



What is special about autocatalysis?

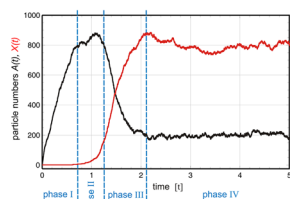
Peter Schuster^{1,2}

Received: 17 January 2019 / Accepted: 20 April 2019 / Published online: 7 May 2019
© The Author(s) 2019

Abstract

The single-step autocatalytic reaction $A + X \leftrightarrow 2X$ in different environments—batch reactor, flow reactor, logistic equation—is studied by means of conventional deterministic kinetics and as a stochastic process. Self-enhancement requires an initial concentration of the autocatalyst X —at least in seeding amounts—for starting the reaction. Deterministic solution curves have sigmoid shapes. At small concentrations, three stochastic phenomena are observed: (1) thermal fluctuations, (2) stochastic delay, and (3) stochastic bifurcations and anomalous fluctuations in case of multiple final states. The introduction of heterogeneous populations containing subspecies with different fitness values gives rise to natural selection in all three environments investigated here. In large populations, *survival of the fittest* is observed, whereas random fluctuations may result in selection of each of the subspecies. Then, the fitness values determine only probabilities of selection. The fittest subspecies, of course, has the largest probability of selection. There is a smooth transition to neutral evolution where the probabilities of selection are the same for all subspecies.

Graphical abstract



Stochastic description of autocatalysis in the flow reactor

Keywords Autocatalysis · Evolution · Master equation · Optimization · Reaction kinetics · Replication · Stochastic simulation algorithm

Introduction

The notion of autocatalysis has been introduced by the German chemist Wilhelm Ostwald [1] for the characterization of reactions that show an acceleration of the rate as

a function of time. Autocatalytic reactions, however, were never popular in conventional chemistry and chemical technology mainly for two reasons: (1) autocatalysis gives rise to positive feedback, which is difficult to control and to handle in large reactors, and (2) autocatalytic systems may show complex dynamical phenomena such as bistability, oscillations, deterministic chaos, and spontaneous formation of spatial patterns or waves, which are not desirable in chemical production. The whole collection of these phenomena has been called *nonlinear chemical dynamics* by Irving Epstein and others [2, 3]. In the second half of the twentieth century, however, complex dynamics and its embedding in thermodynamics have been shifted into the center of physical interest. Experimental chemical reaction systems showing

Dedicated to Heinz Falk on the occasion of the 80th anniversary of his birthday.

✉ Peter Schuster
pks@tbi.univie.ac.at

¹ Institute of Theoretical Chemistry, University of Vienna, Vienna, Austria

² The Santa Fe Institute, Santa Fe, NM, USA

nonlinear phenomena, in particular oscillations and spatial pattern formation, were investigated [4, 5]. Autocatalysis was recognized as a basis for self-organization and has been extensively discussed in the context of models for the origin of life [6]. Oscillations in chemical reactions have been reported earlier [7] but these were singular studies and they took place in heterogeneous media. Beginning around 1960 with the studies of Boris Belousov and Anatol Zhabotinskii on a chemical reaction that exhibits oscillations in homogeneous solution, investigations on chemical self-organization became popular. Since then, a true wealth of experimental works and many theoretical studies on *nonlinear chemical dynamics* were performed [3].

A distinction between first- and higher order autocatalysis turned out to be very useful. First-order autocatalysis comprises all cases where a single autocatalytic particle is involved on the reactant side. The best and simplest example is the reaction $A + X \rightarrow 2X$: a single autocatalyst particle X and a resource A produce together two particles X . This is the essence of asexual reproduction in biology, although X —being, for example, a virus particle or a bacterial cell—represents a highly complex entity and the resource A is not a single molecule but a variety of required building blocks. Simpler systems capable of reproduction such as RNA molecules have been studied by *evolution* in vitro [8, 9], but still reproduction is a complex multi-step process [10]. Nevertheless, single-step mechanisms describing the overall kinetics of the process under suitable conditions can be found (see also the section “[Conclusion](#)”). The most important phenomenon related to first-order autocatalysis is selection in the Darwinian sense: when several types of autocatalysts are present simultaneously all except one are eliminated through competition for resources.

Higher order autocatalysis is observed with several multi-step reactions often involving halogen—chlorine, bromine, and iodine—in different oxidation states. Investigations on higher order autocatalysis suffer even more than first-order autocatalytic systems from the fact that no experimental systems with simple one- or two-step mechanisms were found yet. Largely simplified but still realistic models such as the Oregonator involve five different elementary steps [11, 12]. A comparison of the individual steps in the Brusselator and the Oregonator models shows why the former is unrealistic: The Brusselator involves a termolecular step, $A + 2X \rightarrow 3X$, which has extremely low probability because collisions of three particles occur very rarely. In the Oregonator, the termolecular step is resolved in several bimolecular steps. Oscillations in chemical reactions as well as deterministic chaos in chemistry have been studied in great detail (for an excellent review see, e.g., Ref. [3]). There is also extensive literature on Turing pattern formation [13] as well as spatiotemporal phenomena produced through nonlinear chemical dynamics [14].

Biology other than chemistry is centered around autocatalysis in the special form of reproduction. Multiplication of cells and organisms is in the core of biological thinking since it represents the basis of development and evolution. A second and not less important feature of biology—again in contrast to chemistry—is the necessity to deal with stochastic phenomena caused by small and very small particle numbers. Every mutant, after all, starts out from a single copy.

In this contribution, the focus is laid on first-order autocatalysis and its special role in chemical, biochemical and biological applications. At first, the reaction $A + X \rightarrow 2X$ is analyzed in the batch and in the flow reactors. Deterministic solutions are compared with the results of stochastic modeling at small particle numbers. The process $A + X \rightarrow 2X$ can be understood as a toy model for asexual reproduction. In biology, the logistic equation derived by Pierre-François Verhulst is often used for modeling reproduction and selection at the population level [15, 16]. We shall study it here as a simple alternative to the two classes of common chemical reactors. Broad emphasis is laid on the role of autocatalysis in natural selection. We shall study a population of independently reproducing variants $A + X_j \rightarrow 2X_j$, ($j = 1, 2, \dots, n$), which are coupled only through the exploitation of a common resource A . The deterministic approach leads to selection of the fittest, X_m defined by $f_m = \max\{f_1, f_2, \dots, f_n\}$, which is accompanied by optimization of fitness on the population level. The stochastic process occurring at small particle numbers results also in selection but apart from selection of the fittest, selection of less fit variants can happen as well. We emphasize: selection means only that all variants die out except the variant, which is selected. The stochastic selection process is guided by probabilities of selection that are determined by the fitness values.

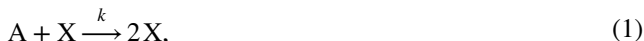
Results and discussion

First-order autocatalysis in its simplest form, $A + X \rightarrow 2X$, is studied under different environmental conditions: (1) the batch reactor, (2) the flow reactor, and (3) a ninetieth century model system for constrained growth, which is still in use and known as logistic equation. Autocatalysis leads to spectacular results when it operates on populations that contain collections of different autocatalysts, which live on the same resources: selection of one variant is observed. In the final part of the “[Results and discussion](#)” section, the relation between autocatalysis and Darwin’s natural selection is illustrated. Each section compares the deterministic solutions of conventional reaction kinetics with the results of the corresponding stochastic process that is modeled by a chemical master equation. The methods applied are described in the “[Methods](#)” section.

Autocatalysis in the batch reactor

The batch reactor [17] is a device that allows for studying chemical reactions in a closed system consisting here of well-mixed solutions with temperature control. The two conditions, spatial homogeneity and constant temperature, are commonly assumed to be fulfilled in the conventional theory of chemical reactions. Simply expressed, performing a reaction in the batch reactor implies that the reactants are perfectly mixed and then left alone in a thermostat.

