The Influence of Mutation on Autocatalytic Reaction Networks

by

PETER F. STADLER

*Institut für Theoretische Chemie
Universität Wien

Währingerstr. 17, A-1090 Wien, Austria
Phone: **43 1 40480 678 Fax: **43 1 4028525
Email: studla@itc.univie.ac.at

The Santa Fe Institute

1660 Old Pecos Trail, Santa Fe, NM 87501, U.S.A.
Phone: **1 (505) 984 8800 Fax: **1 (505) 982 0565
Email: stadler@sfi.santafe.edu

AND

JUAN CARLOS NUÑO

Departamento de Matemáticas, ETSI de Montes
Universidad Politécnica de Madrid

E-28040 Madrid, Spain
Phone: **34 1 3367107 Fax: **34 1 3944159
Email: juanca@solea.quim.ucm.es

*Address for correspondence

Mathematical Biosciences, in press.

Abstract

A particular class of ordinary differential equations (ODE) describing catalyzed, template-induced, and erroneous replication is investigated. The ODEs can be split into a replicator part accounting for the correct replication and a mutation term accounting for all miscopying processes. The set of all species is divided into the catalytically active "viable" species and an error-tail subsuming all other species. Neglecting both the inter-mutation among the viable species and the re-flux from the error-tail allows for an extensive analysis of the autocatalytic network. If mutation rates are small enough a perturbation approach is feasible showing that mutation in general simplifies the qualitative behavior of the dynamical system. Special cases, such as Schlögl's model, the uniform model, and the hypercycle show that the viable species become unstable beyond a critical mutation rate: there is an analogue to the error-threshold of the quasispecies model also in non-linear autocatalytic reaction networks with mutation.

1. Introduction

Two apparently different models for the dynamical behavior of self-replicating macromolecules have been discussed in literature: self-replicating species which are not catalytically active are described by a set of first order replicator equations, known as the *quasispecies* model [1,2]. Self-replication of catalytically active species is usually modeled by second order replicator equation. The most famous special case of this class of dynamical system is the *hypercycle* [3].

Both models may be viewed as particular cases of a general replication-mutation process characterized by an error-prone replication mechanism of the form

$$(\mu) + I_k \xrightarrow{Q_{kk}\mathcal{F}_k} 2I_k + (\xi)$$
$$(\mu) + I_k \xrightarrow{Q_{kj}\mathcal{F}_k} I_k + I_j + (\xi)$$

Here, I_k denotes a self-replicating species, and $\mathcal{F}_k$ represents the sum of the catalytic actions involved in this replication reaction, including the effect of the reaction medium on the replication process. (μ) is the set of materials from which the replicators are built up, and (ξ) denotes the by-products of the reaction. We will assume that (μ) is buffered; therefore the raw material need not be taken into account explicitly in the rate equations, but it is included into the reaction rate constants. The waste (ξ) will not occur in the rate equations at all.

Obviously, the replication of biopolymers (RNA or DNA) consists of a very large number of elementary steps. It is a reasonable approximation, however, to subsume them in over-all reaction rates depending only on the concentrations c_k of the self-replicating species [4]. We can expand $\mathcal{F}$ as power series in terms of concentrations:

$$\mathcal{F}_k(\mathbf{c}) = \alpha_k + \sum_{j=1}^{n} A_{kj} c_j + \sum_{i=1}^{n} \sum_{j=1}^{n} A_{kij} c_j c_i + \dots \qquad k = 1, 2, \dots, n.$$

Truncating this series after the constant term, one obtains a first order replicator model. Retaining also the linear terms $A_{kj} c_j$ we get a second order catalytic model. The matrix $A = [A_{kj}]$ will be referred to as catalytic matrix.

In addition to the above reactions which build up the macromolecules, a more realistic setting will include also the degradation reactions of these macromolecules, which may (but need not) involve specific catalysts:

$$I_k \xrightarrow{\quad \mathcal{G}_k \quad} (\xi)$$

Here $\mathcal{G}_k$ subsumes all degrading reactions pathways and may, like $\mathcal{F}_k$, depend on the concentration $\mathbf{c}$ of the catalytically active species. Again, we can expand $\mathcal{G}_k$ into a power series:

$$\mathcal{G}_k(\mathbf{c}) = \delta_k + \sum_{j=1}^{n} D_{kj} c_j + \sum_{i,j=1}^{n} D_{kij} c_i c_j + \dots$$

The catalytic degradation matrix will be denoted by $D = [D_{kj}]$.

In nature, replication does not work with indefinite fidelity. Replication errors — mutations — are introduced into the model via the probabilities Q_{kl} that the outcome of an attempt to replicate a molecule of type I_l be an (erroneous) "copy" of type I_k. Thus, Q_{ll} denotes the probability for a correct replication. The matrix $Q \equiv [Q_{kl}]$ is referred to as *mutation matrix.*

Eigen [1] proposed a model for the mutation probabilities based on point mutations only. Suppose each species is represented by a string of length ν made up of some alphabet, and for each digit there is a chance p' for being (independently) mutated. Let $q = 1-p'$ be the so-called single digit accuracy. The mutation matrix for Eigen's model is then given by

$$Q_{kl} = q^{\nu - d(k,l)}(1-q)^{d(k,l)}$$

$d(k,l)$ denotes the Hamming distance, i.e., the number of positions in which species I_k differs from species I_l.

Notice that, under this realistic assumption, mutation links all the species together. Of course, this fact prohibits a simple solution of the corresponding selection-mutation equations. Recently, Nuño *et al.* [10] have introduced a model that is the natural simplification of Eigen's selection-mutation equations. This model is based on the assumption that all catalytically active species are well separated in sequence space, i.e., that all pairwise Hamming distances of viable species are large. Then, all components of the mutation matrix that relate viable species with each other may be neglected. Even under this simplification the dynamics of the system remains extremely complex, and hence additional assumptions are needed [10,11].

In the following section, we will describe the details of our model. Section 3 presents some general results derived from the perturbation approach. The first order selection-mutation equation is compared with Eigen's quasispecies model in section 4. Some general properties of second order kinetics are outlined in section 5. In section 6 we analyze the second order model by means of the perturbation theory. In section 7 we will investigate some important special cases in detail. Finally, we will fit our results into the framework of general replicator mutation models.

2. The Model

The particular model we will investigate in this contribution is based on the following simplifying assumptions:

- The variety of all species in the system can be split into two classes:

  The first class consists of the *viable* species I_k which have comparably high replication rates and are in general catalytically active. It is assumed that these species are well separated in sequence space, i.e., that the corresponding mutation probabilities Q_{kj} can be neglected. Each species in this class is characterized by its individual replication and degradation rate functions $\mathcal{F}_k$ and $\mathcal{G}_k$, respectively. If we neglect the catalytic activity of the replicating species the growth function $\mathcal{F}$ reduces to the vector of first order replication constants α. Correspondingly, $\mathcal{G}$ reduces to the vector of degradation rate constants δ.

  The second class consists of the remaining species with low replication rate and negligible activity as specific catalysts. For simplicity, they are assumed to have constant replication and degradation rates independent of the concentrations of the viable species. Thus, they are formally subsumed as a single *error-tail* species I_0.

- Mutation is therefore modeled as additional production of I_0. Strictly speaking there should be mutations from the error-tail species to the viable species. However, the replication rate of I_0 is assumed to be so small that these terms can be neglected. Since for any viable species I_k the only reaction products are I_k itself and the error-tail I_0 we have the condition $Q_{kk} + Q_{0k} = 1$ for all $k = 1, \ldots n$.

- The replication accuracy is the same for all viable species. Therefore, $Q_{kk} = 1 - Q_{0k} = 1 - p$ where p is referred to as mutation rate.
- The reaction takes place in a continuously stirred tank reactor (CSTR) [1,16], see figure 1. A selection pressure is introduced by means of an non-specific dilution flux from the system, Φ. Schematically,

$$I_k \xrightarrow{\quad\Phi\quad} \textit{(outside)}$$

This flux is taken such that the total concentration of macromolecules

$$C = c_0 + \sum_{k=1}^{n} c_k$$

is kept constant. Therefore we may use normalized concentration variables

$$x_k = c_k / C.$$

The phase space of our model is therefore the simplex of normalized concentrations

$$S_{n+1} = \left\{ x \in \mathbb{R}_+^{n+1} \;\middle|\; \sum_{j=0}^{n} x_j = 1 \right\}$$

We will use the notation $x = (x_1, \ldots, x_n)$ for the concentrations of the viable species and x_0 for the concentration of the error-tail, while the symbol $\mathbf{x}$ is reserved for (x_0, x). $\mathbf{1}$ denotes the vector $(1, \ldots, 1)$ of appropriate dimension.

