Complementary Replication

By

Peter F. Stadler

Institut für Theoretische Chemie der Universität Wien

Current Mailing Address: Max Planck Institut für Biophysikalische Chemie
080 Biochemische Kinetik

Am Fassberg D-3400 Göttingen, FRG

Phone: *(0551) 201 / 366 D
Fax: *(0551) 201 / 435

Bitnet: PSTADLE@DGOGWDG1 or A8443GAD@AWIUNI11

Abstract

Differential equations for the kinetics of complementary replicating macromolecules in a flow reactor are derived. It is shown that such a model has many features in common with the differential equation for direct replication, the replicator equation. Two special cases of replication, and the influence of mutation on them, have been studied in detail: In case of first order mass action kinetics – the quasispecies model – complementary replication, like direct replication, exhibits an error threshold for the replication accurary, below which the genetic information is lost. In turns out that the long time behaviour of many special cases of the second order kinetics model can be described in terms of second order replicator equations, although this is not possible in general.

1. Introduction

The replication of polynucleotides, RNA and DNA, consists in a formation of *complementary* strands:

$$I^{\ominus} \longrightarrow I^{\ominus} + I^{\oplus} \quad \text{and} \quad I^{\oplus} \longrightarrow I^{\oplus} + I^{\ominus}. \tag{C}$$

On the other hand the most common theoretical models concerning the dynamics of replicating macromolecules neglect this fact. They are based on a *direct* replication process of the form

$$I \longrightarrow 2I. \tag{D}$$

This seems to be appropriate for minimal models of *in vivo DNA*-replication, which proceeds by *simultanous* leading-strand and lagging-strand synthesis from origins on double-stranded molecules. One may in this case consider the entire *DNA*-double helix as a single self-replicating entity I.

The continuous time formulation of mechanism D leads to well known systems of ordinary differential equations. In the most simple case one obtains the standard model of mathematical ecology, the *Lotka-Volterra* equation

$$\dot{y}_k = y_k(r_k + \sum_{j=1}^{n} b_{kj} y_j). \tag{1.1}$$

The variables y_k denote the population size of species I_k, r_k is its growth rate and the coefficients b_{kj} denote the interaction between different species (b_{kk} represents a self-interaction) [15,24]. Since y_k has to be nonnegative, this ODE is defined on the nonnegative orthant

$$\mathbb{R}^n_+ = \{y \in R^n | y_k \geq 0\}.$$

If one requires the total concentration (population size)

$$\sum_{j=1}^{n} y_j = c$$

to be a constant – this constraint has been called *constant organization* – one obtains the well known *replicator equation* [7,8,17].

$$\dot{x}_k = x_k[f_k(x) - \sum_{j=1}^{n} x_j f_j(x)]. \tag{1.2}$$

The variables x_k are relative abundances and thus the differential equation is defined on the *simplex*

$$S_n = \{x \in \mathbb{R}_+^n | \sum_{j=1}^{n} x_j = 1\}.$$

There is a flood of results on this equation; see e.g. [14] and the references therein. Hofbauer [12] showed that equ.(1.1) is topologically equivalent to the following important special case of equ.(1.2)

$$\dot{x}_k = x_k([Ax]_k - xAx) \tag{1.3}$$

by means of the simple transformation

$$\mathcal{H}: \ S_n \ \longrightarrow \ \mathbb{R}_+^n : y_k = x_k/x_n, \quad k = 1, \ldots, n-1. \tag{1.4}$$

A process of the form

$$mA \ \longrightarrow \ (m+1)A$$

is called a m^{th} order autocatalytic reaction in chemistry for its dynamics is described by the differential equation

$$\frac{d[A]}{dt} \propto [A]^m. \tag{1.5}$$

Consequently, equ.(1.3) is called *second order* replicator equation. In mathematical literature this equation is sometimes called “first order” for the function $f(x)$ in (1.2) accounting for the particular form of the kinetics is linear in this case. In this paper the more chemically motivated nomenclature will be used. Thus “first order” refers to the simplest replicator equation:

$$\dot{x}_k = x_k(f_k - \sum_{j=1}^{n} f_j x_j). \tag{1.6}$$

The replication of *RNA* proceeds by synthesis of a complementary strand from single-stranded templates. It is clear therefore that mechanism (D) can only be a very crude approximation of the actual kinetics. For *in vitro* evolution experiments with *RNA* in an environment containing *Qβ-RNA*-replicase [1,2] mechanism (C) however leads to a reasonable description. It also allows for a more accurate description of phage-host systems *in vivo* if the phage, like Qβ, has a single stranded genome which is not replicated via a double stranded intermediate. Nevertheless, it must remain a minimal model since it does not account for any internal structures of the host cell. These two systems are appropriately described by a first order model.

Some attempts have been made to account for a complementary replication mechanism in the first order model with mutation, the so called *quasispecies equation* by Eigen, McCaskill, Schuster and Swetina [7,10,11,19,23]. Their theory, however, is based on the *a priori* assumption that one can consider a pair of complementary physical entities $(I_k^\oplus, I_k^\ominus)$ as single species (I_k), and that the reaction constants for the replicator-type dynamics of these species can be obtained as the geometric mean of the properties of the underlying unit $(I_k^\oplus, I_k^\ominus)$. They, in fact, obtained very good coincidence of analytic expressions obtained from this model and numerical calculations. It will be shown in this paper that for their choice of the rate constants their expressions yield exact results although their formulas are not correct in general. A computer model accounting for the detailed mechanism of the $Q\beta$-system [3] shows that one may, indeed, not neglect the asymmetry of plus- and minus-strands in general.

It is the purpose of this paper to develop a general theory for a model analogous to equ.(1.2) based on mechanism (C) instead of mechanism (D). Both mechanisms are *catalytic* – the template strands acts a catalyst for the replication process because it remains unchanged. However, there is one major difference: mechanism (D) is *autocatalytic* but the complementary mechanism (C) is not.

For higher order models no experimental systems are available at present, but the replicator models suffer from the same lack of experimental support at least in molecular biology. However, they are of interest in the context of prebiotic *RNA*-world scenarios [4,6]. In sect.5 the analogue to the second order replicator equation is investigated in some detail. It describes a system consisting of *RNA* which is not replicated by some protein enzyme but by a ribozyme-*RNA*-polymerase capable of copying also itself. This system may be viewed as a generalization of the hypercycle [8] to complementary instead of directly replicating entities which are arbitrarily coupled with each other.

The models presented in this paper are an appropriate decription of all kinds of selfreplicating RNA; they are not sufficient, however, to account for the production of enzymes by a gene I since the differential equations contain polynucleotide concentrations as the only variables.

2. Direct Versus Complementary Replication

Let us now derive the differential equations for complementary replication in a flow reactor. There are two independent processes: replication and an unspecific dilution flux which keeps the total concentration constant; the latter yields a contribution $-x_k\Phi$ [8]. We denote the relative concentration of $I_k^\oplus$ and $I_k^\ominus$ by x_k and y_k. The rate of formation of $I_k^\oplus$ has to be proportional to the concentration of its "parent" $I_k^\ominus$ and *vice versa* in the absence of the dilution flux. Thus

$$\dot{x}_k \propto y_k \quad \text{and} \quad \dot{y}_k \propto x_k. \tag{2.1}$$

The replication rate may depend on the medium, i.e. on the concentration of all species present in the reaction vessel. Taking into account the dilution flux mentioned above, we obtain

$$\begin{aligned} \dot{x}_k &= y_k \Psi_k(x,y) - x_k \Phi \\ \dot{y}_k &= x_k \Theta_k(x,y) - y_k \Phi \end{aligned} \tag{2.2a}$$

Since the total concentration shall be constant, we normalize our variables such that

$$\sum_{j=1}^{n} x_j + \sum_{j=1}^{n} y_j = 1. \tag{2.3}$$

Therefore $\sum_{j=1} (\dot{y}_j + \dot{x}_j) = 0$ and we can calculate the dilution flux Φ:

$$\Phi(x,y) = \sum_{j=1}^{n} \big(x_j \Psi_j(x,y) + y_j \Theta_j(x,y) \big). \tag{2.2b}$$

The vector fields Ψ and Θ have to fulfil certain constraints since we assume that a strand *is* produced if its concentration is very low and its complementary strand is present in the reaction vessel. This implies

$$\begin{aligned} x_k = 0 \quad &\text{and} \quad y_k \neq 0 \Longrightarrow \dot{x}_k > 0 \\ y_k = 0 \quad &\text{and} \quad x_k \neq 0 \Longrightarrow \dot{y}_k > 0 \end{aligned} \tag{2.4}$$

The discrimination between *plus* and *minus* strands is simple as long as there are no mutations, i.e. if there are no reactions of the form

$$I_k^{(\alpha)} \longrightarrow I_k^{(\alpha)} + I_m^{(\beta)} \qquad \alpha, \beta \in \{\oplus, \ominus\} \tag{M}$$

Without such a reaction there are n strictly separated pairs. Reactions of type (M) however link any two strands with each other. Therefore it is necessary to find a consistent way of assigning names to the different macromolecules. In general there is no unambiguous or obvious way to do so, as will be shown at the end of this section.

