{E_j t}}. \tag{18}$$

An example of the temporal behaviour for three competing species is given in Fig. 1. In this example, the parameters have been chosen such that $E_1 > E_2 > E_3$ and the initial concentrations have been set to $n_{1,0} = n_{2,0} = n_{3,0} = C/3$. Species $X_1$ is the master species. For the species $X_3$ the relation $E_3 < \bar{E}$ holds from the beginning of the process, i. e., its concentration decreases monotonously. For species $X_2$ the selective value is initially larger than the selection threshold $E_2 > \bar{E}$ but drops below this threshold as $\bar{E}$ increases. This results in an initial increase of the concentration but nevertheless in an extinction in the long run. The selection threshold $\bar{E}$ as a function of the time is depicted in Fig. 2.

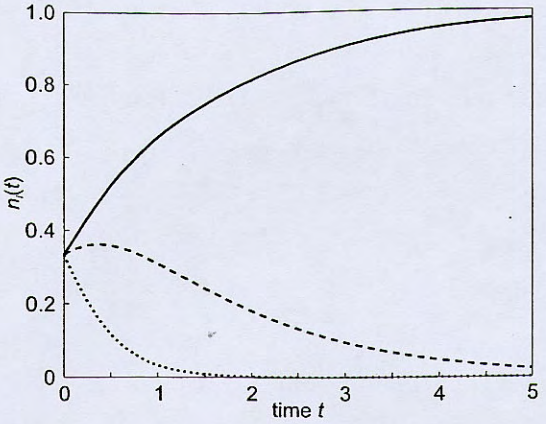


Figure 1: Temporal development of the concentrations $n_1$, $n_2$, and $n_3$ for the parameter values $E_1 = 3.5$, $E_2 = 2.75$, and $E_3 = 0.5$.

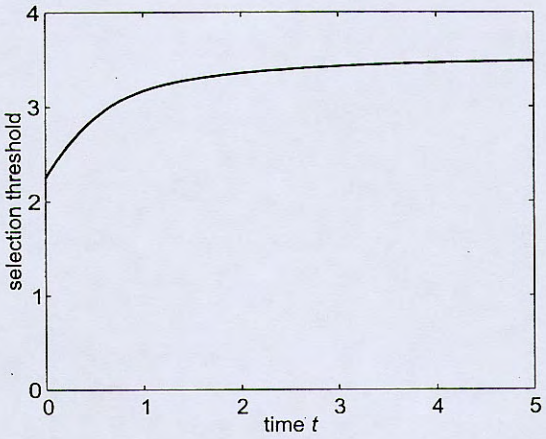


Figure 2: Temporal development of the selection threshold $\bar{E}$. Parameter values are the same as in Fig. 1.

## 4.3 Selection equations under consideration of mutations

The selection equations can be extended to include non-identical (erroneous) replications, i. e. *mutations*. Such processes can be described by the following system of differential equations

$$\frac{dn_i}{dt} = \underbrace{F_i n_i}_{I} - \underbrace{k'_i n_i}_{II} + \underbrace{\sum_j m_{ij} n_j}_{III} - \underbrace{\varphi n_i}_{IV}, \tag{19}$$

where we assume that $m_{ii} = 0$.

Term I describes the *identical self replication*. The corresponding rate constant $F_i$ we write as

$$F_i = k_i Q_i, \tag{20}$$

where $k_i$ denotes the rate constant of all replications and $Q_i$ represents the fraction of error-free replications, i. e., it can be understood as a measure of the *quality* of the reproduction. Term II characterises the *decay rate* of species $X_i$. Term III denotes the rate of the generation of species $X_i$ by mutation of species $X_j$. Thus, the $m_{ij}$ represent the rate constants of the *mutation* processes. Term IV characterises a *dilution* process in analogy to Eq. (6). The quantity $\varphi$ is again determined by the condition that the total concentration remains constant, i. e.

$$0 = \sum_i \frac{dn_i}{dt} = \sum_i k_i Q_i n_i - \sum_i k'_i n_i + \sum_{ij} m_{ij} n_j - \varphi \sum_i n_i. \tag{21}$$

By definition of the quality measure $Q_i$ the following relation must hold:

$$\underbrace{\sum_i k_i (1 - Q_i) n_i}_{\text{total rate of erroneous replications}} = \underbrace{\sum_{ij} m_{ij} n_j}_{\text{total mutation rate}} \tag{22}$$

Using Eq. (22) and considering that

$$\sum_i k_i Q_i n_i = \sum_i k_i n_i - \sum_i k_i (1 - Q_i) n_i \tag{23}$$

transforms Eq. (21) into

$$0 = \sum_i k_i n_i - \sum_i k'_i n_i - \varphi \sum_i n_i \tag{24}$$

yielding for $\varphi$ the same expression as Eq. (10).