The differential equations for such a reaction network can be written in a compact way as

$$\begin{aligned} \dot{x} &= x \circ [(1-p)\mathcal{F}(x) - \mathcal{G}(x) - \Phi\mathbf{1}] \\ \dot{x}_0 &= x_0 \cdot [\alpha_0 - \delta_0 - \Phi] + p \cdot \langle x, \mathcal{F}(x) \rangle \end{aligned} \tag{1}$$

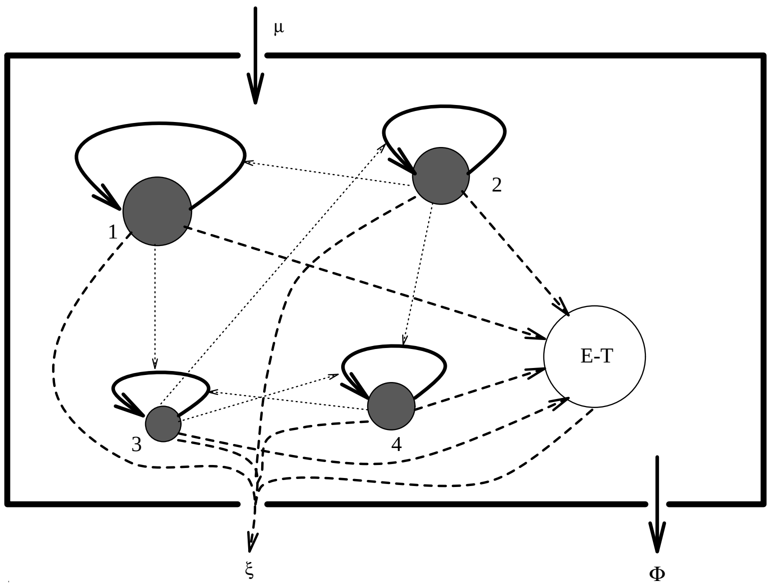

Figure 1: Scheme of the model. A set of viable species (dark balls) and the error-tail (E-T) enclosed in a continuous stirring tank reactor (CSTR) replicate. Catalyzed replication is indicated by bold arrows, the action of catalysts by dotted arrows. The building blocks μ are assumed to be buffered and there are 2 pathways removing the species: an unspecific dilution flux Φ and degradation to non-reactive waste ξ.

From $\sum_{i=0}^{n} \dot{x}_i = 0$ we find for the flux function

$$\Phi = \langle [\mathcal{F}(x) - \mathcal{G}(x)], x\rangle + x_0(\alpha_0 - \delta_0) \tag{2}$$

Here $\circ$ denotes the Schur product $x \circ y = (x_1 y_1, \ldots, x_n y_n)$. For the scalar (dot) product of two vectors we use $\langle\, .\,,\, .\,\rangle$. We will often make use of the general relation

$$\langle \mathbf{1}, x\rangle = 1 - x_0$$

Remark. The set of equations (1-2) can be recast in the general form of a selection mutation model [5,18]

$$\dot{\mathbf{x}} = \mathbf{x} \circ [F(\mathbf{x}) - \langle \mathbf{x}, F(\mathbf{x})\rangle] + \mathcal{M}(\mathbf{x}).$$

With the approximations discussed above we have

$$F(\mathbf{x}) = F(x, x_0) = \begin{pmatrix} \mathcal{F}(x) - \mathcal{G}(x) \\ \alpha_0 - \delta_0 \end{pmatrix}$$

for the effective growth rates and

$$\mathcal{M}(\mathbf{x}) = \mathcal{M}(x, x_0) = p\begin{pmatrix} -x \circ \mathcal{F}(x) \\ \langle x, \mathcal{F}(x)\rangle \end{pmatrix} = (Q - I)\cdot\left[\mathbf{x}\circ\begin{pmatrix}\mathcal{F}(x)\\ \alpha_0\end{pmatrix}\right]$$

for the mutation terms, where Q is the mutation matrix and I is the identity matrix. Note that neither $F(\mathbf{x})$ nor $\mathcal{M}(\mathbf{x})$ depends explicitly on the concentration x_0 of the error tail in our simplified mutation model.

The simplex, S_{n+1}, is forward invariant for the mutation field $\mathcal{M}$. It is clear that the sum of all mutations terms must vanish,

$$\langle \mathbf{1}, \mathcal{M}\rangle = \sum_{j=0}^{n} \mathcal{M}_j = 0.$$

Let $K \subset \{1, \ldots, n\}$ be an arbitrary set of coordinates. Clearly the faces

$$H_K := \{x \in S_{n+1} | x_k = 0 \iff k \in K\}$$

where exactly the species labeled by $k \in K$ vanish, are invariant under the dynamics of the selection mutation equation.

The Q-matrix proposed by Eigen [1,2,12] is irreducible, i.e., there is a pathway between any two viable species via subsequent mutations. In contrast to this, in our model there is no influx into a viable species. Mutation is modeled as an outflux term from the viable species. The invariance of the faces H_K is an immediate consequence of the lack of influx terms. The difference between of Eigen's fully connected sequence model and our simplified ansatz is depicted in figure 2.

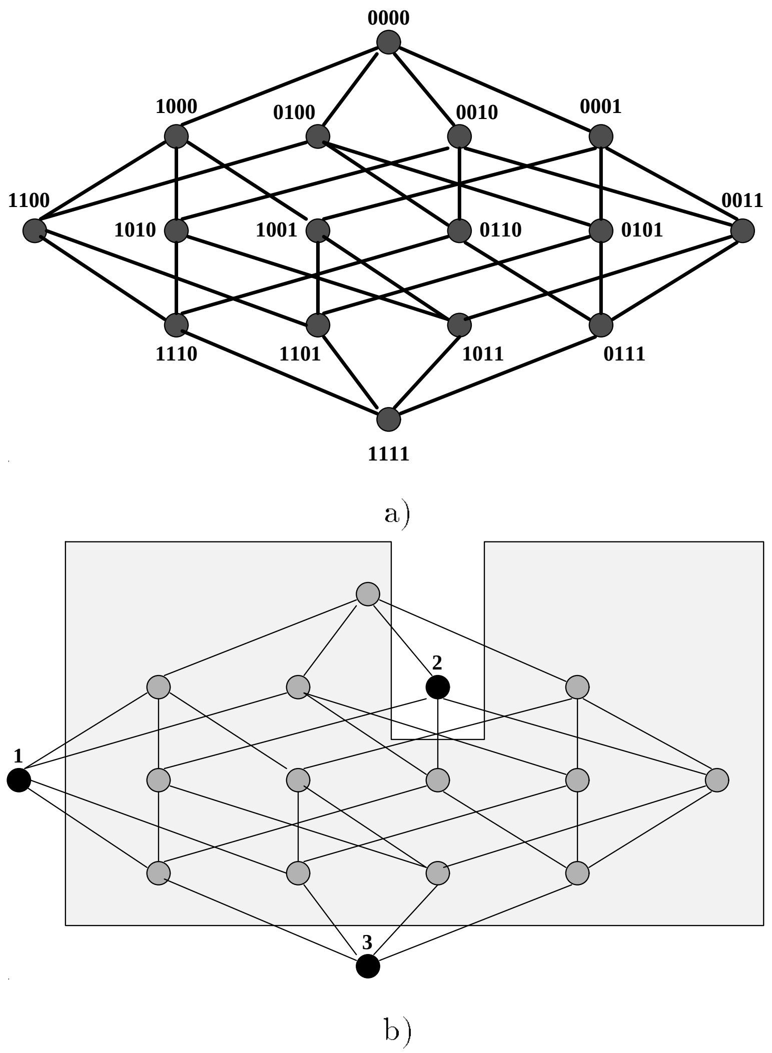

Figure 2: The mutation model introduced by Eigen takes into account all 2^ν possible sequences. Our simplified model subsumes all species except for the n viable species — here denoted by *1,2,3* — in an error-tail (shaded region). This simplification reduces the number of equations from 2^ν to $n+1$, one for each viable species and a single equation for the error-tail.

3. Error-Prone Replication

Let us consider the general selection-mutation equation

$$\dot{x} = x \circ [f(x) - \langle x, f(x)\rangle] + \mathcal{M}(x) \tag{3}$$

with $\langle \mathbf{1}, \mathcal{M}\rangle = 0$. It is defined on the simplex S_n. (There is no distinction between viable and non-viable species in the general model.) In order to simplify notation we have used the net growth function

$$f(x) = \mathcal{F}(x) - \mathcal{G}(x).$$

If $\mathcal{M}$ is small enough on S_n, equation (3) may be considered as a perturbed replicator equation. 'Small' refers here to the norm

$$\|F\| = \sup_{x\in S_n} \left\{ F_k(x), \quad \frac{\partial F_k}{\partial x_l}(x) \quad \middle| \quad 1 \le k, l \le n \right\}$$

for differentiable vector fields F on the simplex.

Let $S[K]$ denote the sub-simplex of S_n which is obtained by omitting all coordinates $k \in K$, and let F_K be the the interior of $S[K]$. It is well known [17] that the replicator equation [7,16]

$$\dot{x} = x \circ [f(x) - \langle x, f(x)\rangle] \tag{4}$$

leaves all subsimplices invariant. Suppose now that $\mathcal{M}$ leaves a family Σ of subsimplices $S[K_1]$ through $S[K_n]$ invariant. It is clear, that for each subsimplex $S[K]$ there exists a unique smallest invariant simplex $S[\tilde{K}]$ in the family Σ. If $S[K]$ itself is invariant under $\mathcal{M}$ then $S[\tilde{K}] = S[K]$.

In general, a physically meaningful mutation field $\mathcal{M}$ will have the following properties:

(i) The simplex is forward invariant for $\mathcal{M}$, i.e., there is no trajectory of $\dot{x} = \mathcal{M}(x)$ with initial condition in the simplex that leaves the simplex.