For simplicity we assume that mutations are mistakes in the copying process which occur with a certain probability. Thus we neglect mutations caused by subsequent chemical changes of a nucleotide. The probability of a reaction of type (M) will be denoted by $p_{mk}^{(\beta\alpha)}$. These probabilities may then be assembled in a *mutation matrix* P which consists of four parts:

- P^{++} contains the probabilities p_{jk}^{++} of producing the plus-strand $I_j^{\oplus}$ from the plus-strand $I_k^{\oplus}$.
- P^{--} contains the probabilities of forming a minus-strand from another minus-strand.
- P^{+-} represents the probabilites of producing a plus-strand from minus-strand and
- P^{-+} contains the probabilities of copying a minus-strand into a plus-strand.

Note that the elements p_{kk}^{+-} and p_{kk}^{-+} contain the probability for *correct complementary* replication of $I_k^{\oplus}$ and $I_k^{\ominus}$, whereas p_{kk}^{++} and p_{kk}^{--} denote the probability of *correct direct* replication of $I_k^{\oplus}$ and $I_k^{\ominus}$. The above four matrices have to fulfil

$$\sum_{k=1}^{n} (p_{kj}^{+-} + p_{kj}^{--}) = \sum_{k=1}^{n} (p_{kj}^{-+} + p_{kj}^{++}) = 1. \tag{2.5}$$

Figure 1: m-error neighborhoods of complementary strands with large Hamming distance (**a**) and of a palindromic sequence (**b**).

Replication of RNA. RNA is replicated in complementary strands with *opposite direction*. We illustrate this in the following example. Suppose the replication of the *plus* strand

AUGGUACCUGGAACCUGGAAAA.

This polynucleotide and its minus-strand may be aligned in the following way:

$$\begin{array}{ccc} 5' & \mathbf{AUGGUACCUGGAACCUGGAAAA} & 3' \\ 3' & \mathbf{UACCAUGGACCUUGGACCUUUU} & 5' \end{array}$$

Thus the *minus* strand, written in the usual $5' \to 3'$ direction, reads

UUUUCCAGGUUCCAGGUACCAU.

Of course there a some RNA molecules which are both *plus* and *minus* strand; for example the polynucleotide

AAACCCUUUCCCGGGAAAGGGUUU

is of this type. We will call such molecules *palindromes*. In order to obtain a unified representation of the differential equations we may assume that palindromes consist of equal concentrations of *plus* and *minus* strands with undistinguishable kinetic properties.

Assuming a given chainlength ν for all RNA molecules in our reaction vessel, there is a natural measure for the difference between each two molecules. The *Hamming distance* $dist(I, J)$ of two sequences I and J is defined by the number of positions along the chain which are different in the two strands. The most simple mutation model takes into account only *point mutation*, i.e. changing only a single nucleotide at once. Let I^* denote the product of correct replication of I and p is the probability for a point mutation. The probability for obtaining J from I is thus given by

$$p_{JI^*} = p^{dist(I^*,J)}(1-p)^{\nu - dist(I^*,J)}$$

for a binary alphabet, i.e. of only the difference of purine bases pyrimidin bases is taken into account. In the general case of κ different nucleotides equ.(2.6) is immediately generalized to

$$p_{JI^*} = \left(\frac{1}{\kappa - 1}\right)^{dist(I^*,J)} \cdot p^{dist(I^*,J)}(1-p)^{\nu - dist(I^*,J)} \tag{2.6'}$$

As a simple approximation we may assume that p^2 is already so small that we can neglect it. Then

$$p_{JI^*} = \begin{cases} 1 - \nu p & \text{if} \quad I^* = J \\ p & \text{if} \quad dist(I^*, J) = 1 \\ 0 & \text{otherwise} \end{cases} \tag{2.7}$$

The mutation matrices defined in equ.(2.6) and equ.(2.7) have been studied in great detail by Rumschitzky [16] for direct replication, i.e. $I = I^*$.

If the Hamming distance of plus- and minus-strand fulfils

$$dist(I_k^{\oplus}, I_k^{\ominus}) > 2m \qquad m = 1, 2, 3, \dots \tag{2.8}$$

then the m-error neighborhoods of $I_k^{\oplus}$ and $I_k^{\ominus}$ do not overlap. We may then call all members of l-error neighborhoods of $I_k^{\oplus}$ *plus* strands for $l \leq m$. In this case we can neglect the contributions of the matrices P^{++} and P^{--}.

If $I_l^{\oplus}$, on the other hand, is a palindrome we must split its one-error neighborhood into $\nu/2$ *plus* and $\nu/2$ *minus* strands, because the complement of a one-error mutant of a palindrome is again one of its one-error mutants (fig.1). Note that such a one-error mutant $I_1^{\oplus}$ must itself have one-error mutants which lie in the one-error neighborhood of other one-error mutants of the palindrome. Some of these may be *minus* strands by previous definition. Thus there is no natural way to distinguish *plus* from *minus* strands. Therefore we may have a situation where the formation of a *plus* strand and a *minus* strand is possible by a single point mutation. In this case we may not neglect the contributions of the matrices P^{++} and P^{--} even in a first order approximation.

æ

3. Fixed Points and Eigenvalues

Let us return to the general selection-only model

$$\begin{aligned} \dot{x}_k &= y_k \cdot \Psi_k(x, y) - x_k \Phi \\ \dot{y}_k &= x_k \cdot \Theta_k(x, y) - y_k \Phi \\ \Phi &= \sum_{j=1}^{n} y_k \Psi_k(x, y) + x_k \Theta_k(x, y) \end{aligned} \tag{3.1}$$

Condition (2.4) implies that whenever $x_k + y_k > 0$ and $(x, y) \in S_{2n}$

$$\begin{aligned} x_k = 0 &\Longrightarrow \Psi_k(x, y) > 0 \\ y_k = 0 &\Longrightarrow \Theta_k(x, y) > 0 \end{aligned} \tag{3.2}$$

Let us define the interior of a *boundary subsimplex* for which the pairs (x_k, y_k) vanish for all indices $k \in K \subset \{1, 2, \dots, n\}$ by

$$F_K = \{(x, y) \in S_{2n} | x_k = 0 \wedge y_k = 0 \iff k \in K\}. \tag{3.3}$$

The closure of F_K will be denoted by $\bar{F}_K$.

The following lemma — the analogue to the *reduction lemma* for replicator equations — is straightforeward:

Lemma 3.1. The subsimplices $\bar{F}_K$ are invariant. The restriction of equ.(3.1) to $\bar{F}_K$ is of the same functional form as equ.(3.1) itself only with the functions Ψ and Θ restricted to $\bar{F}_K$.

Lemma 3.2. Let $\hat{x}$ be a rest point of equ.(3.1). Then for all k holds: either $\hat{x}_k = \hat{y}_k = 0$ or $\hat{x}_k \neq 0 \wedge \hat{y}_k \neq 0$.

