Synchronization phenomena in protocell models

Filisetti, Alessandro; Serra, Roberto; Carletti, Timoteo; Poli, Irene; Villani, Marco

Published in:
Biophysical Reviews and Letters

Publication date:
2008

Document Version
Early version, also known as pre-print

SYNCHRONIZATION PHENOMENA IN PROTOCELL MODELS

ALESSANDRO FILISETTI *
Dipartimento di Scienze Sociali, Cognitive e Quantitative
Università di Modena e Reggio Emilia
via Allegri 9, 42100 Reggio Emilia, Italy
E-mail: alessandro.filisetti@unimore.it

ROBERTO SERRA
Dipartimento di Scienze Sociali, Cognitive e Quantitative
Università di Modena e Reggio Emilia
via Allegri 9, 42100 Reggio Emilia, Italy
E-mail: rserra@unimore.it

TIMOTEO CARLETTI
Département de Mathématique, Université Notre Dame de la Paix
8 rempart de la Vierge, B5000 Namur, Belgium
E-mail: timoteo.carletti@undp.ac.be

IRENE POLI
Dipartimento di statistica, Università Ca' Foscari
San Giobbe - Cannareggio 873, 30121 Venezia, Italy
E-mail: poli@unive.it

MARCO VILLANI
Dipartimento di Scienze Sociali, Cognitive e Quantitative
Università di Modena e Reggio Emilia
via Allegri 9, 42100 Reggio Emilia, Italy
E-mail: marco.villani@unimore.it

keywords: protocell, self-replication, evolution , synchronization

*This author will present the full paper if accepted.

Introduction

Almost all life forms known today, are composed by cells, fundamental constituting units able to *self–replicate* and *evolve* through changes in genetic information; it is generally believed that this was not the case when first life–forms emerged on Earth almost 4 billion years ago. These *protocells* were much simpler, probably exhibiting only few simplified functionalities, that required a primitive embodiment structure, a protometabolism and a rudimentary genetics, so to guarantee that offsprings were similar to their parents [1,2,3].

Artificial protocells have not yet been reproduced and intense research programs are being established aiming at developing reference models [4,5] to capture the essence of the first protocells appeared on earth and enable to monitor their subsequent evolution. The interest for these researches is motivated either by the quest to understand which are the minimal requirements for a life form to exist and evolve, or by the search for indications about the way in which primitive life might have developed on earth. Moreover besides from their interest for the origin–of–life problem, protocells may be of practical interest in applications [2]: obtain populations of protocells that grow and reproduce, specialized for useful tasks, like drug synthesis and reduce pollution.

Because protocells didn't yet exist, in order to study how they can develop researchers have considered simplified models able to capture general behaviors, without carefully adding complicating details [6]. It is widely accepted that a protocell should comprises at least one kind of "container"(typically an amphiphile) and one kind of replicator molecule, typically a linear polymer able to self–replicate or a system of two or more kinds of polymers which catalyze each others synthesis. One can thus emphasize the existence of two kinds of *key reactions* which are crucial for the good functioning of the protocell: those which synthesize the container molecules and those which synthesize the replicators.

Let us observe that these key reactions may take place at different rates, however, to achieve sustained protocell growth and avoiding death by dilution [7], it is necessary that the two are proceed at equal rate, i.e. that the genetic material has to double when the protocell splits into two offsprings, a condition referred to as *synchronization*. Of course the requirement of duplication of the genetic material at duplication time refers to the average behavior, while each single event is affected by noise and fluctuations.

Let us finally observe that synchronization leads to *exponential growth*

of the population of protocells (a straightforward consequence of constant doubling time) and therefore to strictly *Darwinian selection* among protocells, even if the kinetic equations for the replicators have sub linear growth terms [8]. The introduction of a mesoscopic structure (the container) and emergent synchronization change the features of the process. Selection is moved from replicators level to protocells level.

In this paper we consider protocell models where the key reactions take place on the surface [a] of the protocell membrane, that's way they are called *Surface Reaction Models* [9,11]. Our models are inspired by the so–called "Los Alamos bug" [10,5], however due to their abstraction they can cover a larger class of protocell models. In the Los Alamos bug replicators are PNAs which can be found in the membrane, either in its interior or on its surface, but replication takes place when a single–stranded PNA on the surface ligates (via Watson–Crick pairing) the corresponding nucleotides which are supposed to be freely available. The PNA's on the membrane surface also catalyze the formation, from lipid precursors, of amphiphiles which are then incorporated in the membrane.