(ii) Whether a species is produced by mutation depends only on the *presence* of suitable parents (and catalysts), but not on their actual concentrations. Technically this means that for each subsimplex S[K] and all $k \in K$ holds
either $\mathcal{M}_k(x) > 0$ for all $x \in F_K$,
or $\mathcal{M}_k(x) = 0$ for all $x \in F_K$ (and hence even for all $x \in S[K]$).

It is easy to check that these conditions are fulfilled for all mutation fields of the form

$$\mathcal{M} = (Q - I) \cdot (x \circ \mathcal{F}(x))$$

with a positive replication rate functions $\mathcal{F}(x)$ and arbitrary mutation matrix. All mutation model studied so far (including the system discussed in this contribution) are of the form [1,5,10,18].

A rest point $\hat{x}$ lying on a certain face F_K has been called *saturated* [16] if all its transversal eigenvalues are non-positive, i.e., if

$$\lambda_k^{\text{trans}} = f_k(\hat{x}) - \langle \hat{x}, f(\hat{x}) \rangle \leq 0 \qquad \text{for all } k \in K$$

It has been shown recently [18] that saturated regular rest points are exactly those which are kept in the simplex under the effect of mutational perturbations $\mathcal{M}$ with irreducible mutation matrix Q. In the following we will extend this results to selection-mutation models with arbitrary mutation matrices. We will need the following

Definition. Let $\hat{x}$ be a rest point in F_K and let $S[\tilde{K}]$ be the smallest simplex containing F_K which is invariant under $\mathcal{M}$. $\hat{x}$ is *saturated with respect to* $S[\tilde{K}]$ if all transversal eigenvalues of $\hat{x}$ in plane spanned by the simplex $S[\tilde{K}]$ are non-positive, i.e.,

$$\lambda_k^{\text{trans}} = f_k(\hat{x}) - \langle \hat{x}, f(\hat{x}) \rangle \leq 0 \qquad \text{for all } \; k \in (K \setminus \tilde{K})$$

Saturation with respect to a subsimplex $S[\tilde{K}]$ is just saturation in the framework of a replicator system restricted to the simplex $S[\tilde{K}]$, i.e., where all species in $\tilde{K}$ are omitted.

We have the following version of the rest point migration theorem [18]:

Theorem 1. Let $\mathcal{M}$ be a mutation field, and let $\hat{x} \in F_K$ be a regular rest point of the replicator equation and let $S[\tilde{K}]$ be the smallest simplex invariant under $\mathcal{M}$ which contains F_K. Then for sufficiently small $\|\mathcal{M}\|$ holds

(i) There is a unique regular rest point $\hat{y}$ of the perturbed equation in a neighborhood of $\hat{x}$.

(ii) $\hat{y}$ is contained in the interior of $S[\tilde{K}]$ if and only if $\hat{x}$ is saturated with respect to $S[\tilde{K}]$. Otherwise $\hat{y}$ lies in the unphysical outside of the simplex.

(iii) If $\hat{x}$ is hyperbolic then $\hat{y}$ is also hyperbolic with the same stability properties.

The proof parallels exactly the proof in [18] and is omitted here.

This theorem can be immediately applied to the dynamics of our simplified selection mutation model (see also figure 3).

Corollary. A rest point $(\hat{x}, \hat{x}_0)$ of the system (1-2) leaves the simplex whenever the transversal eigenvector pointing in the direction of the error-tail is positive, i.e., for $p > 0$ but sufficiently small holds

$$\hat{x}_0 < 0 \iff \lambda_0 = e_0 - \Phi(\hat{x}) > 0$$

with $e_0 = \alpha_0 - \delta_0$ being the selective value of the error-tail.

Proof. For the simplified selection-mutation dynamics (1-2) all subsimplices H_K are invariant, and a rest point in the interior of H_K has only one transversal eigenvector, which is pointing towards the error-tail. Saturation with respect to H_K is therefore determined by this single eigenvalue. The corollary is now an immediate consequence of the rest point migration theorem 1. ∎

The position of a regular rest point $\hat{x}(0)$ of the replicator equation equ.(4) is changed under the influence of the mutation field $\mathcal{M}$. Up to first order in the strength p of the mutation term, one obtains the expansion

$$\hat{x}(p) = \hat{x}(0) + p \cdot d + \mathcal{O}(p^2) \tag{5}$$

where the shift d is given by

$$d = -J^{-1}(\hat{x}(0))(\hat{x}(0)) \tag{6}$$

where $J(x)$ denotes the Jacobian of the replicator field at the position x [18].

The vector of the first-order shifts, d, can be calculated explicitly from the above perturbation equation (6).

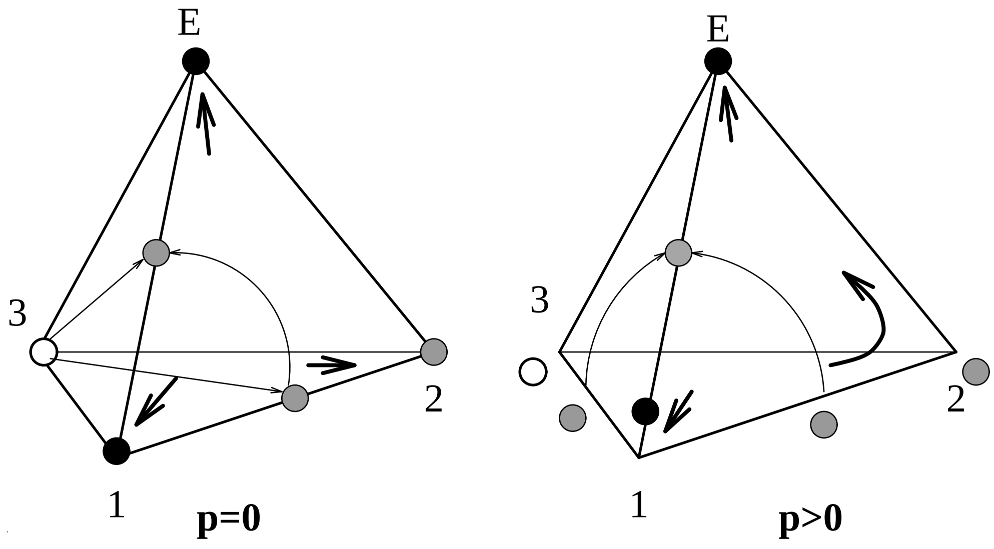

Figure 3: Perturbed and unperturbed phase portrait for

$$\alpha = \begin{pmatrix} 8 \\ 8 \\ 8 \end{pmatrix} \qquad A = \begin{pmatrix} 6 & 0 & 3 \\ 1 & 1 & 3 \\ 1 & 0 & 1 \end{pmatrix} \qquad e_0 = 10$$

Rest point 2 is stable in the absence of the error-tail, but unstable against invading error-tail species. All rest points with positive transversal eigenvalue in the direction towards the error-tail (E) are driven out of the simplex. Only the stable equilibrium 1 is in the physically accessible range for $p>0$.

Theorem 2. Let $\hat{x}$ be a regular rest point of the replicator equation (4). Then the first order shift $\mathbf{d} = (d, d_0)$ for the model (1-2) with $p > 0$ is given by

$$\begin{aligned} d &= [\partial f(\hat{x})]^{-1} \left(\mathcal{F}(\hat{x}) + r \cdot \mathbf{1}\right) \\ d_0 &= \frac{\langle \hat{x}, \mathcal{F}(\hat{x})\rangle}{\Phi(\hat{x}) - e_0} \\ &\text{with} \\ r &= -\frac{\langle \hat{x}, \mathcal{F}(\hat{x})\rangle + (\Phi(\hat{x}) - e_0)\langle \mathbf{1}[\partial f(\hat{x})]^{-1}\mathcal{F}(\hat{x})\rangle}{(\Phi(\hat{x}) - e_0)\langle \mathbf{1}[\partial f(\hat{x})]^{-1}\mathbf{1}\rangle} \end{aligned} \tag{7}$$

Proof. Let $\vec{e}$ denote the unit vector in direction of the error-tail. Suppose $\hat{x}_0 = 0$, i.e., the error-tail vanishes for $p = 0$. The Jacobian J at the position of an interior rest point is then easily calculated as

$$J(\hat{x}) = \partial\left\{\hat{x} \circ [f(\hat{x}) - \langle \hat{x}, f(\hat{x})\rangle \mathbf{1} - e_0 \hat{x}_0 \mathbf{1}]\right\} =$$
$$= \hat{x} \circ [\partial f(\hat{x}) - (\hat{x}\partial f(\hat{x})) \otimes \mathbf{1} - f(\hat{x}) \otimes \mathbf{1} - e_0 \vec{e} \otimes \mathbf{1}]$$

Here $\otimes$ denotes the Kronecker (direct) product. The equation for d becomes thus $\hat{x} \circ [\partial f(\hat{x})d - \langle \hat{x}\partial f(\hat{x})d\rangle \mathbf{1} - \langle f(\hat{x}), d\rangle \mathbf{1} - e_0 d_0 \mathbf{1}] = \hat{x} \circ \mathcal{F}(\hat{x})$ for the non-zero subsystem. Let us introduce the abbreviation

$$r = \langle f(\hat{x}), d\rangle + e_0 d_0 - \langle \hat{x}\partial f(\hat{x})d\rangle$$