The selection equations considering mutation processes can therefore be written as

$$\frac{dn_i}{dt} = \left(k_i Q_i - k'_i - \bar{E}\right) n_i + \sum_j m_{ij} n_j = \left(W_i - \bar{E}\right) n_i + \sum_j m_{ij} n_j, \tag{25}$$

where we defined the selective value $W_i = k_i Q_i - k'_i$.

This equation differs from Eq. (6) by an additional coupling term describing the mutations. Assuming that replications are performed with a high accuracy results in small coupling constants $m_{ij}$. For $m_{ij} \ll 1$ it can be assumed that during the selective process initially the concentrations of those species will increase for which $W_i > \bar{E}$. However, not a single species will be selected because of continuously occurring mutations. These mutations continue to occur even after the *selection equilibrium* characterised by $dn_i/dt = 0$ has been reached. In the special case that the selective value of one species $X_m$ is much larger than the selective values of all other species ($W_m \gg W_j$ for $j \neq m$) the selection process will result in a high concentration of the master species $X_m$. Additionally, however, those species will maintain a positive concentration which are closely related to the master species, i. e., which can be generated from the master species by a small number of mutations. The distribution of the concentrations in the selection equilibrium is called

the *master quasi species distribution.* One can think of this distribution as of a "cloud" within the *space of molecules*, which surround the molecule with the highest selection value. Ansatz (12) is also appropriate for the explicit solution of Eq. (25). One again gets a system of linear differential equations for the new variables $z_i(t)$, which in this case are not uncoupled. The further treatment of this differential equation system requires the calculation of eigenvectors and eigenvalues of a matrix B with the components

$$B_{ij} = W_i\delta_{ij} + m_{ij}. \quad (26)$$

The eigenvector with the largest eigenvalue represents the quasi species, which establishes during the selection process. The number of possible mutants of a given sequence depends on the sequence lengths $N$. Taking into account that RNA and DNA sequences consist of four different kinds of monomers, then for every sequence there exist $3N$ single-error mutants. For mutants with $f$ errors there are

$$N_f = 3^f\binom{N}{f} \quad (27)$$

possibilities. Therefore the number of possible sequences, which belong to a certain quasi species, is normally very high. The quality factor $Q$ of a polymer sequence of length $N$ in Eq. (20) can be expressed as probabilities $q_v(v = 1, \ldots, N)$ of error-free replication of the single monomers. Assuming these probabilities to be independent of the position and kind of replicated monomer we get

$$Q = q^N. \quad (28)$$

A more detailed mathematical analysis demonstrates that the information content within a polymer sequence can only be passed on in a stable manner if the quality factor lies above a certain threshold ($1 > Q > Q_{min}$). Though mutations are necessary to generate novel structures allowing for a progression towards higher organisms during evolution, too high mutation rates destroy the information content that was already accumulated. The closer $Q$ (respectively $q$) is to 1 the longer chains can be replicated in a stable manner. For the maximum length of a polymer sequence the following equation holds approximately

$$N_{max} \approx \frac{1}{1-q}. \quad (29)$$

During evolution the self replication systems increased in complexity and as a result the DNA sequences, which were needed to code for these systems, had to become long. A decisive prerequisite of this evolution was an increase in the *accuracy* of the replication process. Since the number of nucleotides in the DNA of an animal cell is about $N = 3\times10^9$, according to Eq. (29) one expects the error rate per nucleotide, $\tilde{m} = 1-q$, to be about $10^{-9}$. This number is in very good accordance with the latest results concerning the accuracy of the replication machinery in animal cells. This accuracy is directly related to the existence of polymerases catalysing such reactions. Estimations yield that for uncatalysed reactions of RNA molecules the error rate can not be less than $\tilde{m} = 10^{-2}$, which gives a maximum sequence length of $N \cong 100$.