The simple autocatalytic process,



is described by the kinetic equation and its solution curve:

$$\frac{dx}{dt} = kax = -\frac{da}{dt} = v(t); \quad x(t) = \frac{(a(0) + x(0))x(0)}{x(0) + a(0)e^{-k(a(0)+x(0))t}}. \tag{1a}$$

The two variables $a(t)$ and $x(t)$ are constrained by the conservation relation $a(t) + x(t) = c = \text{const}$. The solution curve $x(t)$ is “S” shaped or sigmoid. In reaction kinetics, this shape is typical for autocatalysis (Fig. 1). At low autocatalyst and large resource concentrations— $x(t) \ll a(t)$ —the resource concentration remains approximately constant in the early phase of the reaction: $a(t) \approx a = \text{const}$ for small t , and increasing $x(t)$ leads to self-enhancement since the reaction rate is proportional to the autocatalyst concentration: $v(t) = k a x(t)$. An important feature of autocatalysis is the requirement of—at least—a seeding quantity of autocatalyst for the start of the reaction since the state of extinction S_0 with $x(0) = 0$ implies no reaction, $v(0) = 0$ and ignition of the reaction does not occur.

Random fluctuations become important at small particle numbers when the discrete nature of molecular concentrations comes into play. The irreversible autocatalytic reaction, $A + X \xrightarrow{k} 2X$, has been studied mathematically in form of the chemical master equation [18]:

$$\frac{dP_M}{dt} = k((M + 1)(C - M - 1)P_{M+1} - M(C - M)P_M); \quad M, C \in \mathbb{N}. \tag{2}$$

This equation describes the time dependence of the probability distribution $P_M(t) = \text{Prob}\{A(t) = M\}$ with the conservation relation $C = A(t) + X(t)$. In the stochastic treatment, the probability $P_M(t)$ replaces the concentration variable $a(t)$ and the second variable $X(t)$ is defined by the conservation relation.

Although analytical expressions are available for the probabilities, $P_M(t)$ [18], they are of limited practical use because the results are obtained through rather sophisticated series that do not show universal convergence. In practice, a numerical approach based on sampling of trajectories (section “Methods”) is commonly preferred. Equation (2) has only one running index M and two terms P_{M+1} and P_M , and computation as

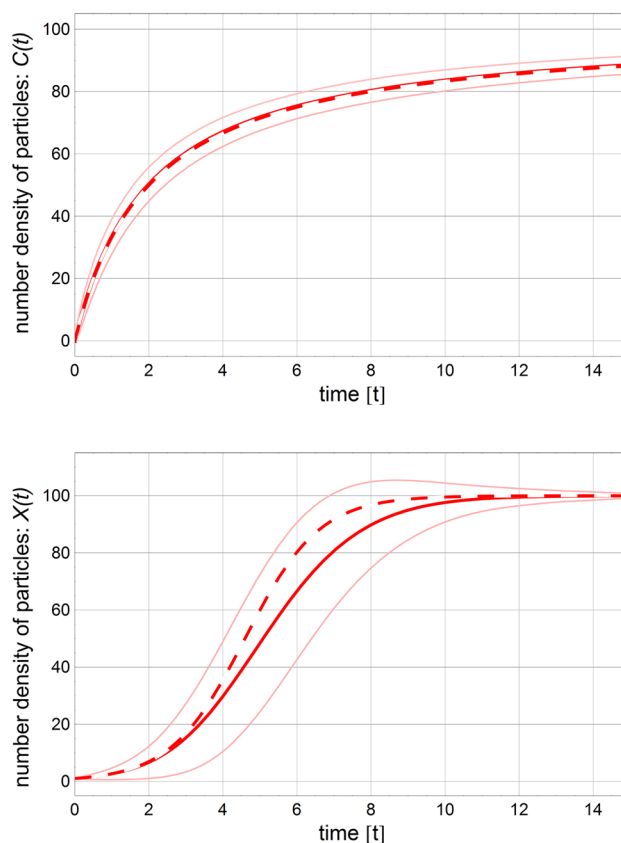


Fig. 1 Stochastic delay in the irreversible first-order autocatalytic reaction $A + X \rightarrow 2X$. The figures compare expectation values (full thick lines), one standard deviation bands (thin light lines) and deterministic solutions (dashed thick lines) of the reactions $A + B \rightarrow 2C$ (upper part) and $A + X \rightarrow 2X$ (lower part). In the upper part, we see the results for the non-autocatalytic chemical reaction. The total number of particles is $N = 100$, the fluctuations in the order of \sqrt{N} , and the deterministic solution almost coincides with the stochastic expectation value $E(C(t))$. The shape of the solution curves is hyperbolic. The lower part presents the analogous curves for the autocatalytic reaction that are “S” shaped or sigmoid. The fluctuations are larger because of two effects: (1) thermal fluctuations are increased because of self-enhancement and (2) stochastic delay, which causes the expectation value $E(X(t))$ to be shifted to longer times. The deterministic curve is well separated from the expectation value (see text). Parameter choice: $k = 0.01 [\text{t}^{-1} \text{C}^{-1}]$, $N = 100$, $A(0) = B(0) = 50$, $C(0) = 0$ and $A(0) = 99$, $X(0) = 1$, respectively. [C] stands here for the unit of number density

well as analysis of trajectories are fast and straightforward. The commonly presented stochastic results are expectation value $E(A(t)) = \sum_{M=1}^{C-1} M P_M(t)$ and standard deviation $\sigma(A(t))$, which are given in terms of one-standard-deviation error bands, $E \pm \sigma(A(t))$. In Fig. 1, we compare a typical example of the autocatalytic process $A + X \xrightarrow{k} 2X$ with a conventional, non-autocatalytic chemical reaction with similar stoichiometry:



Table 1 Stochastic delay ΔX_{\max} in the irreversible autocatalytic reaction $A + X \rightarrow 2X$

N		Initial particle numbers X_0					
		1	2	3	5	7	10
100	$t(\Delta X_{\max})$	6.039	5.125	4.757	4.174	3.922	3.420
	ΔX_{\max}	13.16	6.93	4.57	2.79	1.89	1.34
	δ	0.132	0.139	0.137	0.140	0.132	0.134
400	$t(\Delta X_{\max})$	18.11	15.94	14.690	13.216	12.407	11.471
	ΔX_{\max}	53.94	27.88	18.64	12.03	8.48	5.97
	δ	0.135	0.139	0.140	0.150	0.149	0.149
1000	$t(\Delta X_{\max})$	82.51	73.55	69.15	61.60	59.54	55.42
	ΔX_{\max}	136.58	71.69	47.72	30.13	21.68	14.15
	δ	0.137	0.143	0.143	0.151	0.152	0.142

The deterministic solution is obtained from $dx/dt=kax$ through integration: $x(t)=x(0) \exp(-kct) \exp(kcx_0+ka_0)$. The rate parameters for the three different population sizes $N=100, 400,$ and 1000 are $k=0.01, 0.001,$ and 0.0001 [$t^{-1} C^{-1}$], respectively. The sample size is 10 000 trajectories

In autocatalytic reactions, three types of random fluctuations can be distinguished: (1) thermal fluctuations, (2) stochastic delay, and (3) anomalous fluctuations caused by multiple (quasi)stationary states.¹ With irreversible reactions in the batch reactor, we observe only fluctuations of types (1) and (2), because we are dealing with a single stationary state only.

Stochastic delay is a newly reported stochastic phenomenon. The notion refers to the fact that the expectation value $E(X(t))$ appears delayed relative to the deterministic solution curve $x(t)$ (Fig. 1, lower plot). As a quantitative measure for the delay, we consider the maximal difference between the curves: $\Delta X_{\max} = \max\{x(t) - E(X(t))\} = x(t_{\max}) - E(X(t_{\max}))$ with t_{\max} being the time, where the maximal difference occurs.² The difference ΔX_{\max} can be converted into an incremental quantity δ that is approximately independent of the initial particle density $X_0=X(0)$ and population size N :

$$\delta = \Delta X_{\max} \frac{X_0}{N} [C] \quad \text{or} \quad \Delta X_{\max} = \delta \frac{N}{X_0} [C]. \quad (3)$$

Numerical values are shown in Table 1. The stochastic delay ΔX_{\max} is proportional to the population size and inversely proportional to the initial number of autocatalyst molecules X . The stochastic delay ΔX_{\max} vanishes with increasing X_0 and thus leads to convergence of the

expectation value $E(X(t_{\max}))$ and the deterministic curve $x(t)$. It is worth noticing that the stochastic delay increases with the total population size N within the range shown in Table 1, $100 \leq N \leq 1000$, or in other words the relative stochastic delay $\delta = \Delta X_{\max} X(0)/N$ is insensitive to population size.

