MATCH
Communications in Mathematical and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. **80** (2018) 547-585

ISSN 0340 - 6253

A Mathematical Model of Evolution

Peter Schuster

Institut für Theoretische Chemie der Universität Wien, Währingerstraße 17, 1090 Wien, Austria and The Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA

pks@tbi.univie.ac.at , pks@santafe.edu .

(Received April 30, 2017)

Abstract

Evolution is viewed as a process involving population dynamics on a complex landscape that results from a mapping of genotype space onto phenotype space. Phenotypes are evaluated with respect to fitness, which is interpreted here as the number of offspring. Evolutionary dynamics combines two features: (i) the dynamics of genotype distributions on the population level and (ii) interspecies dynamics on the level of ecosystems. A formal mathematical model of evolution based on three fundamental processes – competition through reproduction, symbiontic cooperation through catalyzed reproduction, and variation through mutation – is introduced here. The intensity parameters of the three processes are plotted along the coordinate axes of a Cartesian coordinate system and the dynamics on the three faces of the Cartesian space is presented, analyzed, and discussed on the deterministic as well as stochastic level. The dynamics in the interior of the parameter space is briefly mentioned.

1 Introduction

Biological evolution is commonly seen in its historical dimension as expressed by the famous quotation of Theodosius Dobzhansky: *"Nothing in biology makes sense except in the light of evolution"* [1]. Alternatively there is the mechanistic aspect of evolutionary processes considering population dynamics on complex fitness landscapes [2]. Here we shall focus on the dynamical aspect and describe biological evolution by a kind of minimal model that is thought to represent a core that can be extended by the currently understood processes of epigenetic inheritance without becoming too sophisticated for

straightforward analysis. Evolution is seen as a special case of chemical reaction networks, which contain either single autocatalytic reaction steps or subsets of reactions forming together an autocatalytic network. In the simplest case the autocatalytic steps are tantamount to reproduction of DNA or RNA molecules. Reproduction in this case is replication, an enzyme-catalyzed template-induced process that copies a nucleic acid molecule nucleotide per nucleotide and hence comprises a large number of reaction steps. Under suitable conditions this multi-step process can be approximated by an overall single step autocatalytic reaction, $\mathbf{A} + \mathbf{X}_i \xrightarrow{f_i} 2\mathbf{X}_i$, giving rise to exponential growth of RNA concentrations [3, 4]. On the population level with several different reproducing variants called *replicators*, $\mathbf{X}_i\,(i = 1 \ldots, n)$, competition for resources $\mathbf{A}$ leads to natural selection of the fittest genotype $\mathbf{X}_m$ with $m:\ f_m = \max\{f_i; i = 1, \ldots, n\}$. In such simple systems, which nevertheless enable evolution, the mean fitness, $\overline{f}(t) = \sum_{i=1}^{n} f_i n_i(t) \ / \ \sum_{i=1}^{n} n_i(t)$ with $n_i(t)$ being the number of replicators $\mathbf{X}_i$ in the population, is optimized during the evolutionary process.

Apart from simple reproduction that leads to competition of we shall also consider catalyzed replication,which leads to cooperation among replicators [5, 6]. Experimental examples of cooperative autocatalytic reaction networks built from RNA molecules that are collectively capable of reproduction have been described in the literature [7–9]. Other cooperative systems constitute the basis of various forms of symbiosis [10–12] and cooperativity is also thought to have played a decisive role in the RNA world [6]. In this contribution molecular symbiotic cooperation is introduced through catalyzed replication in its – mathematically – simplest form:

$$\mathbf{A} + \mathbf{X}_i + \mathbf{X}_j \xrightarrow{h_{ij}} 2\mathbf{X}_i + \mathbf{X}_j \text{ with } i, j = 1, \ldots, n\ ,$$

where $\mathbf{X}_i$ represents the template and $\mathbf{X}_j$ is the catalyst. In the most general form – every molecule $\mathbf{X}_j$ has the potential to act as a catalyst for the replication of every other molecular species $\mathbf{X}_i$ including itself – the number of catalytic terms is n^2. Considering the fact that efficient and specific catalysis is a rare property the number n^2 becomes unrealistically large even for fairly small n. It is possible to reduce straightforwardly the number of catalytic reactions. The simplest example, so far known and analyzed, is the catalytic hypercycle [5, 13–15]: The catalyzed reactions $2\mathbf{X}_i + \mathbf{X}_{i+1} \xleftarrow{h_i} \mathbf{A} + \mathbf{X}_i + \mathbf{X}_{i+1}$ $(i = 1, \ldots, n,\ i \mod n)$ involving n replicators are thought to build a one-dimensional closed cyclic array of mutual dependence,

$$\ldots \leftarrow \mathbf{X}_n \leftarrow \mathbf{X}_1 \leftarrow \mathbf{X}_2 \leftarrow \ldots \leftarrow \mathbf{X}_{n-1} \leftarrow \mathbf{X}_n \leftarrow \ldots \ .$$

The dynamics of hypercycles has been studied in great detail [16–18]. The long-time solutions are asymptotically stable points for $n = 2, 3, 4$ or oscillations resulting from an asymptotically stable closed orbit for $n \geq 5$.

Genetic variation occurs at the level of DNA or RNA genotypes. In nature two forms of variation are observed: mutation and recombination. The simplest form of variation is the point mutation that consists of an exchange of a single nucleotide in the sequence. In the mutation model applied here point mutations arise through incorporation of wrong nucleotides during the replication process. Correct and error-prone replication are considered as parallel reaction channels of one and the same replication mechanism [13]. RNA replication and mutation kinetics under suitable conditions gives rise to simple over-all reaction mechanism:

$$\mathbf{A} + \mathbf{X}_i + \mathbf{E} \overset{Q_{ji} \cdot k_i[\mathbf{A}]}{\rightarrow\rightarrow\rightarrow} \mathbf{X}_j + \mathbf{X}_i + \mathbf{E} \ ,$$

The resolved replication process is initiated through binding of the RNA template molecule to a specific RNA replicase, $\mathbf{X}_i + \mathbf{E} \overset{K_i}{\rightleftharpoons} \mathbf{X}_i \cdot \mathbf{E}$. Starting from the 3'-end of the template $\mathbf{X}_i$ the new RNA molecule $\mathbf{X}_j$ is synthesized nucleotide by nucleotide along the sequence. The dimensionless factors Q_{ji} with $i, j = 1, \ldots, n$ provide the probabilities to obtain $\mathbf{X}_j$ as an (error) copy of the template $\mathbf{X}_i$. Conservation of probabilities requires: $\sum_{j=1}^{n} Q_{ji} = 1$ – a copy has to be either correct, Q_{ii}, or incorrect, $\sum_{j=1, j\neq i}^{n} Q_{ji} = 1 - Q_{ii}$. A very useful approximation is made by the *uniform error rate* assumption: The error per nucleotide and replication event, p, is assumed to be independent of the position and the nature of the nucleotide to be complemented. All elements of the mutation matrix can be expressed then by a simple formula,

$$\mathrm{Q}: \ Q_{ji} \ = \ p^{d_{ij}} \, (1-p)^{l-d_{ij}} \ = \ p^l \varepsilon^{d_{ij}} \ \text{ with } \ \varepsilon = \frac{p}{1-p} \, , \tag{1}$$

with only three parameters: (i) the sequence length l of the RNA molecule, (ii) the mutation parameter p, and (iii) the Hamming distance [19, 20] between the sequences interrelated by the mutation process, $d_{ij} = d_{\mathrm{H}}(X_i, X_j)$. Without changing the results that will turn out being important for the purpose pursued here, the analysis of the model is substantially simplified by the assumption of constant chain lengths l. This assumption is also consistent with the restriction to point mutations since point mutations do not change chain lengths by definition.

Aa an alternative to the Eigen model mutation can be seen, for example, as the result of DNA damage and imperfect damage repair during the whole life time of an organism, which is the idea in the Crow-Kimura mutation model [21, p. 264-266]. Then reproduction and mutation are completely independent processes and appear as additive terms in the kinetic differential equations. Interestingly, the Eigen and the Crow-Kimura model although different with respect to the underlying physics give rise to identical mathematical problems [2, pp.76-78]

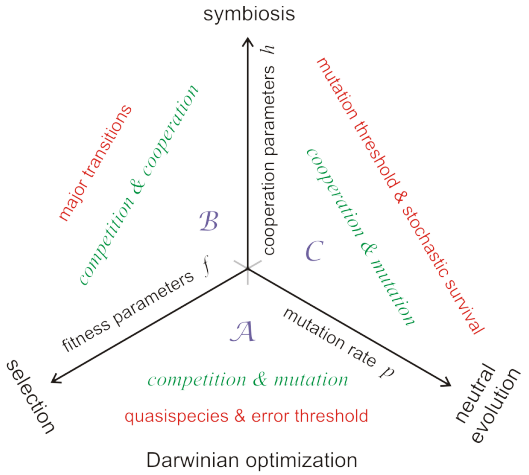

Figure 1. A minimal model for modeling evolution. Evolution is considered as an interplay of three processes: (i) competition through reproduction, (ii) cooperation through symbiosis, and (iii) mutation through error-prone replication. In parameter space the intensity parameters of all three processes, (i) fitness parameters f corresponding to reaction rates for competition, (ii) reaction rates h for catalyzed reproduction, and (iii) an error rate p for mutation are plotted on the axes of a Cartesian coordinate system. On the three two-dimensional faces of the coordinate system we are dealing with the three fundamental evolutionary processes: ($\mathcal{A}$) competitive reproduction and mutation are the basis of Darwinian optimization through natural selection and give rise to the formation of *quasispecies* and eventually to the occurrence of *error thresholds* [13,14]; ($\mathcal{B}$) the interplay of competition and cooperation allows for the description of *major transitions* [22, 23]; ($\mathcal{C}$) the combination of cooperation and mutation enables reintroduction of extinguished species provided the error rate is sufficiently large [24].

2 The evolution model

Evolution is properly illustrated as process in genotype or *sequence space*. Sequence space is a point space – every point is assigned uniquely to one sequence and mutations are moves in sequence space. A population consists of one or more sequences $\mathbf{X}_i\,(i = 1, \ldots, n)$, which are present in $[\mathbf{X}_i] = n_i(t)$ copies at time t. When the sequences are related by mutation we shall denote them as subspecies and call the set of subspecies a (molecular) species. The population size is $N = \sum_{i=1}^{n} n_i(t) = N(t)$. Evolutionary dynamics can be visualized as a superposition of two different processes: (I) an internal dynamics within the population leading to a quasi-stationary state or oscillations of the population variables and (II) the migration of the population in sequence space. In the strong selection - weak mutation scenario mutation is a rare event and the two processes I and II are well separated on the time scale. A useful diagnostic of the two processes considers the *consensus sequence* of the population.[1] Internal dynamics does not necessarily change the consensus sequence and it stays essentially constant at some point in sequence space. Migration of the consensus sequence is just a simplified version of process II replacing the population by its *center of gravity*. Here we are interested in a common mutation scenario that works equally well for weak and strong mutation and in the general case we cannot assume separability of selection and mutation on the time scale. In addition cooperative dynamics complicates the situation and obscures separability. As a result we are dealing with *only one* population dynamics, which describes simultaneously equilibration of genotype distribution and migration of the population in sequence space.

In the model the three processes constituting evolutionary dynamics are determined by a single intensity parameter each: (i) The fitness parameters f determine selection in form of fitness differences Δf, (ii) the cooperation parameter h controls the catalytic efficiency of the catalytic molecule, and (iii) the mutation parameter p determines the frequency of point mutations. Figure 1 sketches the three-dimensional parameter space. Neutral evolution with no mutation in sense of Motoo Kimura [25] is situated at the origin of the coordinate system. The question of independence of the three parameters is worth being discussed. It may be doubtful from the point of view of a structural biologist that catalyzed replication is independent of the uncatalyzed process but sequence

[1] The consensus sequence considers a given positions in all sequences of the population and assigns the most frequent nucleotide – A, U(T), G or C – to this position.

space is so large – it contains κ^l points with $\kappa = 2$ for binary or $\{\mathsf{G}, \mathsf{C}\}$ and $\kappa = 4$ for natural $\{\mathsf{A}, \mathsf{U}, \mathsf{G}, \mathsf{C}\}$ sequences – that independent examples can be found for almost every combination of f and h. Independence of mutation is relatively easy to imagine when, for example, action of antiviral drugs is taken into account.