We have been able to prove under general assumptions [9,11] that synchronization is an emergent property of our models, in contrast to earlier models, like the well–known Chemoton [13] where synchronization was achieved by ad hoc hypotheses concerning the form of kinetic equations. In these papers we considered models with one or two non–interacting replicators in each protocell, here we extend our analysis by considering an arbitrary, but finite, number of replicators inside each protocell, and moreover we allow the replicators to interact between them: by catalyzing or inhibiting the replication of others.

The analytical tools we set up [9] to study such models, are improved to cover these more general cases to answer to the synchronization question; moreover to complete this analysis, dedicated numerical simulations have been performed to deal with cases where the theoretical tools weren't able to provide an answer [14,12].

Our numerical simulations show that synchronization is obtained in a class of models more general than the ones studied previously, but now it is a *fragile phenomenom*: by introducing small changes one can destroy it. Cell duplication time and replicators quantities show an oscillatory behavior in relations to the initial values of parameters. In some cases we observed

[a] Of course this is not the only possibility, and models where the key reactions take place in the interior of the protocell have been proposed [4,7].

a time transient where synchronization seems to be achieved, but then an oscillatory behavior enters.

1. Surface reaction models of protocells

Let us start by briefly recalling the main basic model with some related results, and referring the interested reader to [9] for a more detailed discussion and analysis. In the second part we then generalize this model to the case of N self–replicating molecules interacting each other.

In the case of a single replicator in the protocell lipid phase [b], let its quantity (mass) be denoted by X and let also C be the total quantity of "container"(e.g. lipid membrane in vesicles or bulk of the micelle). Let V be its volume, which is equal to C/ρ (where ρ is the density, which will be assumed constant) and with S we will denote the surface area, which is a function of V (S is approximately proportional to V for a large vesicle with a very thin surface, a condition which will be referred to as the "thin vesicle case", and to $V^{2/3}$ for a micelle).

We assume, according to the Los Alamos bug hypothesis, that X favors the formation of amphiphiles, and that only the fraction which is near the external surface is effective in doing so. That is because precursors are found outside the protocell. We also assume that the replication of X takes place near the external surface, too.

Let us further assume that

- spontaneous amphiphile formation is negligible, so that only the catalyzed term matters;
- the precursors (both of amphiphiles and templates) are buffered;
- S is proportional to V^β, and therefore also to C^β (β ranging between 2/3 for a micelle and 1 for a very thin vesicle);
- diffusion is very fast within the protocell, so concentrations can be assumed to be constant everywhere in the lipid phase;
- the protocell breaks into two identical daughter units when its container size reaches a certain threshold;
- the rate limiting step which may appear in the replicator kinetic equations does not play a significant role when the protocell is smaller than the division threshold.

[b] This model is invariant with respect to the way in which either C or X are measured; for example, if they were measured as number of molecules the equations would retain exactly the same form (of course, the units of the kinetic constants would be different).

Under these hypotheses, as discussed in detail in [9,11], one obtains the following approximate equation which describes the growth of a protocell between two successive divisions:

$$\frac{dX}{dt} = \eta C^{\beta-1} X \quad and \quad \frac{dC}{dt} = \alpha C^{\beta-1} X \,, \tag{1}$$

moreover we assume that once C reaches a critical value, here named θ, the protocell breaks into two equal offsprigs (halving hypothesis), hence at the beginning of each duplicationcycle the initial amount of X equals one half of the value attained at the end of the previous cycle. In between two successive divisions the system is again ruled by Eq. 1.

The generalization to the case where there are more replicators is straightforward. Let:

$$\vec{X} = (X_1, X_2, \ldots, X_N) \,, \tag{2}$$

denote the total quantity (mass) of N different types of replicating molecules in the protocell lipid phase. Obviously, all the X_i's must be real and non negative. The N–dimensional generalization of Eq. 1 is then

$$\frac{d\vec{X}}{dt} = C^{\beta-1} M \vec{X} \quad and \quad \frac{dC}{dt} = C^{\beta-1} \vec{\alpha} \cdot \vec{X} \,, \tag{3}$$

where $\alpha = (\alpha_1, \ldots, \alpha_N)$ and the (constant, real) matrix element M_{ij} represents the contribution of X_j to the growth of X_i. Without loss of generality we will consider the case where $\det M \neq 0$, if this were not the case, some of the differential equations for the X_i's would be redundant (i.e. their values at time t could be expressed as a function of the values of the other variables at t) and they could therefore be removed from the set of differential equations under consideration.