## 4.4 Cooperation and coexistence of self replicating macromolecules

The simple self-replication $I_i \rightarrow 2 \cdot I_i$ is an idealised case. As we have seen, the analysis of this case only leads to a competition between the polymer sequences after introducing a selection pressure (i.e. constant organisation). As we showed in section 4.3, consideration of mutations can lead to a certain *coexistence* of different kinds of macromolecules within a quasi-species. However, this concept is too simple to explain the complex interactions of different macromolecules found in cells within the scope of an evolutionary model. If mutation and selection were restricted to processes described in section 4.2 and 4.3, the evolution would have died down in a very early stage. Thus, the development to more and more complex structures can only be understood if there also exists a *cooperation* between the kinds of molecules beside a *competition.*

### 4.4.1 Complementary self replication of RNA molecules

We consider, as a simple example, the replication of RNA molecules as it occurs during the proliferation of RNA phages. The genetic code of these phages is not located on a DNA but on a single stranded RNA. In this case the reproduction process is also done with complementary base pairing (A $\equiv$ U, C $\equiv$ G). Therefore we can distinguish for each kind of molecule between two complementary states $R_i^+$ and $R_i^-$. A *plus strand* and a *minus strand* is synthesised, where according to the complementary base pairing the following exchanges occur: A$\rightarrow$U, U$\rightarrow$A, G$\rightarrow$C, and C$\rightarrow$G. The formation of one kind of sequences depends on the existence of a second sequence class, i.e., both do not compete with each other.

If we restrict the analysis to one single pair of sequences, we get the reaction scheme depicted in Fig. 3.

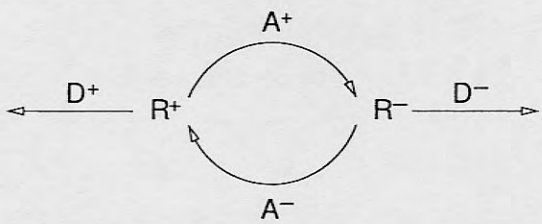


Figure 3: Reaction scheme for one pair of sequences.

Here, $A^-$ and $A^+$ are the kinetic rate constants of the replication and $D^-$ and $D^+$ are the kinetic rate constants of degradation processes. The associated differential equations are

$$\frac{dn^+}{dt} = A^- n^- - D^+ n^+ \tag{30a}$$

$$\frac{dn^-}{dt} = A^+ n^+ - D^- n^- \tag{30b}$$

or in matrix notation

$$\frac{d}{dt}\begin{pmatrix} n^+ \\ n^- \end{pmatrix} = \begin{pmatrix} -D^+ & A^- \\ A^+ & -D^- \end{pmatrix}\begin{pmatrix} n^+ \\ n^- \end{pmatrix}. \tag{31}$$

The solution is

$$n(t) = b^{(1)} e^{\lambda_1 t} + b^{(2)} e^{\lambda_2 t} \tag{32}$$

with the vector of concentrations $n = (n^+, n^-)^T$. $b^{(1)}$, $b^{(2)}$ and $\lambda_1$, $\lambda_2$ are the eigenvectors and the eigenvalues of the matrix in Eq. (31). From

$$Det\begin{pmatrix} -D^+ - \lambda & A^- \\ A^+ & -D^- - \lambda \end{pmatrix} = 0 \tag{33}$$

one gets the eigenvalues

$$\lambda_{1/2} = -\frac{D^+ + D^-}{2} \pm \sqrt{\left(\frac{D^+ - D^-}{2}\right)^2 + A^+ A^-}, \tag{34}$$

which obviously are real numbers.
Whereas $\lambda_2$ (negative root) is always negative, the following holds

$$\lambda_1 < 0 \quad \text{for} \quad A^+ A^- < D^+ D^- \tag{35a}$$

$$\lambda_1 > 0 \quad \text{for} \quad A^+ A^- > D^+ D^-. \tag{35b}$$

According to Eq. (32), case (35a) leads to the extinction of both species, $R^+$ and $R^-$. In case (35b), the right side of the solution (32) contains an increasing and a decreasing term. For longer time scales the first term becomes dominant resulting in an exponential increase of the concentrations of both kinds of molecules. The mutual dependencies of both species are given by the fact that the growth behaviour of one species does not only depend on its own kinetic rate constants, but also on the rate constants of the other species. A particularly simple case is given if the kinetic properties of the species $R^+$ and $R^-$ are identical, i.e. $A^+ = A^- = A$ and $D^+ = D^- = D$. The equation system simplifies to