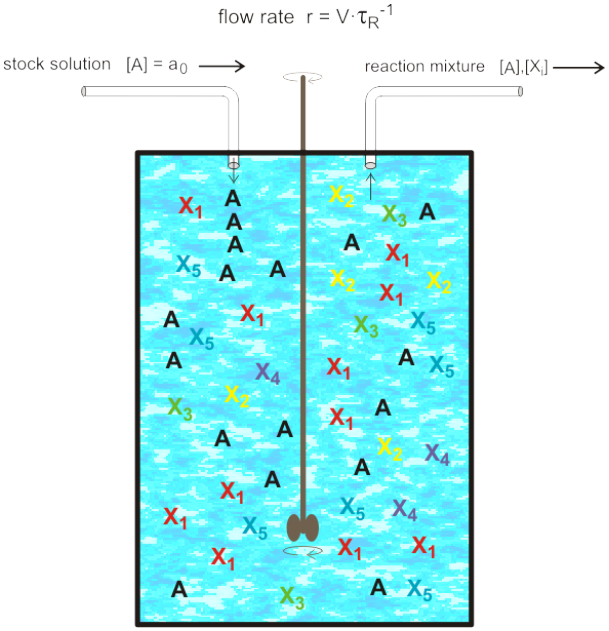

Figure 2. The continuous-flow stirred-tank reactor (CSTR)**.** The figure sketches a device for controlling the environmental condition of evolution experiments. The material needed for reproduction is subsumed by **A**, it flows into the reactor with a (volumetric) flow rate r [V/t] $\equiv$ [volume/time] in form of a solution with concentration $[\mathbf{A}] = a_0$. In the reactor molecules $\mathbf{X}_i$ $(i = 1, \ldots, n)$ are reproduced and **A** is consumed. The volume V of the reactor is constant and hence reaction mixture compensating the volume increase through influx of stock solution has to flow out of the reactor. The mean residence time of a volume element in the CSTR is $\tau_R = V/r$ [t].

2.1 Evolution in the flow reactor

Environmental influence on the outcome of evolution is commonplace. Consistent with a simple model are constant and controllable conditions, which can be realized in flow reactors (Figure 2; see also [26]). The reactor shown in the figure has the advantage of

being relatively unsophisticated for both, kinetic modeling and experimental investigations, and we shall apply it here. More elaborate reactors, which keep, for example, the numbers of bacterial cells constant have been designed and implemented [27–29]. Depending on the specific design and the concentration monitoring system the reactors are called, for example, *chemostat*, *cellstat* or *turbidostat*. More recently flow reactors and other experimental devices for monitoring and controlling evolution in the laboratory were introduced [30–34].

Next a reaction mechanism is conceived and implemented in the flow reactor. In addition to conventional chemical reactions we require $n + 2$ pseudo-reactions, which describe inflow and outflow of material into and from the reactor. The inflow is modeled by a zero-order reaction, $* \xrightarrow{a_0\, r} \mathbf{A}$, and outflow $\mathbf{S} \xrightarrow{r} \varnothing$ with $\mathbf{S} \in \{\mathbf{A}, \mathbf{X}_i (i = 1, \ldots, n)\}$ corresponds to a first-order reaction. Together with the remaining n^2 replication-mutation and n^2 catalyzed replication-mutation reactions we are dealing with the following set of $2n^2 + n + 2$ reactions in the flow reactor:

$$* \xrightarrow{a_0\, r} \mathbf{A}\,, \tag{2a}$$

$$\mathbf{A} + \mathbf{X}_i \xrightarrow{k_i\, Q_{ji}} \mathbf{X}_i + \mathbf{X}_j\,, \tag{2b}$$

$$\mathbf{A} + \mathbf{X}_i + \mathbf{X}_{i+1} \xrightarrow{l_i\, Q_{ji}} \mathbf{X}_i + \mathbf{X}_j + \mathbf{X}_{i+1}\,, \tag{2c}$$

$$\mathbf{A} \xrightarrow{r} \varnothing\,, \quad \text{and} \tag{2d}$$

$$\mathbf{X}_i \xrightarrow{r} \varnothing\,; \quad i, j = 1, \ldots, n;\ i \bmod n\,. \tag{2e}$$

Reaction (2a) supplies the material required for reproduction. A solution with $\mathbf{A}$ at concentration a_0 flows into a continuously stirred tank reactor (CSTR) with a flow rate r [26, p. 87ff.].[2] The reactor operates at constant volume and this implies that the volume per time unit [t] of solution flowing into the reactor is compensated exactly by an outflow, which is described by the equations (2d) and (2e) and concerns all $(n+1)$ molecular species, $\mathbf{A}$ and $\mathbf{X}_i$ $(i = 1, \ldots, n)$. We remark that all chemical reactions come to a standstill when the flow rate r exceeds a certain critical value, $r > r_{\mathrm{cr}}$. Then the mean residence time τ_R is too short for the completion of a single reactions event.

Pseudoreaction (2a) is a zeroth order reaction because there are no reactants, reactions (2d) and (2e) are first order pseudoreactions, reactions (2b) represent second order

[2] We remark that the deterministic kinetic equations (9) were extensively studied under the simplifying assumption of constant population size $\sum_{i=1}^{n} x_i = c_0 = \text{const}$ [5,13,35]. The solution curves formulated in relative concentrations $\xi_i(t) = x_i(t) \big/ \sum_{i=1}^{n} x_i(t)$ are identical for the CSTR and for constant population size but the stochastic system is unstable in the latter case [36].

reactions and finally reactions (2c) is of order three. Third order reaction are often ignored in elementary step reaction kinetics, because simultaneous collision of three or more atoms in the vapor phase are highly improbable and the same is true for ternary or higher encounters in solutions. Here we are not dealing with elementary steps and in overall kinetics the consideration of collision probabilities does not apply. In this contribution we shall classify reactions also according to the nature of the autocatalytic process: Reactions (2b) comprise simple copying processes – producing correct copies or error copies – and will be characterized as autocatalytic processes of first order, whereas reactions (2c) involve two replicator molecules acting as template and catalyst, and hence will be classified as second order autocatalysis. Such a distinction is meaningful because these processes differ with respect to the dynamical phenomena observed (section 3). The characteristic features of first order autocatalysis that are otherwise uncommon in chemistry include selection and optimization. The dynamic repertoire of second order autocatalysis is much richer and covers also oscillations, deterministic chaos, and spontaneous pattern formation in spatially resolved systems. In addition second order autocatalytic processes show much higher sensitivity than their first order counterparts with respect to stochastic phenomena caused by small particle numbers. Projected onto the problems discussed here, competitive reproduction gives rise to selection but catalyzed reproduction is needed for cooperation of competitors.

2.2 Deterministic kinetic equations

The kinetic differential equations of the model mechanism (2) are of the form:

$$\frac{\mathrm{d}a}{\mathrm{dt}} = -a\sum_{i=1}^{n}\left(k_i + l_i x_{i+1}\right)x_i + r\,(a_0 - a) \quad \text{and} \tag{3a}$$

$$\frac{\mathrm{d}x_i}{\mathrm{dt}} = a\left(\sum_{j=1}^{n} Q_{ij}\left(k_j + l_j x_{j+1}\right)x_j\right) - x_i\,r\,; \tag{3b}$$

$$i, j = 1, 2, \ldots, n;\ i \mod n\,.$$

In equation (3a) we made use of the conservation relation $\sum_{i=1}^{n} Q_{ij} = 1$. No analytical solutions are available for equation (3) in general but numerical integration is straightforward as long as n is not too large. In absence of cooperation terms equation (3) can be transformed into an eigenvalue problem of a symmetric matrix, which is readily diagonalized provided n is not very large ($n < 10^6$; section 4). In mutation-free systems, $p = 0$

(section 5), qualitative analysis and determination of stationary states is straightforward, and the dynamics of the complete system can be derived by extrapolation from the error-free results to finite mutations rates. The cooperation system with mutation (section 6) is used here as an example for the study of unconventional consequences of frequent replication errors. In stochastic modeling particle numbers are integers by definition whereas concentrations can adopt arbitrarily small values. Deterministic references are created in these cases by equating some initial variable to zero.

Extensive work related to the evolution model reported here was performed with replicator [37, 38] and Lotka-Volterra equations [39–42], which are equivalent from the mathematical point of view [43]. Indeed if the finiteness constraint providing a limitation for the population size is changed from the flow reactor to constant population size [44] the kinetic ODEs become identical and for vanishing mutation rates, $\lim p \to 0$, we obtain an inhomogeneous replicator equation for nonzero parameters f and h. In the limit $h \to 0$ the replicator equation is of first and for $f \to 0$ of second order.

2.3 Stochastic modeling

In principle the reaction equations (2) can be cast into a chemical master equation. The particle numbers of the molecular species, $[\mathbf{A}] = A(t)$ and $[\mathbf{X}_i] = X_i(t)$ with $i = 1, \ldots, n$, are integers and fulfil the conventional conservation relations, $C(t) = A(t) + \sum_{i=1}^{n} X_i(t) =$ const in absence of flows. The variables of the master equation are the probabilities

$$P_m(t) \;=\; \text{Prob}\Big(A(t) = m\Big) \;\text{ and }\; P_{s_i}(t) \;=\; \text{Prob}\Big(X_i(t) = s_i\Big)\;, \tag{4}$$

and the indices are subsumed in an index vector, $\mathbf{m} = (m, s_1, \ldots, s_n)$, which characterizes the state $S_{\mathbf{m}}$ of the system. The chemical master equation is based on the assumption that chemical processes occur one at a time, all jumps involve single steps and the particle numbers change by ± 1. The jumps $S_{\mathbf{m}'} \to S_{\mathbf{m}}$ or $S_{\mathbf{m}} \to S_{\mathbf{m}'}$ in the entire population are denoted by the shorthand notation

$$\begin{aligned}
\mathbf{m}' \;&=\; (m \pm 1, s_1, \ldots, s_n) \equiv (\mathbf{m}; m \pm 1) \;\text{ or} \\
\mathbf{m}' \;&=\; (m, s_1, \ldots, s_k \pm 1, \ldots, s_n) \equiv (\mathbf{m}; s_k \pm 1)\;.
\end{aligned}$$

Then the master equation of mechanism (2) takes on the form

$$\begin{aligned}\frac{\mathrm{d}P_{\mathbf{m}}}{\mathrm{dt}} &= a_0 r\Big(P_{(\mathbf{m};m-1)} - P_{\mathbf{m}}\Big) + r\Big((m+1)P_{(\mathbf{m};m+1)} - mP_{\mathbf{m}}\Big) + \\ &+ r\sum_{i=1}^{n}\Big((s_i+1)P_{(\mathbf{m};s_i+1)} - s_i P_{\mathbf{m}}\Big) + \\ &+ \sum_{i=1}^{n} Q_{ii}(k_i + l_i s_{i+1})\Big((m+1)(s_i-1)P_{(\mathbf{m};m+1,s_i-1)} - ms_i P_{\mathbf{m}}\Big) + \\ &+ \sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n} Q_{ij}(k_j + l_j s_{j+1})s_j\Big((m+1)P_{(\mathbf{m};m+1,s_i-1)} - mP_{\mathbf{m}}\Big)\ .\end{aligned} \tag{5}$$

Each reaction step changes the probability to be in state $\mathbf{m}$, $P_{\mathbf{m}}$, in two ways: It increases the probability through reactions or pseudoreactions $S_{\mathbf{m}'} \to S_{\mathbf{m}}$ and decreases the probability through the reaction steps $S_{\mathbf{m}} \to S_{\mathbf{m}'}$. The two terms in the first line, for example, describe the two pseudoreactions modeling inflow and outflow of the material **A**, and each reaction is represented by two steps. It is also worth noticing that stoichiometry requires two slightly different replication terms depending on whether the copy is correct or incorrect.

Although it is not difficult to write down a multivariate master equation, the derivation of analytical solutions is successful only in exceptional cases, for example for networks of monomolecular reactions [45,46]. An alternative strategy for studying chemical master equations is computer simulation through sampling of trajectories. The theoretical background for trajectory harvesting has been laid down by Andrey Kolmogorov [47], Willy Feller [48], and Joe Doob [49, 50]. With electronic computers now being generally available simulations of stochastic processes became possible. The conception, analysis, and implementation of a simple and highly efficient algorithm by Daniel Gillespie [51–53] provided a very useful tool for investigations of stochastic effects in chemical kinetics. Here we present results, which illustrate the differences between deterministic and stochastic solutions of the evolution model and use the deterministic solution as reference wherever this is meaningful.