An important simplification can now be considered: as it was demonstrated in [9], in order to determine whether there is synchronization in the asymptotic time limit, one can limit himself to consider the $\beta = 1$ case (the final result does not depend on β, while of course this parameter affects the speed with which it is approached). With this simplification, the basic equations (which are valid between two successive divisions) are then

$$\frac{d\vec{X}}{dt} = M \vec{X} \quad and \quad \frac{dC}{dt} = \vec{\alpha} \cdot X \,. \tag{4}$$

As outlined above, we assume that division takes place when the mass (or equivalently the volume, since density is assumed constant) of the protocell reaches a certain critical size. Without loss of generality we may then assume that the initial size is one half of the final value (indeed, if the size

of the very first protocell were different then it would suffice to consider the evolution from the following generation).

So, starting with an initial quantity of container C at time T_0 equal to $\theta/2$, we assume that once C reaches the critical value θ it will divide into two equal protocells of size $\theta/2$. Let ΔT_0 be the time interval needed to double C from this initial condition, and let $T_1 = T_0 + \Delta T_0$ be the time when the critical mass θ is reached. Since the initial value for C is fixed, ΔT_0 is a function of the initial quantity of replicators, $\vec{X}_0$. The final value of $\vec{X}$, just before the division is denoted by $\vec{X}(T_1)$. Because we assume perfect halving at the division, each offspring will start with an initial concentration of replicators equal to $\vec{X}_1 = \vec{X}(T_1)/2$. The successive doubling time will be denoted by $T_2 = T_1 + \Delta T_1$, and the third generation will start with an initial value $\vec{X}_2 = \vec{X}(T_2)/2$, a.s.o.

We generalize the preceding discussion with the following equations, which refer to the k–th cell division cycle that starts at time T_k and ends at time T_{k+1}:

$$\frac{\theta}{2} = \int_{T_k}^{T_{k+1}} \frac{dC}{dt}(t)\, dt \quad and \quad \vec{X}_{k+1} = \frac{1}{2}\vec{X}(T_{k+1})\,. \tag{5}$$

Note that in general $\vec{X}(T_{k+1}) \neq 2\vec{X}(T_k)$ and $\Delta T_{k+1} \neq \Delta T_k$, however we will prove in the next section that these conditions can be asymptotically approached.

2. Synchronization in linear surface–reaction models

We will now consider under which conditions the system described in the previous section displays synchronization, in the sense that $\lim_{k\to\infty} \vec{X}(T_k) = \vec{X}_\infty$, for some finite positive value $\vec{X}_\infty$, so that, after several cell divisions, the initial quantity of all inner chemicals between successive duplications approaches a constant value. This requires that

$$\lim_{k\to\infty} \left(\vec{X}(T_{k+1}) - \vec{X}(T_k)\right) = 0\,. \tag{6}$$

As observed above, this implies that, as k grows, also the division time approaches a constant value, so that

$$\lim_{k\to\infty} \Delta T_k = \Delta T_\infty\,. \tag{7}$$

Let us therefore consider the behavior of the system in the continuous growth phase between two successive generation, ruled by Eq. 4. From

the linearity of this equation one immediately infers that, during the first replication (i.e. when $0 \leq t \leq T_0$)

$$\vec{X}(t) = e^{M(t-T_0)}\vec{X}_0\,, \tag{8}$$

so that

$$\vec{X}(T_1) = e^{M\Delta T_0}\vec{X}_0 \quad \textit{and} \quad \vec{X}_1 = \frac{1}{2}e^{M\Delta T_0}\vec{X}_0\,. \tag{9}$$

The same reasoning applies to all generations, so

$$\vec{X}(T_{k+1}) = e^{M\Delta T_k}\vec{X}_k \quad \textit{and} \quad \vec{X}_{k+1} = \frac{1}{2}e^{M\Delta T_k}\vec{X}_k\,. \tag{10}$$

