

## Self-Organising Autocatalysis

Nathaniel Virgo<sup>1,3</sup>, Simon McGregor<sup>2,4</sup> and Takashi Ikegami<sup>1,5</sup>

<sup>1</sup>Ikegami Laboratory, University of Tokyo, Japan      <sup>2</sup>University of Sussex, UK

<sup>3</sup>nathanielvirgo@gmail.com      <sup>4</sup>s.mcgregor@sussex.ac.uk      <sup>5</sup>ikeg@sacral.c.u-tokyo.ac.jp

### Abstract

Life on Earth must originally have arisen from abiotic chemistry. Since the details of this chemistry are unknown, we wish to understand, in general, which types of chemistry can lead to complex, life-like behaviour. Our recent work has shown that the inclusion of thermodynamic principles in simple artificial chemistry models can result in the self-organisation of autocatalytic cycles. In this paper we present some new insights into why this happens. Our model is given a more mathematical treatment, allowing us to better understand the assumptions that lead to this phenomenon. The simplest type of autocatalytic cycle results in exponential growth. Through dynamical simulation we demonstrate that when these simple first-order cycles are prevented from forming, the system achieves super-exponential growth through more complex, higher-order autocatalytic cycles. This leads to non-linear phenomena such as oscillations and bistability, the latter of which is of particular interest regarding the origins of life.

### Introduction

How can abiotic chemistry give rise to phenomena such as replication, metabolism and evolution? This question is at the core of the origins of life as a field of inquiry, and there are many approaches to it. One such methodology is *artificial chemistry*, the study of more-or-less abstract artificial chemical reaction networks and their emergent dynamical properties. This field can be traced back to the work of Fontana and Buss (1994) and Kauffman (1986). More recently it has been used to show that evolution can occur without the need for template replication (Vasas et al., 2012), and to investigate the mechanism behind HCN polymerisation, a prebiotically-relevant reaction that is too complex to model with traditional approaches (Andersen et al., 2013).

In these models, thermodynamic principles play at best a minor role. However, the nature of living organisms as far-from-equilibrium structures is fundamental to our understanding of biology at the chemical level. Our work aims to show that new insights can be gained by putting thermodynamics at the heart of artificial chemistry approaches.

Last year we presented a simple yet thermodynamically reasonable artificial chemistry (Virgo and Ikegami, 2013).

When held out of equilibrium, this chemistry formed moderately large and complex autocatalytic cycles. (That is, moderately large sets of species that are collectively capable of exponential growth.) Unlike many previous studies, this chemistry did not include any *a priori* notion of catalysis. Instead, catalytic effects emerged as dynamical properties of the reaction network. The system was allowed to dynamically “choose” the direction of its reactions according to thermodynamic principles, and it was this property that allowed it to self-organise into autocatalytic cycles.

This phenomenon of self-organising autocatalysis happens under some circumstances but not others. In order to fully understand the result we must investigate which classes of reaction networks will lead to autocatalysis and which will not. Moreover, the resulting cycles can be more or less complex, and we wish to understand what factors affect this. A comprehensive answer to this question would be a great help in understanding the origins of life. If we know which types of chemical system are likely to lead to complex, metabolism-like reactions then it will give us clues about what to look for in identifying the types of abiotic chemistry from which life eventually arose.

In this paper we continue this project. We present some new insights into why autocatalysis occurs in far-from-equilibrium systems, and into what is needed for this to happen. As with all non-equilibrium phenomena, the autocatalysis in our systems is caused by the tendency of energy (and other conserved currents) to “flow downhill” from more to less constrained states, i.e. from low to high entropy. Informally, it seems that nature likes to take the “path of least resistance” in order to achieve this. If the easiest path is a simple, direct chemical reaction then no complex behaviour will be observed. Without a “direct” pathway from initial conditions to equilibrium, autocatalytic pathways tend to be observed. Blocking the simplest type of autocatalytic cycle results in the formation of more complex cycles, giving rise to non-linear phenomena such as bistability and oscillations.

Autocatalytic cycles can only emerge if they are possible within the reaction network. However, if the direction in which reactions occur is not fixed *a priori* then even the

simplest reaction schemes contain many autocatalytic sub-networks.

The results we present are very robust. With one minor exception (the oscillatory behaviour shown in Figure 3) they require no parameter tuning at all. Moreover, our mathematical arguments are applicable to a much broader class of chemical systems than the specific models we present. Thus, although our model is relatively simple and abstract, we expect that in future work similar results will be found in much more realistic scenarios.

## Thermodynamics in Artificial Chemistry

A glance at a chemistry text book will reveal a large number of formulae that appear to be largely empirical in nature, and this can give the impression that chemical thermodynamics is something very contingent upon the details of molecular interactions. Perhaps because of this, there is a tendency in artificial chemistry to “abstract away” thermodynamics, choosing a set of reactions arbitrarily, without regard to energetic considerations. This may be justified by saying that some of the reactions are driven by an energy source that is external to the system and is not explicitly modelled. This methodology dates back to the origins of artificial chemistry in the work of Fontana and Buss (1994).