3 First and second order autocatalysis

In the introduction we distinguished simple and catalyzed reproduction, which differed among other things in the formal order of the reaction. First order autocatalysis – reaction (2b) in the model mechanism – is a second order chemical reaction. A reaction

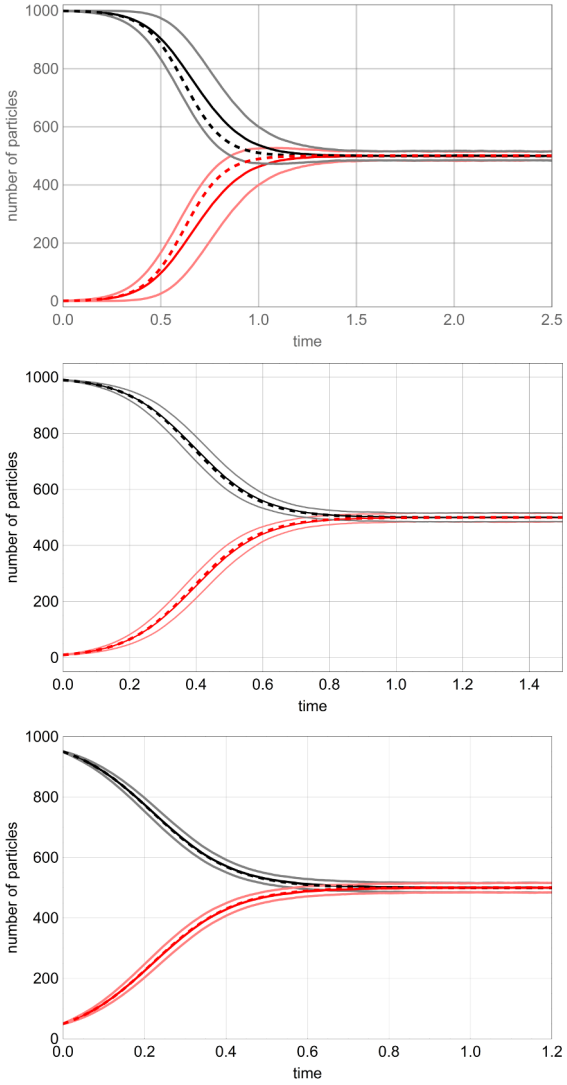

Figure 3. Comparison of fluctuations in the reversible first order autocatalytic reaction $\mathbf{A} + \mathbf{X} \rightleftharpoons 2\mathbf{X}$. Computations leading to the three figures are described in the caption of figure 4. Choice of parameters: $k = g = 0.01\,[\mathrm{M}^{-1}\,\mathrm{t}^{-1}]$, $N = 1000$ with $a_0 = 999$ and $x_0 = 1$ (upper part), $a_0 = 990$ and $x_0 = 10$ (middle part), $a_0 = 950$ and $x_0 = 50$ (lower part).

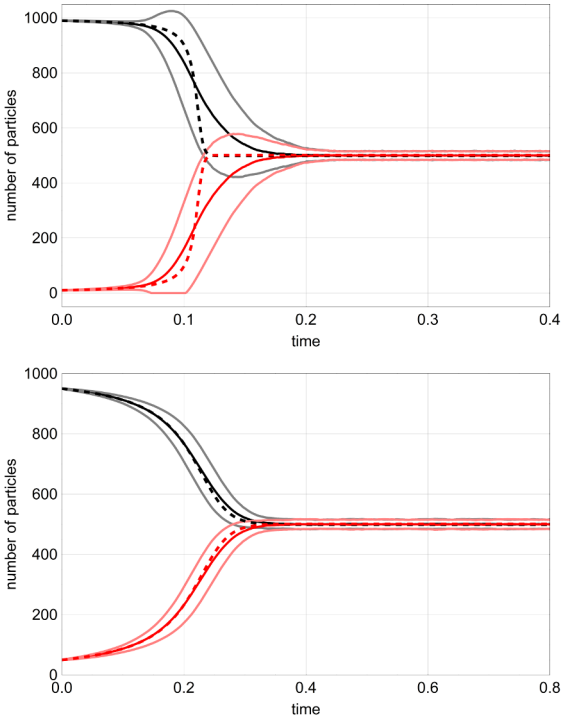

Figure 4. Comparison of fluctuations in the reversible second order autocatalytic reaction $\mathbf{A} + 2\mathbf{X} \rightleftharpoons 3\mathbf{X}$. The reaction is monitored in the closed system: $A(t) + X(t) = a_0 + x_0 = N = 1000$ where $A(0) = a_0$ and $X(0) = x_0$ serve as initial conditions. The figures present the results of statistical evaluation of 1 000 trajectories obtained by computer simulations with Gillespie's method [45]. Choice of parameters: $l = 0.001\,[\mathrm{M}^{-2}\,\mathrm{t}^{-1}]$, $q = 0.001\,[\mathrm{M}^{-2}\,\mathrm{t}^{-1}]$, $a_0 = 990$ and $x_0 = 10$ (upper part), $a_0 = 950$ and $x_0 = 50$ (lower part); color code: $E\big(A(t)\big)$: black, $E\big(A(t)\big) \pm \sigma_A(t)$: gray; $E\big(X(t)\big)$: red, $E\big(X(t)\big) \pm \sigma_X(t)$: pink; deterministic solutions are given as dotted lines, $a(t)$: black and $x(t)$: red. The case with the smallest possible number of x_0, $a_0 = 998$ and $x_0 = 2$ is not shown because of very large computer time required for a converged simulation (see text).

event requires an encounter of individuals from two species, **A** and **X**. Autocatalysis in the single step reaction $\mathbf{A} + \mathbf{X} \rightarrow 2\mathbf{X}$ shows characteristic features when compared to non-autocatalytic bimolecular reactions, for example $\mathbf{A} + \mathbf{B} \rightarrow 2\mathbf{C}$: (i) The probability distributions of $A(t)$ are broader than with non-autocatalytic processes and (ii) the rate is accelerated during the reaction giving rise to a solution curve with sigmoid

shape [54, pp.477-485]. In single step second order catalysis $\mathbf{A}+2\mathbf{X} \rightarrow 3\mathbf{X}$ two particles $\mathbf{X}$ have to meet one particle $\mathbf{A}$. In the reverse reaction even three particles $\mathbf{X}$ are required in an encounter. Any nucleation of a process requiring three partners shows strong stochastic effects at low particle numbers or concentrations. Both above mentioned features of first order autocatalysis are not only observed with second order autocatalytic processes as well but they are much stronger and become fully dominant at low concentrations. In particular an extremely strong dependence on the initial condition, $X(0) = x_0$, is found with small x_0-values.

Table 1. Fluctuations in the autocatalytic reactions $\mathbf{A}+\mathbf{X}\rightleftharpoons 2\mathbf{X}$ and $\mathbf{A}+2\mathbf{X}\rightleftharpoons 3\mathbf{X}$. The table presents the $\sigma(t_c)$-values and the equilibrium fluctuations calculated from sampling of 1 000 trajectories. The standard deviation of the non-autocatalytic reaction $\mathbf{A}+\mathbf{B}\rightleftharpoons 2\mathbf{C}$, $\sigma_{\mathbf{C}}(t)$ starts from $\sigma_{\mathbf{C}}(0) = 0$ and increases monotonously to the equilibrium value. Choice of parameters see figures 3 and 4. For the non-autocatalytic reaction the rate constant $k = l = 0.001\,[\mathrm{M}^{-1}\,\mathrm{t}^{-1}]$ were chosen. Initial conditions, $A(0) = B(0) = 750$ and $C(0) = 0$ were chosen

autocatalysis	**reaction**	**initial conditions**				
order		$x_0 = 1$	$x_0 = 2$	$x_0 = 10$	$x_0 = 50$	$t \rightarrow \infty$
zero	$\mathbf{A}+\mathbf{B}\rightleftharpoons 2\mathbf{C}$	–	–	–	–	18.4
first	$\mathbf{A}+\mathbf{X}\rightleftharpoons 2\mathbf{X}$	112	46	21	19	15.6
second	$\mathbf{A}+2\mathbf{X}\rightleftharpoons 3\mathbf{X}$	–	244	206	52	15.4

For the purpose of illustration two reversible autocatalytic reactions were chosen, one of first order and the other one of second order:

$$\mathbf{A}+\mathbf{X} \underset{g}{\overset{k}{\rightleftharpoons}} 2\mathbf{X} \quad \text{and} \quad \mathbf{A}+2\mathbf{X} \underset{q}{\overset{l}{\rightleftharpoons}} 3\mathbf{X}\,, \text{ respectively.}$$

In order to be able to compare fluctuations we choose equal rate parameters $k = g$ and $l = q$ and accordingly the equilibrium constants and concentrations are the same, $K = k/g = l/q = 1$ and $\overline{A} = \overline{X} = N/2$. In both autocatalytic reactions the two species are strongly correlated by the conservation relation $A(t)+X(t) = a_0+x_0 = N$, and hence variances and standard deviations are the same: $\mathrm{var}\big(A(t)\big) = \mathrm{var}\big(X(t)\big) = \mathrm{var}(t)$ and $\sigma\big(A(t)\big) = \sigma\big(X(t)\big) = \sigma(t)$. For the purpose of comparison a simple measure for scatter

is used: For $a_0 > x_0$ the width of the band

$$\Delta(t) = E\big(A(t)\big) + \sigma(t) - \big(E\big(X(t)\big) - \sigma(t)\big) = E\big(A(t)\big) - E\big(X(t)\big) + 2\sigma(t)$$

is recorded. It starts from $\Delta(0) = a_0 - x_0$, broadens in the intermediated range and eventually converges to the equilibrium value $\overline{\Delta} = 2\overline{\sigma}$. The standard deviation at the crossing point of the other two limits of the $\pm\sigma$-band at time $t = t_c$ fulfils $E\big(A(t_c)\big) - \sigma(t_c) = E\big(X(t_c)\big) + \sigma(t_c)$ and we calculate $\sigma(t_c) = \Delta(t_c)/4$ as a simple measure for the size of fluctuations. Table 1 presents some values obtained from simulations, which illustrate the dramatic differences between the different forms of autocatalysis at small initial values $X(0)$. Typical examples of trajectories are collected in figures 3 and 4. Figure 3 shows three different cases: (i) the largest possible scatter calculated with the initial conditions $a_0 = 999$ and $x_0 = 1$, (ii) an intermediate case with $a_0 = 990$ and $x_0 = 10$, and (iii) $a_0 = 950$ and $x_0 = 50$. As expected the difference in fluctuations concerns only the initial part of the curves and all curves look the same in the approach towards equilibrium. Analogous simulations were recorded for the second order case. The simulation with the smallest initial value of the autocatalyst, $a_0 = 998$ and $x_0 = 2$,[3] is not shown, because the enormous size of the fluctuations would require larger computer resources than available when the simulation is extended to near-equilibrium conditions.

In order to compare the fluctuations with a non-autocatalytic process we consider also the inversion of the bimolecular disproportionation reaction[4]

$$\mathbf{A}+\mathbf{B} \underset{l}{\overset{k}{\rightleftharpoons}} 2\mathbf{C} \ ,$$

which has the closest possible stoichiometry to the first-order autocatalytic reaction. Conditions were chosen such that the equilibrium concentrations are the same, $[\overline{X}] = [\overline{C}]$, and the equilibrium fluctuations of all three reactions are very close with the value for the uncatalyzed reaction being a little larger (table 1). The time courses of the standard deviations of both reactions $\sigma_X(t)$ and $\sigma_C(t)$ are shown in figure 5. Since we are using sharp initial conditions the standard deviation are zero at the beginning of the reaction: $\sigma_X(0) = \sigma_C(0) = 0$. In conventional chemical reactions $\sigma(t)$ approaches the equilibrium value either monotonously or goes through a flat maximum [54, pp. 447-449] whereas it passes a pronounced peak in autocatalytic reactions.

[3]The initial conditions $a_0 = 999$ and $x_0 = 1$ cannot ignite the reaction because two molecules **X** to convert A into X.

[4]An example of a disproportionation reaction in organic chemistry is the *Cannizzaro* reaction: $2\mathbf{C} \rightleftharpoons \mathbf{A} + \mathbf{B}$, where **C**, **A** and **B** are formaldehyde, formic acid and methanol, respectively.

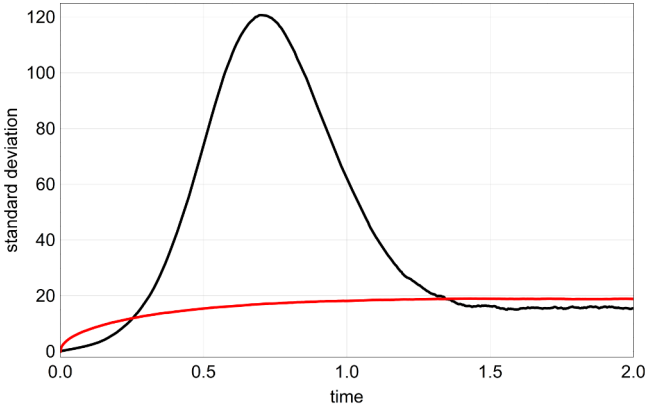

Figure 5. Comparison of the standard deviation in the reversible first order autocatalytic reaction $\mathbf{A} + \mathbf{X} \rightleftharpoons 2\mathbf{X}$ and uncatalyzed inverse disproportionation reaction $\mathbf{A} + \mathbf{B} \rightleftharpoons 2\mathbf{C}$. The reaction is monitored in the closed system, which implies: $A(t) + X(t) = a_0 + x_0 = \text{const} = N$ where $A(0) = a_0$ and $X(0) = x_0$ for the first reaction and $A(t) + B(t) + C(t) = \text{const} = a_0 + b_0 + c_0 = N$. The figures present the results of statistical evaluation of 1 000 trajectories obtained by computer simulations with Gillespie's method [45]. Choice of parameters: $N = 1000$, $k = g = 0.01\,[\mathrm{M}^{-1}\,\mathrm{t}^{-1}]$ for the autocatalytic reaction ($\sigma_X(t) = \sigma(t)$, black) and $k = l = 0.0003\,[\mathrm{M}^{-1}\,\mathrm{t}^{-1}]$ for the inverse disproportionation reaction ($\sigma_C(t)$, red) and hence we have $K = k/g = k/l = 1$ for the equilibrium parameters; initial conditions: $x_0 = 1$, $a_0 = 999$ and $c_0 = 0$, $a_0 = b_0 = 750$ yielding the same equilibrium values $\overline{A} = \overline{X} = 500$ or $\overline{A} = \overline{B} = \overline{C} = 500$.