$$\frac{dn^+}{dt} = An^- - Dn^+ \tag{36a}$$

$$\frac{dn^-}{dt} = An^+ - Dn^-. \tag{36b}$$

We solve this system with the initial conditions $n^+ = n^0$, $n^- = 0$. Addition of Eqs. (36a) and (36b) gives the differential equation

$$\frac{d(n^+ + n^-)}{dt} = (A - D)(n^+ + n^-) \tag{37}$$

with the solution

$$n^+ + n^- = n^0 e^{(A-D)t}. \tag{38}$$

Using Eq. (38), $n^-$ can be eliminated from Eq. (36a). This results in the non-autonomous differential equation

$$\frac{dn^+}{dt} = An^0 e^{(A-D)t} - (A+D)n^+. \tag{39}$$

Applying the method *variation of constants* one finds the solution

$$n^+(t) = \frac{n^0}{2} e^{-Dt}(e^{At} + e^{-At}) = \frac{n^0}{2} e^{-Dt} \cdot cosh(At). \tag{40a}$$

Again using Eq. (38) leads to

$$n^-(t) = \frac{n^0}{2} e^{-Dt}(e^{At} - e^{-At}) = \frac{n^0}{2} e^{-Dt} \cdot sinh(At). \tag{40b}$$

It is obvious that in Eq. (40) for $A > D$ the exponential increasing terms dominate, so that the concentrations of both species increase – see also Eq. (35b).

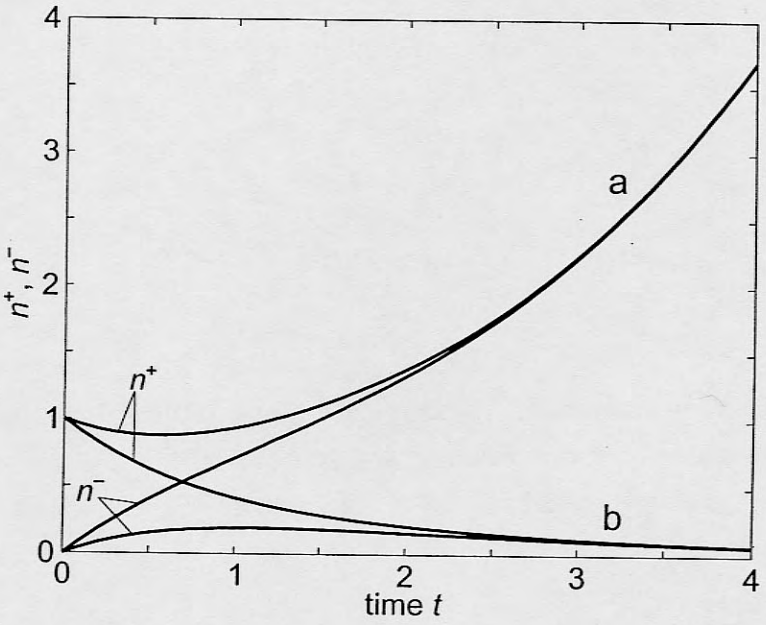


Figure 4: Growth and decay kinetics of the complementary self replication. Case a: $A^+ = A^- = 1$, $D^+ = D^- = 0.5$. Case b: $A^+ = A^- = 0.5$, $D^+ = D^- = 1$.

In Fig. 4 the time course of the concentrations is depicted for two different sets of parameters. It can be seen that after a certain time the differences of the concentrations resulting from different initial values become negligibly small and both species meet the same fate.

### 4.4.2 Selection of hypercycles

Many proteins (in particular enzymes) participate in the self replication of the cellular genetic material. With the help of polymerases the replication process is highly accelerated, resulting in a selective advantage for the respective nucleic acid sequences. Moreover the enzyme catalyzed self replication is much more precise (cf. section 4.3). It is still not known how more and more specific enzymes were added during evolution. Transcription and translation processes require a functioning genetic code. The emergence of a genetic code in an early phase of evolution is another unsolved problem.

A particularly simple form of cooperation between proteins and nuclein acids exists if an enzyme $E$ is the product of translation of the polymer $I$, while $E$ preferentially catalyses the transcription of $I$. The corresponding processes are depicted by the reaction scheme in Fig. 5.