From these last equations one derives a necessary and sufficient condition to ensure synchronization

$$\vec{X}_\infty = \frac{1}{2}e^{M\Delta T_\infty}\vec{X}_\infty\,. \tag{11}$$

Namely $\vec{X}_\infty$ must be an eigenvector of the matrix $e^{M\Delta T_\infty}$, belonging to the eigenvalue 2, i.e. it must be an eigenvector of M belonging to the eigenvalue $\log 2/\Delta T_\infty$:

$$M\vec{X}_\infty = \lambda\vec{X}_\infty \quad \textit{and} \quad \lambda = \frac{\log 2}{\Delta T_\infty}\,. \tag{12}$$

Remember that the X_i's are the quantities of the different replicators, therefore they must be real and non negative, so in order for synchronization to take place in a linear system the (real) matrix M must have a real positive eigenvalue λ with such a real, nonnegative eigenvector. The conditions under which these requirements are satisfied are discussed in the next section 3, where we also discuss which eigenvalue has to be chosen among those of the matrix M. In the rest of this section we will assume that λ is a simple positive eigenvalue of the coefficient matrix M associated with a positive eigenvector. Observe that since eigenvectors are determined up to a multiplicative constant, Eqs. 12 do not suffice to determine a unique solution, and we will now provide the formula which determines the actual values of the X_i's.

Since its determinant is not null, the matrix M is invertible, so from Eqs. 4 we get:

$$\frac{dC}{dT} = \vec{\alpha}\cdot M^{-1}\frac{d\vec{X}}{dt}\,, \tag{13}$$

hence the quantity $Q(t) = C(t) - \vec{\alpha}\cdot M^{-1}\vec{X}(t)$, is a first integral, i.e. a quantity constant during each division cycle (the proof is straightforward,

$dQ/dt = 0$ derives from Eq. 13). Evaluating $Q(t)$ at the beginning and the end of the k–th division we obtain

$$C(T_k) - \vec{\alpha} \cdot M^{-1} \vec{X}(T_k) = C(T_{k+1}) - \vec{\alpha} \cdot M^{-1} \vec{X}(T_{k+1}) \,, \tag{14}$$

recalling that C takes an initial value equal to $\theta/2$ and a final value equal to θ and using the definition of $\vec{X}_k$ (see Eq. 5) we finally get:

$$\frac{\theta}{2} = \vec{\alpha} \cdot M^{-1} \left(\vec{X}_{k+1} - \vec{X}_k \right) \,, \tag{15}$$

in the limit of large k, calling $\vec{X}_k \to \vec{X}_\infty$, we get:

$$\frac{\theta}{2} = \vec{\alpha} \cdot M^{-1} \vec{X}_\infty \,. \tag{16}$$

By multiplying the first relation of Eqs. 12 by M^{-1} and then taking the scalar product with α, from Eq. 16 we get:

$$\Delta T_\infty = \frac{\theta \log 2}{2\vec{\alpha} \cdot \vec{X}_\infty} \,. \tag{17}$$

which is the required relationship. The general approach is now clear: from the matrix of the coefficients M one computes the eigenvalues, λ, which in turn determine the asymptotic interval between two successive divisions ΔT_∞ (Eq. 12b). The components of the eigenvector $\vec{X}_\infty$ are determined except for a constant, which can be determined from Eq. 16.

3. Eigenvalues and eigenvectors

Since the matrix M may have different eigenvalues, it is necessary to find which one should be used in Eq. 12. From Eqs. 9 and 10 one obtains

$$\vec{X}(T_2) = e^{M\Delta T_1} \vec{X}_1 = e^{M\Delta T_1} \frac{\vec{X}(T_1)}{2} = e^{M\Delta T_1} e^{M\Delta T_0} \frac{\vec{X}_0}{2} = e^{M(T_2 - T_0)} \frac{\vec{X}_0}{2} \,, \tag{18}$$

which can be iterated to yield

$$\vec{X}(T_k) = e^{M(T_k - T_0)} \frac{\vec{X}_0}{2^{k-1}} \,. \tag{19}$$

Note that, although $2^k \to \infty$, the r.h.s does not vanish as $k \to \infty$ since, at every generation, the numerator is multiplied by a new term. Recall that T_k measures the total time elapsed from the origin of time to the end of the k–th generation. Indeed as $k \to \infty$, $\vec{X}(T_k)$ tends to $\vec{X}_\infty$ and at each generation the r.h.s of Eq. 19 is multiplied by $e^{M\Delta T_\infty}/2$.