Proof. Suppose $\hat{x}_k = 0$ and $\hat{y}_k \neq 0$, thus $y_k > 0$. Then equ.(3.1) immediately leads to a contradiction: $\dot{x}_k = \hat{y}_k \Psi_k(0, \hat{y}) > 0$.

Thus all equilibria lie in the interior of a uniquely defined minimal invariant (sub)simplex. (Let $\hat{x}$ be a fixed point and $\hat{x}_k = 0$ iff $k \in K$. Then $\hat{x} \in \mathrm{int}\bar{F}_K = F_K$.) As for replicator equations the spectrum of the Jacobian at $(\hat{x}, \hat{y})$ decomposes into three parts:

(i) One *external* eigenvector points outside the simplex S_{2n}. The corresponding eigenvalue does not influence the dynamical behaviour since the simplex is invariant.

(ii) $2\#K$ *transversal* eigenvectors lie in the simplex S_n but point out of the invariant subsimplex $\bar{F}_K$.

(iii) $2(n - \#K) - 1$ eigenvectors lie in the invariant subsimplex $\bar{F}_K$.

Let us now calculate the Jacobian at a fixed point $\hat{x}$:

$$\begin{aligned}
\frac{\partial \dot{x}_k}{\partial x_l} &= y_k \frac{\partial \Psi_k}{\partial x_l} - x_k \frac{\partial \Phi}{\partial x_l} - \delta_{kl}\Phi \\
\frac{\partial \dot{x}_k}{\partial y_l} &= \delta_{kl}\Psi_k + y_k \frac{\partial \Psi_k}{\partial y_l} - x_k \frac{\partial \Phi}{\partial y_l} \\
\frac{\partial \dot{y}_k}{\partial x_l} &= \delta_{kl}\Theta_k + x_k \frac{\partial \Theta_k}{\partial x_l} - y_k \frac{\partial \Phi}{\partial x_l} \\
\frac{\partial \dot{y}_k}{\partial y_l} &= x_k \frac{\partial \Theta_k}{\partial y_l} - y_k \frac{\partial \Phi}{\partial y_l} - \delta_{kl}\Phi
\end{aligned} \tag{3.4}$$

First we observe that $(1, 1, \ldots, 1)^T$ is a left eigenvector:

$$\sum_{k=1}^{n} \left(\frac{\partial \dot{x}_k}{\partial x_l} + \frac{\partial \dot{y}_k}{\partial x_l} \right) = -\Phi + \Theta_l + \sum_{k=1}^{n} \left(y_k \frac{\partial \Psi_k}{\partial x_l} + x_k \frac{\partial \Theta_k}{\partial x_l} \right) - \frac{\partial \Phi}{\partial x_l}.$$

A short calculation shows that all terms but $-\Phi$ cancel. The same holds for the columns containing $\frac{\partial}{\partial y_l}$. Therefore we have the same result as for replicator equations:

Lemma 3.3. The external eigenvalue is $-\Phi(\hat{x}, \hat{y})$.

Let us now assume without loss of generality that $(\hat{x}_r, \hat{y}_r) = (0, 0)$ for $m < r \leq n$ and $x_r, y_r \neq 0$ for $r \leq m$. The Jacobian then takes the form:

$$J = \begin{pmatrix} A & X \\ 0 & \Lambda \end{pmatrix} \tag{3.5}$$

where A denotes the Jacobian of the restriction of equ.(3.1) to $\bar{F}_{\{m+1,\ldots,n\}}$ and Λ is a block-diagonal matrix consisting of $n - m$ 2×2-blocks of the form

$$\Lambda_k = \begin{pmatrix} -\Phi(\hat{x}, \hat{y}) & \Psi_k(\hat{x}, \hat{y}) \\ \Theta_k(\hat{x}, \hat{y}) & -\Phi(\hat{x}, \hat{y}) \end{pmatrix} \tag{3.6}$$

The elements in block X do not contribute to the eigenvalues. Since the matrix Λ corresponds to the transversal eigenvalues, we may readily compute them:

$$\det(J - \lambda E) = \det(A - \lambda E) \cdot \prod_{k=m+1}^{n} \det(\Lambda_k - \lambda E). \tag{3.7}$$

From

$$\det(\Lambda_k - \lambda E) = (\Phi + \lambda)^2 - \Psi_k \Theta_k = 0 \tag{3.8}$$

We find

Lemma 3.4. The transversal eigenvalues at a fixed point $(\hat{x}, \hat{y}) \in F_K$ are real and given by

$$\lambda_k^{\pm} = -\Phi(\hat{x}, \hat{y}) \pm \sqrt{\Psi_k(\hat{x}, \hat{y}) \cdot \Theta_k(\hat{x}, \hat{y})} \tag{3.9}$$

for all $k \in K$.

As for replicator equations we will call a rest point *saturated* iff all its transversal eigenvalues are non-positive. An immediate consequence is

Theorem 3.1. A rest point $\hat{x} \in F_K$ of equ.(2.2) is saturated iff for all $k \in K$ holds

$$\Phi(\hat{x}, \hat{y}) \geq \sqrt{\Psi_k(\hat{x}, \hat{y}) \cdot \Theta_k(\hat{x}, \hat{y})}. \tag{3.10}$$

We might compare this result with the well known expressions for replicator equation [14]: A rest point of a equ.(1.2) is saturated iff $\Phi(\hat{x}) \geq f_k(\hat{x})$ for all transversal directions k.

Since all transversal eigenvalues are real, the following corrollary holds:

Corrollary. Let $\hat{x} \in \mathrm{bd}S_{2n}$ be a regular rest point of equ.(2.2). Then $\hat{x}$ is not a source.

The *rest point migration theorem* [21] holds in the same form as for replicator equations. We omit the proof for it does not differ from the proof of the analogous theorem for replicator equations.

Theorem 3.2. Let $p_0 = (\hat{x}_0, \hat{y}_0)$ be a regular rest point on F_K. Let furthermore $(\mathcal{M}_\epsilon^+, \mathcal{M}^-{}_\epsilon)$ be a C^1 vector-field defined on some compact neighborhood of the simplex S_{2n} depending continuously on a parameter ϵ fulfilling $\mathcal{M}_0(x, y) = 0$ and $\mathcal{M}_k(\hat{x}_0, \hat{y}_0) > 0$ for $\epsilon > 0$ and all $k \in K$. Then for each $\delta > 0$ there is an $\epsilon > 0$ such that

(i) there is a unique rest point $p_\epsilon = (\hat{x}_\epsilon, \hat{y}_\epsilon)$ of the perturbed equation

$$\begin{aligned} \dot{x}_k &= y_k \Psi_k(x, y) - x_k \Theta + \mathcal{M}_k^+(x, y) \\ \dot{y}_k &= x_k \Psi_k(x, y) - y_k \Theta + \mathcal{M}_k^-(x, y) \end{aligned} \tag{3.11}$$

in a δ-neighborhood of $(\hat{x}_0, \hat{y}_0)$,

(ii) The Jacobian of p_ϵ has the same number of positive and negative eigenvalues as p_0, and

(iii) p_ϵ lies in $\mathrm{int}S_{2n}$ iff p_0 is saturated.

æ

4. Quasispecies

The first order autocatalytic model for complementary replicating units reads

$$\begin{aligned} \dot{x}_k &= \alpha_k y_k - x_k \Phi \\ \dot{y}_k &= \beta_k x_k - y_k \Phi \\ \Phi &= \sum_{j=1}^{n} (\beta_j x_j + \alpha_j y_j). \end{aligned} \tag{4.1}$$

It follows from equ.(3.2) that α_k and β_k are positive. It is straight foreward to calculate all rest points: The equilibrium P_r, $r \in \{1, 2, \ldots, n\}$ is given by

$$\begin{aligned} \hat{x}_r &= \frac{\sqrt{\alpha_r}}{\sqrt{\alpha_r} + \sqrt{\beta_r}} \\ \hat{y}_r &= \frac{\sqrt{\beta_r}}{\sqrt{\alpha_r} + \sqrt{\beta_r}} \\ \hat{x}_j &= \hat{y}_j = 0 \qquad \forall j \neq r \end{aligned} \tag{4.2}$$

Under the assumption that $(\alpha_i, \beta_i) \neq (\alpha_j, \beta_j)$ for all $i \neq j$, these fixed points are unique and lie on 2-subsimplices. In the following theory we will neglect the degenerate cases.