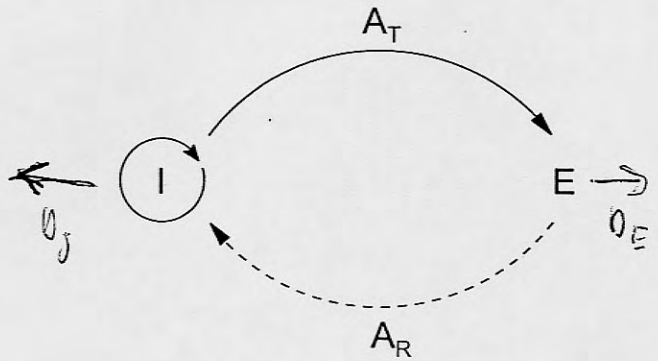


Figure 5: Simple hypercycle.

$A_T$ and $A_R$ are the kinetic constants for the translation and replication, respectively. The circle with the arrow symbolises the identical self replication of $I$. A more general form of cooperative interaction is depicted in Fig. 6. It contains a cycle with $2n$ elements ($n$ polynucleotides and $n$ enzymes), where the translation product $E_i$ of the polynucleotide $I_i$ catalyses the replication of a polynucleotide $I_{i+1}$. The cycle is closed because $E_n$ catalyses the replication of $I_1$. Such an arrangement of reactions is called a *hypercycle.*

In the following we only consider the dynamics of two-component hypercycles. The decay of polymer sequences is included in the description and linear and bilinear kinetic rate equations are assumed. Thus, the dynamic behaviour of a hypercyle can be described by the following system of differential equations:

$$\frac{dn_E}{dt} = A_T n_I - D_E n_E \tag{41a}$$

$$\frac{dn_I}{dt} = A_R n_I n_E - D_I n_I. \tag{41b}$$

In contrast to the replication models in the previous sections a non-linear term occurs ($= A_R n_I n_E$). It shows that the replication velocity rises with increasing polynucleotide

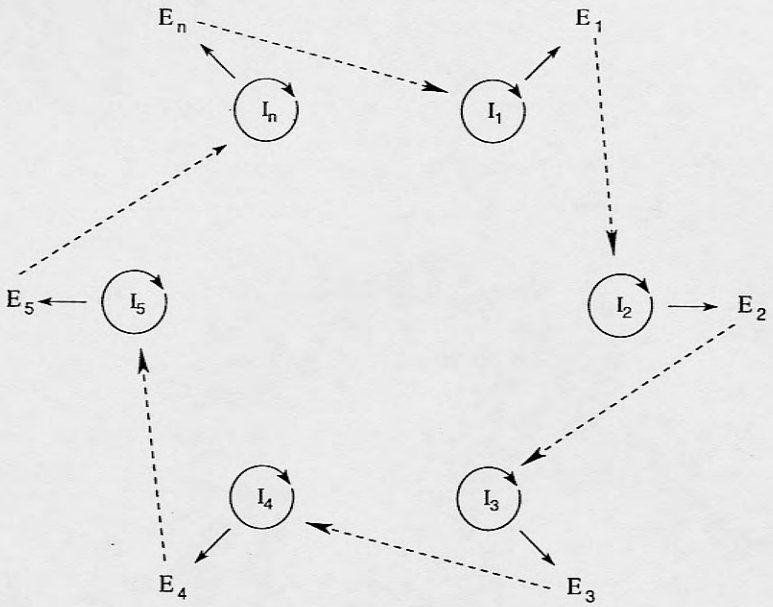


Figure 6: Complex hypercycle with $n$ self replicating polynucleotides and enzymes, respectively.

concentration as well as increasing enzyme concentration. It is more realistic to take into account a saturation of enzymatic activity, but the mathematical treatment is much more complicated. For simplicity we assume the translation to be very fast compared to the replication and the decay of the polynucleotide. In this case we can apply the quasi-steady state approximation (QSSA) for the variable $n_E$ ($dn_E/dt \cong 0$), which leads to

$$n_E = \frac{A_T}{D_E} n_I. \tag{42}$$

Inserting Eq. (42) into Eq. (41b) results in the non-linear differential equation