We will now suppose that M is diagonalizable, i.e. it has N independent eigenvectors. In this case there exists a nonsingular matrix A such that $A^{-1}MA = \Lambda$, where Λ is a diagonal matrix whose diagonal elements are the eigenvalues of M. The columns of A are the corresponding eigenvectors. Recalling that $M = A\Lambda A^{-1}$, and that $e^{AMA^{-1}} = Ae^{M}A^{-1}$, from Eq. 19 one gets:

$$\vec{X}(T_k) = Ae^{\Lambda(T_k-T_0)}A^{-1}\frac{\vec{X}_0}{2^{k-1}} \quad and \quad A^{-1}\vec{X}(T_k) = \frac{1}{2^{k-1}}e^{\Lambda(T_k-T_0)}A^{-1}\vec{X}_0\,. \tag{20}$$

By introducing a new variable $\vec{Y}(T_k) = A^{-1}\vec{X}(T_k)$, one obtains:

$$\vec{Y}(T_k) = e^{\Lambda(T_k-T_0)}A^{-1}\frac{\vec{Y}_0}{2^{k-1}} \quad and \quad Y_i(T_k) = \frac{1}{2^{k-1}}e^{\lambda_i(T_k-T_0)}Y_{0i}\,, \tag{21}$$

where Y_i denotes the i–th component of the vector $\vec{Y}$.

If, for every i, $\Re\lambda_i < 0$, then $\vec{Y}$ asymptotically tends to 0 and so does $\vec{X} = A\vec{Y}$ (recall that $\det A \neq 0$) . The same holds, due to the growing denominators in Eq. 21, if $\Re\lambda_i = 0$ for every i. In all these cases the quantities of replicators asymptotically vanish.

Let us then consider the case where, for some i, $\Re\lambda_i > 0$. Let us also suppose that there is a *single* eigenvalue with *largest real part*, without loss of generality we can suppose that this eigenvalue is the first one, λ_1. We will also suppose that λ_1 is a *simple eigenvalue* and we will denote its eigenvector as $\vec{v}_1$. So $\Re\lambda_1 > \Re\lambda_j$ for every $j \neq 1$. As k increases, T_k goes to infinity so does the ratio Y_1/Y_j for all $j > 1$, see Eq. 21b, hence Y_j becomes negligible with respect to Y_1. Therefore $\vec{Y}_\infty$ is proportional, up to a multiplicative constant, to $(1, 0, \ldots, 0)$. But since A diagonalizes M, its columns are the eigenvectors of M, hence $\vec{Y}_\infty$ is proportional to the first column of A, i.e to the eigenvector $\vec{v}_1$.

By definition we get $\vec{X}_\infty = A\vec{Y}_\infty$ and we come therefore to the conclusion that the long term behavior of the system is ruled by the eigenvalue with the largest real part, and by the corresponding eigenvector. Let us call for brevity *ELRP* the eigenvalue with the largest real part.

As we have seen, if the real part of the ELRP is null or negative, the system dies out, and the asymptotic quantities X_i's vanish in successive generations. We may have sustained growth and synchronization only if the real part of the ELRP is positive.

Let us now analyze the physical conditions ensuring that the matrix M has a single eigenvalue with largest real part and a corresponding positive eigenvector. Let us first discuss the important case where all the matrix

elements are non negative, i.e $M_{ij} \geq 0$, for all $i, j = 1 \ldots, N$. This implies that there is *no negative interference* between different replicators i and j, the only possible alternatives being that either i favors (e.g. catalyzes) the formation of j or that it does not influence it in any way. Moreover, we must also require that at least one of the entries M_{ij} does not vanish, since otherwise there would be no replication at all.

We can therefore apply the Perron–Frobenius theorem [16,15], which states that if the matrix M is non-negative and non–null then the eigenvalue with the largest module is real, positive and unique, and that there is a non–negative eigenvector belonging to that eigenvalue. Since it is real, the eigenvalue with the largest module is also the ELRP, which thus rules the long time behavior of the system.

In Fig. 1 a simulation of a system of this type is reported: note that the cell division time converges, in successive generations, to the value given by Eq. 12. Moreover, one also observes that the quantity of genetic material at the beginning of the protocell growth cycle tends to a constant value as generations follow generations. Similar results are obtained with non–negative matrices of arbitrary size.