We find from equ.(4.1):

$$\Phi(P_r) = \sqrt{\alpha_r \beta_r} > 0. \tag{4.3}$$

For the Jacobian of the subsystem on the egde (x_r, y_r) of S_{2n} we have

$$J_2 = \begin{pmatrix} -\Phi - \beta_r \hat{x}_r & \alpha_r \hat{y}_r \\ \beta_r \hat{x}_r & -\Phi - \alpha_r \hat{y}_r \end{pmatrix}. \tag{4.4}$$

The characteristic polynomial is

$$\lambda^2 + \lambda[2\Phi + \alpha_r \hat{y}_r + \beta_r \hat{x}_r] + [\Phi^2 + \Phi(\alpha_r \hat{y}_r + \beta_r \hat{x}_r)].$$

Thus the eigenvalues are

$$\begin{aligned} \lambda_1 &= -\Phi < 0, \\ \lambda_2 &= -\Phi - (\alpha_r \hat{y}_r + \beta_r \hat{x}_r) < 0. \end{aligned} \tag{4.5}$$

The eigenvalue $\lambda_1 = -\Phi$ is the external one according to Lemma 3.3. The transversal eigenvalues are given by Lemma 3.4:

$$\lambda_j^{\pm}(P_r) = -\sqrt{\alpha_r \beta_r} \pm \sqrt{\alpha_j \beta_j} \qquad \forall\ j \neq r. \tag{4.6}$$

Since we are interested in the generic case only, we may assume that $\alpha_j \beta_j \neq \alpha_i \beta_i$ holds for all $i \neq j$. The following theorem is immediate:

Theorem 4.1. Equ.(4.1) *generically* admits a unique globally stable equilibrium P_m with

$$\alpha_m \beta_m = \max_j \{\alpha_j \beta_j\}. \tag{4.7}$$

Again this result is reminiscent of the theory of replicator equations: The same result hold for first order replicator networks if we replace the geometric mean of the complementary replication constants $\sqrt{\alpha_k\beta_k}$ by the replication constant for direct replication f_k. This formulation was chosen heuristically in [23].

Let us now include mutation into our considerations. We assume that mutation occurs only as miscopying during the replication process. The *quasispecies equation* for complimentary replicating species then reads

$$\begin{aligned}\dot{x}_k &= \sum_{j=1}^n p_{kj}^{+-}\alpha_j y_j + \sum_{j=1}^n p_{kj}^{++}\beta_j x_j - x_k\Phi \\ \dot{y}_k &= \sum_{j=1}^n p_{kj}^{-+}\beta_j x_j + \sum_{j=1}^n p_{kj}^{--}\alpha_j y_j - y_k\Phi\end{aligned} \tag{4.9}$$

A short calculation shows that the flux function Φ does not depend on the mutation term:

$$\Phi(x,y) = \sum_{j=1}^n (\beta_j x_j + \alpha_j y_j) \tag{4.10}$$

As for dynamical system describing erroneous direct replication [13,21,22], replication and mutation may be separated:

$$\begin{aligned}\dot{x}_k &= [\alpha_k y_k - x_k\Phi] + \mathcal{M}_k^+(x,y) \\ \dot{y}_k &= [\beta_k x_k - y_k\Phi] + \mathcal{M}_k^-(x,y)\end{aligned} \tag{4.11}$$

with the mutation contributions

$$\begin{aligned}\mathcal{M}_k^+ &= \sum_{j=1}^n \Big[(p_{kj}^{+-}\alpha_j y_j - p_{jk}^{+-}\alpha_k y_k) + (p_{kj}^{++}\beta_j x_j - p_{jk}^{--}\alpha_k y_k) \Big] \\ \mathcal{M}_k^- &= \sum_{j=1}^n \Big[(p_{kj}^{-+}\beta_j x_j - p_{jk}^{-+}\beta_k x_k) + (p_{kj}^{--}\alpha_j y_j - p_{jk}^{++}\beta_k x_k) \Big]\end{aligned} \tag{4.12}$$

Let us now suppose that $\hat{x}_0 = (x_1, y_1, 0, \ldots, 0)$ is the unique sink of equ.(4.1). Up to first order in mutation probabilities the unique sink of equ.(4.11) is given by

$$\hat{x} = \hat{x}_0 + d^{(1)} + O(p^2) \tag{4.13}$$

where the shift $d^{(1)}$ may be calculated from the Jacobian $J(\hat{x}_0)$ of equ.(4.1) at the position $\hat{x}_0$ [21]

$$J(\hat{x}_0)\cdot d^{(1)} = -\mathcal{M}(\hat{x}_0). \tag{4.14}$$

For $k \neq 1$ the right side of equ.(4.14) reads

$$\begin{aligned}\mathcal{M}_k^+ &= \sqrt{\alpha_k\beta_k}\cdot\frac{p_{k1}^{+-}\sqrt{\alpha_1} + p_{k1}^{++}\sqrt{\beta_1}}{\sqrt{\alpha_1}+\sqrt{\beta_1}} \\ \mathcal{M}_k^- &= \sqrt{\alpha_k\beta_k}\cdot\frac{p_{k1}^{-+}\sqrt{\beta_1} + p_{k1}^{--}\sqrt{\alpha_1}}{\sqrt{\alpha_1}+\sqrt{\beta_1}}\end{aligned} \tag{4.15}$$

Because of the special form of the Jacobian (*cf.* section 3), equ.(4.14) decomposes into $n-1$ linear equations of the form

$$\begin{pmatrix} -\Phi(\hat{x}_0) & \alpha_k \\ \beta_k & -\Phi(\hat{x}_0) \end{pmatrix} \cdot \begin{pmatrix} d_k^+ \\ d_k^- \end{pmatrix} = -\begin{pmatrix} \mathcal{M}_k^+(\hat{x}_0) \\ \mathcal{M}_k^-(\hat{x}_0) \end{pmatrix} \tag{4.16}$$

These are readily solved:

$$\begin{aligned} d_k^{(1)+} &= \frac{\Phi\mathcal{M}_k^+ + \alpha_k\mathcal{M}_k^-}{\Phi^2 - \alpha_k\beta_k} \\ d_k^{(1)-} &= \frac{\Phi\mathcal{M}_k^- + \beta_k\mathcal{M}_k^+}{\Phi^2 - \alpha_k\beta_k} \end{aligned} \tag{4.17}$$

The explicit solution is rather clumsy, thus we introduce the abbreviations

$$\begin{aligned} P_k^+ &= \frac{p_{k1}^{+-}\sqrt{\alpha_1} + p_{k1}^{++}\sqrt{\beta_1}}{\sqrt{\alpha_1}+\sqrt{\beta_1}} \\ P_k^- &= \frac{p_{k1}^{-+}\sqrt{\beta_1} + p_{k1}^{--}\sqrt{\alpha_1}}{\sqrt{\alpha_1}+\sqrt{\beta_1}} \end{aligned} \tag{4.18}$$

which might be regarded as weighted mutation probabilities of the master pair $I_1^\oplus, I_1^\ominus$. The shifts are given by

$$\begin{aligned} d_k^{(1)+} &= \frac{\sqrt{\alpha_1\beta_1}}{\sqrt{\alpha_1\beta_1}-\sqrt{\alpha_k\beta_k}} \cdot \frac{\sqrt{\alpha_1\beta_1}P_k^+ + \alpha_k P_k^-}{\sqrt{\alpha_1\beta_1}+\sqrt{\alpha_k\beta_k}} \\ d_k^{(1)-} &= \frac{\sqrt{\alpha_1\beta_1}}{\sqrt{\alpha_1\beta_1}-\sqrt{\alpha_k\beta_k}} \cdot \frac{\sqrt{\alpha_1\beta_1}P_k^- + \beta_k P_k^+}{\sqrt{\alpha_1\beta_1}+\sqrt{\alpha_k\beta_k}}. \end{aligned} \tag{4.19}$$