The difference between first and second order autocatalysis is the basis of the different dynamics observed with competitive and cooperative systems. At the same time it provides the explanation of the different size of fluctuations that we shall encounter in sections 5 and 6.

4 Competition, mutation and quasispecies

The bottom face of the three-dimensional parameter space – $\mathcal{A}$ in figure 1 – is dealing with the combination of competitive selection and mutation. Natural selection and mutation are sufficient for a Darwinian scenario of competitive evolution that consists in selection of a fittest distribution of subspecies characterized as *quasispecies*. *In vitro* evolution has been verified experimentally in a great number of investigations (For an overview of

early works on this subject see [3, 55], as a recent review we mention [56]). The kinetic equations describing replication and replication induced mutation were introduced 1971 by Manfred Eigen in his scholarly written paper [13]. Eigen's selection equations describe the evolution of the distribution of asexually reproducing genotypes in a population of constant size N. Correct replication and mutation are seen as parallel chemical reactions and replication independent mutations are neglected. The selection-mutation scenario was found to provide the molecular basis for an understanding of virus evolution (A recent survey is found in the contributed volume [57]).

A population consists of several genotypes that are present in time dependent concentrations,

$$\boldsymbol{\Upsilon}(t) = x_1(t)\mathbf{X}_1 \oplus x_2(t)\mathbf{X}_2 \oplus \ldots \oplus x_n(t)\mathbf{X}_n \tag{6}$$

and the stationary solution, $\lim_{t\to\infty} \boldsymbol{\Upsilon}(t) = \overline{\boldsymbol{\Upsilon}} = \overline{x}_1\mathbf{X}_1 \oplus \overline{x}_2\mathbf{X}_2 \oplus \ldots \oplus \overline{x}_n\mathbf{X}_n$, is called the *quasispecies* [14]. The quasispecies $\overline{\boldsymbol{\Upsilon}}$ is the unique deterministic long-time solution of the replication mutation problem for infinite population size. In the continuous description the quasispecies contains all species at finite positive concentrations, although they may be unrealistically small since numbers of molecules are positive integers and can't be less than one. In reality we are dealing therefore with a truncated distribution denoted as *discrete quasispecies*:

$$\widetilde{\boldsymbol{\Upsilon}} \;=\; \overline{\eta}_1\mathbf{Y}_1 \oplus \overline{\eta}_2\mathbf{Y}_2 \oplus \ldots \oplus \overline{\eta}_n\mathbf{Y}_n \;\text{ with }\; \overline{\eta}_i \;=\; \begin{cases} \lfloor \overline{x}_i \rfloor & \text{if } \overline{x}_i \geq 1 \\ 0 & \text{if } \overline{x}_i < 1 \end{cases} . \tag{7}$$

The major influence of the mutation parameter p concerns the width of the quasispecies distribution. In the mutation-free system, $p = 0$, *survival of the fittest* in its pure form is observed and the quasispecies is of the form: $\overline{\boldsymbol{\Upsilon}} = \overline{x}_m\mathbf{X}_m$ with m being defined by $f_m = \max\{f_i; i = 1, \ldots, n\}$. Nonzero mutation rates have a distinct influence on the stationary state of the mutation-selection problem: Not a single genotype is selected but a family of genotypes $\widetilde{\boldsymbol{\Upsilon}}$, which consist of a most frequent sequence, the *master sequence*, together with a distribution of mutants that have sufficiently high fitness and are situated close by the master in sequence space. The quasispecies becomes broader with increasing mutation parameter p until a threshold value p_{cr} is reached above which error propagation is too strong for sustaining a stationary state. Interestingly, the critical mutation rate can be derived from the continuous quasispecies theory.

The replication-mutation mechanism follows from equation (2) by elimination of (2c).

Then the kinetic differential equations are obtained from equation (3) by setting $l_i = 0\ \forall\, i = 1, \ldots, n$:

$$\frac{\mathrm{d}a}{\mathrm{dt}} \;=\; -a \sum_{i=1}^{n} k_i x_i \;+\; r(a_0 - a) \tag{8a}$$

$$\frac{\mathrm{d}x_i}{\mathrm{dt}} \;=\; a \sum_{j=1}^{n} Q_{ij} k_j \, x_j \;-\; r \, x_i \,, \;\; i = 1, \ldots, n \;. \tag{8b}$$

Equation (8) describes the formation of quasispecies in the flow reactor. Solutions can be expressed in terms of eigenvectors of the *value matrix* $\mathrm{W}(t) = \{W_{ij} = Q_{ij} k_j a(t) - r\delta_{ij};\, i, j = 1, \ldots, n\}$[5] and are essentially the same as obtained for constant population size: $\sum_{i=1}^{n} x_i(t) = c = \text{const}$ with $a(t) = a_0 = \text{const}$ and $f_i = k_i a_0$ [44].

Depending on the distribution of fitness values f_i in sequence space quasispecies may show a sharp transition in plots of $\Upsilon(p)$ against p. From $p = 0$ to $p = p_{\mathrm{cr}}$ the distribution changes smoothly from a single sequence to a broad mutant cloud around the master sequence. At the critical mutation parameter $p = p_{\mathrm{cr}}$ the distribution shows a transition to the uniform distribution $\mathcal{U}$: $\{\widetilde{x}_1 = \widetilde{x}_2 = \ldots = \widetilde{x}_{\kappa^l} = 1/\kappa^l\}$. Above the critical error rate mutations occur too often to allow for a faithful reproduction of the template sequence and the result is *random replication* – in the long run every sequence is obtained with the same probability. The transition has been characterized as *error threshold* since evolutionary dynamics does not sustain a structured long-time population at higher error rates, $p > p_{\mathrm{cr}}$. On typical fitness landscapes the error threshold sharpens with increasing chain length l and becomes a first order phase transition in the limit $l \to \infty$ [58, 59]. We mention that certain unrealistically smooth landscapes do not show error thresholds [60].

Darwinian evolution is modified in the *strong mutation scenario*.[6] Not a single fittest genotype is selected but a uniquely defined distribution of genotypes, which is determined by the largest eigenvector of the value matrix W that represents the long-time or stationary solution of equation (8). The mean fitness of populations is not always optimized since situations can be constructed in which the fitness is decreasing during the approach towards the stationary state. Such situations, however, are rather rare and Darwinian optimization still remains a powerful heuristic that applies to almost all scenarios. For

[5] By δ_{ij} with $i, j \in \mathbb{N}$ we denote Kronecker's δ: $\delta_{ij} = 1$ for $i = j$ and $\delta_{ij} = 0$ otherwise.

[6] Biologists [61, 62] and computer scientists commonly distinguish strong and weak mutation. The weak mutation scenario assumes that adaptive mutations are sufficiently rare and do not interfere with the selection process but initiate replacements of currently fittest genotypes by still fitter variants. The strong mutation scenario is characterized by sufficiently large values of p that give rise to quasispecies dynamics mentioned here (for details see [63]) and to mutation induced cooperative dynamics (section 6).

error rate parameters exceeding a critical value p_{cr} the largest eigenvector approaches the uniform distribution over the entire sequence space, which is the exact solution for the value $p = p_{\text{u}}$ leading to incorporations of correct and incorrect nucleotides with equal probabilities – for binary sequences this happens at $p_{\text{u}} = 1 - p_{\text{u}} = \frac{1}{2}$. In realistic populations the uniform distribution is incompatible with a discrete quasispecies (7). Instead randomly migrating populations are observed.

5 Competition and cooperation

Face $\mathcal{B}$ contains the combinations of parameters f and h, which determine competition and cooperation. In the flow reactor they can be expressed through kinetic rate parameters multiplied by the time dependent concentration of material **A**: $f_i = k_i a(t)$ and $h_i = l_i a(t)$. The reaction mechanism is obtained from (2) through neglecting mutation or formally setting $Q_{ij} = \delta_{ij}$.

5.1 Deterministic solutions

The differential equations are of the form

$$\frac{\mathrm{d}a}{\mathrm{dt}} \;=\; -a \sum_{i=1}^{n} x_i\big(k_i \,+\, l_i x_{i+1}\big) \;+\; r\,(a_0 - a) \quad \text{and} \tag{9a}$$

$$\frac{\mathrm{d}x_i}{\mathrm{dt}} \;=\; x_i\Big(a(k_i \,+\, l_i x_{i+1}) \,-\, r\Big);\; i = 1, 2, \ldots;\; i \mod n\,. \tag{9b}$$

Inspection of the terms that are critical for reproduction, $k_i + l_i x_{i+1}\,(i \bmod n)$, allows for a straightforward explanation of the concentration dependence. In the small concentration limit we have $k_i \gg l_i \bar{x}$ with $\bar{x} = \sum_{i=1}^{n} x_i/n$ and the system shows competition selecting for the species with largest fitness. In the opposite situation, in the high concentration limit, $k_i \ll l_i \bar{x}$, cooperative dynamics – here pure hypercycle dynamics [16, 44] – is observed. Hypercycle dynamics shows an interesting dependence on the number of cooperating species: (i) asymptotically stable stationary states are found for $n = 2$ and $n = 3$, (ii) the concentrations of four species oscillate with very slow convergence to the stable steady state, and (iii) for $n \geq 5$ the systems oscillate. The existence of a stable closed orbit has been proven for $n = 5$ [18]. With increasing number of subspecies, n, the shape of the oscillations approaches the form of a rectangular pulse [64] and the concentrations at the minima becomes smaller and smaller. This fact becomes particularly relevant in the

stochastic description, because subspecies are more and more likely to die out when the concentration come close to one molecule or less in the reaction volume. Since mutations play an important role in preventing systems from extinction, oscillating systems with $n = 4$ and $n = 5$ will be discussed in section 6.

Equating (9a) and (9b) to zero yields equations for the positions of the potentially stable stationary states in concentration space that can be solved analytically [24, 65]. Stability analysis of the maximal 2^n stationary states can be carried out straightforwardly: As simple examples we consider the positions of the stationary states as functions of the two external parameters a_0 and r. Results for the cases $n = 2$ and $n = 3$ are shown in figures 6 and 7. In case of $n = 2$ four states are possible: (i) the state of extinction S_0 where both species are extinct($\overline{x}_1 = \overline{x}_2 = 0$), (ii) the state of selection of $\mathbf{X}_1$, $S_1^{(1)}$ ($\overline{x}_1 \neq 0, \overline{x}_2 = 0$), (iii) the state of selection of $\mathbf{X}_2$, $S_1^{(2)}$ ($\overline{x}_1 = 0, \overline{x}_2 \neq 0$), (iv) the cooperative state S_2 where both species are present ($\overline{x}_1 \neq 0, \overline{x}_2 \neq 0$). Only one of the selection states is asymptotically stable, and according to what had been said before this is the state $S_1^{(m)}$ with $m : k_m = \max\{k_i; i = 1, \ldots, n\}$. Here we were choosing $k_2 > k_1$ and $S_1^{(2)}$ is asymptotically stable in the range $r/k_2 < a_0 < r/k_2 + (k_2 - k_1)/l_1$. The coordinates of the stationary states S_0, $S_1^{(1)}$, and $S_1^{(2)}$ are constants or linear functions of a_0. In case of S_2 the concentration of $\mathbf{A}$ is obtained as one root of the quadratic equation [24]:

$$\overline{a} = \tfrac{1}{2}\big(a_0 + \psi - \sqrt{(a_0 + \psi)^2 - 4r\phi}\big) \text{ with } \psi = k_1/l_1 + k_2/l_2,\ \phi = l_1^{-1} + l_2^{-1}.$$

From the square root follows that a flow rate $r < (a_0 + \psi)^2/(4\phi)$ is required for the existence of the state S_2. There is a second solution of the quadratic equation,

$$\overline{a}^* = \tfrac{1}{2}\big(a_0 + \psi + \sqrt{(a_0 + \psi)^2 - 4r\phi}\big),$$

which corresponds to an unstable stationary state S_2^* that separates the basins of attraction between S_2 and S_0. The sequence of stable stationary states as a function of increasing a_0 – representing the availability of resources – is

$$\textbf{extinction} \rightarrow \textbf{selection} \rightarrow \textbf{cooperation}\,.$$

The relative size of the reproduction parameters k_1 and k_2 determines the species which is selected, and $k_1 < k_2$ leads to a state with $\mathbf{A}$ and $\mathbf{X}_2$ present and $\mathbf{X}_1$ absent. It is worth noticing that the initiation of cooperation requires more resources than natural selection [66].