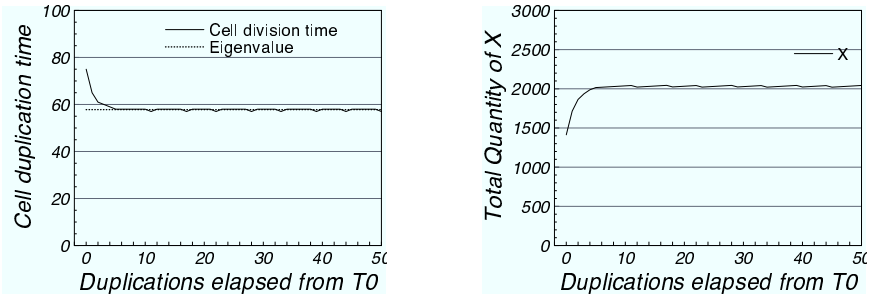

Figure 1. Data refer to simulations of the system described by Eq. 3 with matrix $M_{11} = M_{22} = 0$, $M_{12} = M_{21} = K$, whose eigenvalues are $\pm K$. On both panel, we report on the x axis the number of generations elapsed from T_0, while on the y axis, on the left panel, cell division time si shown (the level predicted by the theory is also shown), on the right one, the initial amount of X_k at generation k is reported.

Let us now consider the case where some entries of the real matrix M can be negative, while it still possesses N independent eigenvectors. In this case M can still be diagonalized and therefore the eigenvalue(s) with the largest real part determine the long term behavior of the system, recall that the results following Eq. 21 have been obtained by supposing only that M is diagonalizable.

If it happens that the eigenvalue of M with the largest real part is real and positive, and that its eigenvector is non negative, then the behavior of the system is exactly the same as described above (as confirmed by numerical simulations). However now in general: i) the eigenvector of the ELRP may have negative components and ii) the ELRP may be complex, so the previous equations loose their physical meaning.

Let us first consider the case of a real eigenvalue whose eigenvector has positive and negative components [c]. A possible ansatz could be to try to extend the theory to deal with these cases by assuming that, whenever one of X_i's becomes negative, it has to be interpreted as being actually equal to zero (the non–physical negative value indicating some limitation of the model used). The rationale is that if X_i, starting from a positive value, "becomes negative", it must have passed through the value 0: in this case there is no more replicator in the system, and it is justified to set its value equal to 0. The value of X_i may become positive again at a later time if it is produced by reactions involving other replicators .

Since the analytical theory is not applicable we resort to simulations which show that in this case it often happens that some components get permanently extinguished. If we drop from the matrix M those components which the simulation shows go to extinction, we obtain a reduced matrix M'. If its ELRP is positive and its eigenvector non–negative then the previous analytical theory applies and correctly predicts asymptotic duplication time and quantities of replicators, see Fig. 2.

It may however happen that this latter condition is not satisfied. While several simulations show synchronization, we have indeed also found some different behaviors, where the duplication time does not reach a constant value but seems to oscillate periodically in time, see Fig. 3.

Let us now consider an example where the long time behavior is ruled by a complex conjugate pair of eigenvalues. A simple 2×2 example is given

[c]Note that, since the components of eigenvectors are determined up to a multiplicative constant, if $\vec{v}$ is an eigenvector so it is also $-\vec{v}$: there is no absolute sign attached to the components, saying that the eigenvector is non–negative means that all its components have the same sign. Therefore the case we are considering now is indeed that of components of both signs. Nonetheless, for brevity, we will sometimes refer to it as the case with "negative components".

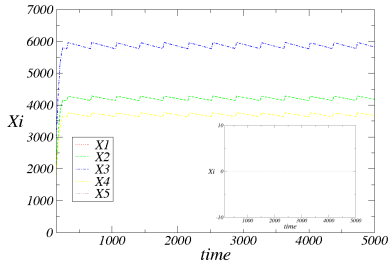

Figure 2. An example of a 5×5 matrix M with negative entries: the replicators which survive are those which might have been predicted by inspection of $\vec{v}_1$.