The equation for the shifts d simplifies to $[\partial f(\hat{x})]d = \mathcal{F}(\hat{x}) + r\mathbf{1}$. Solving for d and multiplying by $\mathbf{1}$ yields

$$\langle \mathbf{1}, d\rangle = -d_0 = \langle \mathbf{1}[\partial f(\hat{x})]^{-1}\mathcal{F}(\hat{x})\rangle + r\langle \mathbf{1}[\partial f(\hat{x})]^{-1}\mathbf{1}\rangle$$

hence we can calculate r in terms of d_0:

$$r = -\frac{d_0 + \langle \mathbf{1}[\partial f(\hat{x})]^{-1}\mathcal{F}(\hat{x})\rangle}{\langle \mathbf{1}[\partial f(\hat{x})]^{-1}\mathbf{1}\rangle}$$

The first order approximation of the shift of the concentration of the error-tail is readily obtained from the 0th component of the perturbation equation

$$[e_0 - \Phi(\hat{x})]d_0 = -\langle \hat{x} g(\hat{x})\rangle$$

We obtain therefore

$$r = -\frac{\langle \hat{x}\mathcal{F}(\hat{x})\rangle + (\Phi(\hat{x}) - e_0)\langle \mathbf{1}[\partial f(\hat{x})]^{-1}\mathcal{F}(\hat{x})\rangle}{(\Phi(\hat{x}) - e_0)\langle \mathbf{1}[\partial f(\hat{x})]^{-1}\mathbf{1}\rangle}$$

∎

Only a few results can be obtained beyond the regime of the perturbation theory. The most interesting is

Theorem 3. For all systems with $e_0 > 0$ there is a critical error rate $p^{\mathrm{cr}} < 1$ such that for all p with $p^{\mathrm{cr}} \leq p \leq 1$ the pure error-tail state, $x_0 = 1$, is locally stable.

Proof. The stability of the error-tail is determined by the transversal eigenvalues

$$\lambda_k^e = \alpha_k - \delta_k - e_0 - p\alpha_k$$

which are negative for

$$p > 1 - \frac{\delta_k + e_0}{\alpha_k}$$

if $\alpha_k > 0$ and they are always if $\alpha_k = 0$. ∎

Corollary. (i) If $\alpha_k > 0$ for some k then

$$p^{\mathrm{cr}} = 1 - \min_k\{\frac{\delta_k + e_0}{\alpha_k}\}$$

where the minimum is to be taken over all k such that $\alpha_k > 0$.

(ii) If $\alpha_k = 0$ for all k then the error-tail is locally stable for all error rates.

4. First Order Kinetics

If we consider a system without any specific catalytic action of the viable species we obtain our version of the quasispecies model [2] defined by the following set of first order reactions

$$\begin{aligned}
I_k &\xrightarrow{(1-p)\alpha_k} 2I_k \\
I_k &\xrightarrow{p\alpha_k} I_k + I_0 \\
I_k &\xrightarrow{\delta_k} \emptyset \\
I_k &\xrightarrow{\Phi} \textit{(outside)}
\end{aligned}$$

Note that in this case the function $\mathcal{F}$ and $\mathcal{G}$ are just the vectors of the replication constants, α, and degradation constants, δ. The differential equations (1-2) therefore become

$$\dot{x}_k = x_k(e_k - \Phi) - p\alpha_k x_k$$

$$\dot{x}_0 = x_0(e_0 - \Phi) + p\langle x, \alpha\rangle$$

where $e_k = \alpha_k - \delta_k$ for all $k = 1, \ldots, n$.

A stochastic version of this model with a single viable species I_1 has been studied recently [9]. In the following we will restrict ourselves to the generic case $e_i \neq e_j$ and $\alpha_i \neq \alpha_j$ for all $i \neq j$ and $e_0 \ll e_k$ for all $k \neq 0$.

Following Eigen [2] we define the fitness of a species by

$$w_k = \alpha_k(1-p) - \delta_k = e_k - p\alpha_k. \tag{8}$$

We have then

Theorem 4. (*Survival of the Fittest.*) All trajectories converge to a state in which only the viable species with maximal fitness $w_{\max}$ and the error-tail are present. For almost all choices of p the survivor is a unique viable species together with its error-tail.

Proof. In the generic case all equilibria contain only a single viable species [20]. Let I_m be this master species with fitness $w_m = w_{\max}$. The transversal eigenvalues [7] corresponding to the directions of the other viable species are given by $\lambda_k^{\mathrm{trans}}(I_m) = e_k - \Phi - p\alpha_k$ for all $k \neq 0, m$. Substituting $\Phi = e_m - p\alpha_m$ we have

$$\lambda_k^{\mathrm{trans}}(I_m) = (e_k - p\alpha_k) - (e_m - p\alpha_m) = w_k - w_m,$$

i.e., a rest point is saturated if and only if the corresponding species has the maximal fitness value. Recall that only saturated equilibria can be stable. The remaining eigenvalue describes the stability against the error tail. It is negative for $e_k > e_0$, i.e., under the assumptions of the model. Hence if there is a unique species I_m with maximal fitness then the corresponding rest point is globally stable. This is the case for almost all choices of the mutation rate p. ∎

The unique surviving viable species is referred to as the master species. It is instructive to compare this result with the selection process in Eigen's more elaborate quasispecies model. As shown by Swetina and Schuster [21] the quasispecies is not always centered around the fittest species, i.e., the master species need not be the fittest. The actual outcome of selection in Eigen's quasispecies model depends critically not only on the fitness values of the sequences themselves but also on the fitness values of their neighbors. For sufficiently small mutation rates the master species is always the fittest. Swetina and Schuster [21] showed that, as the mutation rate increases, a transition can occur in which the fittest species (according to the definition in equ.(8)) is lost. The population becomes then dominated by a new master species which has a lower fitness by itself but which is surrounded by fitter neighbors than the previous one. This effect is due to the mutational influx terms from the error-tail. We have neglected these terms, therefore our model does not show such kind of transition.

We find, however, another type of transitions between different master species which has received little attention so far. At a critical value p^* of the mutation rate the fitness of the master species I_m equals the fitness of some competitor I_t, $(1-p)\alpha_m-\delta_m=(1-p)\alpha_t-\delta_t$. In this case, the master species at low mutation rates, I_m, must have both higher replication rate and higher degradation rate than the competitor I_t which takes over for $p>p^*$. With increasing mutation rate the relative effect of the degradation rate becomes more important since more and more replication products are lost to the error-tail. At the critical mutation rate p^* a degenerate bifurcation occurs: there is a line of fixed points connecting the solutions

$$(1-x_0^m,0,\ldots,x_0^m) \quad \text{and} \quad (0,1-x_0^t,0,\ldots,x_0^t)$$

where x_0^m and x_0^t are the relative concentrations of the error-tail at the transition, given that only the master species (or its competitor) is present in the

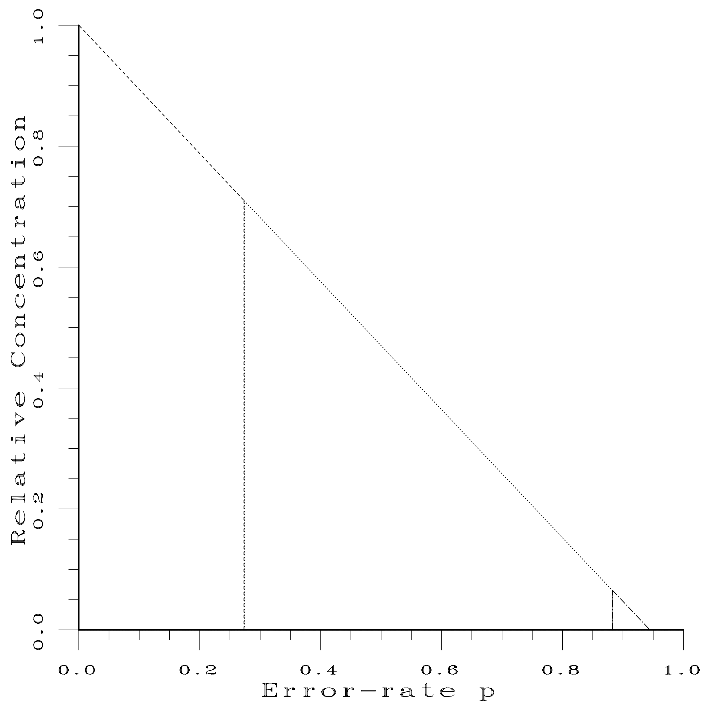

Figure 4: Concentrations of master species as a function of error- rate in system consisting of 10000 species with randomly assigned replication rates α_k and degradation rates δ_k. In this example, three different masters denoted by $\diamond$, $\circ$ and $\star$ appear before the pure error-tail becomes stable.

system. Note that x_0^m and x_0^t are the right and left limit of $x_0(p)$, respectively, when p approached the critical value p^*.

Theorem 5. The concentration x_0 of the error-tail increases linearly with p between transitions. It decreases, however, at all transition points.