Introduction of the inverse reaction,



leads to little change in the reaction scenario. Instead of going into full turnover of A , the reaction converges to the thermodynamic equilibrium:

$$\frac{[X]^2}{[A][X]} = \frac{[X]}{[A]} = \frac{k}{h} = K. \quad (4)$$

The stochastic delay ΔX_{\max} shows essentially the same behavior in the reversible and the irreversible cases.

Thermal fluctuations are present in all chemical reactions and their amplitudes are typically in the order of \sqrt{N} , where N is the total number of molecules. Considering a single-step autocatalytic reaction with sharp initial conditions, $P_M(0) = \delta_{M,X_0}$, the initial standard deviation is zero, $\sigma(A(0))=0$. It increases with time, passes through a maximum, decreases, and vanishes for the irreversible reaction. For the reversible reaction, it converges to its equilibrium value either after having passed a maximum or monotonously. The problem of extinction in the one-step autocatalytic reaction appears in different forms in the deterministic and the stochastic models. Because of the conservation relation $a(t) + x(t) = c > 0$, Eq. (1a) can be written in a single variable and we obtain for the reversible reaction:

$$\frac{dx}{dt} = k(c-x)x - hx^2 = kcx - (k+h)x^2. \quad (1a')$$

¹ In stochastic dynamical systems, it is appropriate to distinguish *absorbing* and *quasistationary* states. When a system reaches an absorbing state, it will stay there forever. A quasistationary state behaves like a true stationary or absorbing state for all practical purposes but a gigantic fluctuation of extremely low probability may transfer the system into an absorbing state.

² It might seem more natural to define the stochastic delay as the maximal time difference between the two curves, $\Delta t_{\max} = \max\{t_{E(X(t))=\gamma} - t_{x(t)=\gamma}\}$. The results of this calculation are qualitatively the same as those obtained with ΔX_{\max} .

The positive term dominates for small x , $dx/dt > 0$, x increases and hence the autocatalyst X cannot go extinct. In the stochastic system, the analysis gives a similar result for a different reason. The reaction step reducing the number of autocatalyst molecules, $2X \rightarrow A + X$, requires at least two molecules X , a single X molecule cannot undergo this conversion and hence X cannot die out. In other words, the state $S_1 = \{(C-1)A, X\}$ is a reflecting barrier and cannot be surmounted by the single-step autocatalytic reaction.

Autocatalysis in the flow reactor

A continuous-flow stirred-tank reactor (CFSTR) is sketched in Fig. 2. For a reaction in the reactor, inflow and outflow are modeled as *pseudoreactions*.³ This yields the five-step reaction mechanism:



The two chemical species in mechanism (5), A and X , are independent and hence the system is described by the differential equations:

$$\frac{da}{dt} = -kax + hx^2 + (c_0 - a)r \quad \text{and} \quad \frac{dx}{dt} = kax - hx^2 - xr, \tag{6}$$

and sustains two stationary states: (1) the state of extinction S_0 and (2) the reaction state S_1 . The values of variables at stationary states are readily calculated through equating to zero the two expressions in Eq. (6):

$$S_0 : \bar{a}^{(0)} = c_0, \bar{x}^{(0)} = 0 \quad \text{and} \\ S_1 : \bar{a}^{(1)} = \frac{hc_0 + r}{k + h}, \bar{x}^{(1)} = \frac{kc_0 - r}{k + h}. \tag{7}$$

In the flow reactor, the conservation relation, $a(t) + x(t) = c_0$, is not valid in general, but it is fulfilled at the steady states, $\bar{a} + \bar{x} = c_0$, since the concentration in the reactor becomes equal to the concentration in the stock solution, $\lim_{t \rightarrow \infty} (a(t) + x(t)) = c_0$, for sufficiently long time. Stability

³ Inflow into and outflow from the reactor—reactions (5a) and (5d, 5e), respectively—are no proper chemical reactions since no molecules are changed but they are incorporated into mechanisms as zeroth-order (5a) or first-order (5d, 5e) reactions.

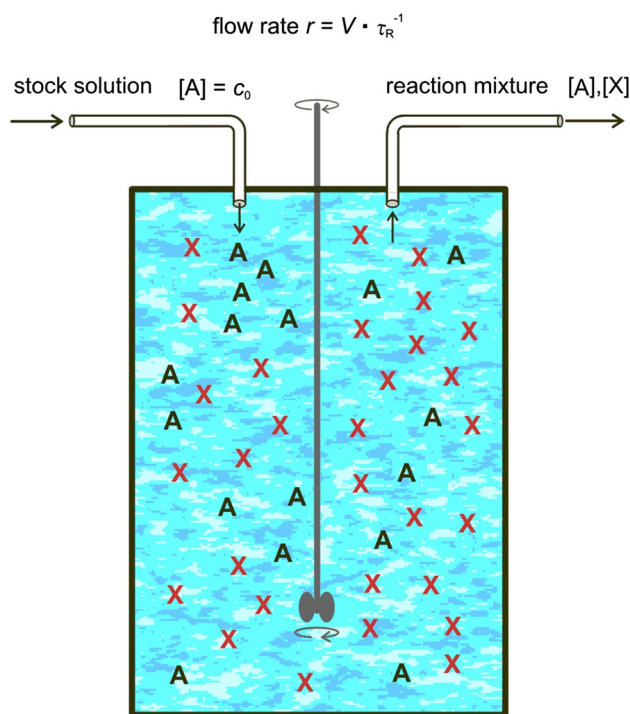
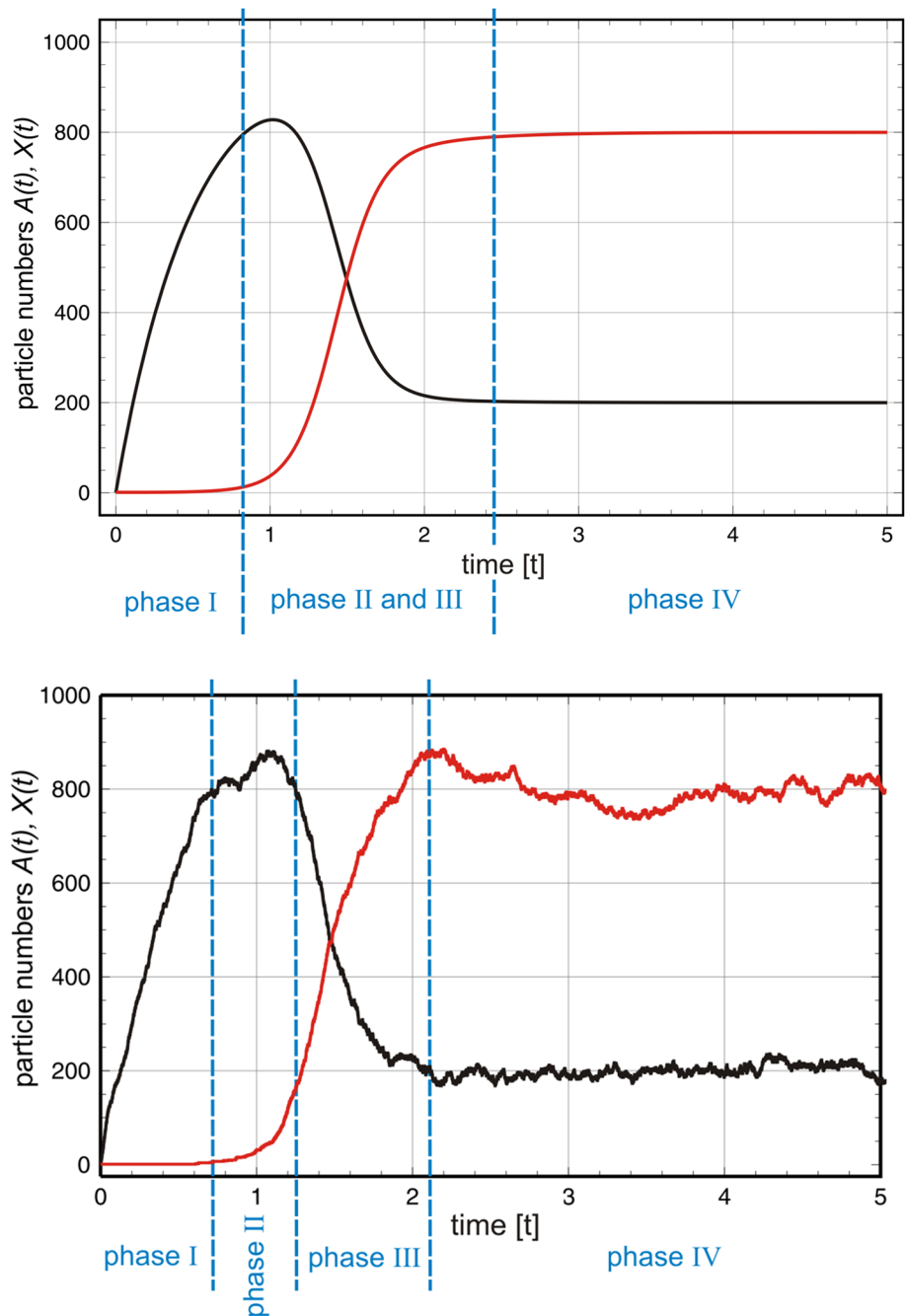