$$\frac{dn_I}{dt} = (\frac{A_R A_T}{D_E}) n_I^2 - D_I n_I. \tag{43}$$

This equation has the two steady states

$$n_I^{(1)} = 0 \tag{44}$$

and

$$n_I^{(2)} = \frac{D_I D_E}{A_R A_T}. \tag{45}$$

Eq. (43) can be solved by separation of variables and expansion into partial fractions, yielding

$$n_I(t) = \frac{n_I^0 n_I^2}{n_I^0 - (n_I^0 - n_I^2)e_I^D t} \tag{46}$$

with $n_I^0 = n_I(0)$. Equation (46) describes a decay process for $n_I^0 < n_I^{(2)}$ and a growth process for $n_I^0 > n_I^{(2)}$. In case of growth the polynucleotide concentrations grow unbounded for

$$t \to t_{crit} = \frac{1}{D_I} ln \frac{n_I^0}{n_I^0 - n_I^{(2)}}. \tag{47}$$

Therefore Eq. (46) is only valid for $0 \leq t < t_{krit}$. The temporal behaviour of $n_I$ is depicted in Fig. 7.

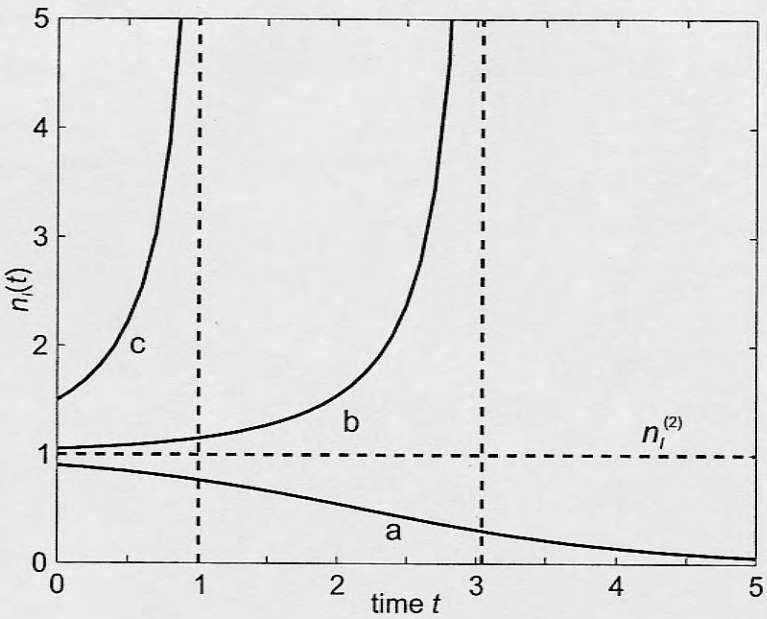


Figure 7: Growth kinetics of one single hypercycle according to Eq. (46). $A_I = A_E = D_I = D_E = 1$. Initial values: a) $n_I^0 = 0.9$, b) $n_I^0 = 1.05$, c) $n_I^0 = 1.5$.

Particularly interesting effects occur if there is a *competition of more than one hypercycle.* Under the condition of constant organisation for the polynucleotides as well as for the enzymes one gets the differential equation system

$$\frac{dn_{E,i}}{dt} = A_{T,i} n_{I,i} - \varphi_E(t) n_{E,i} \tag{48a}$$

$$\frac{dn_{I,i}}{dt} = A_{R,i} n_{I,i} n_{E,i} - \varphi_I(t) n_{I,i}. \tag{48b}$$

The decay terms were neglected since they influence the dynamics only marginally in the present case. From

$$\sum_i \frac{dn_{E,i}}{dt} = 0 \quad , \quad \sum_i \frac{dn_{I,i}}{dt} = 0 \tag{49}$$

it follows that

$$\varphi_E = \frac{\sum_i A_{T,i} n_{I,i}}{\sum_k n_{E,k}} \quad , \quad \varphi_I = \frac{\sum_i A_{R,i} n_{I,i} n_{E,i}}{\sum_k n_{I,k}}. \tag{50}$$

Considering the competition of only two two-component hypercycles one finds that two different steady states exist:

$$1. \quad n_{E,1}^{(1)} = 0 \quad , \quad n_{I,1}^{(1)} = 0 \quad , \quad n_{E,2}^{(1)} = C_E \quad , \quad n_{I,2}^{(1)} = C_I \tag{51a}$$

$$2. \quad n_{E,1}^{(2)} = C_E \quad , \quad n_{I,1}^{(2)} = C_I \quad , \quad n_{E,2}^{(2)} = 0 \quad , \quad n_{I,2}^{(2)} = 0 \tag{51b}$$

with the total concentration $C_E = n_{E,1} + n_{E,2} = const.$ and $C_I = n_{I,1} + n_{I,2} = const.$ These steady states correspond to the survival of one hypercycle and the extinction of the other. The dynamics of the system can be depicted in a two-dimensional phase diagram.