Such an abstraction is of course perfectly valid, but we hope to show that important insights can be gained if we choose not to make it. The dynamics of a system where there are only a few energy sources and the majority of reactions must “flow downhill” are quite different from those in which the reactions proceed in arbitrarily-chosen directions. Moreover, the flow of free energy through the network can be tracked, giving us insights into its dynamics that would be missing if the energetics were not modelled.

Although chemical thermodynamics might appear to be a rather empirical set of results, its core principle (the second law) is an extremely fundamental property of physical systems. Essentially, it boils down to a requirement that for a system with no external supply of energy, there must exist a global Lyapunov function, which may be called the entropy or the free energy, depending on whether the system is connected to a heat bath. The existence of such a function is derived from statistical mechanics, and is a consequence of the time-reversibility of physics on the microscopic level. The details of this formalism are beyond the scope of this paper and may be found, for example, in Kondepudi and Prigogine (1998). We mention it in order to emphasise that our results are less contingent upon the details of molecular interactions than they might at first appear.

The key feature of a thermodynamically reasonable artificial chemistry is that the reaction network doesn’t specify the *direction* of the reactions. If the network includes a reaction, such as  $A + B \rightarrow C$ , it must also include the reverse reaction,  $C \rightarrow A + B$ . In general, these two reactions may proceed at different rates. We refer to the difference in their

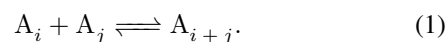
rates as the *net rate* of the bidirectional reaction,  $A + B \rightleftharpoons C$ .

The dynamics must be such that, in the absence of any energy sources, the system will eventually reach a state of *detailed balance* in which the net rate of every reaction is zero. That is, every forward reaction must be balanced, pairwise, by its corresponding reverse reaction. (This point is of fundamental importance. In a general artificial chemistry a dynamical equilibrium can be formed by a cycle, such as  $A \rightarrow B$ ;  $B \rightarrow C$ ;  $C \rightarrow A$ . In a thermodynamic system with no driven reactions this is a physical impossibility.) Every net reaction must always proceed in the direction that takes the system closer to this equilibrium state. Thus, no reaction can occur at a non-zero net rate unless the system is either in a transient state or is being driven by an external power source.

In our model we use a simple implementation of these dynamics, known as *mass action kinetics*. Essentially this corresponds to a model of a well-mixed reactor in which molecules collide with each other at random. When two molecules collide they react with a probability determined by the “rate constant” of the reaction. Thermodynamics puts constraints on the values of the rate constants; these constraints are satisfied by the dynamics described below. (See Kondepudi and Prigogine (1998) for details of the formalism, and Virgo and Ikegami (2013) for its application to models of the type described below.)

## The A-polymerisation model

When implemented in a thermodynamically plausible way, even very simple reaction schemes can give rise to interesting behaviour. Our model consists of *unary strings* A, AA, AAA etc., which we write as  $A_1, A_2, A_3, \dots$ . In principle these are countably infinite, but we impose a limit for computational reasons. These species can be thought of as (non-oriented) polymers, with  $A_1$  (or simply A) representing a monomer, and  $A_i$  a polymer (or oligomer) of length  $i$ . All reactions are of the form



That is, two oligomers may join together into a longer one, or an oligomer may split into two smaller ones. Reactions of this form conserve the total number of A monomers. We take the forward and reverse reaction rates to be constants  $k_s$  and  $k_d$ , (for “synthesis” and “decomposition”) with the same value for every reaction.

This corresponds to an assumption that a constant amount of energy  $E_{AA}$  is involved in every A–A bond (Virgo and Ikegami, 2013). Briefly, the free energy of formation of species  $A_i$  is given by  $G_{A_i}^\circ = (i - 1)E_{AA}$  (plus other terms that cancel), and the equilibrium constant for reaction  $A_i + A_j \rightleftharpoons A_{i+j}$  is given by  $\frac{k_s}{k_d} = \exp \frac{G_{i+j}^\circ - G_i^\circ - G_j^\circ}{k_B T} = \exp \frac{E_{AA}}{k_B T}$ , which is indeed the same for every reaction.

We follow the slightly unusual convention of treating  $A_i + A_j \rightleftharpoons A_{i+j}$  and  $A_j + A_i \rightleftharpoons A_{i+j}$  as two distinct reactions when  $i \neq j$ . This is merely a convenience that allows us to write the equations below in a more compact form. (Without this, reactions of the form  $2A_i \rightleftharpoons A_{2i}$  need to be treated as a special case, since there are two distinct bonds that can be broken to split an  $A_5$  into an  $A_2$  and an  $A_3$ , but only one way to split  $A_4$  into two  $A_2$  dimers.)

If every reaction of the form in Equation 1 is included in the network then the results are not especially interesting. However, we can construct A-polymerisation chemistries with more complex dynamics by disallowing particular reactions. In real chemistry reactions may be disallowed (or happen at very slow rates) for a variety of reasons, such as conservation laws or steric effects. Our goal here is not to model such effects specifically, but rather to understand, in general, the effects of imposing constraints on the reaction network.