The first term describes the selective advantage of the master pair over the pair $I_k^\oplus, I_k^\ominus$; the second term again is a weighted mutation probability. From the above it seems justified to consider a pair of complementary strands as a single species. Thus we define

$$\begin{aligned} d_k^{(1)} &:= d_k^{(1)+} + d_k^{(1)-} = \\ &= \frac{\sqrt{\alpha_1\beta_1}}{\sqrt{\alpha_1\beta_1}-\sqrt{\alpha_k\beta_k}} \cdot \frac{\sqrt{\alpha_1\beta_1}(P_k^+ + P_k^-) + \alpha_k P_k^- + \beta_k P_k^+}{\sqrt{\alpha_1\beta_1}+\sqrt{\alpha_k\beta_k}} = \\ &=: \frac{\sqrt{\alpha_1\beta_1}}{\sqrt{\alpha_1\beta_1}-\sqrt{\alpha_k\beta_k}} \cdot \tilde{p}_k \end{aligned} \tag{4.20}$$

Equ.(4.20) has a counterpart in the theory of direct replicating quasispecies [21]:

$$d_k^{(1)} = \frac{f_1}{f_1 - f_k} \cdot p_{k1}. \tag{4.21}$$

Again we find the correspondence $f_j \iff \sqrt{\alpha_j\beta_j}$ but instead of the simple mutation probability we have a rather complicated weighted mean in the complementary case.

Example: Let us consider Eigen's mutation model – which has been outlined in section 2 for the special case of a four digits alphabet – in its general form with κ different symbols [5,7,8,10,11,16,23]. As a first case we assume that plus and minus-strand of the master pair have sufficiently large Hamming distance $d(I_1^\oplus, I_1^\ominus) > 3$. Then the sets of one-error mutants ($U_1^\oplus$ and $U_1^\ominus$) of the plus and the minus-strand of the master pair are disjoint; furthermore the complementary strand of each $I_k^{\oplus} \in U_1^{\oplus}$ lies in $U_1^{\ominus}$ and *vice versa*. Since we are interested in a first order approximation, equ.(4.18) reduces to

$$\begin{aligned} P_k^{+} &= \frac{\sqrt{\alpha_1}}{\sqrt{\alpha_1}+\sqrt{\beta_1}}p \\ P_k^{-} &= \frac{\sqrt{\beta_1}}{\sqrt{\alpha_1}+\sqrt{\beta_1}}p \end{aligned} \tag{4.22}$$

Equ.(4.20) then becomes with $\tilde{p}_k = p \cdot \mu_k$

$$d_k^{(1)} = \begin{cases} \frac{\sqrt{\alpha_1\beta_1}}{\sqrt{\alpha_1\beta_1}-\sqrt{\alpha_k\beta_k}} \cdot \mu_k p & \text{for } k \in U_1^{\oplus} \\ 0 & \text{otherwise} \end{cases} \tag{4.23}$$

where the "asymmetry factor" μ_k is defined by

$$\mu_k = \frac{\sqrt{\alpha_1\beta_1}(\sqrt{\alpha_1}+\sqrt{\beta_1}) + \beta_k\sqrt{\alpha_1} + \alpha_k\sqrt{\beta_1}}{(\sqrt{\alpha_1\beta_1}+\sqrt{\alpha_k\beta_k})(\sqrt{\alpha_1}+\sqrt{\beta_1})}. \tag{4.24}$$

This term is the consequence of asymmetric replication constants; it reduces to $\mu_k = 1$ for $\alpha_k = \beta_k$. Numerical investigations presented in the papers of Eigen, McCaskill and Schuster [10,11] and Swetina and Schuster [19,23] have been performed with this simplification. Thus their analytical expressions based on equ.(4.37) agree perfectly with the simulation results. In general μ_k is of the order of unity:

$$\frac{1}{2} < \frac{\sqrt{\alpha_1\beta_1}}{\sqrt{\alpha_1\beta_1}+\sqrt{\alpha_k\beta_k}} \le \mu_k < 1 + \frac{\alpha_k+\beta_k}{2} / \sqrt{\alpha_k\beta_k}. \tag{4.25}$$

As a second case we consider a *palindromic* master sequence $I_1^{\oplus} = I_1^{\ominus}$. Of course there is a single one-error neighborhood consting of $(\kappa-1)\nu/2$ plus and $(\kappa-1)\nu/2$ minus strands. In this case we have $f_1 := \alpha_1 = \beta_1$ and the asymmetry factor reduces to

$$\mu_k = \frac{2f_1 + \alpha_k + \beta_k}{f_1 + \sqrt{\alpha_k\beta_k}}. \tag{4.27}$$

For $\alpha_k \to \beta_k$, μ_k approaches 2 in this case.

Now we are in the position to calculate a first order approximation for the *error threshold*. To facilitate comparison with previous papers on this topic [10,11,18,19,23] we begin with the results for direct replicating quasispecies. The approach presented here is based on a perturbation theory for the coordinates of an equilibrium and is completely different from previous derivations. The condition for the errorthreshold is given in terms of coordinates x_k:

$$\sum_{k\neq 1} x_k = 1. \tag{4.27}$$

Up to first order this formula can be approximated by

$$\sum_{k\neq 1} d_k^{(1)} = 1, \tag{4.28}$$

where $d_k^{(1)}$ is given in equ.(4.21). As first order approximation we find

$$d_k = \begin{cases} p \cdot \frac{f_1}{f_1 - f_k} & \text{for} \quad d(k,1) = 1 \\ 0 & \text{otherwise} \end{cases} \tag{4.29}$$

and thus

$$1 = p(\kappa - 1)\nu \cdot \frac{1}{(\kappa - 1)\nu} \sum_{d(k,1)=1} \frac{f_1}{f_1 - f_k} =: p(\kappa - 1)\nu \cdot A_1 \tag{4.30}$$

The parameter A_1 may be called "average first order shift". With the single digit accuracy $q = 1 - (\kappa - 1)p$ we get

$$q^\nu \approx (1 - \nu \cdot (\kappa - 1)p) = \frac{A_1 - 1}{A_1}. \tag{4.31}$$

With the definition of the *superiority parameter*

$$\sigma_1 = \frac{A_1}{A_1 - 1} \tag{4.32}$$

we find finally

$$\nu \approx -\frac{\ln \sigma_1}{\ln q}. \tag{4.33}$$

It is easy to verify that

$$\sigma_1 = \frac{\sum_{d(k,1)=1} \frac{f_1}{f_1 - f_k}}{\sum_{d(k,1)=1} \frac{f_k}{f_1 - f_k}} \tag{4.34}$$

is just the first order approximation of Eigen's [7] definition of the superiority parameter

$$\sigma_0 = f_1 / \bar{E}_{k \neq 1} = f_1 \cdot \frac{\sum_{k \neq 1} d_k}{\sum_{k \neq 1} f_k d_k}. \tag{4.35}$$

Thus equ.(4.32) is the correct first order approximation for the errorthreshold.

The analogous calculation for complementary replication is straight foreward. Instead equ.(4.27) we use

$$\sum_{j \neq 1} (x_j + y_k) = 1 \tag{4.27$'$}$$

as condition for the errorthreshold. Up to first order we approximate this by equ.(4.28) with $d_k^{(1)}$ defined in equ.(4.20) instead of equ.(4.21). Direct and complementary replication differ only by the definition of the parameter A_1:

$$A_1 = \frac{1}{(\kappa - 1)\nu} \sum_{d(k,1)=1} \frac{\sqrt{\alpha_1 \beta_1}}{\sqrt{\alpha_1 \beta_1} - \sqrt{\alpha_k \beta_k}} \cdot \mu_k \tag{4.36}$$

Note that A_1 differs by the asymmetry factor μ_k from expressions based on formulas for direct replication and the substitution

$$f_k \longrightarrow \sqrt{\alpha_k \beta_k}. \tag{4.37}$$

Note that A_1 is a mean value also in the case of a palindromic master sequence. The sum in equ.(4.36) runs only over $\nu/2$ elements, but this is compensated by the larger values of μ_k.