Fig. 2 The flow reactor as a device for studying autocatalytic reactions in an open system. The sketch shows a continuous-flow stirred-tank reactor (CFSTR). The interior of the reactor is assumed to be well mixed and at constant controllable temperature. A stock solution containing the resource A at concentration c_0 is assumed to flow into the reactor at a volume flow rate $r = V \cdot \tau_R^{-1}$ where V is the volume of the reactor and τ_R the mean residence time of a volume element in the reactor. The volume increase is exactly and instantaneously compensated by an outflow of the reaction mixture in the reactor

analysis of the steady states by means of the eigenvalues of the Jacobian matrix is straightforward: the reaction state S_1 is stable for volume flow rates $r < k c_0$, whereas extinction occurs for larger flow rates. The stationary concentrations for the irreversible autocatalytic reaction are readily obtained by putting $h = 0$. Interestingly, the stationary concentrations at the steady states depend on the reaction parameter h but the eigenvalues of the Jacobian and the stabilities are the same in the irreversible and in the reversible cases.

Typical deterministic and stochastic trajectories converging to S_1 for *empty reactor* initial conditions, no resource A ($A(0) = 0$) and the autocatalyst X in seeding quantities ($X(0) = 1, 2, 3, \dots$) are shown in Fig. 3. These initial conditions allow for the identification of different phases in the approach towards the long-time solutions and facilitate the analysis of anomalous fluctuations (Fig. 4). In phase I, the reactor is filled with resource A ; the concentration $A(t)$ increases until it is sufficiently large for making the reaction observable. Then, $X(t)$ increases and $A(t)$ decreases in phase II and approaches the long-time state in phase III. In phase IV, the system has eventually reached the steady state.

Fig. 3 The irreversible autocatalytic reaction $A + X \rightarrow 2X$ in a flow reactor of the type is shown in Fig. 2. The initial conditions are an empty reactor with seeding amounts of the autocatalyst: $A(0)=0$ and $X(0)=1, 2, 3, \dots$. The upper part of the figure shows the deterministic solution $a(t)$ (black) and $x(t)$ (red); the lower part presents a stochastic trajectory with $A(t)$ (black) and $X(t)$ (red). Four phases of the stochastic trajectory can be distinguished: (1) phase I: filling the reactor with the resource A, (2) phase II: decision to which (quasi) stationary state the trajectory is going to converge, (3) phase III: convergence towards the neighborhood of the (quasi)stationary state, and (4) phase IV: fluctuations around the (quasi)stationary state. In the deterministic trajectory phase II and phase III cannot be distinguished because of the uniqueness of solutions, and there are no fluctuations around the final state. Choice of parameters and initial conditions: $N=1000$, $k=0.01$ [$t^{-1} M^{-1}$], $r=2.0$ [$t^{-1} V^{-1}$], $A(0)=0$, $X(0)=1$



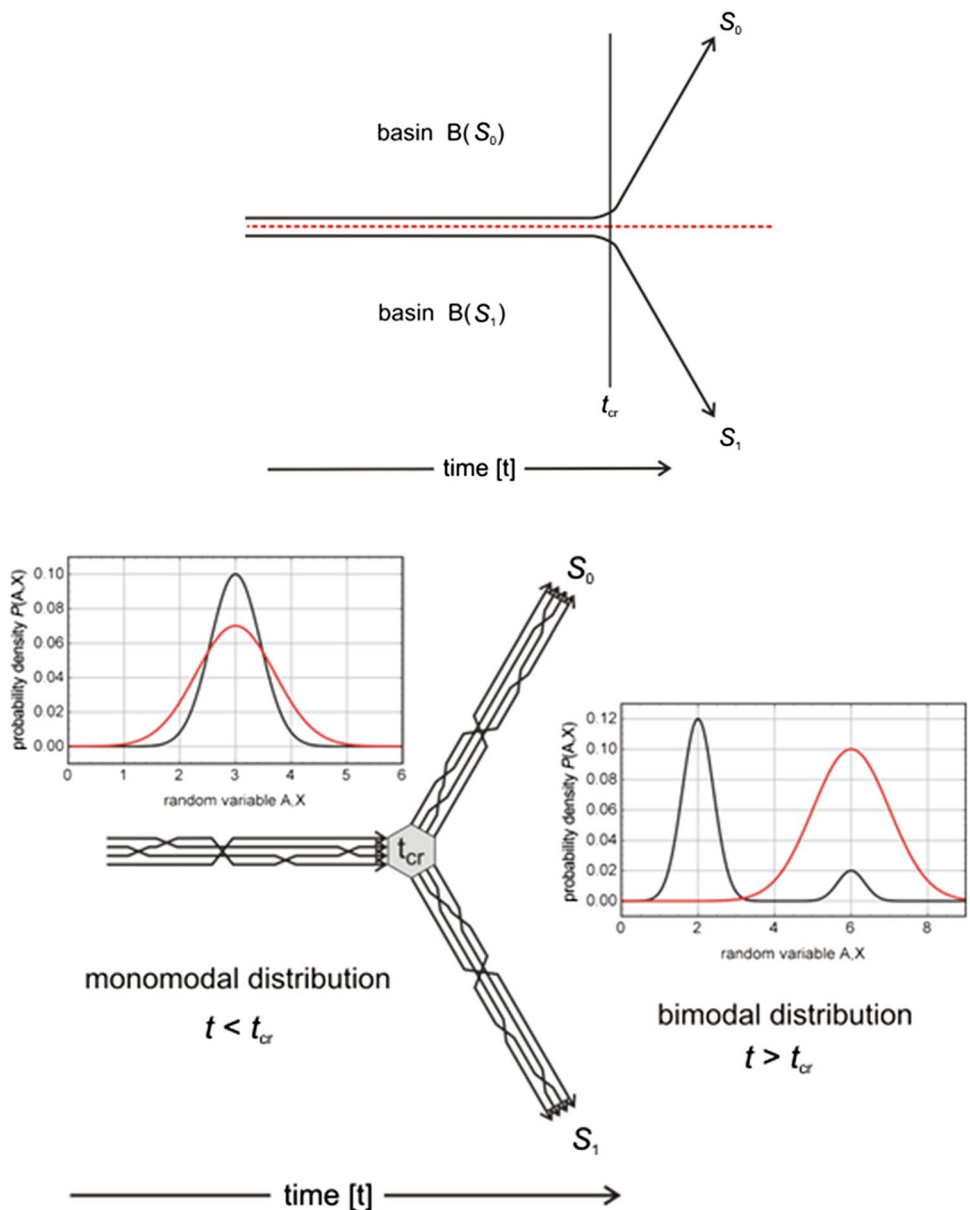
The trajectories fluctuate around the expectation value corresponding to the error band $E \pm \sigma$. For multiple long-time states in the stochastic system, a random decision in phase II determines the state towards which convergence occurs in phase III.

In case of multiple stationary states, the deterministic and the stochastic system behave differently. The uniqueness theorem for the solutions of differential equations requires that trajectories have well-defined α - and ω -limits within

a single basin of attraction $B(S)$.⁴ Basins are separated by *separatrices*. Alternatively, we could state: deterministic trajectories do not cross. No uniqueness theorem of trajectories is valid for stochastic dynamics: arbitrarily many different trajectories may have identical α -limits and in most cases there are no simple ω -limits, for example, point limits. Instead, the trajectories are fluctuating around an expectation

⁴ α - and ω -limits characterize the beginning and the end of a(n infinite) trajectory.