Let  $r_{i,j} = 1$  if the reaction  $A_i + A_j \rightleftharpoons A_{i+j}$  is included in the network, and 0 otherwise, with  $r_{i,j} = r_{j,i}$ . Mass action kinetics then leads to the ordinary differential equations

$$\begin{aligned} \frac{da_i}{dt} = & \sum_{j=1}^{i-1} r_{i,j-i} (k_s a_j a_{i-j} - k_d a_i) \\ & + \sum_{j=1}^{\infty} (r_{i,j} + r_{j,i}) (k_d a_{i+j} - k_s a_i a_j) + \phi_i, \end{aligned} \quad (2)$$

where  $a_i$  represents the concentration of  $A_i$ . The term  $k_s a_j a_{i-j}$  represents the formation of  $A_i$  through the joining of  $A_j$  to  $A_{j-i}$ ; the  $k_d a_i$  term is the loss of  $A_i$  through splitting into  $A_j$  and  $A_{j-i}$ ; the  $k_d a_{i+j}$  term is the gain in  $A_i$  due to  $A_{i+j}$  splitting into  $A_i$  and  $A_j$ ; and  $k_s a_i a_j$  represents the loss of  $A_i$  by joining to  $A_j$  to make  $A_{i+j}$ .

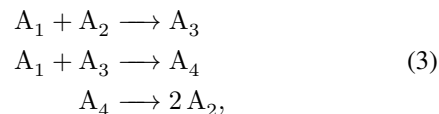
The  $\phi_i$  terms represent the flux of each species  $A_i$  in or out of the system. These terms may be set to non-zero values in order to model an open system with external energy sources. In most of this paper these terms will be set to zero, representing a closed system whose free energy comes purely from a non-equilibrium initial state.

## The linear case: self-organisation of first-order autocatalysis

In this section we present a simple non-equilibrium A-polymerisation chemistry that can be solved semi-analytically. This gives some insight into some sufficient conditions for exponential growth, and the reasons why it emerges under these conditions.

In this chemistry we set  $r_{1,1} = 0$  and all other  $r_{i,j} = 1$ . That is, two monomers ( $A_1$ ) are not allowed to directly join to form a dimer ( $A_2$ ) and, *vice versa*, a dimer cannot directly split into monomers. All other reactions are permitted. There are multi-step reactions that can convert monomers

into dimers in this system, such as the autocatalytic cycle



which has the net reaction  $2A_1 + A_2 \longrightarrow 2A_2$ . (Here, as below, we use unidirectional arrows to represent the net direction in which these reversible reactions proceed.) This is called a first-order autocatalytic cycle, because there is only one  $A_2$  on the left-hand side. First-order autocatalysis leads to exponential growth, as we will see below. (See King (1978, 1982) for background on the theory of autocatalytic cycles, and Andersen et al. (2012) for a formal definition.)

Let us suppose that the initial conditions contain a very high concentration of  $A_1$ , and trace amounts of every other species. That is,  $a_1$  is large and  $a_i$  very small for  $i > 1$ . Under these assumptions, the rate of change of  $a_1$  will be small in comparison to its value, so that we can treat it as constant. Using the approximation  $a_i a_j = 0$  for  $i, j > 0$ , Equations 2 become

$$\begin{aligned} \frac{da_2}{dt} &= 2 \sum_{j=1}^{\infty} (k_d a_{2+j}) - 2k_s a_1 a_2 \\ \frac{da_i}{dt} &= 2k_s a_1 a_{i-1} - (i-1)k_d a_i \\ &+ 2 \sum_{j=1}^{\infty} (k_d a_{i+j}) - 2k_s a_1 a_i \end{aligned} \quad (4)$$

for  $i > 2$ . This may be written  $da/dt = \mathbf{R}\mathbf{a}$ , where  $\mathbf{a}$  is a column vector with elements  $a_2, a_3, \dots$ , and  $\mathbf{R}$  is the matrix

$$k_d \begin{pmatrix} -2K & 2 & 2 & 2 & \dots \\ 2K & -2K - 2 & 2 & 2 & \dots \\ 0 & 2K & -2K - 3 & 2 & \dots \\ 0 & 0 & 2K & -2K - 4 & \dots \\ \vdots & & & & \ddots \end{pmatrix}, \quad (5)$$

with  $K = a_1 k_s / k_d$ .

It should be noted that although this approximation leads to a linear system of equations, it is not a ‘‘near-equilibrium’’ approximation in the thermodynamic sense. The thermodynamic equilibrium of this type of system will in general contain a positive concentration of every species. We are linearising the system around the state where most of the species are at zero concentrations; this state is a dynamical fixed point, but it is very far from thermodynamic equilibrium.