Proof. The concentration of the error-tail for almost all mutation rates p is easily obtained from the parameters of the dominating species I_m:

$$x_0 = p \cdot \frac{\alpha_m}{e_m - e_0},$$

hence x_0 increases linearly between transitions. Suppose I_m dominates for $p < p^*$ and I_t dominates for $p > p^*$. Then $\alpha_m > \alpha_t$ and $e_m > e_t$. The transition occurs for

$$1 > p^* = \frac{e_m - e_t}{\alpha_m - \alpha_t} > 0$$

The second inequality is always fulfilled. The first inequality can be rewritten as

$$\alpha_t \left(\frac{e_t - e_0}{\alpha_t} - 1 \right) = \alpha_t (\frac{p^*}{x_0^t} - 1) > \alpha_m \left(\frac{e_m - e_0}{\alpha_m} - 1 \right) = \alpha_m (\frac{p^*}{x_0^m} - 1)$$

Dividing by α_t and using that $\alpha_m/\alpha_t > 1$ yields $1/x_0^t > 1/x_0^m$ and hence $x_0^t < x_0^m$, i.e., the concentration of the error tail decreases by a finite step at the critical mutation rate p^*. ∎

The effect of a transition between master species is in general fairly small in large systems, as can be seen in figure 4.

5. Second Order Kinetics

The dynamics become by orders of magnitude more complicated if we allow for linear functions of the variable

$$\mathcal{F}(x) = \alpha + Ax$$

$$\mathcal{G}(x) = \delta + Dx$$

The corresponding mass action kinetics mechanism is

$$\begin{aligned}
I_k &\xrightarrow{\alpha_k} 2I_k \\
I_j + I_k &\xrightarrow{A_{kj}} 2I_k + I_j \\
I_k &\xrightarrow{\delta_k} \emptyset \\
I_j + I_k &\xrightarrow{D_{kj}} I_j \\
I_k &\xrightarrow{\Phi} \textit{(outside)}
\end{aligned}$$

The first two reactions are first and second order replication of I_k, the other two lines give first and second order degradation of I_k. As before, the last line represents the outflux from the system. In the replicator part only the net rates

$$E_{kj} = A_{kj} - D_{kj} \quad \text{and} \quad e_k = \alpha_k - \delta_k$$

play a role, whereas the mutation field depends on the replication reactions alone. This model has been introduced recently by Nuño *et al.* [10,11]. Neglecting specific degradation reactions, i.e., setting $D = 0$ and $\delta_k = \delta$ for all $k = 0, \dots n$ simplifies the mathematics considerably. Non-specific degradation, which affects all viable species and the error-tail in the same way, is equivalent to an increased dilution flux and can formally be removed from the differential equation [19]. We will restrict attention therefore to the following system of differential equations:

$$\begin{aligned} \dot{x} &= x \circ [\alpha + Ax - \Phi(x)] - p \cdot x \circ [\alpha + Ax] \\ \dot{x}_0 &= x_0 [e_0 - \Phi(x) + p \{\langle \alpha, x \rangle + \langle xAx \rangle\} \\ \Phi(x) &= e_0 x_0 + \langle \alpha, x \rangle + \langle xAx \rangle \end{aligned}$$

Next we will investigate the behavior of the system in case of vanishing mutation rate $p = 0$. We will see that in general there are two types of stationary solutions involving a given (sub)set of viable solutions. For $p = 0$ the system reduces to a second order replicator equation

$$\begin{aligned} \dot{x} &= x \circ [\alpha + Ax - \Phi] \\ \dot{x}_0 &= x_0 [e_0 - \Phi]. \end{aligned} \tag{9}$$

These equations have been shown to be topologically equivalent to the well known Lotka-Volterra Equations [6].

An interior rest point on a given face fulfills

$$\hat{x} = A^{-1} [\phi \mathbf{1} - \alpha] \tag{10}$$

where ϕ is the flux at equilibrium and the equation is understood as restricted to the non-zero coordinates. For $x_0 = 0$, i.e., if no error-tail is present in the system, we obtain

$$\langle \mathbf{1}, \hat{x} \rangle = 1 = -\langle \mathbf{1} A^{-1} \alpha \rangle + \phi_+ \langle \mathbf{1} A^{-1} \mathbf{1} \rangle$$

from the normalization of the coordinates and thus

$$\phi_+ = \frac{1 + \langle \mathbf{1} A^{-1} \alpha \rangle}{\langle \mathbf{1} A^{-1} \mathbf{1} \rangle}$$

On the other hand, for $x_0 \neq 0$ we obtain a second solution from equ.(9), namely $\phi_- = e_0$. Substituting this solution into equ.(10), multiplying by $\mathbf{1}$, and using $\langle \mathbf{1} x \rangle = 1 - x_0$ yields

$$x_0 = 1 + \langle \mathbf{1} A^{-1} \alpha \rangle - e_0 \langle \mathbf{1} A^{-1} \mathbf{1} \rangle$$

We will refer to the solutions with ϕ_+ and ϕ_- as the *low mutant* branch and the *high mutant* branch, respectively. The low mutant branch corresponds the pure replicator dynamics of the viable species. In the high mutant branch, however, the error-tail is populated although there is no influx into it. This reflects the fact that the species in the error-tail have a non-zero first order growth rate and thus may well compete with the other species under certain circumstances. The distinction of low and high mutant branches is, of course, not restricted to the second order model. In the general system, however, no explicit calculations are possible. From a physical point of view it seems that the low mutant branch is by far more important.

A third branch of equilibria is given by the trivial solution $x_0 = 1,\ x = 0$. It was termed 0-coexistence state in [10,11].

6. Perturbation Theory for the Second Order Model

Equation (7) simplifies considerably if $\mathcal{F}$ and $\mathcal{G}$ are linear. In order to simplify notation we will use the following characteristic quantities of the system (1-2):

$$\begin{aligned} \varphi &= \frac{1+\langle \mathbf{1}A^{-1}\alpha\rangle}{\langle \mathbf{1}A^{-1}\mathbf{1}\rangle} \\ \psi &= 1+\langle \mathbf{1}A^{-1}\alpha\rangle \\ \vartheta &= e_0\langle \mathbf{1}A^{-1}\mathbf{1}\rangle \end{aligned} \tag{11}$$

For the low mutant branch we obtain $\langle \hat{x}, \mathcal{F}(\hat{x})\rangle = \varphi$, $\langle \mathbf{1}[\partial f(\hat{x})]\mathbf{1}\rangle = \psi$

$$r = -\varphi(1+\frac{1}{\psi-\vartheta})$$

The shifts are now easily obtained from (6) noting that $[\partial f(\hat{x})]^{-1}\mathcal{F}(\hat{x}) = A^{-1}\alpha + \hat{x}$ and substituting for $\hat{x}$

$$\begin{aligned} d &= -\frac{\phi_+}{\phi_+ - e_0}\cdot\frac{A^{-1}\mathbf{1}}{\langle \mathbf{1}A^{-1}\mathbf{1}\rangle} \\ d_0 &= \frac{\phi_+}{\phi_+ - e_0} \end{aligned} \tag{12}$$

where ϕ_+ denotes the flux at fixed point at the low mutant branch

It is instructive to compare this with the exact solutions. As in the previous section we assume $E = A$ and $e = \alpha$. The equation for an interior rest point thus reads

$$\begin{aligned} (\alpha + Ax)(1-p) &= \Phi(x) \\ 0 &= x_0(e_0 - \Phi(x)) + p(\alpha x + \langle xAx\rangle) \end{aligned} \tag{13}$$

The first equation allows us to compute $\Phi(x)$ by solving for x and multiplying by $\mathbf{1}$:

$$\Phi(x) = \frac{1-p}{\langle \mathbf{1}A^{-1}\mathbf{1}\rangle}\left[1+\langle \mathbf{1}A^{-1}\alpha\rangle - x_0\right]$$

The second part of equ.(13) can be rewritten as

$$x_0(e_0 - \Phi(x)) + p\Phi(x) - pe_0x_0 = 0$$

Note that $\varphi = \phi_+$ is the flux at the unperturbed low mutant branch equilibrium. The concentrations of the catalytic species are obtained from equ.(13) as

$$x = (\psi - x_0)\frac{A^{-1}\mathbf{1}}{\langle \mathbf{1}A^{-1}\mathbf{1}\rangle} - A^{-1}\alpha$$

The equilibrium concentration of the error-tail, x_0, fulfills the quadratic equation

$$\hat{x}_0^2 - (\psi - \vartheta + p)\hat{x}_0 + p\psi = 0$$

with the explicit solutions

$$\begin{aligned} \hat{x}_0 &= \frac{\psi - \vartheta + p}{2} \pm \sqrt{\frac{(\psi - \vartheta + p)^2}{4} - p\psi} \\ &= \frac{\psi - \vartheta + p}{2} \cdot \left[1 \pm 1 \mp \frac{2p\psi}{(\psi - \vartheta + p)^2}\right] + \mathcal{O}(p^2) \end{aligned} \tag{14}$$

The signs correspond to the low and high mutant branch, respectively. For the low mutation branch we recover from this

$$\hat{x}_0 = \frac{\psi p}{\psi - \vartheta + p} + \mathcal{O}(p^2) = p \cdot \frac{\psi}{\psi - \vartheta} + \mathcal{O}(p^2)$$

Inserting the explicit expressions we recover, as expected, equ.(12)

$$\hat{x}_0 = p \cdot d_0 = p \cdot \frac{\varphi}{\varphi - e_0} \tag{15}$$

Remark. For second order model with $E = A$, $e = \alpha$ and $e_0 = 0$ the first order perturbation result coincides with the exact solution.