Fig. 4 The course of deterministic and stochastic bifurcations in the reversible autocatalytic reaction $A + X \leftrightarrow 2X$ in the flow reactor. In the deterministic process (upper part) uniqueness of trajectories implies the existence of basins of attraction, $B(S)$: Trajectories from all points within the basin converge to the stationary state S . The autocatalytic reaction in the flow reactor has two basins, $B(S_0)$ for the state of extinction and $B(S_1)$ for the reaction state. The bifurcation occurs at time $t = t_{cr}$. The two basins are separated by a *separatrix* (red dotted line). In the stochastic approach deterministic trajectories are replaced by the probability distributions $P(A)$ and $P(X)$ (lower part). The distributions are monomodal before the bifurcation at time $t = t_{cr}$ where they become bimodal. In the example sketched here, only A is present at the extinction state S_0 whereas both molecular species A and X are observed at the reaction state S_1



value and the ω -limit may be characterized best by an error band $E \pm \sigma$. As we shall see in the section on “Selection”, there may be quasistationary states in the stochastic system that have no counterparts in the deterministic scenario.

Figure 4 sketches a deterministic and a stochastic bifurcation. Two features in which the stochastic approach differs from the deterministic solution are important here: (1) deterministic trajectories are unique and different trajectories cannot cross but stochastic trajectories can and (2) the deterministic trajectories have simple ω -limits, stochastic trajectories in general converge only in terms of the expectation value E , whereas the trajectory itself continues to fluctuate around E with a one-standard-deviation error band $E \pm \sigma$. Deterministic trajectories are confined to one basin of attraction and stochastic trajectories are defined

by probability distributions P that are extended over the entire domain of the stochastic variable. At a bifurcation point, deterministic trajectories coming from different basins of attraction separate, whereas the probability distribution of stochastic trajectories switches from monomodal to bimodal. Examples of stochastic trajectories converging to two different states, extinction S_0 and reaction S_1 are shown in Fig. 5.

Anomalous fluctuations [19] are the result of convergence of stochastic trajectories to alternative (quasi)stationary states. Then, the expectation value E calculated in the conventional way represents a weighted mean of the expectation values for the two quasistationary states covered by a bimodal distribution and accordingly, the standard deviation σ is very large. More illustrative than the broad distribution

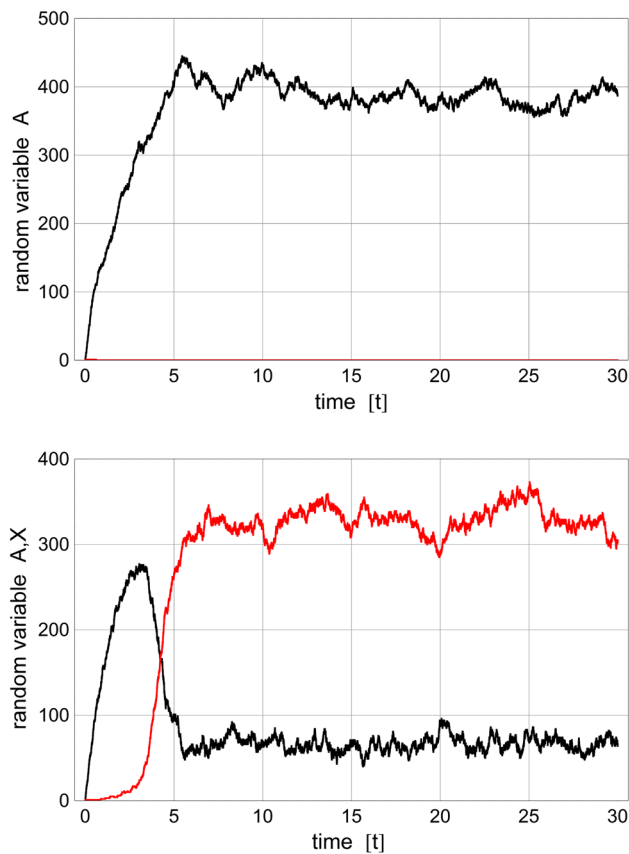


Fig. 5 Two stochastic trajectories of the reversible autocatalytic reaction $A + X \leftrightarrow 2X$ in the flow reactor with identical α -limits converging to two different states, extinction S_0 (upper plot) and reaction S_1 (lower plot). The two trajectories are computed with identical sets of parameters and initial conditions except different seeds s for the (pseudo)random number generator. Choice of parameters: $N=400$, $r=0.5$ [t^{-1} V], $k=0.01$ [t^{-1} M $^{-1}$], $h=0.0005$ [t^{-1} M $^{-2}$], seeds of the (pseudo)random number generator (Wolfram Mathematica, ExtendedCA): $s=491$ (upper plot) and $s=919$ (lower plot)

is the result of counting final states of stochastic trajectories (Table 2).

The results of trajectory counting are shown in Table 2 and reveal the expected dependence on initial conditions: the state of extinction S_0 is most commonly approached for $X(0)=1$ where it occurs with a frequency of about 78% for a population size of $N=100$ that decreases to 28% when the population size is raised to $N=1000$. At larger population sizes N , the frequency of extinction decreases faster with increasing number of initially present autocatalysts, $X(0)$: for $X(0)=1$ and $N=100$ the percentage of extinctions lies around 78% and is reduced to 47% for $X(0)=3$, whereas for $N=1000$ it decreases from 27.9 to 1.3%. The standard deviations σ_A and σ_X illustrate the result of anomalous fluctuations in case of multiple (quasi)stationary states: the standard deviations become very large in cases of small initial values of $X(0)$, which give rise to a high percentage of extinction.

Logistic equation

The logistic equation has been introduced in the first half of the nineteenth century by Pierre-François Verhulst as a quantitative model for exponential growth in systems with limited resources [20, 21]. He had been inspired by Thomas Robert Malthus who had discussed the disastrous consequences of uncontrolled exponential growth of human populations [22]. Malthus' work has also been influential for the development of the concept of natural selection through Charles Darwin [23] and Alfred Russel Wallace [24]. Verhulst's work was apparently forgotten and rediscovered several times in the early twentieth century [15, 25]. Despite its simplicity, Verhulst's differential equation is still in use, for example, in population ecology for modeling microbial growth.

The logistic equation does not explicitly consider the resource A but limits the growth of a (homogeneous) population, $\Pi = \{X\}$, by means of a negative quadratic term:

$$\frac{dx}{dt} = fx \left(1 - \frac{x}{C}\right) \quad \text{and} \quad x(t) = \frac{x(0)C}{x(0) + (C - x(0))\exp(-ft)}. \quad (8)$$

The parameter f is called the fitness of the species or subspecies⁵ and the parameter C , the carrying capacity of the ecosystem, represents the maximal populations size N that is sustained stably by the environmental conditions. It is worth reconsidering the nature of the logistic approach to reproduction: The multiplication process is implicitly modeled by the ansatz $dx/dt = fx$, whereas growth limitation is represented in abstract mathematical form by the term $-fx^2/C$. No hint is given on the physical nature of this limitation. It could be finite resources but also the consequence of overpopulation of the ecosystem without lack of nutrition.

In the deterministic approach, the logistic equation supports two stationary states, the extinction state S_0 with $\bar{x}^{(0)} = 0$, and the saturation state S_1 with $\bar{x}^{(1)} = C$. At saturation, the population size adopts its maximal value, $x=C$. The logistic growth curve is sigmoid: It starts out like an exponential function and then after having passed an inflection point goes into saturation. Stability analysis is straightforward: S_0 is unstable because small amounts of X introduced into the system lead to exponential growth in the early phase, S_1 is asymptotically stable and will be approached from x -values above ($x > C$) and below ($x < C$) the carrying capacity.

The interpretation of logistic growth as a stochastic process is less straightforward. Here, we mention only one

⁵ Fitness, in general, counts the number of fertile offspring that carry the genetic information into the next generation. A subspecies is a variant within a species consisting of individuals with the same fitness.