Since this is a linear dynamical system, its asymptotic behaviour is determined by the eigenvectors and eigenvalues of  $\mathbf{R}$ . In particular, since the only negative elements of  $\mathbf{R}$  are on the diagonal,  $\mathbf{R}$  has only one eigenvector  $\mathbf{v}$  with all positive entries (by application of the Perron-Frobenius theorem to

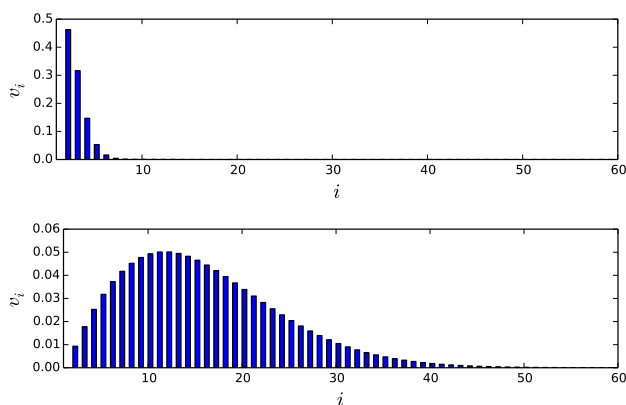


Figure 1: The elements of the leading eigenvector of the matrix  $R$ , for  $K = 1.0$  (top) and  $K = 100.0$  (bottom). The corresponding eigenvalues depend on  $k_d$ , but are positive in both cases, meaning that the concentrations grow exponentially while remaining proportional to the elements of the leading eigenvector. Both represent exponential growth, but through different mechanisms.

the irreducible matrix  $e^{\varepsilon R} \approx I + \varepsilon R$  for sufficiently small  $\varepsilon$ ), whose corresponding eigenvalue  $\lambda$  is real. For sufficiently large  $t$  we have  $a_i(t) \approx Av_i e^{\lambda t}$ , with  $A$  a real constant that depends on the initial conditions.  $A = 0$  in the case where every  $a_i = 0$  (for  $i > 1$ ). For all other initial conditions,  $A$  is positive.

The eigenvector and corresponding eigenvalue can readily be found by power iteration of  $e^{R(N)}$ , where  $R^{(N)}$  is an  $N \times N$  matrix whose entries are the same as those of  $R$  (with the exception that  $R_{NN}^{(N)} = -N$ , due to the absence of the reaction  $A_N + A_1 \rightarrow A_{N+1}$ . For sufficiently large  $N$  this makes very little difference).

Figure 1 shows two examples. In both, the leading eigenvalue of  $R$  is positive, meaning that the species other than  $A_1$  will grow exponentially. In the case where  $K = 1$ , this growth is mostly due to the minimal autocatalytic cycle shown in Equation 3, with a few side reactions producing longer polymers.

However the case with a larger  $K$  (meaning proportionally faster reaction rates in the synthesis direction, or just a higher concentration of  $A_1$  initially) is dominated instead by larger molecules, and its growth is therefore due to much larger autocatalytic cycles. There is no one cycle that dominates; instead the microscopic picture is of a collection of polymers of many different lengths, which grow at their ends due to monomer addition, but which also occasionally split somewhere in the middle, giving rise to two new polymers.

It is a theorem in mass action kinetics that there is a global Lyapunov function given by the Gibbs energy, whose minimum is at the thermodynamic equilibrium point. Since we are linearising around a different fixed point, this guarantees

that it will not be stable; at least one of its eigenvalues must be non-negative. Because of this, there are no parameters that have to be tuned in order to observe autocatalysis in this model. For any value of  $K$  there will always be a region around the point  $a_2 = a_3 = \dots = 0$  in which autocatalytic growth occurs.

To derive these results we assumed that  $a_i$  was small for  $i > 1$ , but since the concentrations of these are increasing exponentially, this assumption will be violated eventually. This means that either  $a_1$  will start to decrease at a substantial rate, changing the value of  $K$ , or terms of the form  $a_i a_j$  will start to be important, changing the dynamics. In a real system, this may or may not happen before the exponential growth becomes manifest enough to be clearly observable.

It is worth examining the assumptions we used to obtain this result. One important part of the model is that the reaction  $2A_1 \rightleftharpoons A_2$  is removed from the reaction network. The reason for this is that this reaction would rapidly produce enough  $A_2$  that the assumption of small  $a_2$  would be violated. We propose that in general the absence of such a “direct route” to product formation will be necessary in order for autocatalysis to be observed.

The other important assumption was that the polymer species have (initially) very low concentrations, so that the  $a_i a_j$  terms vanish. Physically, this corresponds to a solution so dilute that two polymer molecules are vanishingly unlikely to meet and react. ( $A_1$  monomers, however, are highly concentrated, so the probability of reacting with a monomer is high.) Under these conditions, exponential growth can occur through the molecules (a) “growing” into larger molecules by reacting with monomers, and (b) spontaneously splitting into two smaller molecules.

Such dynamics can be expected in much more complex thermodynamically reasonable chemistries, as long as these conditions are obeyed; our reasoning is applicable to a much broader class of models than the one presented here. The chemistry need not take the form of A-polymerisation; this reasoning will hold as long as there is some set of “food” species that cannot directly react to form products, and so long as there are no energetically irreversible “sinks” in the network. No autocatalytic cycles can arise unless they already exist in the network of potential reactions, but our example shows that for even such a simple chemistry as A-polymerisation there are many such potential loops. We suspect these conditions will be satisfied by many chemistries, as long as they are sufficiently densely connected and all the reactions have some degree of reversibility.