æ

5. Second Order Catalytic Networks

The *second order catalysis* model for complementary replication reads

$$\begin{aligned} I_k^{\oplus} + I_j^{(\alpha)} &\longrightarrow I_k^{\ominus} + I_k^{\oplus} + I_j^{(\alpha)} \\ I_k^{\ominus} + I_j^{(\beta)} &\longrightarrow I_k^{\oplus} + I_k^{\ominus} + I_j^{(\beta)} \end{aligned}$$

with $\alpha, \beta \in \{\oplus, \ominus\}$. Assuming mass action kinetics [9] one obtains the following system of differential equations:

$$\begin{aligned} \dot{x}_k &= y_k([Ax]_k + [By]_k) - x_k\Phi \\ \dot{y}_k &= x_k([Cx]_k + [Dy]_k) - y_k\Phi \\ \Phi(x,y) &= yAx + yBy + xCx + xDy \end{aligned} \tag{5.1}$$

From equ.(3.2) it follows immediately that A, B, C and D are positive matrices.

Remark. Like the second order replicator equations equ.(1.3), the above model equ.(5.1) contains also the inhomogeneous case [20].

$$\begin{aligned} \dot{x}_k &= y_k(r_k + [Ax]_k + [By]_k) - x_k\Phi \\ \dot{y}_k &= x_k(s_k + [Cx]_k + [Dy]_k) - y_k\Phi \\ \Phi(x,y) &= yAx + yBy + xCx + xDy + sx + ry \end{aligned} \tag{5.2}$$

since one can rewrite equ.(5.2) in the form of equ.(5.1) using

$$a'_{kj} = a_{kj} + r_k, \quad b'_{kj} = b_{kj} + r_k, \quad c'_{kj} = c_{kj} + s_k, \quad d'_{kj} = d_{kj} + s_k. \tag{5.3}$$

Lemma 5.1. The restriction of equ.(5.1) onto an edge (x_k, y_k) of S_{2n} admits a unique globally stable rest point [22].

We have seen in the previous section that we may consider a pair of complementary strands as a single species. This idea suggests the following change of coordinates

$$z = x + y \qquad \text{and} \qquad \vartheta = x - y. \tag{5.4}$$

Equ.(5.1) becomes

$$\begin{aligned} \dot{z}_k &= z_k([Mz]_k + [R\vartheta]_k) - \vartheta_k([Sz]_k + [T\vartheta]_k) - z_k\Phi(z,\vartheta) \\ \dot{\vartheta}_k &= -\vartheta_k([Mz]_k + [R\vartheta]_k]) + z_k([Sz]_k + [T\vartheta]_k) - \vartheta_k\Phi(z,\vartheta) \\ \Phi(z,\vartheta) &= (zMz) + (zR\vartheta) - (\vartheta Sz) - (\vartheta T\vartheta) \end{aligned} \tag{5.5}$$

with the matrices

$$\begin{aligned} M &= \frac{1}{4}(A + B + C + D) \\ R &= \frac{1}{4}(A - B + C - D) \\ S &= \frac{1}{4}(A + B - C - D) \\ T &= \frac{1}{4}(A - B - C + D) \end{aligned} \tag{5.6}$$

Figure 2: Invariant planes in the S_4 according to equ.(5.9)

In the *fully symmetric case* — plus and minus-strand differ neither in catalytic activity nor in their quality as templates — we have $R = S = T = 0$ and equ.(5.5) becomes extremely simple. It decomposes into a second order replicator equation describing the selection of the pairs

$$\dot{z}_k = z_k([Mz]_k - zMz) \tag{5.7a}$$

and a second equation decribing the equilibration of plus and minus strands

$$\vartheta_k = -\vartheta([Mz]_k + zMz). \tag{5.7b}$$

Since M is a positive matrix we find

$$\lim_{t\to\infty} \vartheta_k(t) = 0. \tag{5.8}$$

Let us now suppose that $S = T = 0$ but $R \neq 0$. This means that $A = C$ and $B = D$, i.e. that plus and minus strand have equal properties as templates but not as catalysts. Equ.(5.5) reduces to

$$\begin{aligned}
\dot{z}_k &= z_k([Mz]_k + [R\vartheta]_k - \Phi) \\
\dot{\vartheta}_k &= -\vartheta_k([Mz]_k + [R\vartheta]_k + \Phi) \\
\Phi &= (zMz) + (zR\vartheta)
\end{aligned} \tag{5.9}$$

Note that the bracket in equ.(5.9b) is always positive since $z_k \geq |\vartheta_k|$ and $m_{ij} > r_{ij}$ for all i, j. Thus ϑ_k decreases exponentially. In this case — and in the previous one — the hyperplanes

$$H_k = \{(x, y) \in S_{2n} | x_k = y_k\} \tag{5.10}$$

are invariant and attractive, and so is their intersection

$$S_n^* = \{(x, y) \in S_{2n} | x = y\}. \tag{5.11}$$

Thus the assymptotic behaviour is defined by equ.(5.9a) alone. Because all terms containing ϑ decrease exponentially we have the following

Theorem 1. Let M be such that the second order replicator equation equ.(5.7a) is robust. Then the ω-limits of equ.(5.7a) and equ.(5.9) are the same.

The *reciprocal* case is physically more interesting. Suppose the two reactions

$$\begin{aligned}
I_k^{(\alpha)} + I_j^{(\beta)} &\longrightarrow I_k^{(\bar\alpha)} + I_k^{(\alpha)} + I_j^{(\beta)} \qquad \text{and} \\
I_k^{(\bar\alpha)} + I_j^{(\bar\beta)} &\longrightarrow I_k^{(\alpha)} + I_k^{(\bar\alpha)} + I_j^{(\bar\beta)}
\end{aligned}$$

(where $I_k^{(\bar{\alpha})}$ denotes the complementary strand of $I_k^{(\alpha)}$) have the same reaction constant. This means that plus and minus strands act like mirror images. Therefore we have $A = D$ and $B = C$ or $R = S = 0$ but $T \neq 0$. We obtain

$$\begin{aligned}\dot{z}_k &= z_k([Mz]_k - zMz) - \vartheta_k[T\vartheta]_k + z_k(\vartheta T\vartheta)\\ \dot{\vartheta} &= -\vartheta_k([Mz]_k + zMz - \vartheta T\vartheta) + z_k[T\vartheta]\end{aligned} \tag{5.12}$$

In this case the hyperplanes H_k are no longer invariant. Their intersection S_n^* however is invariant.

Theorem 5.2. The invariant set $S_n^* \subset S_{2n}$ is attractive for the reciprocal case equ.(5.12).

Proof. Consider z as parameter in equ.(5.12b). Then 0 is a rest point of this differential equation for ϑ for all values of the parameter z. The Jacobian matrix a $\vartheta = 0$ is given by

$$J_{kl} = \frac{\partial\dot{\vartheta}_k}{\partial\vartheta_l}(0) = -\delta_{kl}([Mz]_k + zMz) + z_k t_{kl} \tag{5.13}$$

With $D = \mathrm{diag}(z_1, \ldots, z_n)$ it follows that JD is negative diagonally dominant:

$$z_k([Mz]_k + zMz) + z_k^2 t_{kk} > \sum_{j=1}^{n} |z_k t_{kj}| z_j$$

since

$$[Mz]_k > [Tz]_k \quad \text{for all} \quad k.$$

Thus all eigenvalues of J are negative.

Again the vector field on S_n^* converges exponentially towards the replicator equation equ.(5.7a).

The above special case is immediately generalized to $S = 0$ because a contribution $R \neq 0$ does not enter into the Jacobian matrix (5.13). This is the most general case which can be treated by reduction to a replicator equation for pairs on S_n^*. In terms of reaction constants $S = 0$ means $a_{ij} + b_{ij} = c_{ij} + d_{ij}$.