Table 2 Frequency of convergence towards the two (quasi)stationary states of the autocatalytic reaction $A + X \leftrightarrow 2X$ in the flow reactor

<i>N</i>		Initial particle numbers <i>X</i> (0)						
		1	2	3	4	5	7	10
100	<i>S</i> ₀	78.4 ± 4.6	59.2 ± 2.3	47.3 ± 4.8	35.4 ± 3.2	27.3 ± 5.8	18.0 ± 3.5	7.9 ± 1.5
	<i>S</i> ₁	21.6 ± 4.6	40.8 ± 2.3	52.7 ± 4.8	64.6 ± 3.2	72.7 ± 5.8	82.0 ± 3.5	92.1 ± 1.5
	σ_A	22.5	24.11	24.16	24.16	22.09	19.32	15.31
	σ_X	20.9	24.09	23.98	23.92	22.20	19.87	16.29
400	<i>S</i> ₀	39.1 ± 5.2	16.9 ± 4.7	7.5 ± 4.2	2.8 ± 1.7	1.5 ± 1.2	0.2 ± 0.4	0
	<i>S</i> ₁	60.9 ± 5.2	83.1 ± 4.7	92.5 ± 4.2	97.2 ± 1.7	98.5 ± 1.2	99.8 ± 0.4	100
	σ_A	161.7	123.3	87.42	53.31	37.94	28.63	13.31
	σ_X	161.8	124.3	88.98	56.85	42.80	34.55	22.98
1000	<i>S</i> ₀	27.9 ± 6.1	7.2 ± 3.4	1.3 ± 1.0	0.2 ± 0.4	0.1 ± 0.3	0	0
	<i>S</i> ₁	72.1 ± 6.1	92.8 ± 3.4	98.7 ± 1.0	99.8 ± 0.4	99.9 ± 0.3	100	100
	σ_A	399.4	224.7	123.7	57.89	42.11	31.05	9.79
	σ_X	399.8	226.9	127.4	65.13	51.66	42.64	31.40

Shown are the results of counting the numbers of trajectories leading to state *S*₀ or state *S*₁, respectively, and the standard deviations σ_A and σ_X in phase IV (Fig. 3). Choice of parameters: $k=0.01$ [t⁻¹ C⁻¹], $h=0.0005$ [t⁻¹ C⁻¹], $r=0.5$ [t⁻¹ V]. Sample size: 10 × 100 trajectories. The standard deviations were calculated as mean values in the time ranges $60 \leq t \leq 100$ for $N=100$ and $15 \leq t \leq 20$ for $N=400$ and $N=1000$

Additional values are (no values are given for *S*₀: 0 and *S*₁: 100)

$N=100$: $X(0)=15$, $S_0: 1.8 \pm 1.3$, $S_1: 98.2 \pm 1.3$, $\sigma_A=11.77$, $\sigma_X=13.27$

$X(0)=20$, $\sigma_A=10.02$, $\sigma_X=11.79$; $X(0)=30$, $\sigma_A=9.32$, $\sigma_X=11.26$

$X(0)=40$, $\sigma_A=8.95$, $\sigma_X=11.03$; $X(0)=100$, $\sigma_A=9.01$, $\sigma_X=11.02$

$N=400$: $X(0)=15$, $\sigma_A=8.45$, $\sigma_X=20.36$

$X(0)=20$, $\sigma_A=8.43$, $\sigma_X=20.14$; $X(0)=30$, $\sigma_A=8.45$, $\sigma_X=20.50$

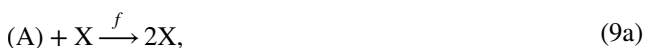
$X(0)=40$, $\sigma_A=8.44$, $\sigma_X=20.09$; $X(0)=100$, $\sigma_A=8.43$, $\sigma_X=20.08$

$N=1000$: $X(0)=15$, $\sigma_A=9.78$, $\sigma_X=31.21$

$X(0)=20$, $\sigma_A=9.78$, $\sigma_X=31.71$; $X(0)=30$, $\sigma_A=9.74$, $\sigma_X=31.61$

$X(0)=40$, $\sigma_A=9.81$, $\sigma_X=31.34$; $X(0)=100$, $\sigma_A=9.76$, $\sigma_X=31.68$

approach through modeling by means of a two-step chemical reaction, which gives to a rate equation with a positive linear term and a quadratic negative term as required for the limitation of growth. A negative quadratic term would also result, for example, from the reverse autocatalytic process, $A + X \leftarrow 2X$, but as shown in the section on the “Batch Reactor” this mechanism would introduce a reflecting barrier at the state $X=1$. Choosing an annihilation reaction for the introduction of the quadratic term turns out to be more suitable:



The expectation value embedded in the one standard deviation error band, $E(X(t)) \pm \sigma(X(t))$, looks very similar to the lower part of Fig. 1; stochastic delay is observed and quantitatively the δ -value it is very close to the value calculated for the reaction in the batch reactor.

The stochastic model in Eq. (9) has two stationary states that correspond to *S*₀ and *S*₁ of the deterministic system (8).

The state of extinction *S*₀ is an absorbing barrier and the state of saturation *S*₁ represents a quasistationary state.

Autocatalysis and natural selection

The replacement of the homogeneous population $\Pi(t) = \{X(t)\}$ by a population that is heterogeneous with respect to the fitness values of subspecies, $\Pi(t) = \{X_1, X_2, \dots, X_n\}$, results in a simple mathematical model for natural selection [16]. Concentrations and fitness values of subspecies are denoted by $x = (x_1, x_2, \dots, x_n)$ and $f = (f_1, f_2, \dots, f_n)$. With different fitness values for different subspecies, we obtain a Verhulst equation, which is generalized for heterogeneous populations:

$$\frac{dx_j}{dt} = x_j \left(f_j - \frac{\sum_{i=1}^n x_i}{C} \phi(t) \right) \quad \text{with} \tag{10a}$$

$$\phi(t) = \frac{1}{\sum_{i=1}^n x_i} \sum_{i=1}^n f_i x_i; \quad j = 1, \dots, n.$$

Table 3 Frequency of occurrence of (quasi)stationary states of selection between three species in the flow reactor

$\Delta f/f$	t_{fin}	$N=100$				$N=200$			
		$A(t_{\text{fin}})$	$X_1(t_{\text{fin}})$	$X_2(t_{\text{fin}})$	$X_3(t_{\text{fin}})$	$A(t_{\text{fin}})$	$X_1(t_{\text{fin}})$	$X_2(t_{\text{fin}})$	$X_3(t_{\text{fin}})$
0	600	1.5 ± 1.3	30.5 ± 3.9	34.2 ± 4.6	33.4 ± 4.1	0.5 ± 0.9	30.6 ± 4.6	30.9 ± 5.0	32.0 ± 4.7
0.02	600	1.8 ± 1.4	41.8 ± 4.8	32.9 ± 3.8	23.4 ± 4.0	0.6 ± 0.8	50.4 ± 5.7	27.7 ± 4.9	17.3 ± 2.6
0.04	400	2.4 ± 2.1	45.4 ± 5.0	31.3 ± 4.5	19.9 ± 2.5	0.7 ± 0.8	58.3 ± 4.6	25.6 ± 4.5	11.0 ± 2.6
0.1	400	2.1 ± 1.7	59.8 ± 5.5	28.0 ± 4.1	10.0 ± 2.9	0.4 ± 0.5	73.9 ± 4.1	20.6 ± 3.5	4.8 ± 1.9
0.2	400	1.9 ± 1.1	68.3 ± 4.5	23.1 ± 3.7	6.7 ± 2.8	0.5 ± 0.7	76.6 ± 4.1	19.3 ± 2.8	3.6 ± 1.7
0.4	400	2.3 ± 1.8	71.7 ± 6.0	20.8 ± 5.2	5.2 ± 2.4	0.9 ± 0.6	82.0 ± 4.2	13.8 ± 3.8	3.3 ± 1.7
1.0	200	2.7 ± 2.4	78.4 ± 4.7	15.8 ± 3.3	3.1 ± 1.5	0.9 ± 0.9	83.6 ± 4.0	12.6 ± 3.2	2.9 ± 1.5
1.8	200	4.3 ± 1.1	80.8 ± 2.9	13.6 ± 2.9	1.3 ± 1.2	1.5 ± 1.3	83.8 ± 3.3	12.7 ± 2.5	2.0 ± 1.7