The mechanism as described so far results in *exponential* growth. Equations 4 are a linear dynamical system, and therefore the only possibilities are exponential growth, exponential decay or neutral stability. (Oscillation around the equilibrium point is impossible, because it would require concentrations to become negative.) The more general assumptions outlined in the previous paragraph will also lead

to linear dynamical systems, resulting either in exponential growth or decay. In the next section we explore the importance of non-linearity in chemical networks, and we demonstrate that with some changes to the reaction network, self-organisation can produce super-exponential (or “hyperbolic”) growth in a simple A-polymerisation chemistry.

### Non-linearity and super-exponential growth

The constituents of a living cell are not capable of exponential growth at low concentrations; one of the many functions of the cell membrane is to keep the concentrations high enough for growth to occur. If the components of a living cell were to be placed in a dilute solution, it would be unlikely for an enzyme to collide with its substrate in order to bind to it, and therefore enzyme-catalysed reactions would occur at vanishingly low rates.

Exponential autocatalysis is interesting from an origins of life point of view (Virgo and Ikegami, 2013), precisely because it doesn’t require high concentrations, and therefore could exist before the evolution of the cell membrane or other forms of compartmentalisation. However, the linear analysis in the previous section makes it clear that the concentrations will always converge towards the *same* profile, regardless of the initial conditions (as long as they obey the low concentration assumption). This seems to leave no room for evolution: there is no way one exponential replicator can out-compete another.

Many models of prebiotic chemistry are nonlinear, because they involve reactions where two non-food molecules must collide with each other. These include all models based on explicit catalysis, such as the work of Kauffman (1986) or the hypercycle model of Eigen and Schuster (1979).

For these reasons it is worth exploring under which circumstances non-linear, super-exponential growth can emerge spontaneously. In this section we demonstrate circumstances in which this can happen in A-polymerisation chemistries. The key is to constrain the reaction network such that first-order, exponential autocatalytic cycles cannot occur. Under these circumstances, growth occurs if we assume reactions of the form  $a_i a_j$  (for  $i, j > 0$ ) cannot occur. However, when integrating the model numerically, we find autocatalytic cycles that rely on these reactions. Their kinetics are non-linear (in fact super-exponential), because they involve these quadratic terms. We show that this non-linearity can give rise to behaviours such as bistability and oscillations.

The exponential growth in the previous section occurs because a molecules can split into two shorter molecules, each of which can then grow through monomer addition until it reaches the length of the original molecule, giving a net reaction  $A_n + nA_1 \rightarrow 2A_n$ . Now we ask what happens if we prevent this simple first-order autocatalysis mechanism while retaining many of the possible reactions from the original network. We show that autocatalysis still self-organises,

but it does so through a more complex, second-order mechanism that results in super-exponential growth.

In order to prevent this, we constrain the reaction network in the following way: we define a set of “banned species”  $B$ . A reaction  $A_i + A_j \rightleftharpoons A_{i+j}$  is included in the reaction network *unless* either  $A_i \in B$ ,  $A_j \in B$ , or  $A_{i+j} \in B$ .

This technique of “banning” species is merely a convenient way to prevent the formation of first-order autocatalytic cycles while retaining much of the original network; it is not supposed to directly represent something one would expect to happen in polymerisation chemistry. However, in more complex chemistries a species may be stoichiometrically possible but have such a high free energy that it is never formed, and so networks with broadly this type of structure might not be entirely absent from real chemistry. We suspect that there are also many other reasons why a complex network would fail to contain first-order loops.

To see how banning species can prevent first-order autocatalysis, let us imagine that  $B = \{A_{2^{i+1}} : i \in \mathbb{Z}\}$ . That is, molecules whose lengths are powers of two are banned, apart from  $A_1$ . A molecule of  $A_n$  (with  $n$  not a power of two) may split into two smaller molecules,  $A_m$  and  $A_{n-m}$ , as long as neither  $m$  nor  $n-m$  is a power of two. However, at least one of these molecules is no longer able to grow by monomer addition until it reaches length  $n$ . For let  $r = 2^{\lfloor \log_2 n \rfloor}$  be the largest power of two less than  $n$ . Then either  $m < r$  or  $n-m < r$ , since  $r \geq n/2$ . A molecule of size less than  $r$  cannot grow to size  $n$  by monomer addition alone due to the lack of the reaction  $A_{r-1} + A_1 \rightarrow A_r$ . Therefore net reactions of the form  $A_n + nA_1 \rightarrow 2A_n$  are impossible in this chemistry.

By excluding even more species we can obtain quite complex behaviour. Here we present results from a system that excludes the species  $B = \{A_{3^i} : i \in \mathbb{Z}\}$ , and also excludes the reaction  $2A_1 \rightleftharpoons A_2$ . For computational reasons we also exclude species longer than 60 monomer units.

If, as in the previous section, we linearise this system around the point where all non-food species have zero concentration, we find that the leading eigenvalue is zero. In such cases the stability of the fixed point is determined by the nonlinear terms, according to neutral manifold theory. Since we know there is a global Lyapunov function (the Gibbs energy) with a minimum at a different point, we should expect the nonlinear terms to result in a second-order instability around this fixed point. Our numerical results in this section show that this is indeed the case.