In the three species system we are dealing with $2^3 = 8$ stationary states, twice as many

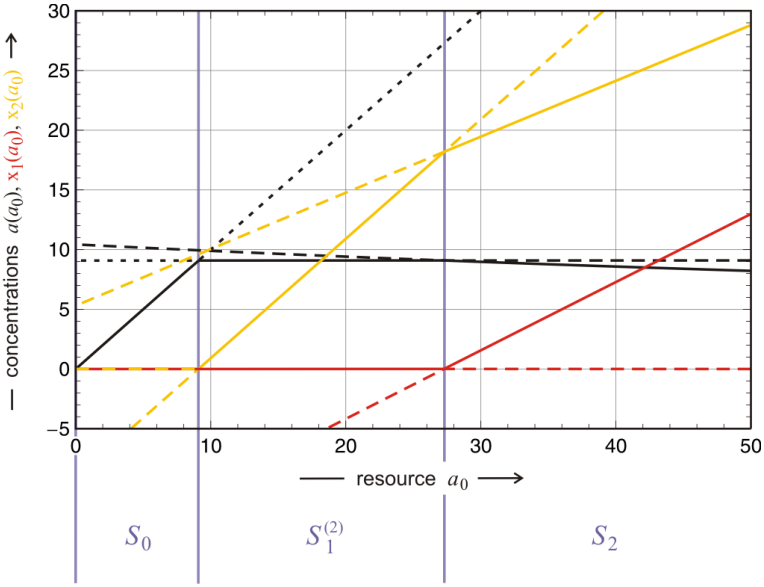

Figure 6. Sequence of bifurcations in the competition-cooperation system with two subspecies $n = 2$**.** The plot shows the stationary concentrations as functions of the input concentration: $\overline{a}(a_0)$ (black), $\overline{x}_1(a_0)$ (red), and $\overline{x}_2(a_0)$ (yellow). Two transcritical bifurcations $S_0 \leftrightarrow S_1^{(2)}$ and $S_1^{(2)} \leftrightarrow S_2$ occur at $a_0 = 100/11 \approx 9.091$ and $a_0 = 300/11 \approx 27.273$, respectively. Full lines refer to the asymptotically stable states, broken lines to unstable states. Choice of parameters: $k_1 = 0.09$, $k_2 = 0.11\,[\mathrm{M}^{-1}\mathrm{t}^{-1}]$; $1_1 = 0.0011$, $l_2 = 0.0009\,[\mathrm{M}^{-2}\mathrm{t}^{-1}]$

states as in case $n = 2$, but the repertoire of asymptotically stable states is enlarged only by one, by a state where two species are present and one is eliminated or excluded – it is called the *state of exclusion* therefore:

extinction → selection → exclusion → cooperation .

As in the previous case, $n = 2$, the sequence of stable states depends on the relative size of the rate parameters k_1, k_2, k_3, l_1, l_2, and l_3. In figure 7 we sketch the transcritical bifurcation for special values of the parameters with relative sizes $k_1 < k_2 < k_3$ and $l_1 > l_2 > l_3$. Then the selection leads to a state with only $\mathbf{A}$ and $\mathbf{X}_3$ present and in the exclusion state $\mathbf{X}_1$ is absent. Calculation of the transcritical bifurcations yields the analogous results for the stability range as given above for $n = 2$:

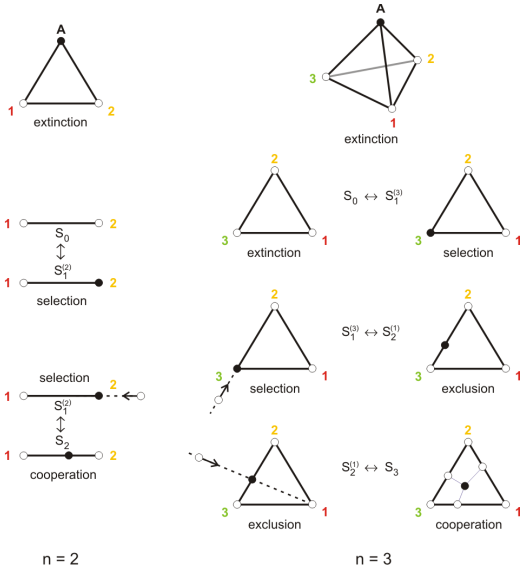

Figure 7. The bifurcations in the competition-cooperation system with $n = 2$ and $n = 3$. The top row shows the states of extinction on the complete concentration spaces for three ($\mathbf{A}$, $\mathbf{X}_1$, $\mathbf{X}_2$) and four ($\mathbf{A}$, $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{X}_3$ species). All other diagrams refer to projections onto reduced spaces containing only the reproductive species, the unit simplices $\mathbb{S}_n^{(1)}$: $\Xi_n = (\bar{\xi}_i; i = 1, \ldots, n)$ with $\bar{\xi}_i = \bar{x}_i / \sum_{j=1}^{n} \bar{x}_j$, $0 \leq \bar{\xi}_i \leq 1$, and $\sum_{i=1}^{n} \bar{\xi}_i = 1$. The sketch on the left hand side shows the concentration space for n=2 with two transcritical bifurcations $S_0 \leftrightarrow S_1^{(2)}$ and $S_1^{(2)} \leftrightarrow S_2$ constituting the sequence *extinction→selection→cooperation* with increasing values of a_0 (figure 6). The right hand side shows the analogous diagrams for $n = 3$ with three transcritical bifurcations: $S_0 \leftrightarrow S_1^{(2)}$, $S_1^{(2)} \leftrightarrow S_2^{(1)}$, and $S_2^{(1)} \leftrightarrow S_3$ giving rise to *extinction→selection→exclusion→cooperation*. Choice of parameters: $k_1 = 0.09$, $k_2 = 0.11$ ($n = 2$) and $k_2 = 0.10$ ($n = 3$), $k_3 = 0.11\,[\mathrm{M}^{-1}\mathrm{t}^{-1}]$; $1_1 = 0.0011$, $l_2 = 0.0009$ ($n = 2$) and $l_2 = 0.0010$ ($n = 3$), $l_3 = 0.0009\,[\mathrm{M}^{-2}\mathrm{t}^{-1}]$.

Table 2. Probabilities to reach quasi-stationary states in the cooperative regime with $n = 2$ and different initial conditions. The table provides probabilities of occurrence for all four possible long-term states: extinction S_0, selection of $\mathbf{X}_1$ in $S_1^{(1)}$, selection of $\mathbf{X}_2$ in $S_1^{(2)}$, or cooperation S_2. The counted numbers of events are sample means and unbiased standard deviations calculated from ten packages, each of them containing 10 000 trajectories computed with identical parameters and initial conditions, and differing only in the sequence of random events determined by the seeds of the pseudorandom number generator (*Extended CA, Mathematica*). Choice of parameters: $k_1 = 0.09\,[\mathrm{M^{-1}t^{-1}}]$, $k_2 = 0.11\,[\mathrm{M^{-1}t^{-1}}]$, $l_1 = 0.0011\,[\mathrm{M^{-2}t^{-1}}]$, $l_2 = 0.0009\,[\mathrm{M^{-2}t^{-1}}]$, $a_0 = 200$, $r = 0.5\,[\mathrm{V\,t^{-1}}]$. Initial value $A(0) = 0$. Probabilities are obtained by multiplication by 10^{-4}.

Initial values		**Counted numbers of states in final outcomes**			
$X_1(0)$	$X_2(0)$	N_{S_0}	$N_{S_1^{(1)}}$	$N_{S_1^{(2)}}$	N_{S_2}
1	1	385.1 ± 23.6	1481.0 ± 36.8	1719.6 ± 37.8	6414.3 ± 53.8
2	1	77.4 ± 9.1	1822.6 ± 41.6	367.6 ± 17.0	7733.3 ± 38.3
1	2	71.6 ± 8.5	280.6 ± 20.0	2075.8 ± 28.9	7572.0 ± 39.2
3	1	15.0 ± 2.9	1900.4 ± 30.9	74.6 ± 10.0	8009.0 ± 35.3
1	3	14.0 ± 3.7	53.1 ± 4.8	2180.5 ± 48.4	7752.3 ± 53.8
2	2	14.9 ± 2.6	303.7 ± 16.0	354.5 ± 23.8	9326.8 ± 44.9
3	3	0	70.2 ± 10.0	106.2 ± 10.9	9823.4 ± 15.7
4	4	0	12.1 ± 2.6	28.0 ± 5.0	9959.9 ± 6.4
5	5	0	2.5 ± 1.1	6.3 ± 2.6	9991.2 ± 3.0

	stability range
extinction	$0 < a_0 < \frac{r}{k_3}$,
selection of $\mathbf{X}_3$	$\frac{r}{k_3} < a_0 < \frac{r}{k_3} + \frac{k_3-k_2}{l_2}$,
exclusion of $\mathbf{X}_1$	$\frac{r}{k_3} + \frac{k_3-k_2}{l_2} < a_0 < \frac{r}{k_3} + \frac{k_3-k_2}{l_2} + \frac{k_3-k_1}{l_1}$,
cooperation	$\frac{r}{k_3} + \frac{k_3-k_2}{l_2} + \frac{k_3-k_1}{l_1} < a_0$.

A more detailed discussion of the bifurcation pattern can be found in [24, pp. 8-10].

5.2 Stochastic solutions

The stochastic system describing competition and cooperation, equations (2) with $Q_{ij} = \delta_{ij}$, sustains one asymptotically stable stationary state corresponding to an absorbing boundary, the state of extinction S_0 defined by $\lim_{t\to\infty} A(t) = a_0$, and $\lim_{t\to\infty} X_1(t) = 0$, and $\lim_{t\to\infty} X_2(t) = 0$. For $n = 2$ we are dealing in addition with three quasi-stationary

states:[7]

(i) selection of $\mathbf{X}_1$ denoted by $S_1^{(1)}$ with $\lim_{t\to\infty}\overline{A}(t) = r/k_1$,
$\lim_{t\to\infty}\overline{X}_1(t) = a_0 - r/k_1$, and $\lim_{t\to\infty}\overline{X}_2(t) = 0$,

(ii) selection of $\mathbf{X}_2$ denoted by $S_1^{(2)}$ with $\lim_{t\to\infty}\overline{A}(t) = r/k_2$,
$\lim_{t\to\infty}\overline{X}_1(t) = 0$, and $\lim_{t\to\infty}\overline{X}_2(t) = a_0 - r/k_2$,

(iii) cooperation denoted by S_2 with $\lim_{t\to\infty}\overline{A}(t) = \alpha$,
$\lim_{t\to\infty}\overline{X}_1(t) = (r - k_2\alpha)/(l_2\alpha)$, and $\lim_{t\to\infty}\overline{X}_2(t) = (r - k_2\alpha)/(l_2\alpha)$,
where $\alpha = \frac{1}{2}\big(a_0 + \psi - \sqrt{(a_0+\psi)^2 - 4r\phi}\big)$ as before.

The stochastic system with $n = 3$ has an absorbing boundary at the state of extinction, $\overline{A} = a_0$, $\overline{X}_1 = \overline{X}_2 = \overline{X}_3 = 0$, which is the only asymptotically stable stationary state, three selection states, $S_1^{(1)}$, $S_1^{(2)}$ and $S_1^{(3)}$, three exclusion states $S_2^{(1)}$, $S_2^{(2)}$ and $S_2^{(3)}$, and one cooperative state S_3. It is worth noticing that all stationary states of the deterministic system show up as stochastic quasi-stationary states, no matter whether they are asymptotically stable or not in the deterministic approach.

The rate parameters, which determine the stability in the deterministic case, provide the basis for probabilities of occurrence in the stochastic system. In table 2 we present the probabilities for the case $n = 2$ with a set of parameters belonging to the cooperative regime in the deterministic case. As expected we observe very strong dependence on initial conditions. Since the state of extinction is an absorbing boundary initial conditions $X(0) = 1$ require only a single death event in order to eliminate the corresponding subspecies irreversibly. The results shown by the table are remarkable in two aspects: (i) initially five copies of each subspecies are sufficient for reducing the probability of not ending up in the cooperative state S_2 to a few thousandth and (ii) the probability of going extinct, P_{S_0}, is essentially determined by the sum of the initially present particles, $X_1(0) + X_2(0)$, and not by the distribution. Stochastic solutions of the competition-cooperation system were derived through simulation of individual trajectories [53]. Examples of trajectories for $n = 2$ and $n = 3$ are shown in figures 8 and 9. A typical trajectory allows for the distinction of different phases in the approach towards a long-time state (figures 8): (i) an initial phase I during which the stationary concentration of $\mathbf{A}$ is established, (ii) a stochastic phase II during which the system is canalized towards

[7] A stochastic quasi-stationary state is a state towards which the system converges stochastically in the long-time limit and around which it fluctuates. It is not an absorbing state, and if true asymptotically stable stationary states exist the system converges to one of them it in the limit $t \to \infty$.