There are no results in the case $S \neq 0$. So far, there are also no results concerning existence and uniqueness of rest points for equ.(5.1). Because of the reduction lemma 3.1 this is equivalent with the problem of finding positive solutions to the following system of algebraic equations:

$$\begin{aligned}[Ax]_k + [By]_k &= x_k/y_k\\ [Cx]_k + [Dy]_k &= y_k/x_k\end{aligned}$$

where A, B, C, D are positive matrices.

As for the quasispecies equation let us now introduce mutation [13] into our model. We assume that the reaction

$$I_k^{(\alpha)} + I_j^{(\gamma)} \longrightarrow I_k^{(\alpha)} + I_l^{(\beta)} + I_j^{(\gamma)}, \quad \alpha,\beta,\gamma \in \{\oplus,\ominus\}$$

occur with a probability $p_{lk}^{(\beta\alpha)}$. This assumption leads to the same splitting of the mutation matrix as described in the previous section. We find

$$\begin{aligned}
\dot{x}_k &= \sum_{i=1}^{n} \big[p_{ki}^{+-} y_i([Ax]_i + [By]_i) + p_{ki}^{++} x_i([Cx]_i + [Dy]_i)\big] - x_k\Phi \\
\dot{y}_k &= \sum_{i=1}^{n} \big[p_{ki}^{-+} x_i([Cx]_i + [Dy]_i) + p_{ki}^{--} y_i([Ax]_i + [By]_i)\big] - y_k\Phi \\
&\Phi(x,y) = (yAx) + (yBy) + (xCx) + (yDy)
\end{aligned} \tag{5.14}$$

Again it is possible to separate selection and mutation terms

$$\begin{aligned}
\dot{x}_k &= y_k([Ax]_k + [By]_k) - x_k\Phi(x,y) + \mathcal{M}_k^+(x,y) \\
\dot{y}_k &= x_k([Cx]_k + [Dy]_k) - y_k\Phi(x,y) + \mathcal{M}_k^-(x,y)
\end{aligned} \tag{5.15}$$

The mutation fields are given by

$$\begin{aligned}
\mathcal{M}_k^+(x,y) = \sum_{i=1}^{n} \Big\{ &\big[p_{ki}^{+-} y_i([Ax]_i + [By]_i) - p_{ik}^{+-} y_k([Ax]_k + [By]_k\big] + \\
&+ \big[p_{ki}^{++} x_i([Cx]_i + [Dy]_i) - p_{ik}^{--} y_k([Ax]_k + [By]_k\big] \Big\} \\
\mathcal{M}_k^-(x,y) = \sum_{i=1}^{n} \Big\{ &\big[p_{ki}^{-+} x_i([Cx]_i + [Dy]_i) - p_{ik}^{-+} x_k([Cx]_k + [Dy]_k\big] + \\
&+ \big[p_{ki}^{--} y_i([Ax]_i + [By]_i) - p_{ik}^{++} x_k([Cx]_k + [Dy]_k\big] \Big\}
\end{aligned} \tag{5.16}$$

In order to compare the dynamics of erroneous complimentary replication with the dynamics of erroneous direct replication we calculate the first order approximation for the coordinates of a rest point on an edge of the S_{2n}. Without losing generality we assume $x_1, y_1 \neq 0$ and $x_j = y_j = 0$ otherwise. For $k \neq 1$ we find

$$\begin{aligned}
\mathcal{M}_k^+(P_1) &= p_{k1}^{+-} y_1(a_{11}x_1 + b_{11}y_1) + p_{k1}^{++} x_1(c_{11}x_1 + d_{11}y_1) \\
\mathcal{M}_k^-(P_1) &= p_{k1}^{-+} x_1(c_{11}x_1 + d_{11}y_1) + p_{k1}^{--} y_1(a_{11}x_1 + b_{11}y_1)
\end{aligned} \tag{5.17}$$

With $([Ax]_1 + [By]_1])y_1 = x_1\Phi$ and $([Cx]_1 + [Dx]_1)x_1 = y_1\Phi$ at the position P_1 we may write equ.(5.17) as

$$\begin{aligned}
\mathcal{M}_k^+(P_1) &= (p_{k1}^{+-} x_1 + p_{k1}^{++} y_1)\Phi \\
\mathcal{M}_k^-(P_1) &= (p_{k1}^{-+} y_1 + p_{k1}^{--} x_1)\Phi
\end{aligned} \tag{5.17}$$

Analogously to equ.(4.16) we have to solve

$$\begin{pmatrix} -\Phi & (a_{k1}x_1 + b_{k1}y_1) \\ (c_{k1}x_1 + d_{k1}y_1) & -\Phi \end{pmatrix} \cdot \begin{pmatrix} d_k^+ \\ d_k^- \end{pmatrix} = -\Phi \begin{pmatrix} p_{k1}^{+-} x_1 + p_{k1}^{++} y_1 \\ p_{k1}^{-+} y_1 + p_{k1}^{--} x_1 \end{pmatrix}$$

The solutions are

$$\begin{pmatrix} d_k^+ \\ d_k^- \end{pmatrix} = -\frac{\Phi}{(c_{k1}x_1 + d_{k1}y_1)(a_{k1}x_1 + b_{k1}y_1) - \Phi^2} \cdot$$
$$\cdot \begin{pmatrix} \Phi(p_{k1}^{+-}x_1 + p_{k1}^{++}y_1) + (a_{k1}x_1 + b_{k1}y_1)(p_{k1}^{-+}y_1 + p_{k1}^{--}x_1) \\ \Phi(p_{k1}^{-+}y_1 + p_{k1}^{--}x_1) + (c_{k1}x_1 + d_{k1}y_1)(p_{k1}^{+-}x_1 + p_{k1}^{++}y_1) \end{pmatrix}$$

If we regard a pair of complementary strands as a single species we have to sum up $d_k^+ + d_k^- := d_k$. Finally we obtain

$$d_k = \frac{-\Phi}{(c_{k1}x_1 + d_{k1}y_1)(a_{k1}x_1 + b_{k1}y_1) - \Phi^2} \cdot$$
$$\left\{ \Phi[(p_{k1}^{+-} + p_{k1}^{--})x_1 + (p_{k1}^{-+} + p_{k1}^{++})y_1] + (a_{k1}x_1 + b_{k1}y_1)(p_{k1}^{-+}y_1 + p_{k1}^{--}x_1) + (c_{k1}x_1 + d_{k1}y_1)(p_{k1}^{+-}x_1 + p_{k1}^{++}y_1) \right\}$$

Suppose now that $S = 0$ and $P_1 \in S_n^*$. In this case the above expressions simplifies considerably since we have $x_1 = y_1 = 1/2$ and $a_{ij} + b_{ij} = c_{ij} + d_{ij} = m_{ij}/2$. With

$$\Phi = \frac{1}{4}(a_{11} + b_{11} + c_{11} + d_{11}) = m_{11} \tag{5.18}$$

we find

$$d_k = \frac{m_{11}(m_{11} + m_{k1})}{m_{k1}^2 - m_{11}^2} \cdot (p_{k1}^{+-} + p_{k1}^{-+} + p_{k1}^{++} + p_{k1}^{--})$$

or

$$d_k = \frac{m_{11}}{m_{k1} - m_{11}} \cdot (p_{k1}^{+-} + p_{k1}^{-+} + p_{k1}^{++} + p_{k1}^{--}) \tag{5.19}$$

This formula is the same as for direct replication with

$$p = p_{k1}^{+-} + p_{k1}^{-+} + p_{k1}^{++} + p_{k1}^{--} \tag{5.20}$$

Again we find a very close relation between the dynamics of complementary pairs and replicator dynamics for direct replication.

æ

6. Conclusions

This contribution shows that there are many close relations between direct and complementary replication although there are fundamental differences. In general the direct replication model is much easier to treat mathematically. Direct and complementary replication models have in common that the state-space is a simplex, that the spectrum of the Jacobian at a rest point can be decomposed in the same way and that the invariant submanifolds are of a similar form. For many purposes one can consider a pair of complementary strands as a single species, especially in the case of DNA replication. Its kinetic parameters may be obtained – at least approximately – by an appropriate averaging procedure from the kinetic parameters of its plus- and minus-strand. There are however additional effects corresponding to the differences of kinetic parameters of plus- and minus-strand which become important in the case of RNA replication.