Figure 2 shows the results of numerically integrating this system, using Equations 2 with no approximations, from an initial condition consisting of a high concentration of  $A_1$  and small amounts of everything else. The result is that at around  $t = 450$  there is a sudden transition to the equilibrium state, in which all species are present, exponentially distributed according to their lengths. These dynamics indicate that the species other than  $A_1$  grow (super-)exponentially at low

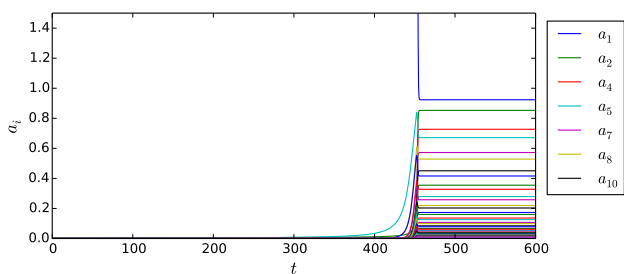


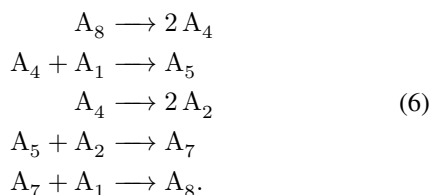
Figure 2: The results of numerically integrating an A-polymerisation chemistry over time. This chemistry excludes the species  $B = \{A_{3i} : i \in \mathbb{Z}\} \cup \{A_i : i > 60\}$ , and in addition excludes the reaction  $2A_1 \rightleftharpoons A_2$ . The parameters used are  $k_s = k_d = 1$ . The initial concentration of  $A_1$  is 100, and the initial concentrations of all other species are  $10^{-5}$ . Prior to the sudden transition at around  $t = 450$ , the species' concentrations grow super-exponentially through second and third-order autocatalysis.

Reaction	Rate
$A_4 + A_1 \longrightarrow A_5$	$3.194 \times 10^{-5}$
$A_7 + A_1 \longrightarrow A_8$	$1.888 \times 10^{-5}$
$A_8 \longrightarrow 2A_4$	$1.882 \times 10^{-5}$
$A_4 \longrightarrow 2A_2$	$1.366 \times 10^{-5}$
$A_5 + A_2 \longrightarrow A_7$	$1.084 \times 10^{-5}$
$2A_5 \longrightarrow A_{10}$	$8.050 \times 10^{-6}$
$A_{10} + A_1 \longrightarrow A_{11}$	$8.021 \times 10^{-6}$
$A_{11} \longrightarrow A_7 + A_4$	$7.942 \times 10^{-6}$
$A_{13} + A_1 \longrightarrow A_{14}$	$9.351 \times 10^{-8}$
$A_{10} \longrightarrow A_8 + A_2$	$8.598 \times 10^{-8}$

Table 1: The top ten reactions in the system shown in Figure 2, at time 100. The rates shown are absolute net rates, in moles per unit time.

rates until their concentrations become high enough for the growth's effects to be felt. This is similar to the results from first-order autocatalytic systems presented in (Virgo and Ikegami, 2013). However, the mechanism is rather more complicated.

At time  $t = 100$ , the growth in the system is largely due to the following set of reactions:



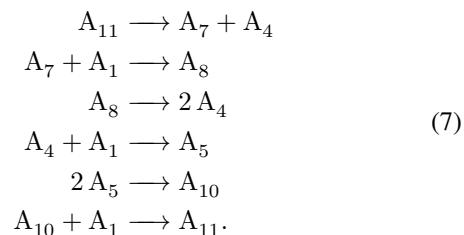
These reactions may be stoichiometrically balanced to form the net reaction  $3A_8 + 8A_1 \longrightarrow 4A_8$ . It can therefore be classified as third-order autocatalysis. (Though these are not the only reactions in the system, and they are not necessarily happening at stoichiometrically balanced rates, because the

Reaction	Rate
$A_4 + A_1 \longrightarrow A_5$	$3.745 \times 10^{-3}$
$A_7 + A_1 \longrightarrow A_8$	$1.549 \times 10^{-3}$
$A_8 \longrightarrow 2A_4$	$1.429 \times 10^{-3}$
$A_{10} + A_1 \longrightarrow A_{11}$	$1.202 \times 10^{-3}$
$2A_5 \longrightarrow A_{10}$	$1.150 \times 10^{-3}$
$A_{11} \longrightarrow A_7 + A_4$	$1.088 \times 10^{-3}$
$A_5 + A_2 \longrightarrow A_7$	$3.669 \times 10^{-4}$
$A_4 \longrightarrow 2A_2$	$2.466 \times 10^{-4}$
$A_{13} + A_1 \longrightarrow A_{14}$	$7.950 \times 10^{-5}$
$A_8 + A_5 \longrightarrow A_{13}$	$3.669 \times 10^{-5}$

Table 2: The top ten reactions in the system shown in Figure 2, at time 400. The overall difference in magnitudes compared to Table 1 is due to increases in concentrations over time.

concentrations are changing over time. See Table 1 for the true reaction rates.) The nonlinearity comes from the reaction  $A_5 + A_2 \longrightarrow A_7$ , which requires an  $A_5$  to collide with an  $A_2$  dimer in order to proceed; it will therefore increase in rate as the concentrations of these species increase.