Proof. We observe that the quadratic equation simplifies to

$$\hat{x}_0^2 - (\psi + p)\hat{x}_0 + p\psi = 0$$

which has the solutions

$$\hat{x}_0 = \begin{cases} p \\ \psi \end{cases}$$

Thus we have for the low mutant case

$$\hat{x}_0^{\mathrm{low}} = p \cdot d_0 = p$$

coinciding with the perturbation result. It remains to show that $\hat{x}^{\mathrm{low}}(p) = \hat{x}^{\mathrm{low}} + pd$ also holds exactly. The explicit formula (with $e_0 = 0$) reads

$$\hat{x}^{\mathrm{low}}(p) = (1 + \langle \mathbf{1} A^{-1} \alpha \rangle - p) \frac{A^{-1}\mathbf{1}}{\langle \mathbf{1} A^{-1} \mathbf{1} \rangle} - A^{-1}\alpha$$

Using $1 + \langle \mathbf{1} A^{-1} \alpha \rangle = \varphi \langle \mathbf{1} A^{-1} \mathbf{1} \rangle$ one easily verifies

$$\hat{x}^{\mathrm{low}}(p) = \hat{x}^{\mathrm{low}} - p \cdot \frac{A^{-1}\mathbf{1}}{\langle \mathbf{1} A^{-1} \mathbf{1} \rangle}$$

∎

While an analysis of other branches than the low mutant branch seems to be very difficult for general models, some interesting expressions can be derived for second order kinetics. The exact solution is given by the upper sign in equation (14). A first order approximation for small p reads

$$\hat{x}_0^{\mathrm{high}}(p) = \psi - \vartheta + p - \frac{p}{\psi - \vartheta} + \mathcal{O}(p^2)$$

This can be written in terms of the interior rest point of the catalytic subsystem and the replication rate of the error-tail as

$$\hat{x}_0^{\mathrm{high}}(p) = \langle \mathbf{1} A^{-1} \mathbf{1} \rangle (\varphi - e_0) - p \cdot \frac{e_0}{\varphi - e_0} + \mathcal{O}(p^2)$$

We note that, with $e_0 = 0$, we have independent of p

$$\hat{x}_0^{\mathrm{high}}(p) = \langle \mathbf{1} A^{-1} \mathbf{1} \rangle) \varphi = 1 + \langle \mathbf{1} A^{-1} \alpha \rangle$$

7. Special Systems

In this section we will study three special models under the hypotheses stated in section 2.

7.1 Schlögl's model

The Schlögl's model of independently replicating units in its most simple form has the parameters

$$A = I \qquad \text{and} \qquad \alpha = 0,$$

I being the identity matrix. One obtains immediately from equ.(11)

$$\varphi = 1/n^2, \qquad \psi = 1, \qquad \vartheta = e_0 \cdot n^2$$

Let us first calculate the Jacobian matrix at a rest point $\hat{x}$. We assume that the first m coordinates of $\hat{x}$ are 0 and the remaining viable species are present with non-zero frequencies. We will make no assumption on the concentration of error-tail species, i.e., we will not fix the branch at the moment. A simple calculation shows that $J(\hat{x})$ is of the following form:

$$J(\hat{x}) = \begin{pmatrix} A & \emptyset & 0 \\ Y & B & s \\ 0 & r & a \end{pmatrix}$$

where A is a diagonal matrix containing the transversal eigenvectors

$$\lambda^{\text{trans}} = -\Phi(\hat{x})$$

The entries in B are of the form

$$B_{kl} = \hat{x}_k[(1-p)\delta_{kl} - 2\hat{x}_l] = -\frac{2\Phi^2(\hat{x})}{(1-p)^2} + \Phi\delta_{kl}$$

The vectors r and s defined by $r_l = J_{0l}$ and $s_k = J_{k0}$ are given by

$$r = (p - x_0)\frac{\Phi}{1-p}\mathbf{1} \qquad s = -e_0\frac{\Phi}{1-p}\mathbf{1}$$

and finally

$$a = J_{00} = (1 - x_0)e_0 - \Phi$$

Lemma 1. All transversal eigenvalues corresponding to the viable species are non-positive. If $e_0 > 0$, then they are strictly negative.

Proof. The flux Φ is always non-negative. It is strictly positive if $e_0 > 0$. ∎

We may thus restrict attention to the in-plane part $\begin{pmatrix} B & s \\ r & a \end{pmatrix}$ of the Jacobian. Before we proceed with its analysis we will prove the following

Lemma 2. The characteristic polynomial of a $(n+1) \times (n+1)$ matrix of the form

$$Q_{n+1} = \begin{pmatrix} a & 0 & 0 & \dots & 0 & c \\ 0 & a & 0 & \dots & 0 & c \\ 0 & 0 & a & \dots & 0 & c \\ & & & \vdots & & \\ 0 & 0 & 0 & \dots & a & c \\ b & b & b & \dots & b & d \end{pmatrix}$$

is given by

$$\wp(\lambda) = (a-\lambda)^{n-1}[(a-\lambda)(d-\lambda) - n \cdot bc]$$

Proof. For simplicity of notation we write a and d instead of $a-\lambda$ and $d-\lambda$. Expanding with respect to the first row yields

$$\det Q_{n+1} = a \det Q_n + (-1)^{n+1} c \det \begin{pmatrix} 0 & a & 0 & \dots & 0 \\ 0 & 0 & a & \dots & 0 \\ & & & \vdots & \\ 0 & 0 & 0 & \dots & a \\ b & b & b & \dots & b \end{pmatrix}$$

Let us call this last matrix W_n. Expanding W_n with respect to the first column leaves us with $(-1)^n b \cdot \det[aI]$ where $[aI]$ is an $(n-1) \times (n-1)$ diagonal matrix. Thus

$$\det Q_{n+1} = a \det Q_n - bc \cdot a^{n-1}$$

Repeating this expansion $n-2$ times yields

$$\det Q_{n+1} = a^{n-1}[\det Q_2 - (n-1)bc]$$

Finally, $\det Q_2 = \det \begin{pmatrix} a & c \\ b & d \end{pmatrix} = ad - bc$ which completes the proof. ∎

Lemma 3. All rest points containing more than one viable species are unstable.

Proof. First we observe that a physically meaningful eigenvector η of J must fulfill $\langle \mathbf{1}, \eta \rangle = 0$ because it do not point out of the simplex, and can hence be represented as the difference of two points in S_n. Let $\mathcal{I}$ denote the matrix will all entries 1. We have $\mathcal{I}\eta = 0$, therefore we can add $c\mathcal{I}$ to Jacobian without affecting the physically meaningful eigenvalues.

The in-plane part of the Jacobian is thus equivalent to the following matrix, which is of the form of the matrix in lemma 2 with the abbreviations

$$\begin{aligned} a &= \Phi \\ b &= (p - x_0)\frac{\Phi}{1-p} + \frac{2\Phi^2}{(1-p)^2} \\ c &= -e_0\frac{\Phi}{1-p} + \frac{2\Phi^2}{(1-p)^2} \\ d &= (1 - x_0)e_0 - \Phi + \frac{2\Phi^2}{(1-p)^2} \end{aligned}$$

Lemma 2 implies that we have the $(n-1)$-fold eigenvalue $+\Phi$, indicating that all equilibria with more than one viable species are unstable for all values of p. ∎

Theorem 6. For the symmetric Schlögl-model with $0 < e_0 < 1$ there are $n + 1$ stable rest points corresponding to the n viable species and the pure error-tail if $p < p^{\rm cr}$ and a single globally stable rest point, the pure error-tail if $p > p^{\rm cr}$. All rest points containing at least two and lacking at least one viable species are saddle points. The rest points containing all viable species are a saddle on the low mutant branch and a source on the high mutant branch.

Proof. Lemmas 1 and 3 imply that all equilibria with more than one viable species and at least one transversal direction are saddle points. For the rest points with all species present we know that they have at least $n-1$ positive eigenvalues, and hence one of them is a source and the other one a saddle with a single stable direction. Continuity ensures that the low mutant branch of the interior equilibrium has a stable direction since the flux on the boundary points inwards.