Table 1: Comparison of differential equation models for direct and complementary replication.

	Complementary	Direct
	Replication	
General system:		
Dynamics	$\dot{x}_k = y_k\Psi_k(x,y) - x_k\Phi(x,y)$	$\dot{x}_k = x_k f_k(x) - x_k\Phi(x)$
	$\dot{y}_k = x_k\Theta_k(x) - y_k\Phi(x,y)$	
	$\Phi(x,y)=\sum\left[x_i\Theta_i(x,y)+y_i\Psi_i(x,y)\right]$	$\Phi(x) = \sum x_i f_i(x)$
State Space S^*	S_{2n}	S_n
Invariant Manif.	$F_K=\{x\vert(x_k,y_k)=(0,0) \iff k\in K\}$	$F_K=\{x\vert x_k=0 \iff k\in K\}$
External Eigenv.	$-\Phi$	
Transv. Eigenv.	$-\Phi \pm \sqrt{\Psi_k\Theta_k}$	$-\Phi + f_k$
with mutation:		
$P(\epsilon) \in \text{int}S^*$ iff	$P(0)$ saturated	
Quasispecies		
Generic rest p.	$(0,\ldots,0,\hat{x},\hat{y},0,\ldots,0)$	$(0,\ldots,0,1,0,\ldots,0)$
with Mutation	unique globally stable equilibrium	
Shift $d_k^{(1)}$	$\frac{\sqrt{\alpha_1\beta_1}}{\sqrt{\alpha_1\beta_1}-\sqrt{\alpha_k\beta_k}}\tilde{p}_{k1}$	$\frac{f_1}{f_1-f_k}p_{k1}$
Errorthreshold	$v_{\max} = -\frac{\ln\sigma_1}{\ln q}, \quad \sigma_1 = A_1/(A_1-1)$	
	$A_1 = \frac{1}{(\kappa-1)\nu}\sum_{d(k,1)=1} d_k^{(1)}$	
Second Order		
Interior rest p.	**??**	$0, 1, \infty$
Inhomogeneous case	included	
Special cases reducible to replicator dynamics		

In the simplest case – first order catalysis – complementary and direct replication differ only in minor details. Both models show a characteristic behaviour known as error-threshold and even the formalism for its mathematical treatment is very similar in both cases.

In the second order catalytic case, however, the models differ considerably. Only some special cases – although some physically very important ones – can be compared with replicator dynamics.

Table 1 summarizes the comparison of direct and complementary replication.

Acknowledgements

The author wishes to thank P. Schuster and J. Swetina (Vienna) and M. Gebinoga (Göttingen) for many useful discussions and one of the referees for proof-reading of the manuscript. Preparation of the figures by J. König is gratefully acknowledged.

References

[1] Biebricher C. Darwinian Evolution of Self-Replicating RNA. *Chemica Scripta* 26B: 51-57 (1986).

[2] Biebricher C. and M. Eigen. Kinetics of RNA Replication by $Q\beta$-replicase. In *RNA Genetics, Vol.I* , Domingo, Holland and Ahlquist (eds.) p.1-21. CRC Press, Baton Rouge (1988).

[3] Biebricher C., M. Eigen and W.C. Gardiner. Kinetics of RNA replication. *Biochemistry* 22: 2544-2559 (1983).

[4] Cech T. RNA as an enzyme. *Scientific American* 255(5): 76-84 (1986).

[5] Demetrius L., P. Schuster and M. Eigen. Polynucleotide Evolution and Branching Processes. *Bull.Math.Biol.*, 47: 239-262 (1985).

[6] Doudna J.A, J.A. Szostak. RNA-catalysed synthesis of complementary-strand RNA. *Nature* 399: 519-522 (1989).

[7] Eigen M. Self-organization of Matter and the Evolution of Biological Macromolecules. *Naturwissenschaften*, 58: 465-526 (1971).

[8] Eigen M. and P. Schuster *The Hypercycle. A Principle of Natural Self-organization.* Springer-Verlag, Berlin, Heidelberg New York, 1979.

[9] Feinberg M. Mathematical aspects of mass action kinetics *in Chemical Reactor Theory: A Review*, L. Lapidus and N. Amundson, eds. Prentice-Hall, Englewood Cliffs, New Jersey, 1977.

[10] Eigen M., J. McCaskill and P. Schuster. The Molecular Quasi-Species, *J.Phys.Chem.*, 92: 6881-6891 (1988).

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

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

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

[14] Hofbauer J. and K. Sigmund. *The Theory of Evolution and Dynamical Systems*, Cambridge Univerity Press, 1988.

[15] Lotka A.J. Undamped Oszillations Derived from the Law of Mass Action, J.Amer.Chem.Soc., 42: 1595-1598 (1920).

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

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

[18] Schuster P. The Physical Basis of Molecular Evolution. *Chemica Scripta*, 26B: 27-41 (1986).

[19] Schuster P. and J. Swetina. Stationary Mutant Distribution and Evolutionary Optimization, *Bull.Math.Biol.*, 50: 635-660 (1988).

[20] Stadler P. Dynamics of Small Autocatalytic Reaction Networks IV: Inhomogeneous Replicator Networks. (submitted to *Biosystems*)

[21] Stadler P. and P. Schuster. Mutation in Autocatalytic Reaction Networks. An Analysis Based on Perturbation Theory. *SFI preprint series* 90-022, submitted to *J. Math. Biol.*

[22] Stadler P., W. Schnabl, C. Forst and P. Schuster. Small Autocatalytic Reaction Networks II: Analytically Treatable Special Cases. (submitted to *Bull. Math. Biol.*)

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

[24] Volterra V. *Lecons sur la theorie mathematique de la lutte pour la vie* Gauthier-Villars, Paris (1931).

Figure Captions

Figure 1. m-error neighborhoods of complementary strands with large Hamming distance (**a**) and of a palindromic sequence (**b**).

Figure 2. Invariant planes in the S_4 according to equ.(5.9).

Table 1: Comparison of differential equation models for direct and complementary replication.

	Complementary	Direct
	Replication	
General system:		
Dynamics	$\dot{x}_k = y_k\Psi_k(x,y) - x_k\Phi(x,y)$	$\dot{x}_k = x_k f_k(x) - x_k\Phi(x)$
	$\dot{y}_k = x_k\Theta_k(x) - y_k\Phi(x,y)$	
	$\Phi(x,y)=\sum\left[x_i\Theta_i(x,y)+y_i\Psi_i(x,y)\right]$	$\Phi(x) = \sum x_i f_i(x)$
State Space S^*	S_{2n}	S_n
Invariant Manif.	$F_K=\{x\mid(x_k,y_k)=(0,0) \iff k\in K\}$	$F_K=\{x\mid x_k=0 \iff k\in K\}$
External Eigenv.	$-\Phi$	
Transv. Eigenv.	$-\Phi \pm \sqrt{\Psi_k\Theta_k}$	$-\Phi + f_k$
with mutation:		
$P(\epsilon) \in \mathrm{int}S^*$ iff	$P(0)$ saturated	
Quasispecies		
Generic rest p.	$(0,\dots,0,\hat{x},\hat{y},0,\dots,0)$	$(0,\dots,0,1,0,\dots,0)$
with Mutation	unique globally stable equilibrium	
Shift $d_k^{(1)}$	$\frac{\sqrt{\alpha_1\beta_1}}{\sqrt{\alpha_1\beta_1}-\sqrt{\alpha_k\beta_k}}\tilde{p}_{k1}$	$\frac{f_1}{f_1-f_k}p_{k1}$
Errorthreshold	$v_{\max} = -\frac{\ln\sigma_1}{\ln q}, \quad \sigma_1 = A_1/(A_1-1)$	
	$A_1 = \frac{1}{(\kappa-1)\nu}\sum_{d(k,1)=1} d_k^{(1)}$	
Second Order		
Interior rest p.	??	$0,1,\infty$
Inhomogeneous case	included	
Special cases reducible to replicator dynamics		