Bold numbers indicate the scenarios closest to neutral evolution and survival of the fittest, respectively
 Given are the expectation values of the frequencies of convergence towards the state of extinction S_0 and the quasistationary states of selection $S_1, S_2,$ or S_3 for population sizes $N=100$ and $N=200$. The fitness is given in terms of differential fitness: $f_2=f, f_1=f+\Delta f/2f,$ and $f_3=f-\Delta f/2f$ with $f=0.1$ [$t^{-1} C^{-1}$], and t_{fin} [t] is the final time of the computer simulation. Further parameters: $r=0.5$ [$t^{-1} V$], $c_0=50$ or 100 [C]

The function $\phi(t)$ is the mean fitness of the population. We use $N = \sum_{i=1}^n x_i$ and reformulate the differential Eq. (10a) to get an equation for the growth of the total population:

$$\frac{dN}{dt} = N \left(1 - \frac{N}{C} \right) \phi(t). \tag{10b}$$

This equation can be solved analytically and yields:

$$N(t) = N(0) \frac{C}{N(0) + (C - N(0)) \exp(-\Phi(t))} \quad \text{with} \tag{10c}$$

$$\Phi(t) = \int_0^t \phi(\tau) d\tau.$$

Equation (10c) is not autonomous because the distribution of subspecies and its time dependence are required for the calculation of $\Phi(t)$. The complete solution is derived easiest in terms of normalized concentrations for the subspecies, $\xi = x/C$ with $\sum_{i=1}^n \xi_i = 1$, which fulfills the differential equation:

$$\frac{d\xi_j}{dt} = \xi_j (f_j - \phi(t)) \quad \text{with} \quad \phi(t) = \sum_{i=1}^n f_i \xi_i; \quad j = 1, 2, \dots, n. \tag{10d}$$

Equation (10d) can be solved exactly by conventional techniques—for example, by integrating factor transformation—and finally we obtain:

$$\xi_j(t) = \frac{\xi_j(0) \exp(f_j t)}{\sum_{i=1}^n \xi_i(0) \exp(f_i t)}. \tag{10e}$$

Insertion of (10e) into (10d) completes the solution.

Two properties of the solution of the extended Verhulst equation can be readily derived from Eq. (10e) and are of primary relevance for evolution: (1) selection of the fittest subspecies, and (2) optimization of the mean fitness of

populations. Property (1) follows directly from Eq. (10e): we consider the non-neutral case—all f values are different—in the limit of long times. The sum in the denominator converges to the term containing the largest exponential and this is the term resulting from the subspecies X_m with the largest fitness, $f_m = \max\{f_1, f_2, \dots, f_n\}$:

$$\sum_{i=1}^n \xi_i(0) \exp(f_i t) \approx \xi_m(0) \exp(f_m t) \quad \text{for large } t.$$

Insertion into Eq. (10e) leads to the long-time concentrations:

$$\bar{\xi}_j = \frac{\xi_j(0)}{\xi_m(0)} \lim_{t \rightarrow \infty} \{ \exp((f_j - f_m)t) \},$$

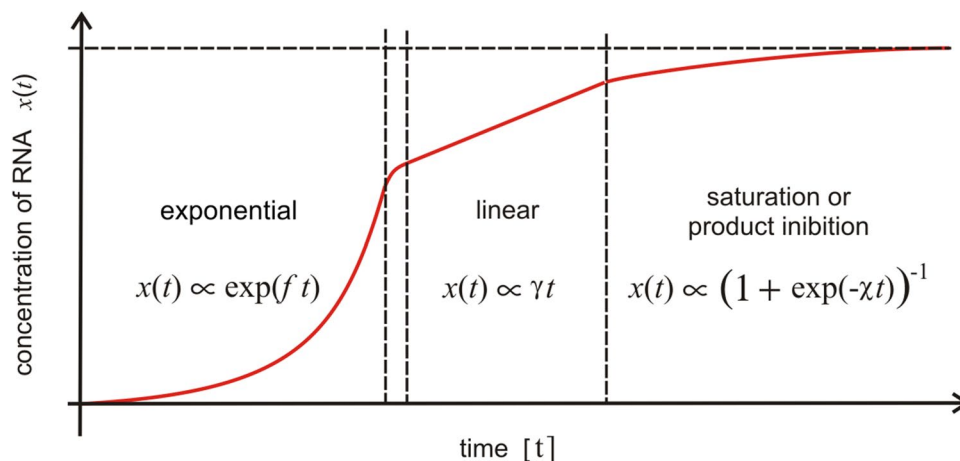
and provides the desired result: all $\bar{\xi}_j$ -values are zero except $\bar{\xi}_m = 1$ and selection of the fittest has occurred. A proof for property (2) can be given straightforwardly through the calculation of the time derivative of the mean fitness:

$$\frac{d\phi}{dt} = \sum_{i=1}^n f_i \dot{\xi}_i = \sum_{i=1}^n f_i \left(f_i \xi_i - \xi_i \sum_{j=1}^n f_j \xi_j \right) = \langle f^2 \rangle - \langle f \rangle^2 = \text{var}\{f\}. \tag{11}$$

Since a variance is always nonnegative, the mean fitness $\phi(t)$ is a non-decreasing function of time, it is optimized during the selection process and $d\phi/dt$ vanishes if and only if $\text{var}\{f\} = 0$, i.e., when all subspecies have the same fitness and the population has become homogeneous through the selection process.

The stochastic effects on selection are rather spectacular: instead of selection of the fittest every subspecies can be selected. Stochastic dynamics has one absorbing barrier corresponding to extinction of all subspecies, $S_0=(A=C, X_j=0 \forall j=1, 2, \dots, n)$, and n quasistationary states of selection,

Fig. 6 Overall kinetics of complementary RNA replication by Q β -replicase. The replication process in the batch reactor allows for a distinction of three different phases: (1) an exponential growth phase with excess enzyme, (2) a linear growth phase where practically all enzyme molecules are bound to RNA templates, and (3) a phase of saturation where very little or no RNA is produced because of product inhibition. Drawn after Fig. 10 in [9]



$S_j = (A = rf_j, X_j = C - rf_j, X_i = 0 \forall i, i \neq j)$. The fitness values of the individual subspecies determine the probabilities of selection. A specific example of selection among three subspecies is shown in Table 3. Fitness values cover the entire range from neutrality, $f_1 = f_2 = f_3 = f$, with equal probability of all three states S_1, S_2 , and S_3 to strong selection, where $X_1 = X_m$ is chosen preferentially. Under the initial condition of an empty reactor, the decision on the state towards which the trajectory converges is done by random fluctuations in phase II (Fig. 3).

It is worth noticing that natural selection and optimization of mean fitness occur in the population no matter whether one considers the logistic equation or autocatalysis in the batch or in the flow reactor. Remarkable indeed is the fact that during the nineteenth century no attempt was made of casting the principle of natural selection into rigorous mathematical form, although all mathematical tools including Verhulst’s equation have been available already 20 years before the publication of Darwin’s “Origin of Species”. The first true milestone of mathematics in evolution is Ronald Fisher’s book published as late as 1930 [26].

Conclusion

The first-order autocatalytic single-step reaction, $A + X \leftrightarrow 2X$, can be analyzed by mathematical tools in great detail and in diverse environments. Although pure single-step autocatalysis is extremely rare, the results derived here are of general importance since—as already mentioned in the introduction—complex many-step mechanisms may follow simple kinetics under certain conditions. As an example, we mention here the kinetics of plus–minus replication of RNA which was studied in great detail by Christoph Biebricher in the laboratory of Manfred Eigen [9, 10, 27]:



where X_+ and X_- are RNA molecules with complementary sequences. The growth curve of total RNA concentration in a batch reactor is sketched in Fig. 6. At low RNA concentration, perfect exponential growth as expected for the reaction $A + X \xrightarrow{k} 2X$ with $k = \sqrt{k_+ k_-}$ is recorded [27] and all properties of autocatalysis discussed in this contribution can be observed provided the conditions are chosen such that the reaction remains in the range of the exponential growth phase. A simplified mechanism of complementary replication has been proposed and analyzed [28].