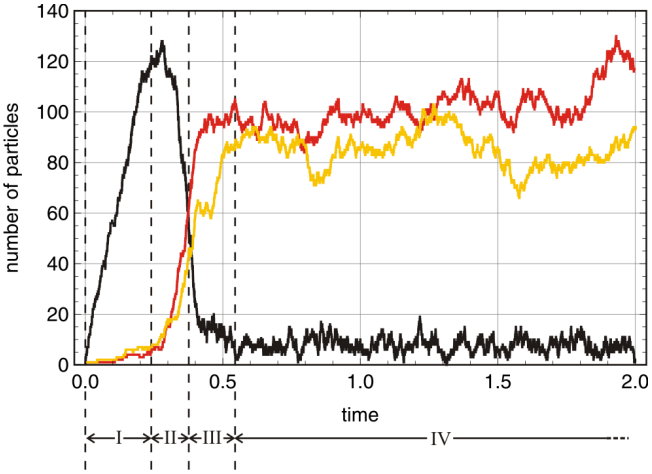

Figure 8. Sequence of phases in the approach towards a quasi-stationary state for $n = 2$. A stochastic trajectory simulating competition and cooperation of two species in a flow reactor is shown in the plot above. The stochastic process is assumed to start with an empty reactor except seeds for the two autocatalysts and can be partitioned into four phases: (I) fast raise in the concentration of A, (II) a random phase where the decision is made into which final state – S_0, $S_1^{(1)}$, $S_1^{(2)}$ or S_2 – the trajectory progresses, (III) the approach towards the final state, and (IV) fluctuations around the values of the (quasi)stationary state. Parameter values: $k_1 = 0.099$, $k_2 = 0.101\,[\mathrm{M^{-1}t^{-1}}]$, $l_1 = 0.0050$, $l_2 = 0.0045\,[\mathrm{M^{-2}t^{-1}}]$, $a_0 = 200$, $r = 4.0\,[\mathrm{V\,t^{-1}}]$, pseudorandom number generator: *ExtendedCA, Mathematica*, seeds $s = 631$. Initial conditions: $A(0) = 0$, $X_1(0) = X_2(0) = 1$. Color code: $\mathsf{A}(t)$ black, $\mathsf{X}_1(t)$ red, and $\mathsf{X}_2(t)$ yellow.

one of the quasi-stationary states, (iii) phase III dealing with the statistical convergence to the quasi-stationary state and (iv) phase IV describing random fluctuations around the steady state. Phase I is determined by the initial conditions for material $\mathbf{A}$ and for the initially empty reactor, which are $A(0) = 0$ and seeding quantities for $\mathbf{X}_i$ $(i = 1, \ldots, n)$, in particular $X_i(0) \leq 10$, and consists of overshooting of $A(t)$ followed by the decrease to the stationary value. Phase II contains the various extinction patterns of subspecies from extinction of all of them to cooperation of all of them. Phase I and phase II can show substantial overlap or may even coincide on the time axis. Figure 9 presents trajectories for the system with three subspecies ($n = 3$). Parameters for the upper plot were chosen to lie in the cooperative regime and extension of the computations to longer

times did not lead to extinction of subspecies. Calculation of the position of bifurcations yields $(a_0)_{cr} = 46.4$ and with $a_0 = 400$ the point in parameter space is well situated in the interior of the cooperative regime. For the parameters of the lower plot we calculate $(a_0)_{cr1} = 351.5$ and $(a_0)_{cr2} = 957.6$ and $a_0 = 400$ falls in the exclusion regime with $\overline{X}_1 = 0$. Indeed this variable stays at low values and vanishes around $t = 8\,[\mathrm{t}]$ and we may think we are dealing with the stochastic analogue of the stable deterministic state $S_2^{(1)}$. After some time at $t \approx 100\,[\mathrm{t}]$, however, $\mathbf{X}_2$ dies out as well and the long-time result is selection of $\mathbf{X}_3$. Without mutation subspecies that have died out cannot be reintroduced into the system and therefore every quasi-stationary state can be transformed into another one only through extinction of a subspecies by a sequence of death events or a sufficiently large fluctuation.

Comparing stochasticity in the selection system and the cooperation systems we observe a striking difference that boils down to the impact of catalysis in replication dynamics, which manifests itself in the difference between first and second order catalysis and is already evident from the deterministic systems (see, for example, [67, pp.18-27 and pp.70-75], figures 3 and 4). Fluctuations along the transients in the first order autocatalytic reaction $\mathbf{A} + \mathbf{X} \rightarrow 2\mathbf{X}$ are larger than with non-autocatalytic reactions (see figure 5 and [54, pp.477-485]) but the difference between first order and second order autocatalysis shown, for example, by the reaction $\mathbf{A} + 2\mathbf{X} \rightarrow 3\mathbf{X}$ is much more dramatic. This is already seen in the approach towards equilibrium of the two autocatalytic processes and has the ultimate consequence that second order autocatalysis leads to complex dynamical phenomena like oscillations, deterministic chaos and spontaneous pattern formation in homogeneous solutions. Indeed oscillations and deterministic chaos were also found with generalized hypercycles [68]. The molecularity of the process provides a strong hint for the dependence on the resource parameter a_0. Catalyzed replication is – at least – termolecular, and termolecularity implies an encounter of three agents, template, replicase and catalyst that has very low probability at low concentrations and accordingly is highly sensitive to fluctuations.

The scenario leading from competition to cooperation is induced by increasing the cooperation parameter $h = l \cdot a_0$ and can be studied, for example, through variation of h at constant a_0 or variation of a_0 at constant h. We chose here the former strategy and progress from symbiosis to selection. At sufficiently large values of h the system follows

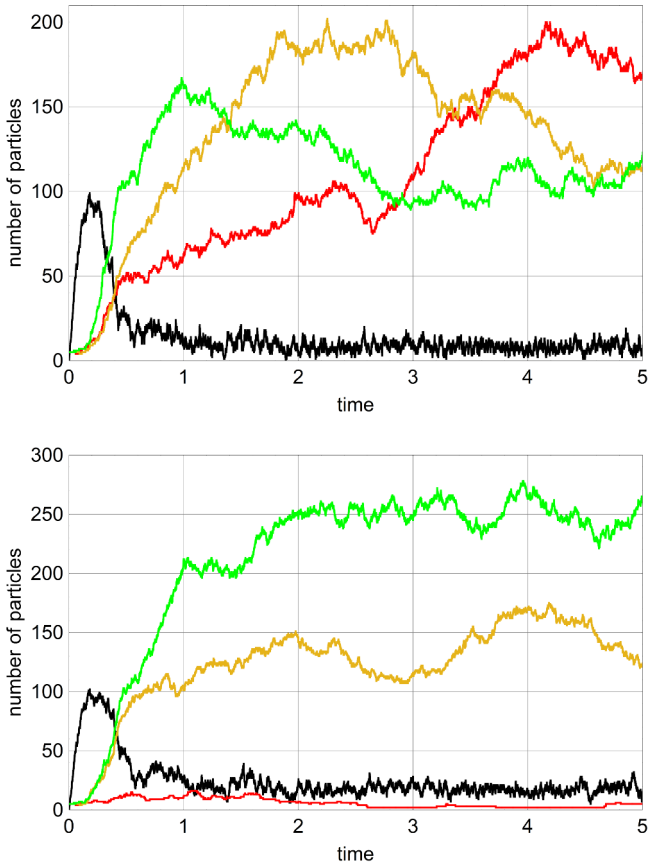

Figure 9. Sequence of phases in the approach towards a quasi-stationary state for $n = 3$. Two stochastic trajectories simulating competition and cooperation of three species in a flow reactor are shown in the plots above. The stochastic process is assumed to start with an empty reactor except seeds for the three autocatalysts. Parameter values: $k_1 = 0.09$, $k_2 = 0.10$, $k_3 = 0.11\,[\mathrm{M^{-1}t^{-1}}]$, $l_1 = 0.0011$, $l_2 = 0.0010$, $l_3 = 0.0009\,[\mathrm{M^{-2}t^{-1}}]$ (upper plot), $l_1 = 0.000033$, $l_2 = 0.000030$, $l_3 = 0.000027\,[\mathrm{M^{-2}t^{-1}}]$ (lower plot), $a_0 = 800$, $r = 4.0\,[\mathrm{V\,t^{-1}}]$, pseudorandom number generator: *ExtendedCA, Mathematica*, seeds $s = 491$. Initial conditions: $A(0) = 0$, $X_1(0) = X_2(0) = X_3(0) = 5$. Color code: $\mathsf{A}(t)$ black, $\mathsf{X}_1(t)$ red, $\mathsf{X}_2(t)$ yellow, and $\mathsf{X}_3(t)$ green.

hypercycle dynamics of dimension n with the characteristically large fluctuations because of second order self-enhancement. Reducing the parameter h increases the fluctuations and then the stochastic systems with $n = 2$ and $n = 3$ pass a series of quasi-stationary states with decreasing number of non-vanishing subspecies: $n \to n-1 \to n-2 \to \ldots \to 1$. At $h = l = 0$ simple selection behavior of the fittest subspecies is observed, increasing h leads to longer and longer survival times of the variant with lowest fitness until a quasi-stationary state with $n = 2$ is obtained. This is the cooperative state for $n = 2$ and the state of exclusion of the least fit subspecies in case of $n = 3$. Further increase in h for $n = 3$ results in the state of cooperation (figure 9). The concentrations in systems with $n \geq 4$ oscillate and will be discussed in the next section 6. In essence, the scenario is the same as in the deterministic case with the already mentioned change that all quasi-stationary states are possible as long time solutions and instability of the corresponding deterministic states results only in a smaller probability of occurrence.

6 Cooperation and mutation

The third coordinate face of the parameter space ($\mathcal{C}$) contains symbiotic or hypercyclic systems with mutation. The major role of mutation is the reintroduction of subspecies that had died out previously into the system that is particularly important for hypercycles with $n \geq 4$, which sustain oscillations. The model is obtained through elimination of reaction (2b) from the mechanism (2) and the kinetic differential equations are of the form:

$$\frac{\mathrm{d}a}{\mathrm{d}t} \;=\; -a \sum_{i=1}^{n} l_i\, x_i x_{i+1} \;+\; r\,(a_0 - a) \quad \text{and} \tag{10a}$$

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} \;=\; a \sum_{j=1}^{n} Q_{ij}\, l_j\, x_j x_{j+1} \;-\; x_i\, r;\; i = 1, 2, \ldots;\; i, j \mod n\,, \tag{10b}$$

where the mutation rate is given by equation (1). Equation (10) cannot be solved analytically but numerical integration is straightforward. In a way the dynamical system is simpler than the competition-cooperation case because it sustains only two asymptotically stable stationary states, which are easily obtained from the conditions $\mathrm{d}a/\,\mathrm{d}t = 0$ and $\mathrm{d}x_i/\,\mathrm{d}t = 0$ for $i = 1, \ldots, n$. One state is the state of extinction S_0 with $\overline{a} = a_0$ and $\overline{x}_i = 0$ $(i = 1, \ldots, n)$ and the second state S_n is the cooperative state with $\overline{x}_i = r/(l_{i-1}\overline{a})$

Table 3. Long-time behavior in the cooperation-mutation system with $n = 3$ and different initial conditions. The table provides expectation values at the time of the end of the simulation ($t = 30$) for different mutation rate parameters p and different initial conditions. Choice of parameters: $l_1 = 0.011$, $l_2 = 0.010$, $l_3 = 0.009\,[\mathrm{M}^{-2}\mathrm{t}^{-1}]$, $a_0 = 400$, $r = 0.5\,[\mathrm{V\,t}^{-1}]$. Initial value $A(0) = 0$.

Mutation	**Initial values**			**Expectation values at** $t = 30$				**Counts**
p	$X_1(0)$	$X_2(0)$	$X_3(0)$	$E(A)$	$E(X_1)$	$E(X_2)$	$E(X_3)$	$P(S_3)$
0	1	1	1	279.3	44.56	36.06	39.58	0.301
0	2	2	2	81.53	117.7	94.9	105.4	0.814
0	3	3	3	14.56	142.5	115.5	128.3	0.971
0	4	4	4	2.023	146.1	119.6	132.2	0.996
0	5	5	5	1.153	146.0	120.2	132.8	1.0
0	10	10	10	0.376	147.0	120.0	132.6	1.0
0	deterministic			0.377	147.0	120.3	132.3	1
0.05	1	1	1	177.2	81.66	68.16	72.64	0.432
0.05	2	2	2	28.57	136.1	113.8	121.5	0.937
0.05	3	3	3	0.782	147.1	121.9	130.5	0.999
0.05	4	4	4	0.383	147.0	122.2	130.4	1.0
0.05	5	5	5	0.382	146.8	122.2	130.4	1.0
0.05	10	10	10	0.379	147.0	122.6	130.3	1.0
0.05	deterministic			0.377	146.7	122.3	130.5	1

where $\overline{a}$ is a root of a quadratic equation

$$\left(\overline{a}\right)_{1,2}^{(2)} \;=\; \frac{1}{2}\left(a_0 \mp \sqrt{a_0^2 \,-\, 4\,\phi\, r}\right) \quad \text{with} \quad \phi \;=\; \sum_{i=1}^{n} l_i^{-1}\;. \tag{11}$$

The root $\left(\overline{a}\right)_1^{(2)}$ with the minus sign corresponds to the asymptotically stable state S_n whereas the second root belongs to a saddle point S_n^* separating the basins of attraction of S_0 and S_n. For $0 < r < a_0^2/4\phi$ we have two asymptotically stable states whereas S_n and S_n^* do not exist at higher values of r (see also section5.1). Then, the only long-time solution of the system is extinction.