Figure 8 shows trajectories in the $(n_{I,1}, n_{E,1})$-plane. The concentrations of the other hypercycle can be obtained from $n_{I,2} = C_I - n_{I,1}$ and $n_{E,2} = C_E - n_{E,1}$.

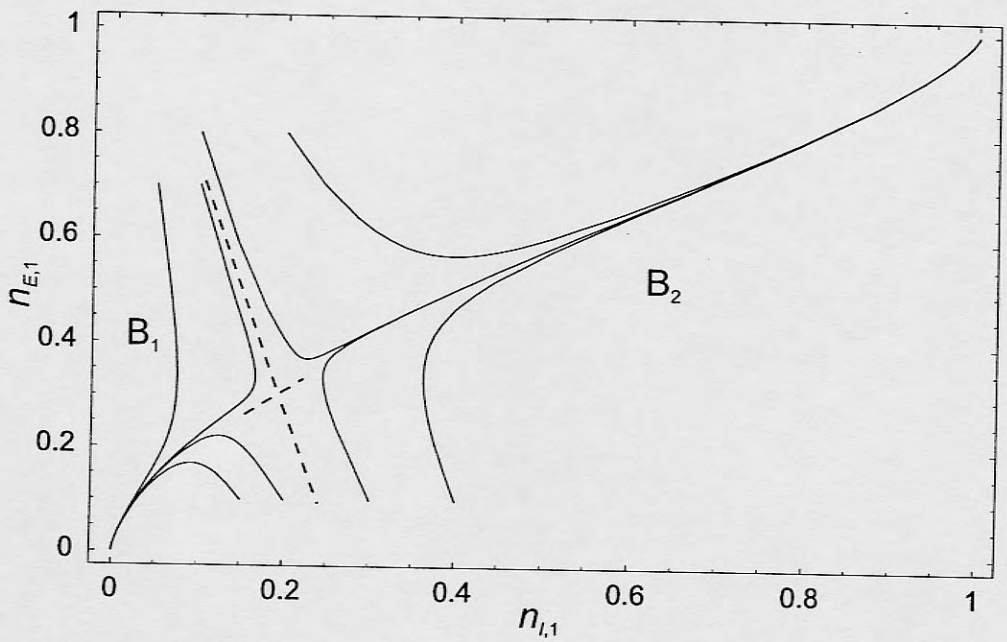


Figure 8: Competition of two hypercycles for different initial values, depicted in the $(n_{I,1}, n_{E,1})$-plane. Parameter values: $A_{T,1} = 1, A_{R,1} = 1, A_{T,2} = 0.5, A_{R,2} = 0.5$.

The kinetic parameters were chosen such that the first hypercycle shows a better growth behaviour: $A_{T,1} = 2A_{T,2}, A_{R,1} = 2A_{R,2}$. One finds that the system's behaviour depends strongly on the initial values. Both steady states, (51a) and (51b), have a *basin of attraction*. Trajectories with initial values within $B_1$ ($B_2$) will end up in the steady state 1 (2). Although the second hypercycle shows a bad growth behaviour, it can be competitively successful provided that its initial values are more convenient (high initial

concentrations $n^0_{I,2}$ and $n^0_{E,2}$). Due to the more efficient growth rate, the first hypercycle has a larger basin of attraction.

Because of the dependency on the initial values, the selection of systems organised in an hypercyclic manner differ fundamentally from those of identical or complementary self replication without the help of specific catalysts. Let's assume that in a population of hypercycles during a growth phase or in the selection equilibrium a new hypercycle occurs due to incorrect replication. Since the associated molecules $I$ and $E$ initially are present in very low concentrations (at first only one copy), this hypercycle has only a very small chance to prevail. This is also valid if it is more efficient than all the other existing hypercycles. The disadvantageous influence of bad initial values can only be compensated with a very high reproduction rate. This growth characteristic brings about that in evolution also non-optimal systems could have prevailed if coincidentally they had convenient initial values. Because of this possibility of a "*once-for-ever*" selection, one has to assume, that many properties of living systems can be regarded as *fixed coincidences*. Hypercycles combine in a convenient manner the properties of polynucleotides (self reproduction) with those of enzymes (acceleration of the processes by catalyses, increasing accuracy of replication).