It remains to consider the subsystems consisting of a single viable species and the error-tail. Using $z = 1 - x_0$ we have to analyze the one-dimensional ODE

$$\dot{z} = z[(1-p)z - z^2 - (1-z)e_0] = h(z)$$

on the unit interval. Clearly the 0-coexistence state $z = 0$ is a rest point. Its stability is determined by

$$\frac{\partial h}{\partial z}(0) = -e_0$$

Therefore the pure error-tail is stable. On the other hand we find $h(1) = -p$, i.e., the flow points inwards for $p > 0$. For $p = 0$ we find that $z = 1$ is a rest point and its stability is given by the transversal eigenvalue $\lambda_0^{\text{trans}} = e_0 - 1$, thus it is stable for $e_0 < 1$ and unstable for $e_0 > 1$. Only the first case — error-tail replicates slower than the viable species — is in accordance with the specification of our model. It follows directly from continuity that only two generic phase portraits are possible:

$$\longrightarrow \bullet \longleftarrow \circ \longrightarrow \bullet \qquad \text{and} \qquad \xrightarrow{\hspace{6em}}\bullet$$

(They are separated by a phase portrait containing a single degenerate equilibrium in the interior of the unit interval corresponding to a saddle-node bifurcation.) The first case exhibits two stable solutions ($\bullet$), the pure error-tail and the low mutant branch. In the second phase portrait only the pure error-tail is stable. Whether there are two or no rest points in the interior of the unit interval is governed by $z^2 - (1 + e_0 - p)z + e_0$. The discriminant of this quadratic equation vanishes for

$$p^{\text{cr}} = 1 + e_0 - 2\sqrt{e_0}$$

where a saddle-node bifurcation is observed. It is easy to see that for any choice $0 < e_0 < 1$ the critical mutation rate is also between 0 and 1. For very small p we know (either from direct calculation or by applying the rest point migration theorem) that the low mutant branch is stable. ∎

7.2 Uniform Cooperation Model

The uniform cooperation model

$$\alpha_k = \alpha \qquad A_{kl} = A(1-\delta_{kl})$$

has been studied extensively by Nuño *et al.* [10,11]. We will keep this subsection short and omit formal proofs because most of the results summarized here are immediate consequences of section 2.1 of reference [11]. For completeness we give the explicit form of the characteristic parameters (11):

$$\varphi = \frac{n-1}{n}A + \alpha; \qquad \psi = 1 + \frac{n}{n-1}\frac{\alpha}{A}; \qquad \vartheta = \frac{n}{n-1}\frac{e_0}{A}$$

These expressions are easily derived from the following

Lemma 4. Let $\mathcal{I}$ denote the matrix with all entries 1. Then $\langle \mathbf{1}(\mathcal{I}-I)^{-1}\mathbf{1}\rangle = \frac{n}{n-1}$.

Proof. With the ansatz $(\mathcal{I}-I)^{-1} = a\mathcal{I} + bI$ and $\mathcal{I}^2 = n\mathcal{I}$ we obtain

$$(\mathcal{I}-I)(a\mathcal{I}+bI) = an\mathcal{I} + b\mathcal{I} - a\mathcal{I} + bI = I$$

and form this $b=-1$ and $a = 1/(n-1)$. The proof is completed by observing that $\langle \mathbf{1}\mathcal{I}\mathbf{1}\rangle = n^2$ and $\langle \mathbf{1}I\mathbf{1}\rangle = n$. ∎

Figure 5: Bifurcation diagrams for Schlögl's model with mutation. The simplified mutation model (1-2) **(a)** is compared with the results [19] for Eigen's Q-matrix **(b)**.

While the symmetric model undergoes a cascade of bifurcations eventually leading to a single globally stable rest point in the center of the simplex in case of Eigen's mutation matrix, our model exhibits a single bifurcation leaving the pure error-tail as unique globally stable rest point in the system. Of course, the interpretation is consistent: At very high mutation rates there is a uniform distribution of genes in the systems, i.e., no organization. This may be viewed as the analogue to the error-threshold: below a critical mutation rate there are quite a lot of stable equilibria (species) and beyond the threshold p^{cr} there is only random replication.

Full lines denote stable rest points, dashed lines refer to saddle points. The only source in the system is the interior rest point for low mutation rates in Eigen's model and the high mutant branch of the interior rest point in case of our model. They are represented by dotted lines.

As for Schlögl's system, the uniform model may exhibit up to $2^{n+1}-1$ rest points. Only the pure error-tail and the full coexistence state, however, can be stable. The most interesting feature of this model is, that for an appropriate relation of α and e_0 there is bistability between the pure error-tail and the full coexistence state in case of error-free replication and (by virtue of the rest point migration theorem) also for sufficiently small error-rate p. The actual condition for $p=0$ reads

$$\alpha < e_0 < \varphi \tag{16}$$

If $\alpha > e_0$ then the low mutant branch becomes globally stable while the pure error-tail becomes unstable. At $\alpha = e_0$ we have a degenerate transcritical bifurcation. A non-degenerate transcritical bifurcation occurs for $e_0 = \varphi$. Beyond this point the error-tail is globally stable.

In case of error-prone replication we encounter in general two critical mutation rates: at

$$p_{\mathrm{tc}} = \frac{\alpha - e_0}{\alpha}$$

a transcritical bifurcation occurs. For

$$p_{\mathrm{sn}} = \vartheta\left[\frac{\varphi + e_0 - 2\sqrt{\varphi e_0}}{e_0}\right]$$

a saddle node bifurcation occurs. The bifurcation point lies in the physical part of the state space, if additionally

$$e_0 > \alpha A \frac{\varphi - 1}{\varphi} \tag{17}$$

is fulfilled.

If $p > p_{\mathrm{sn}}$ then the error-tail is globally stable. If the condition (17) is not fulfilled, then the error-tail is globally stable beyond p_{tc}. If (16) holds then there is bistability between the full coexistence state and the pure error-tail for $p < p_{\mathrm{sn}}$. On the other hand, if (17) holds, the region of bistability is the interval $p_{\mathrm{tc}} < p < p_{\mathrm{sn}}$ and for still smaller values of p the full coexistence state is globally stable.

7.3 The Hypercycle

The hypercycle [3] equation with mutation reads

$$\dot{x}_k = x_k[(1-p)x_{k-1} - \Phi]$$
$$\dot{x}_0 = x_0[e_0 - \Phi] + p\sum_{i=1}^{n} x_i x_{i-1}$$

Lemma 5. Let H be defined by $H_{kl} = \delta_{k-1,l}$. Then $\langle \mathbf{1}H^{-1}\mathbf{1}\rangle = n$.

Proof. The inverse of H is given by $K_{jk} = \delta_{j+1,k}$ since

$$\sum_k K_{jk}H_{kl} = \sum_k \delta_{j+1,k}\delta_{k-1,l} = \sum_k \delta_{j+1,k}\delta_{k,l+1} = \delta_{jl}$$

and clearly $\langle \mathbf{1}K\mathbf{1}\rangle = n$. ∎

The characteristic parameters are therefore given by

$$\varphi = \frac{1}{n}; \qquad \psi = 1; \qquad \vartheta = ne_0$$

Lemma 6. There are two interior rest points provided $e_0 < 1/n$ and

$$p < p^{\mathrm{cr}} = 1 + ne_0 - 2\sqrt{ne_0}$$

At the critical value p^{cr} of the mutation rate we have a saddle-node bifurcation. If $e_0 \geq 1/n$ then there is no interior rest point.

Proof. As a consequence of equ.(14) there is no interior rest point for any $p > 0$ if $e_0 \geq 1/n$, and the interior rest point of the unperturbed hypercycle is stable against the error-tail if $e_0 < \Phi$, i.e., if $e_0 < 1/n$. With the abbreviation $a = \psi - \theta = 1 - ne_0$ equ.(14) becomes

$$\hat{x}_0(p) = \frac{a+p}{2} \pm \sqrt{\frac{(a+p)^2}{4} - p}$$

A saddle node bifurcation occurs when the square root above vanishes. Some simple algebra shows that this happens for $p = p^{\mathrm{cr}}$. ∎

Lemma 7. The symmetry line

$$\sigma = \{\mathbf{x} | x_i = (1 - x_0)/n\}$$

is invariant for all p in the hypercycle model.

Proof. Substituting $x_i = (1 - x_0)/n$ into the perturbed hypercycle equation yields

$$\dot{x}_i = (1 - x_0/n)[(1 - p)(1 - x_0/n) - \Phi]$$

independent of i. Thus the ratio of any two x_i, x_j remains constant over time. ∎

Lemma 8. The high mutant branch of the interior rest points of the hypercycle model is unstable whenever its exists.

Proof. The dynamics along σ is governed by the ODE

$$\dot{x}_0 = \frac{1}{n} x_0 (1 - x_0)(x_0^2 - (1 - ne_0 + p)x_0 + p)$$

which has a stable equilibrium for $x_0 = 1$ corresponding to the zero coexistence state and for $p > 0$ we have $\dot{x}_0(0) > 0$. Therefore there are either no interior rest-points or two interior rest points corresponding to low and high mutant branch of the interior equilibrium of the hypercube. As an immediate consequence of continuity the high mutant branch is always an unstable rest point. ∎

Theorem 7. The low mutant branch of the interior rest points of the hypercycle model has the following stability properties as a function of the number of viable species if it exists, i.e., if $p < p^{\mathrm{cr}}$:

(i) For $n \leq 3$ the rest point is stable.

(ii) For $n = 4$ it is marginally stable.

(iii) For $n \geq 5$ it is unstable.

The pure error-tail is always stable, and all other rest points are always unstable.

Proof. The Jacobian of the low mutant branch of the interior rest point is given by

$$J(\hat{x}) = \frac{1-x_0}{n}\begin{pmatrix}(1-p)Z & r\cdot\mathbf{1}\\ s\cdot\mathbf{1} & t\end{pmatrix} + 2\left(\frac{1-x_0}{n}\right)^2 \mathcal{I}$$

where Z denotes the circulant matrix $Z_{kl} = \delta_{k-1,l}$. Because of the invariance of the simplex we are interested only in eigenvalues corresponding to eigenvectors fulfilling $\langle\eta, \mathbf{1}\rangle = 0$. Thus we may omit $\mathcal{I}$. Some algebra shows that the characteristic polynomial of the interesting part reads

$$\begin{aligned}\wp(\lambda) &= \det\begin{pmatrix}(1-p)Z & r\cdot\mathbf{1}\\ s\cdot\mathbf{1} & t\end{pmatrix} = \\ &= \det\begin{pmatrix}(1-p)-\lambda & r\\ s & t-\lambda\end{pmatrix}\cdot\sum_{j=0}^{n-1}(1-p)^{n-j-1}\lambda^j\end{aligned}$$

The eigenvalues of the 2×2 matrix correspond to the non-physical external direction and to the symmetry line σ, see lemma 8. The characteristic polynomial corresponding to the remaining directions, i.e., the sum in the above equation, can be simplified using the new variables $\xi = \lambda(1-p)$. We find then

$$\wp^\star(\xi) = (1-p)^{n-1}\sum_{j=0}^{n-1}\xi^j$$

The solutions are the $n-1$ roots of unity, $\xi_\ell = \exp(\frac{2\pi\ell}{n})$, with $1 \le \ell \le n-1$. The first part of theorem follows now immediately.