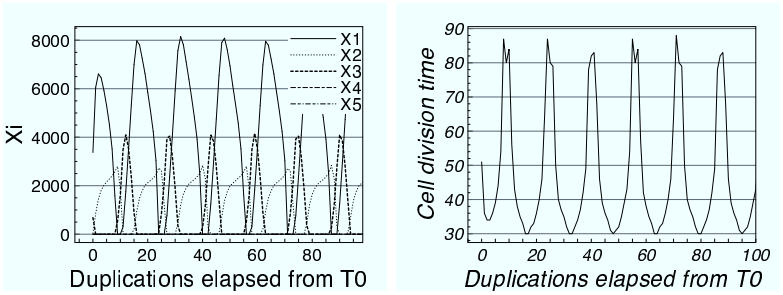

Figure 3. An example of a 5×5 matrix M with negative entries where simple synchronization is not achieved: the graphs show (left panel) the time behavior of the values of the components of $\vec{X}_\infty$ and (right panel) the duplication time in function of the generation number.

by the following

$$\begin{cases} \frac{dX}{dt} &= aX - qY \\ \frac{dY}{dt} &= qX + aY \\ \frac{dC}{dt} &= \alpha X + \alpha' Y \, . \end{cases} \tag{22}$$

The eigenvalues and eigenvectors are

$$\lambda_1 = a + iq \Rightarrow (1, -i) \quad \bar{\lambda}_1 = a - iq \Rightarrow (1, i) \, , \tag{23}$$

and the continuous time solution (between two successive divisions) is

$$(X(t), Y(t)) = 2e^{at} \left(d \cos qt - b \sin qt, b \cos qt + d \sin qt\right) \, , \tag{24}$$

where d and b are real coefficients determined by the initial conditions

$$(X(0), Y(0)) = 2 \, (d, b) \, . \tag{25}$$

The system described by Eq. 24 oscillates in time, and it is impossible to guarantee that both X and Y remain positive. It is possible to simulate the behavior of this system and to prove that Y survives while X gets extinct, even if the quantity of X is greater at the beginning (see Fig. 4).

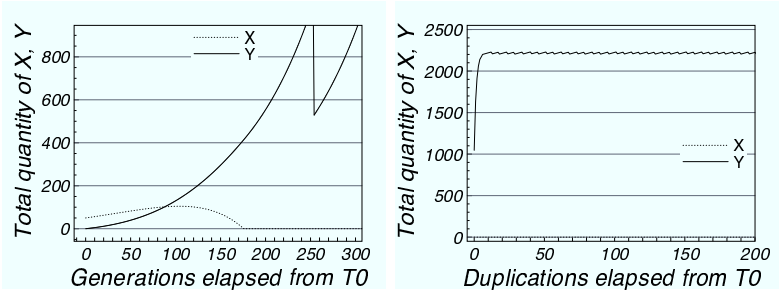

Figure 4. The behavior of the system described by Eq. 22: left panel, variation in time of the quantities of the two replicators in the first steps of the simulation; right panel, the value of the quantity of X at duplication time approaches a constant value.

Briefly, one can conclude that the analytical method precisely describes the system behavior when the ELRP is real and positive and its eigenvector non–negative. In different cases one has to resort to simulations, however the analytical theory may still help in understanding the system's behavior (like in the case where the reduced matrix M' has the properties required to apply it).

The above analysis could be extended also to non–diagonalizable matrices, which can be reduced to Jordan normal form. The idea being to compute the exponential of the matrix M as we did in Eq. 21 by using some standard linear algebra computation, and to observe that the first component of $\vec{Y}_k$, associated to the single eigenvalue with largest real part grows faster that all the other, and thus in the long run it prevails over the remaining components.

Let us now consider in detail the case where M is not diagonalizable. We will consider here the case where there is a single Jordan block and we left to the Appendix 4 the analysis of the more general situation involving several Jordan blocks. In the present case there is one real eigenvalue with algebraic multiplicity n and then there exists a non-singular matrix A such

that:

$$AMA^{-1} = M_j = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \ddots & \ddots & 1 \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & \lambda \end{pmatrix} \tag{26}$$

where M_j is the standard Jordan form. For all division event k, let us introduce the auxiliary variables, Y_k such that:

$$\vec{Y}_k = A\vec{X}_k\,, \tag{27}$$

hence (riferimento 3.5 di serra) can be rewritten as:

$$\vec{Y}_{k+1} = \frac{1}{2}Ae^{M\Delta T_k}A^{-1}\vec{Y}_k = \frac{1}{2}Ae^{AMA^{-1}\Delta T_k}\vec{Y}_k = \frac{1}{2}e^{M_j\Delta T_k}\vec{Y}_k\,. \tag{28}$$