Stochastic systems modeling competition and cooperation with $n = 2, 3$ have the selection possibility to choose between 2^n stationary states, one absorbing boundary and $2^n - 1$ quasi-stationary states corresponding to n deterministic states, out of which one or two can be stable for a given set of parameters. In the cooperation-mutation system the

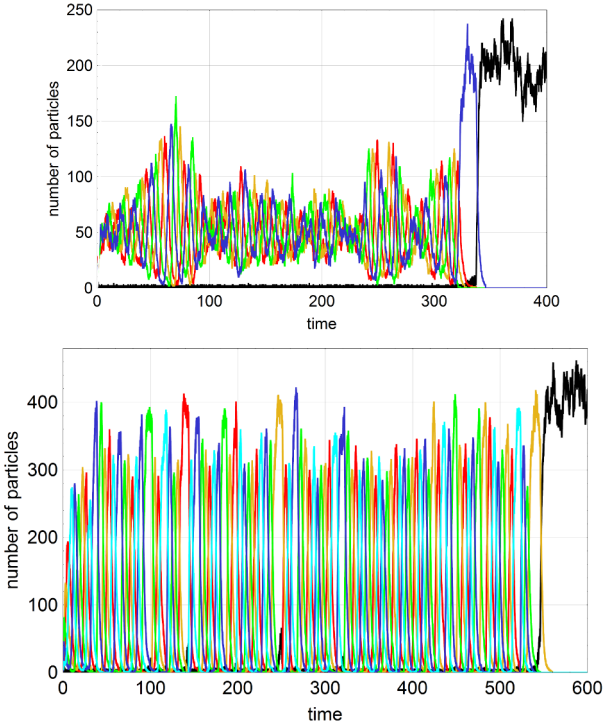

Figure 10. Extinction of oscillating hypercycles with $n = 4$ and $n = 5$. The figures show single stochastic trajectories computed by means of the Gillespie algorithm [53] for mechanism (2) in the flow reactor with $k_1 = k_2 = k_3 = k_4 = (k_5 =) 0$, $n = 4$ (upper plot) and $n = 5$ (lower plot). In both cases the oscillations grow until one species dies out, then the hypercycle is extinguished by the ratchet mechanism and only compound A remains. Choice of parameters, upper plot: $a_0 = 200$, $r = 0.5\,[\mathrm{V^{-1}t^{-1}}]$, $l_1 = l_2 = l_3 = l_4 = 0.1\,[\mathrm{M^{-2}t^{-1}}]$ and $p = 0.001$; lower plot: $a_0 = 400$, $r = 0.5\,[\mathrm{V^{-1}t^{-1}}]$, $l_1 = l_2 = l_3 = l_4 = l_5 = 0.01\,[\mathrm{M^{-2}t^{-1}}]$ and $p = 0.002$. Pseudorandom number generator: *Extended CA* (Mathematica 10), seed: $s = 089$ (upper plot) and $s = 919$ (lower plot). Initial conditions: $A(0) = 0$, $X_1(0) = X_2(0) = X_3(0) = X_4(0) = 4$ (upper plot) and $A(0) = 0$, $X_1(0) = X_2(0) = X_3(0) = X_4(0) = X_5(0) = 5$ (lower plot). Color code: $\mathsf{A}(t)$ black, $\mathsf{X}_1(t)$ red, $\mathsf{X}_2(t)$ yellow, $\mathsf{X}_3(t)$ green, $\mathsf{X}_4(t)$ blue, and $\mathsf{X}_5(t)$ cyan.

stochastic trajectories can only choose between two alternatives: the state of extinction S_0 and the state of cooperation S_n. The system with three-subspecies ($n = 3$) is easier to interpret and we start with it therefore. Considering the regime where the cooperative state S_n exists, $r < a_0^2/4\phi$, the decision whether a trajectory approaches S_0 or S_n in the long run falls in phase II (figure 8) and is determined by the initial conditions, $X_1(0)$, $X_2(0)$, and $X_3(0)$, as well as the sequence of random events – as controlled in the simulation by the seeds of the pseudorandom number generator. In table 3 we illustrate the long-time behavior of the non-oscillating cooperation-mutation system by the long-time expectation values of the variables $A(t)$ and $X_i(t)$ with $i = 1, 2, 3$ taken at a predefined time ($t = 30$) as well as the counted ratios of final states. The stochastic results calculated by Gillespie simulations are compared with the deterministic values obtained from integration of the kinetic ODEs. The mutation-free system ($p = 0$) shows significant differences only for the initial conditions $X_1(0) = X_2(0) = X_3(0) = X(0) = 1, 2, 3$. Interestingly, the approximate stationary resource concentration, $E\big(A(30)\big)$, is more sensitive to the initial conditions than the other three expectation values. The probability to end up in the state of cooperation, $P(S_3)$, shows precisely the same trend as the expectation values: for $X(0) = 1$ only 30 % of the trajectories go into S_3, for $X(0) = 2$ the percentage has increased to 91 %, for $X(0) = 3$ to 97 %, for $X(0) = 4$ it reaches almost 100 % and for $X(0) > 4$ we find that practically all trajectories converge to the state of cooperation S_3. The numerical values for the expectation values calculated with $X(0) = 10$ and the results of ODE-integration are almost the same. All values calculated for a mutation rate parameter $p = 0.05$ show an increases propensity to end up in the cooperative state. Already at $X(0) = 3$ we calculated almost exclusive convergence to S_3. The interpretation is straightforward: All systems escape from dying out when during the time necessary for hypercycle extinction the missing subspecies is reintroduced by mutation. The more members the hypercycle has the longer is time span that allows for a *successful repair of the system*. In the two-membered cooperative system the effect of mutation on survival of the system is very small or not detectable. This can be made plausible by recalling that for $n = 2$ the other subspecies of the system is simultaneously the previous and the next member in the cycle. In summary we have two stochastic effects in the small cooperative systems with mutation: (i) initial conditions with very small particle numbers have very strong influence on the solution curves and are the cause of very large fluctuations as we

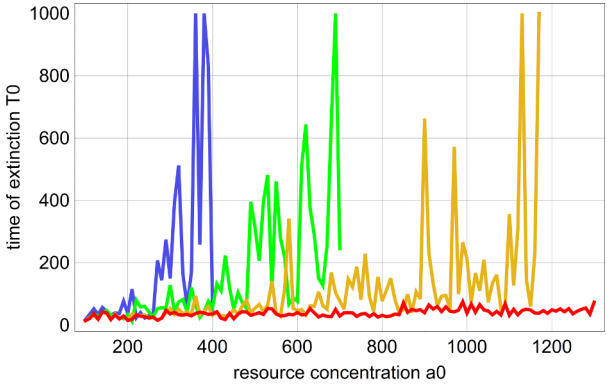

Figure 11. Time to extinction T_0 of oscillating hypercycles with $n = 5$ and mutation rates p. The plot shows times to hypercycle extinction as a function of available resources, $T_0(a_0)$ calculated by means of the Gillespie algorithms. Individual points coarse grained to $\Delta a_0 = 10$ are connected by straight lines. Choice of parameters: $r = 0.5\,[\mathrm{V^{-1}t^{-1}}]$, $l_1 = l_2 = l_3 = l_4 = l_5 = 0.01\,[\mathrm{M^{-2}t^{-1}}]$. Pseudorandom number generator: *Extended CA* (Mathematica 10), seed: $s = 491$. Initial conditions: $A(0) = 0$, $X_1(0) = X_2(0) = X_3(0) = X_4(0) = X_5(0) = 5$. Color code: $p = 0.0$ red, $p = 0.0005$ yellow, $p = 0.0010$ green, and $p = 0.0020$ blue.

observed in case of second order autocatalysis (section 3), and (ii) mutation can partially compensate for the extinction of subspecies.

In the stochastic approach the oscillating systems are particularly sensitive to fluctuation when one or more components adopt very small values of concentrations in the low tide phase of the period and sufficiently large values of the mutation rate parameter p may substantially increase the lifetime of the system. For $n = 4$ sufficiently large populations sizes N can prevent the system from extinction but for $n \geq 5$ the relaxation oscillations [64] lead sooner or later to the elimination of one species, say $\mathbf{X}_k$. After $\mathbf{X}_k$ has vanished the whole mutation-free system dies out stepwise like driven by a ratchet as can be easily verified from equation (10) by putting $Q_{ij} = \delta_{ij}$ for error-free reproduction:

$$x_k = 0 \;\rightarrow\; \frac{\mathrm{d}x_{k-1}}{\mathrm{dt}} = -\,x_{k-1}\,r \;\rightarrow\; x_{k-1} = 0\,,$$
$$x_{k-1} = 0 \;\rightarrow\; \frac{\mathrm{d}x_{k-2}}{\mathrm{dt}} = -\,x_{k-2}\,r \;\rightarrow\; x_{k-2} = 0\,,\ \text{etc.},$$

until all subspecies have disappeared. Figure 10 illustrates extinction for two examples with small p-values and $n = 4$ and $n = 5$. In the former case the oscillations show an

interesting kind of beat.

The role of mutation is to replace extinguished subspecies $\mathbf{X}_k$ sufficiently fast in order to prevent the onset of the ratchet. We consider the time of extinction of the system, T_0, for $n \geq 5$. Without mutation the T_0-values show large random scatter but as a function of the resource concentration a_0 they are confined to values below a certain limit that corresponds to a fairly short lift time of the hypercycles. Despite enormous random scatter the extinction times show an interesting regularity (figure 11). On introduction of mutation the time of extinction of the hypercycle, T_0, becomes longer with increasing mutation rate parameter p. Large scatter is to be expected since we are dealing with second order autocatalysis (section 3). In addition the dependence on a_0 indicates some threshold behavior: With increasing a_0 the extinction time $T_0(a_0)$ stays little higher than the values for $p = 0.0$ but above some critical resource concentration $(a_0)_{cr}(p)$ the hypercycle lifetime shows fast increase with a_0 and above a certain a_0 value the hypercycle survives to long time. The critical resource concentration becomes smaller with increasing p. The behavior of the extinction times $T_0(a_0)$ is similar in the cases with $n = 4$ but the critical concentrations $(a_0)_{cr}(p)$ for the different p-values lie much closer together.

7 The complete model

Completion of the model requires bringing together the three faces of the coordinate system in figure 1 and analysing the interior. An appropriate strategy consists in choosing certain type of behavior on one of the three faces and increasing the third parameter from zero to the value of interest. An example is the influence of mutation, which is commonly negligible for small value of p, and to raise the mutation rate until the effect of the third parameter, reintroduction of already extinguished subspecies becomes evident. The complete collection of scenarios observed in the full three-dimensional parameter space is very rich and describing all features would go beyond the scope of this contribution. Therefore we shall postpone the complete description to a forthcoming publication [69] and illustrate here by one example, the transition from competition to cooperation with increasing rate parameters $h = l \cdot a$ only. For $p = 0$ and small numbers of subspecies ($n = 2, 3$) the transition has been discussed in section 5 and leads from selection of the fittest to a cooperative state with all subspecies present. For oscillating systems ($n > 4$) the hypercycles are unstable in the stochastic approach and raising the rate parameter h

leads from selection to extinction. In the intermediate parameter range where exclusion occurs in the deterministic system we find highly irregular oscillation of different numbers of subspecies whereby the number of species present increases with increasing h-values. Nonzero mutation rates p stabilize the hypercycles and at sufficiently frequent mutations the vanished subspecies are reintroduced as mutants of other subspecies fast enough so that a kind of mutation supported hypercycle dynamics originates. Then the scenario of a transition from low to high h-values leads from selection of the fittest to oscillating hypercycles. Similar procedures are suitable for variation of p at constant f and h or variation of f at constant p and h.

The model has been presented here in its most simple form. The three modules, competition, cooperation and variation, can be made arbitrarily complex. Variation, for example, can be extended to include more elaborate mutations, recombination, and environmental influences. In case of viruses the reproduction mechanism might consider also the defense system of the host, epigenetic phenomena may be taken into account through considering several generations at a time, and for higher organism the real challenge is to deal in simple form with the enormous complexity of development. Cooperation at the molecular level may also include reproductive autocatalytic networks whereas social phenomena in reproductive groups or societies represent the currently highest step in the open ended complexity of biological evolution. There is no limitation to make the model more complex, the problem evidently is to include the desired phenomenon but to keep the model simple enough for mathematical analysis.

At the end we present a few examples, which show that experimental implementations of the evolution model are quite straightforward in certain cases. The flow reactor introduced here has the advantage to be a suitable device for both theoretical computations and evolution experiments. Other setups were also successfully applied to evolution experiments based on competition and mutation, for example, reaction diffusion systems in capillaries [70, 71] or rather sophisticated machines [72]. Quasispecies formation and evolution have been extensively studied and in this case predictions by theory can be readily checked by proper experiments [73]. Application of the theory to virus evolution gave important insights into virus population dynamics *in vitro* and *in vivo* and provided important hints for the development of novel antiviral strategies [63, 74]. Although symbiosis and its negative counterpart in form of predator-prey systems are common in field

biology and in bacterial communities, easy to analyse molecular systems showing cooperation through second order autocatalysis are not at hand despite an impressive number of designing attempts. The systems coming closest to the ideal kinetics are described in publications representative for others [7, 8, 75, 76].