## Exercises

### Exercise 1. Condition of constant organisation without mutations

Demonstration of the temporal development of three molecule concentrations by numerical integration of the equation system:

$$\frac{d}{dt}n_i = (A_i - D_i - \bar{E}(t))n_i \quad \text{with} \quad \bar{E} = \frac{\sum_{k=1}^{3}(A_K - D_k)n_k}{\sum_{k=1}^{3} n_k}$$

it holds $\sum_{k=1}^{3} n_k = 1$ .

(a) What are possible steady states of the system?

(b) Plot the temporal development of the molecule concentrations for the parameters $A_1 = 4, A_2 = 4, A_3 = 3, D_1 = 0.3, D_2 = 0.2, D_3 = 0.5$ and different initial values.

(c) Plot $n_2$ and $n_1$ in a phase plane.

### Exercise 2. Allowing for mutations

Taking into account mutations the differential equation system becomes

$$\frac{d}{dt}n_i = (A_iQ_i - D_i - \bar{E}(t))n_i + \sum_{j \neq i} m_{ij}n_j \quad \text{with} \quad \bar{E} = \frac{\sum_{k=1}^{3}(A_K - D_k)n_k}{\sum_{k=1}^{3} n_k}$$

Again it holds $\sum_{k=1}^{3} n_k = 1$ .

(a) Figure out the correlation between the parameters $Q_i$ and $m_{ij}$.

(b) Plot the temporal development of all molecule concentrations and also $n_2$ and $n_1$ in a phase plane.

(c) Calculate the eigenvalues $\lambda_i$ and the eigenvectors $v_i$ of the matrix

$$\begin{pmatrix} A_1Q_1 - D_1 & m_{12} & m_{13} \\ m_{21} & A_2Q_2 - D_2 & m_{23} \\ m_{31} & m_{32} & A_3Q_3 - D_3 \end{pmatrix}$$

with MATHEMATICA (Eigenvalues[matrix] and Eigenvectors[matrix]) for a fixed set of parameters.

(d) Prove that for large values of time $t$ it holds:

- mean overproduction $\bar{E}(t) \rightarrow \max\{\lambda_i\}$ and
- the components of the eigenvectors belonging to the largest eigenvalue give the distribution of the polynucleotides.

The reason for this is, that the differential equation can be simplified via a linear transformation:

$$n_i \to \sum_{j=i}^{3} c_{ij} y_j \quad (\text{with} \quad \sum_i n_i = \sum_i y_i = const = 1) \longrightarrow \frac{d}{dt} y_i = (\lambda_i - \bar{E}(t)) y_i.$$

which can be treated analogously to the case "selection without mutation".

**Exercise 3. Cooperation of species**

Three species are given, each catalysing the replication of one other species according to the scheme (mutations and decay processes are neglected):

$$\left.\begin{array}{l} x_1 + x_3 \xrightarrow{f_1} 2x_1 + x_3 \\ x_2 + x_1 \xrightarrow{f_2} 2x_2 + x_1 \\ x_3 + x_2 \xrightarrow{f_3} 2x_3 + x_2 \end{array}\right\} \Rightarrow \tfrac{d}{dt} n_i^{QS} = f_i n_i^{QS} n_j^{QS} - n_i^{QS} \varphi \quad (i = 1,2,3 \quad j = i - 1 + 3\delta_{i1})$$

(a) Determine $\varphi$ under the condition of constant organisation.

(b) Calculate the steady states of the system.

(c) Plot the solutions of the numerical integration.

(d) By considering an example, show that the concentrations of $n \geq 5$ species, which are arranged in a hypercycle, oscillate.

**Exercise 4. Competition of hypercycles**

Competing hypercycles can be formulated mathematically as:

$$\frac{d}{dt} n_i^{HZ} = k_i (n_i^{HZ})^2 - n_i^{HZ} \varphi$$

(a) Determine $\varphi$ under the condition of constant organisation.

(b) Plot the solution and investigate the expansion of the (three) basins of attraction in the phase plane with appropriate choice of different initial values.