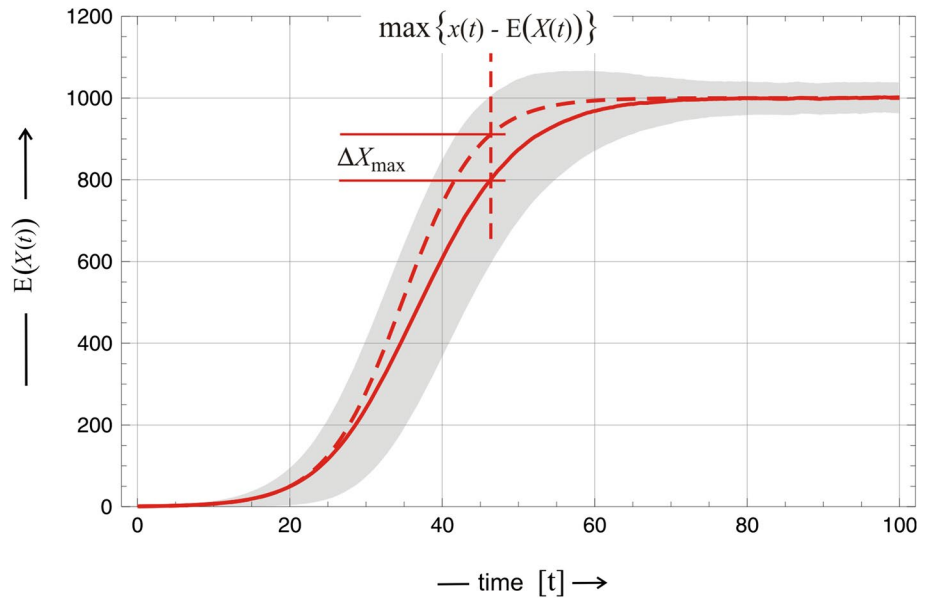
In the high concentration limit of the deterministic approach autocatalysis shows two specific features that are absent in other reactions: (1) due to self-enhancement sigmoid solution curves are observed with single-step reactions and (2) autocatalytic reactions require (at least) seeding quantities of the autocatalyst for ignition of the reaction, because $x(0) = 0$ leads to $x(t) = 0 \forall t$ —no reaction takes place.

Three stochastic features are observed with autocatalytic processes at low particle numbers: (1) thermal fluctuations that are the same as in every other chemical reaction and lead to $\delta N \propto \sqrt{N}$ at thermodynamic equilibrium, (2) stochastic delay at small initial concentrations, and (3) anomalous fluctuations in case of multiple long-time states.

Stochastic delay with autocatalytic reactions to our knowledge has not been reported before in this form. It is a consequence of self-enhancement together with discretization at low particle numbers. Interestingly, the dependence of stochastic delay on initial particle numbers and population size is rather simple and an increment δ has been derived that is fairly constant and independent of the parameters mentioned.

Anomalous fluctuations are dominant at very small initial particle numbers and result from stochastic bifurcations in case of multiple (quasi)stationary states. Such small particle numbers are commonly not important in chemistry but they

Fig. 7 A sketch of an unambiguous computation of the stochastic delay in different stochastic models. The difference $\Delta X_{\max}(t) = x(t) - E(X(t))$ is scanned as a function of time and the maximum value of this difference, ΔX_{\max} is determined



play an important role in evolution. Every mutant after all starts out from a single copy.

Methods

All numerical computations reported here were performed with Wolfram Mathematica, versions 7-11. Deterministic solution curves of kinetic differential equations were calculated with the Mathematica ODE integration routine.

All quantities with physical dimensions are given in arbitrary units, which are written in square brackets. This implies that in comparison with experimental data the appropriate unit element can be inserted. For example, time can be expressed as $[t] = \text{min, sec, msec, } \mu\text{s, etc.}$ The unit elements used are time $[t]$, volume $[V]$, concentration $[M] = [\text{mol V}^{-1}]$, and number density $[C] = [\text{particles V}^{-1}]$. Molecules are denoted by upper case letters: A, X, etc. The notation used for the most frequently used variables is:

concentrations: $[A(t)] = a(t)$, $[X(t)] = x(t)$, ...

number densities (particle numbers per unit volume):

$$[A(t)] = A(t), \dots$$

Stochastic trajectories were computed by the stochastic simulation algorithm (SSA) developed by Daniel Gillespie (for reviews see Refs. [29, 30]). We used here a biologically motivated simulation routine for Gillespie's algorithm, which has been developed by Bruce Shapiro within the xCellerator software design project [31].

The standard deviation is derived from the second moment of the distribution and, therefore, more sensitive to random fluctuations than the first moment, the expectation

value. Accordingly, the corresponding quantities in Tables 2 and 3 are less accurate.

Calculation of the increment δ of the stochastic delay: the calculation of the stochastic delay is sketched in Fig. 7. The quantity that is extracted from the deterministic solution curve $x(t)$ and the stochastic expectation value $E(X(t))$ is the maximum difference ΔX_{\max} at some time t_{\max} .⁶ This quantity turns out to be proportional to the population size N and inversely proportional to the initial number of autocatalysts $X(0)$:

$$\delta = \Delta X_{\max} \frac{N}{X(0)}.$$

Acknowledgements Open access funding provided by University of Vienna. Support of the work by the University of Vienna and the Santa Fe Institute is gratefully acknowledged. The author acknowledges many fruitful discussions with Professors Ivo L. Hofacker and Christoph Flamm. The public accessibility of the stochastic simulation algorithm (SSA) within the xCellerator project by Bruce Shapiro is acknowledged.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

⁶ It could seem more natural to use the maximum difference along the time axis, Δt_{\max} for the calculation of the stochastic delay. The qualitative results obtained are the same as those shown in Table 1 but the numerical computations are more involved.

References

1. Ostwald W (1890) *Ber Verh Kgl Sächs Ges Wiss Leipzig. Math Phys Class* 42:189
2. Epstein IR, Pojman JA (1998) *An introduction to nonlinear chemical dynamics Oscillations, waves, patterns, and chaos*. Oxford University Press, New York
3. Sagués F, Epstein IR (2003) *J Chem Soc Dalton Trans* 2003:1201
4. Zhabotinsky AM (1991) *Chaos* 1:379
5. Castets V, Dulos E, Boissonade J, De Kepper P (1990) *Phys Rev Lett* 64:2953
6. Eigen M (1971) *Naturwissenschaften* 58:465
7. Ostwald W (1900) *Z Phys Chem* 35:33
8. Joyce GF (2007) *Angew Chem Int Ed* 46:6420
9. Biebricher CK (1983) *Evol Biol* 16:1
10. Biebricher CK, Eigen M, William C, Gardiner WC Jr (1983) *Biochemistry* 22:2544
11. Field RJ, Körös E, Noyes RM (1972) *J Am Chem Soc* 94:8649
12. Field RJ, Noyes RM (1974) *J Chem Phys* 60:1877
13. Turing AM (1952) *Philos Trans R Soc Lond B* 237:37
14. Walgraef D (2012) *Spatio-temporal pattern formation. With examples from physics, chemistry, and materials science*. Springer, New York
15. Kingsland S (1982) *Quart Rev Biology* 57:29
16. Schuster P (2010) *Theory Biosci* 130:71
17. Schmidt LD (2005) *The engineering of chemical reactions*, 2nd edn. Oxford University Press, New York
18. Arslan E, Laurenzi IJ (2008) *J Chem Phys* 128:e015101
19. de Pasquale F, Tartaglia P, Tombesi P (1980) *Lett Nuovo Cimento* 28:141
20. Verhulst P-F (1838) *Corresp Math Phys* 10:113
21. Verhulst P-F (1847) *Mèm Acad R Sci Lett Beaux-Arts Belg* 20:1
22. Malthus TR (1798) *An essay of the principle of populations as it affects the future improvement of society*. J Johnson, London
23. Darwin C (1859) *On the origin of species by means of natural selection or the preservation of favoured races in the struggle for life*. John Murray, London
24. Wallace AR (1870) *Contributions to the theory of natural selection*, 2nd edn. Macmillan, London
25. Lloyd PJ (1967) *Popul Stud* 21:99
26. Fisher RA (1930) *The genetical theory of natural selection*. Oxford University Press, Oxford
27. Biebricher CK, Eigen M, Luce R (1981) *J Mol Biol* 148:391
28. Gassner B, Schuster P (1982) *Monatsh Chem* 113:237
29. Gillespie DT (2007) *Annu Rev Phys Chem* 58:35
30. Gibson MA, Bruck J (2000) *J Phys Chem A* 104:1876
31. Shapiro BE, Levchenko A, Meyerowitz EM, World BJ, Mjolsness ED (2003) *Bioinformatics* 19:677

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.