Acknowledgements: The authors wishes to acknowledge fruitful discussions with Professors Ivo L. Hofacker and Christoph Flamm.

References

[1] T. Dobzhansky, F. J. Ayala, G. L. Stebbins, J. W. Valentine, *Evolution*, Freeman, San Francisco, 1977.

[2] P. Schuster, Quasispecies on fitness landscapes, in: E. Domingo, P. Schuster (Eds.), *Quasispecies: From Theory to Experimental Systems*, Springer, Berlin, 2016, pp. 61–120.

[3] C. K. Biebricher, Darwinian selection of self-replicating RNA molecules, in: M. K. Hecht, B. Wallace, G. T. Prance (Eds.), *Evolutionary Biology*, Plenum Pub. Corporation, 1983, pp. 1–52.

[4] C. K. Biebricher, M. Eigen, W. C. Gardiner, Jr, Kinetics of RNA replication, *Biochemistry* **22** (1983) 2544–2559.

[5] M. Eigen, P. Schuster, The hypercycle. A principle of natural self-organization. Part B: The abstract hypercycle, *Naturwissenschaften* **65** (1978) 7–41.

[6] P. G. Higgs, N. Lehman, The RNA world: Molecular cooperation at the origins of life, *Nat. Rev. Genet.* **16** (2015) 7–17.

[7] T. A. Lincoln, G. F. Joyce, Self-sustained replication of an RNA enzyme, *Science* **323** (2009) 1229–1232.

[8] N. Vaidya, M. L. Manapat, I. A. Chen, R. Xulvi-Brunet, E. J. Hayden, N. Lehman, Spontaneous network formation among cooperative RNA replicators, *Nature* **491** (2012) 72–77.

[9] D. P. Horning, G. F. Joyce, Amplification of RNA by an RNA polymerase ribozyme, *Proc. Natl. Acad. Sci. USA* **113** (2016) 9786–9791.

[10] G. G. Dimijian, Evolving together: The biology of symbiosis, part 1, *Baylor Univ. Med. Center Proceed.* **13** (2000) 217–226.

[11] G. G. Dimijian, Evolving together: The biology of symbiosis, part 2, *Baylor Univ. Med. Center Proceed.* **13** (2000) 381–390.

[12] S. Paracer, V. Ahmadjian, *Symbiosis. An Introduction to Biological Associations*, Oxford Univ. Press, Oxford, 2000.

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

[14] M. Eigen, P. Schuster, The hypercycle. A principle of natural self-organization. Part A: Emergence of the hypercycle, *Naturwissenschaften* **64** (1977) 541–565.

[15] M. Eigen, P. Schuster, The hypercycle. A principle of natural self-organization. Part C: The realistic hypercycle, *Naturwissenschaften* **65** (1978) 341–369.

[16] P. Schuster, K. Sigmund, R. Wolff, Dynamical systems under constant organization I. Topological analysis of a family of non-linear differential equations – A model for catalytic hypercycles, *Bull. Math. Biol.* **40** (1978) 734–769.

[17] J. Hofbauer, P. Schuster, K. Sigmund, R. Wolff, Dynamical systems under constant organization II: Homogenoeous growth functions of degree $p = 2$, *SIAM J. Appl. Math.* **38** (1980) 282–304.

[18] J. Hofbauer, J. Mallet-Paret, H. L. Smith, Stable periodic solutions for the hypercycle system, *J. Dyn. Diff. Eqs.* **3** (1991) 423–436.

[19] R. W. Hamming, Error detecting and error correcting codes, *Bell Syst. Tech. J.* **29** (1950) 147–160.

[20] R. W. Hamming, *Coding and Information Theory*, Prentice–Hall, Englewood Cliffs, 1986.

[21] J. F. Crow, M. Kimura, *An Introduction to Population Genetics Theory*, Sinauer Associates, Sunderland, 1970.

[22] J. Maynard Smith, E. Szathmáry, *The Major Transitions in Evolution*, Freeman, Oxford, 1995.

[23] P. Schuster, How does complexity arise in evolution. Nature's recipe for mastering scarcity, abundance, and unpredictability, *Complexity* **2/1** (1996) 22–30.

[24] P. Schuster, Increase in complexity and information through molecular evolution, *Entropy* **18** (2016) #e397.

[25] M. Kimura, *The Neutral Theory of Molecular Evolution*, Cambridge Univ. Press, Cambridge, 1983.

[26] L. D. Schmidt, *The Engineering of Chemical Reactions*, Oxford Univ. Press, New York, 2005.

[27] A. Novick, L. Szillard, Description of the chemostat, *Science* **112** (1950) 715–716.

[28] A. Novick, L. Szillard, Experiments with the chemostat on spontaneous mutations of bacteria, *Proc. Natl. Acad. Sci. USA* **36** (1950) 708–719.

[29] V. Bryson, W. Szybalski, Microbial selection, *Science* **116** (1952) 45–46.

[30] Y. Husimi, K. Nishigaki, Y. Kinoshita, T. Tanaka, Cellstat – A continuous culture system of a bacteriophage for the study of the mutation rate and the selection process at the DNA level, *Rev. Sci. Instrum.* **53** (1982) 517–522.

[31] D. E. Dykhuizen, D. L. Hartl, Selection in chemostats, *Microbiol. Rev.* **46** (1983) 150–168.

[32] Y. Husimi, Selection and evolution of a bacteriophages in cellstat, *Adv. Biophys.* **25** (1989) 1–43.

[33] A. Koltermann, U. Kettling, Principles and methods of evolutionary biotechnology, *Biophys. Chem.* **66** (1997) 159–177.

[34] G. Strunk, T. Ederhof, Machines for automated evolution experiments *in vitro* based on the serial-transfer concept, *Biophys. Chem.* **66** (1997) 193–202.

[35] M. Eigen, J. McCaskill, P. Schuster, The molecular quasispecies., *Adv. Chem. Phys.* **75** (1989) 149–263.

[36] B. L. Jones, H. K. Leung, Stochastic analysis of a non-linear model for selection of biological macromolecules, *Bull. Math. Biol.* **43** (1981) 665–680.

[37] P. Schuster, K. Sigmund, Replicator dynamics., *J. Theor. Biol.* **100** (1983) 533–538.

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

[39] A. J. Lotka, Analytical note on certain rhythmic relations in organic systems, *Proc. Natl. Acad. Sci. USA* **6** (1920) 410–415.

[40] A. J. Lotka, *Elements of Physical Biology*, Williams and Wilkins, Baltimore, 1925.

[41] V. Volterra, Variazioni e fluttuazioni del numero d'individui in specie animal conviventi, *Memoria della Regia Accademia Nationale dei Lincei* **ser.6, 2** (1926) 31–113, translated as an Appendix entitled *Variations and Fluctuations of the Number of Individuals in Animal Species Living Together* to R. N. Chapman. *Animal Ecology.* McGraw-Hill, New York, 1931.

[42] V. Volterra, Fluctuations in the abundance of a species considered mathematically, *Nature* **118** (1926) 558–560.

[43] J. Hofbauer, On the occurrence of limit cycles in the Volterra-Lotka equation, *Nonlin. Anal. Theory Meth. Appl.* **5** (1981) 1003–1007.

[44] P. Schuster, K. Sigmund, Dynamics of evolutionary optimization., *Ber. Bunsenges. Phys. Chem.* **89** (1985) 668–682.

[45] T. Jahnke, W. Huisinga, Solving the chemical master equation for monomolecular reaction systems analytically, *J. Math. Biol.* **54** (2007) 1–26.

[46] P. Deuflhard, W. Huisinga, T. Jahnke, M. Wulkow, Adaptive discrete Galerkin methods applied to the chemical master equation, *SIAM J. Sci. Comput.* **30** (2008) 2990–3011.

[47] A. N. Kolmogorov, Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung, *Math. Ann.* **104** (1931) 415–458, in German.

[48] W. Feller, On the integro-differential equations of purely discontinuous Markoff processes, *T. Am. Math. Soc.* **48** (1940) 488–515.

[49] J. L. Doob, Topics in the theory of Markoff chains, *T. Am. Math. Soc.* **52** (1942) 37–64.

[50] J. L. Doob, Markoff chains – Denumerable case, *T. Am. Math. Soc.* **58** (1945) 455–473.

[51] D. T. Gillespie, A general method for numerically simulating the stochastic time evolution of coupled chemical reactions, *J. Comp. Phys.* **22** (1976) 403–434.

[52] D. T. Gillespie, Exact stochastic simulation of coupled chemical reactions, *J. Phys. Chem.* **81** (1977) 2340–2361.

[53] D. T. Gillespie, Stochastic simulation of chemical kinetics, *Ann. Rev. Phys. Chem.* **58** (2007) 35–55.

[54] P. Schuster, *Stochasticity in Processes. Fundamentals and Applications in Chemistry and Biology*, Springer, Berlin, 2016.

[55] S. Spiegelman, An approach to the experimental analysis of precellular evolution, *Quart. Rev. Biophys.* **4** (1971) 213–253.

[56] G. F. Joyce, Forty years of *in vitro* evolution, *Angew. Chem. Int. Edit.* **46** (2007) 6420–6436.

[57] E. Domingo, P. Schuster, What is a quasispecies? Historical origins and current scope, in: E. Domingo, P. Schuster (Eds.), *Quasispecies: From Theory to Experimental Systems*, Springer, Berlin, 2016, pp. 1–22.

[58] P. Tarazona, Error thresholds for molecular quasispecies as phase transitions: From simple landscapes to spin glasses, *Phys. Rev. A* **45** (1992) 6038–6050.

[59] J.-M. Park, E. Muñoz, M. W. Deem, Quasispecies theory for finite populations, *Phys. Rev. E* **81** (2010) #e011902.

[60] T. Wiehe, Model dependency of error thresholds: The role of fitness functions and contrasts between the finite and infinite sites models, *Genet. Res. Camb.* **69** (1997) 127–136.

[61] A. M. Dean, J. W. Thronton, Mechanistgic approaches to the study of evolution: The functional synthesis, *Nat. Rev. Genet.* **8** (2007) 675–688.

[62] P. D. Sniegowski, P. J. Gerrish, Beneficial mutations and the dynamics of adaptation in asexual populations, *Phil. Trans. Roy. Soc. B* **356** (2010) 1255–1263.

[63] E. Domingo, P. Schuster (Eds.), *Quasispecies: From Theory to Experimental Systems*, Springer, Berlin, 2016.

[64] P. E. Phillipson, P. Schuster, F. Kemler, Dynamical machinery of a biochemical clock, *Bull. Math. Biol.* **46** (1984) 339–355.

[65] P. Schuster, Some mechanistic requirements for major transitions, *Phil. Trans. Roy. Soc. London B* **371** (2016) #e20150439.

[66] P. Schuster, Major transitions in evolution and in technology. What they have in common and where they differ, *Complexity* **21/4** (2016) 7–13.

[67] P. E. Phillipson, P. Schuster, *Modeling by Nonlinear Differential Equations. Dissipative and Conservative Processes*, World Sci., Singapore, 2009.

[68] W. Schnabl, P. F. Stadler, C. Forst, P. Schuster, Full characterizatin of a strang attractoe. Chaotic dynamics on low-dimensional replicator systems, *Physica D* **48** (1991) 65–90.

[69] P. Schuster, Molecular evolution between chemistry and biology. A minimal model for competition, cooperation, and mutation, *Eur. Biophys. J.* **46** (2017) submitted.

[70] G. J. Bauer, J. S. McCaskill, H. Otten, Travelling waves of *in vitro* evolving RNA, *Proc. Natl. Acad. Sci. USA* **86** (1989) 7937–7941.

[71] J. S. McCaskill, G. J. Bauer, Images of evolution: Origin of spontaneous RNA replication waves, *Proc. Natl. Acad. Sci. USA* **90** (1993) 4191–4195.

[72] A. Watts, G. Schwarz (Eds.), *Evolutionary Biotechnology – From Theory to Experiment*, Elesvier, Amsterdam, 1997, pp. 67–284.

[73] C. K. Biebricher, Quantitative analysis of mutation and selection in self-replicating RNA, *Adv. Space Res.* **12** (1992) 191–197.

[74] E. Domingo, J. Sheldon, C. Perales, Viral quasispecies evolution, *Microbiol. Mol. Biol. R.* **76** (2012) 159–216.

[75] J. S. McCaskill, Spatially resolved *in vitro* molecular ecology, *Biophys. Chem.* **66** (1997) 145–158.

[76] B. Wlotzka, J. S. McCaskill, A molecular predator and its prey: Coupled isothermal amplification of nucleic acids, *Chem. Biol.* **4** (1997) 25–33.