The 0-coexistence state is always stable since the tranversal eigenvalues are $\lambda_k^{\mathrm{trans}} = -\Phi = -e_0$.

It remains to consider the rest points on the boundary of the simplex which contain at least one viable species. The unperturbed hypercycle equation exhibits a complicated geometry of faces consisting entirely of rest points [15]. Suppose $\hat{y}$ is a rest point on the boundary, i.e., at least one coordinate, say y_1 is zero. Then from

$$\dot{y}_2 = 0 = y_2[(1-p)y_1 - \Phi]$$

implies that $\Phi = 0$ and furthermore $y_i y_{i+1} = 0$ for all i. Furthermore from $e_0 > 0$ we find $y_0 = 0$ as well This leads to the same geometrical arrangement of fixed point faces as for the unperturbed hypercycle. All these rest points are non-regular and have at least one positive transversal eigenvalue. The proofs for these statements parallel those for the corresponding result for the unperturbed hypercycle in [18] and are therefore omitted here.

∎

Theorem 7 implies that from the physical point of view the only relevant rest points are the interior rest points and the 0-coexistence state $x_0 = 1$. The eigenvalues of the interior rest points perpendicular to the symmetry line σ are (up to a scaling factor) the same as for the unperturbed hypercycle. Hence there is no Hopf-bifurcation as p increases. It is well known that for $n \geq 5$ there are stable limit cycles [3]. A complete proof of their existence has be given only recently [8]. In [18] it has been shown that an analogue to the rest point migration theorem also hold for limit cycles. Therefore we expect for sufficiently small p a stable limit cycle. For lack of other bifurcations involving the interior rest point it is sensible to expect the typical limit cycle of the hypercycle for $p \lesssim p^{\mathrm{cr}}$. Figure 6 shows numerical experiments for $n = 5$. The limit cycle vanishes for $p^\star \approx 0.0857$ with $e_0 = 0.1$; the analytic estimate $p^{\mathrm{cr}} = 0.085786\ldots$ agrees very well with the numerical result. The stable limit cycle vanishes at this mutation rate in a saddle-node like bifurcation colliding with the unstable limit cycle surrounding the high mutant branch rest point.

a) b)

c) d)

Figure 6: The five-membered hypercycle. A projection onto the simplex consisting of 3 subsequent viable species (bottom face) and the error-tail (top) is shown for **a)** p=0.0, **b)** p=0.08, **c)** p=0.0857, **d)** p=0.1,

8. Conclusions

For a description of evolutionary processes in biology mutation cannot be neglected in most cases. Despite of the purely stochastic nature of mutation, deterministic models, as the one presented here, may be used to investigate, at least qualitatively, the influence of mutation super-imposed on a selection mechanism.

However, studying the role of mutation for the behavior of self-replicating systems is a very difficult task if all parts of the system are strongly interacting. Recently, an attempt to include mutations into catalytic networks formed by completely separated species has been presented [10]. Even using this simplification, only particular situations can be treated analytically in full detail [11]. In this contribution, we show that at least in a simplified model the influence of mutation can be understood without any further hypotheses. In general, there is an *error-threshold-like* behavior also in systems with strong cross-catalysis, and not simply in models close to the classical quasispecies. Furthermore, if the replication rate of the error-tail species is much smaller than the replication rate of the viable species, then catalytic networks are stable even for fairly large mutation rates.

Our simplified version of the quasispecies model exhibits transitions between different master species as a function of the mutation rate solely as a consequence of the different influence of replication and degradation rate. The concentration of the error-tail increases linearly between the transitions but decreases at the transition points. Although these transitions are reminiscent a similar phenomena in the classical quasispecies model, they rely on an entirely different mechanism: they do not depend on the fitness of close-by neighbors in sequence space, as it has been observed by Swetina and Schuster [21].

A comparison of our description of mutation with models based on Eigen's Q-matrix highlights that dynamics qualitatively simplify with increasing mutation rate [13,14,18,19] in both models. In the framework of our model, the error-tail becomes globally stable for sufficiently high mutation rates, while models based on the Q-matrix exhibit a globally stable rest-point close to the center of the simplex in this case.

Acknowledgements

Computational assistance by Christian Forst is greatfully acknowledged. The numerical experiments have been performed the IBM3090 mainframe of the Computer Center of the University of Vienna, supported by the IBM EASI initiative. J.C.N. thanks the *Universidad Politécnica de Madrid* and the *Comisión Asesora de Investigación Científica y Técnica*, projects No. PB88-0140 and PB089-0108 for the financial support he received to visit the Institut für Theoretische Chemie of the University of Vienna whose hospitality during his stay is gratefully acknowledged. P.F.S. thanks the German *Max Planck Gesellschaft* for supporting his stay at the Santa Fe Institute.

References

[1] Eigen M. Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften* **58**, 465-523 (1971).

[2] Eigen M., J. McCaskill and P. Schuster. The Molecular Quasispecies. *Adv. Chem. Phys.* **75** 149-263 (1989).

[3] Eigen M. and P. Schuster. *The Hypercycle.* Springer-Verlag: New York, Berlin 1979 and *Naturwissenschaften* **64** 541-565 (1977), **65**, 7-41 and 341-369 (1978).

[4] Eigen M., P. Schuster, K. Sigmund and R. Wolff. Elementary Step Dynamics of Catalytic Hypercycles. *BioSystems* **13**, 1-22 (1980).

[5] Hofbauer J. The Selection Mutation Equation. *J. Math. Biol.* **23**, 41-53 (1985).

[6] Hofbauer J. On the Occurrence of Limit Cycles in Lotka-Volterra Equations. *Nonlinear Analysis*, **5** 1003-1007 (1981).

[7] Hofbauer J., K. Sigmund. *Dynamical Systems and the Theory of Evolution.* Cambridge Univ. Press, Cambridge, U.K. (1988).

[8] Hofbauer J., J. Maller-Paret and H.L. Smith. Stable Periodic Solutions for the Hypercycle System. *J. Dynam. Diff. Eqns.* **3**, 423-436 (1991).

[9] Nowak M. and P. Schuster. Error Threshold in Finite Populations: Mutation Frequencies and the Onset of Muller's Ratchet. *J. Theor. Biol.* **137**, 375-395 (1989)

[10] Nuño J.C., M.A. Andrade, F. Morán and F. Montero. A model of an autocatalytic network formed by error-prone self-replicative species. *Bull. Math. Biol.* **55**, 385-415 (1993).

[11] Nuño J.C., M.A. Andrade and F. Montero. Non-Uniformities and Superimposed Competition in a Model of an autocatalytic network formed by Error-Prone Selfreplicative Species *Bull. Math. Biol.* **55**, 417-449 (1993).

[12] Rumschitzky D.S. Spectral Properties of Eigen Evolution Matrices, *J.Math.Biol.* **24** 667-680 (1987).

[13] Schnabl W., P.F. Stadler, C. Forst and P. Schuster. Full Characterization of a Strange Attractor. Chaotic Dynamics in Low-Dimensional Replicator Systems. *Physica D* **48**, 65-90 (1991).

[14] Andrade, M.A., J.C. Nuño, F. Morán, F. Montero and G. Mpitsos. Complex dynamics of a catalytic network having faulty replication into error-species. *Physica D* **63**, 21-40 (1993)

[15] Schuster P., K. Sigmund and R. Wolff. Dynamical Systems under Constant Organization I: Topological Analysis of a Family of Nonlinear Differential Equations — A model for Catalytic Hypercycles. *Bull. Math. Biol.* **40**, 743-769 (1978).

[16] Schuster P., K. Sigmund, R. Wolff. Mass Action Kinetics of Selfreplication in Flow Reactors. *J. Math. Anal. Appl.* **78**, 88-112 (1980).

[17] Schuster P., K. Sigmund. Replicator Dynamics. *J. Theor. Biol.* **100**, 533-538 (1983).

[18] Stadler P.F. and P. Schuster. The Selection Mutation Equation. An Approach Based on Perturbation Theory. *J. Math. Biol.* **30** 597-632 (1992).

[19] Stadler P., W. Schnabl, C. Forst and P. Schuster. Small Autocatalytic Reaction Networks II: Analytically Treatable Special Cases. (in preparation)

[20] Stadler P.F. Dynamics of Autocatalytic Reaction Networks IV: Inhomogeneous Replicator Networks. *BioSystems* **26**, 1-19 (1991)

[21] Swetina J. and P. Schuster. Self Replication with Errors: A Model for Polynucleotide Replication. *Biophys.Chem.*, **16**, 329-345 (1982).