By  $t = 400$  the system is dominated by the following overlapping but different set of autocatalytic reactions, as can be seen in Table 2:



This is a second-order autocatalytic cycle with the net reaction  $2A_{11} + 11A_1 \longrightarrow 3A_{11}$ . Where the cycle in Equations 6 produces  $A_7$  by splitting  $A_4$  into  $A_2$  and then joining it with  $A_5$ , this network instead produces  $A_7$  by forming and then splitting  $A_{11}$ . The transition from Equations 6 to 7 is probably the result of the general increase in concentrations, making the reaction  $2A_5 \longrightarrow A_{10}$  more likely to proceed. In this network,  $2A_5 \longrightarrow A_{10}$  is the nonlinear step.

We reiterate that this nonlinear autocatalytic increase in concentrations is a logical consequence of the reaction network's structure. The values of the numerical parameters are more or less arbitrary; the only tuning required is to set the initial concentrations small enough for the effect to be observed.

This nonlinearity can lead to bistability. Suppose that additional reactions are added such that the polymer species decay exponentially into inert products. Then at low concentrations the decay reactions will be faster than the autocatalytic cycles and the polymer species will not be able to persist in the system. However, at higher concentrations the autocatalytic species will grow at a faster rate than their decay. Thus the polymer species can persist in the system

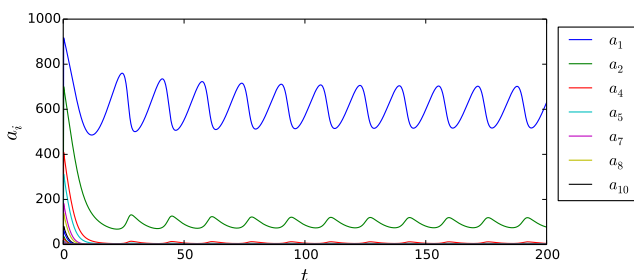


Figure 3: Convergence to a limit cycle due to a higher-order autocatalytic cycle in a flow reactor. The reaction network is the same as for Figure 2. The flow reactor is modelled with the flux terms  $\phi_1 = 0.4(10000 - a_1)$  and  $\phi_i = -0.01a_i$  for  $i > 1$ . The parameters used are  $k_s = 1.0$  and  $k_d = 1200$ . The autocatalysis is primarily due to the reactions shown in Equation 6.

only if they are initially present in concentrations above a certain threshold. (The technical details of this result are not especially enlightening, so we omit them here.) This is in contrast to the case of exponential growth, where the zero-concentrations state is an unstable fixed point, meaning that even a single molecule can eventually grow to dominate the system.

The existence of a concentration threshold leads to a kind of “precariousness” that is an important property of life-like systems, since it is analogous to biological death (Virgo, 2011; Barandiaran and Egbert, 2013). Precariousness means there are some perturbations from which a system cannot recover; in this case some perturbations push the concentrations below the threshold, and the concentrations must then eventually decay to nothing.

This same kind of nonlinearity can also lead to a more easily visualised phenomenon, namely temporal oscillations, as shown in Figure 3. To obtain this result, the reaction is modelled as if in a flow reactor. There is a constant input of  $A_1$  and exponential decay of every other species, which can be thought of as the species leaving the system in the reactor’s outflow. The oscillations occur through a similar mechanism to the famous lynx-hare cycle in ecology: the autocatalytic cycle consumes too much of its food and as a result its constituent species decay somewhat. This gives the food concentration a chance to recover (since it is being continually fed in to the system), and the autocatalytic species once again increase in concentration, completing the cycle.

This oscillatory behaviour does not occur for every choice of parameter values. However, it demonstrates nicely that the nonlinearity caused by higher-order autocatalysis can lead to phenomena that are not possible in the linear, first-order case. We reiterate that all of the other results in this paper are very robust and require no parameter tuning at all.

## Analysis and Conclusions

The second law of thermodynamics implies that physical systems must always “flow downhill”, where “downhill” means in the direction of increasing entropy, or decreasing free energy. We have implemented this feature in a very simple class of chemical reaction networks and found that it can lead to self-organisation of complex autocatalytic cycles.

This is in some way analogous to the self-organisation of dissipative structures in physical systems, such as convection cells. These also involve the formation of cycles, and also increase the rate at which energy is dissipated.

Our informal, intuitive understanding of these results is as follows: under some (but by no means all) circumstances, physical systems far from equilibrium not only “flow downhill”, but also tend to seek out faster and faster ways to do this. The formation of convection cells is again an example of this; the breaking of a dam is perhaps a more vivid one. The “full” A-polymerisation reaction network (with all reactions included) will flow to its equilibrium rather rapidly. However, if we block the most direct route (the  $2A_1 \rightleftharpoons A_2$  reaction) then an indirect route must be found, by taking advantage of catalytic cycles. Blocking the zeroth-order route leads to the self-organisation of first-order autocatalytic cycles, because (in a very informal sense) these lead to faster rates of dissipation.

Blocking these first-order solutions then leads to higher-order solutions. It is as if the system tries to find the easiest route towards its equilibrium, and only once the simplest paths are blocked does it find the more complex ones. Having tried to find autocatalytic networks in these chemistries using pen and paper, the solutions found by the simulation seem rather elegant and novel.

From a more formal point of view, these results occur because the initial conditions are near a fixed point that is not the thermodynamic equilibrium point, and therefore cannot be stable. By preventing first-order instability, we force this fixed point to exhibit a second-order instability.

The nonlinearity of the resulting systems leads to phenomena such as precariousness and oscillations. A nonlinear system with many equilibria is more interesting from an origins of life point of view, because it is closer to a system in which multiple different autocatalytic systems can compete with one another, leading to evolution.

It is interesting to note that we found complex behaviour in this system by removing reactions rather than by adding them. If we consider the full A-polymerisation chemistry, including both forward and reverse reactions, all of these autocatalytic cycles — both first and higher-order — are already present as its subnetworks. Whether the more complex networks become manifested or not is a matter of whether the right constraints exist to prevent simpler networks from being manifested instead.

In this respect we are reminded of Terrence Deacon’s (2011) comments about constraints in relation to the origins

of life. Our systems do not yet go as far as Deacon's ideas, since the constraints are imposed from outside rather than being generated by the system itself, but nevertheless it is an interesting demonstration of the idea that constraints can lead to complexity.

Many authors in the origins of life have worried about the problem of "side-reactions" that can prevent autocatalysis from occurring (e.g. Szathmary, 2000; King, 1982; Virgo et al., 2012). Our results suggest that under the right energetic conditions, such issues might not arise. Our reaction networks contain many possible reactions besides the ones that eventually form the autocatalytic cycles, yet these do not seem to disrupt the autocatalysis to any meaningful extent. The reason is that any molecules produced by side reactions eventually end up participating in autocatalytic cycles of their own.

A closely related issue in the origins of life is the combinatorial explosion that can arise from several small molecular components combining in different ways. Our model is incomplete in regards to this issue — we have only one type of monomer, and they combine only linearly — but nevertheless our results suggest that the combinatorial explosion might not be as serious an issue as it might at first seem. One can imagine a class of molecules that can grow in multiple ways, leading to a combinatorial explosion, but which can also undergo reactions that decompose them into smaller molecules. With such a chemistry, at low concentrations one should still expect exponential growth through the mechanism that arises in the first of our models. We therefore suggest that the future of origins of life models might be not in "taming the combinatorial explosion" (Schuster, 2000) but in embracing it. If every molecule forms part of an autocatalytic cycle then the combinatorial explosion represents an efficient way to explore the space of possible replicators, and it may be that it was through such a process that the first high-fidelity replicators arose.

This research represents a step towards understanding what properties a chemical reaction network must have in order for complex, life-like dynamics to emerge. There is still a lot of work to do to understand the effects of network topology and energetics on the emergence of complex dynamics, but the results in this paper represent a substantial step in the right direction.

## References

- Andersen, J. L., Andersen, T., Flamm, C., Hanczyc, M., Merkle, D., and Stadler, P. F. (2013). Navigating the chemical space of hcn polymerization and hydrolysis: Guiding graph grammars by mass spectrometry data. *Entropy*, 15:4066–4083.
- Andersen, J. L., Flamm, C., Merkle, D., and Stadler, P. (2012). Maximizing output and recognizing autocatalysis in chemical reaction networks is np-complete. *Journal of Systems Chemistry*, 3(1):1–9.
- Barandiaran, X. and Egbert, M. (2013). Norm-establishing

and norm-following in autonomous agency. *Artificial Life*, 20(1):5–28.

- Deacon, T. (2011). *Incomplete Nature: How Mind Emerged from Matter*. W. W. Norton & Company.
- Eigen, M. and Schuster, P. (1979). *The Hypercycle: A principle of natural self-organisation*. Springer.
- Fontana, W. and Buss, L. W. (1994). What would be conserved if "the tape were played twice"? *Proceedings of the National Academy of Sciences*, 91:757–761.
- Kauffman, S. (1986). Autocatalytic sets of proteins. *Journal of Theoretical Biology*, 119:1–24.
- King, G. A. M. (1978). Autocatalysis. *Chemical Society Reviews*, 7(2):297–316.
- King, G. A. M. (1982). Recycling, reproduction and life's origins. *BioSystems*, 15:89–97.
- Kondepudi, D. and Prigogine, I. (1998). *Modern Thermodynamics: from heat engines to dissipative structures*. Wiley.
- Schuster, P. (2000). Taming combinatorial explosion. *PNAS*, 97(14):7678–7680.
- Szathmary, E. (2000). The evolution of replicators. *Philosophical Transactions of the Royal Society, Series B.*, 355:1669–1676.
- Vasas, V., Fernando, C., Santos, M., Kauffman, S., and Szathmary, E. (2012). Evolution before genes. *Biology Direct*, 7(1).
- Virgo, N. (2011). *Thermodynamics and the Structure of Living Systems*. PhD thesis, University of Sussex, UK.
- Virgo, N., Fernando, C., Bigge, B., and Husbands, P. (2012). Evolvable physical self-replicators. *Artificial Life*, 18:129–142.
- Virgo, N. and Ikegami, T. (2013). Autocatalysis before enzymes: The emergence of prebiotic chain reactions. In *Advances in Artificial Life, ECAL*. MIT Press.