It is a standard result of linear algebra the computation of the exponential of a $n \times n$ matrix in standard Jordan form:

$$e^{M_j\Delta T_k} = \begin{pmatrix} e^{\lambda\Delta T_k} & \Delta T e_k^{\lambda\Delta T_k} & \frac{(\Delta T_k)^2}{2}e^{\lambda\Delta T_k} & \dots & \frac{(\Delta T_k)^{n-1}}{(n-1)!}e^{\lambda\Delta T_k} \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \frac{(\Delta T_k)^2}{2}e^{\lambda\Delta T_k} \\ \vdots & \ddots & \ddots & \ddots & \Delta_k e^{\lambda\Delta T_k} \\ 0 & \dots & \dots & 0 & e^{\lambda\Delta T_k} \end{pmatrix}. \tag{29}$$

Let us observe also that for all k and m the matrices $M_j\Delta T_k$ and $M_j\Delta T_\infty$ do commute thus:

$$e^{M_j\Delta T_k}e^{M_j\Delta T_m} = e^{M_j(\Delta T_k+\Delta T_m)} = e^{M_j\Delta T_m}e^{M_j\Delta T_k}. \tag{30}$$

Eq. 28 can be iterated back in such a way we can express $\vec{Y}_{k+1}$ in terms of $\vec{Y}_0$, it follows thus from Eq. 30 that:

$$\vec{Y}_{k+1} = \frac{1}{2^{k+1}}e^{M_j\sum_{m=0}^{k}\Delta T_m}\vec{Y}_0\,, \tag{31}$$

This relation is the key point to conclude that also in this case, the long-term behaviour can be explicitly determined. In fact let us call for short $S_k = \sum_{m=0}^{k}\Delta T_m$, then from Eq. 28, Eq. 29 and Eq. 31 we get:

$$Y_{k+1}^{(i)} = \frac{1}{2^{k+1}}e^{\lambda S_k}\left(Y_0^i + S_k Y_0^{(i+1)} + \frac{(S_k)^2}{2}Y_0^{(i+2)} + ... + \frac{(S_k)^{n-1}}{(n-1)!}Y_0^{(n)}\right), \tag{32}$$

for all $i = 1, \ldots, n$. Let us now compute the following ratios for all $i = 2, \ldots, n$:

$$\frac{Y_{k+1}^{(i)}}{Y_{k+1}^{(1)}} = \frac{Y_0^{(i)} + S_k Y_0^{(i+1)} + \frac{(S_k)^2}{2} Y_0^{(i+2)} + \cdots + \frac{(S_k)^{n-1}}{(n-1)!} Y_0^{(n)}}{Y_0^{(i)} + S_k Y_0^{(2)} + \frac{(S_k)^2}{2} Y_0^{(3)} + \cdots + \frac{(S_k)^{n-1}}{(n-1)!} Y_0^{(n)}}$$
$$= \frac{1}{(S_k)^{i-1}} \frac{\frac{Y_0^{(i)}}{(S_k)^{n-i}} + \frac{Y_0^{(i+1)}}{(S_k)^{n-i-1}} + \frac{1}{2}\frac{Y_0^{(i+2)}}{(S_k)^{n-i-2}} + \cdots + \frac{1}{(n-i)} Y_0^{(n)}}{\frac{Y_0^{(1)}}{(S_k)^{n-1}} + \frac{Y_0^{(2)}}{(S_k)^{n-2}} + \frac{1}{2}\frac{Y_0^{(3)}}{(S_k)^{n-3}} + \cdots + \frac{1}{(n-1)} Y_0^{(n)}} \quad (33)$$

hence, observing that $S_k \to \infty$, being [d] $\Delta T_k \to \Delta T_\infty > 0$, we can conclude that for all $i = 2, ..., n$:

$$\frac{Y_{k+1}^{(i)}}{Y_{k+1}^{(1)}} \underset{k\to\infty}{\longrightarrow} 0\,, \quad (34)$$

and excluding the unboundedness of $Y_\infty^{(1)}$, we can conclude that:

$$Y_{k+1}^{(1)} \to Y_\infty^{(1)} \; and \; Y_{k+1}^{(i)} \to 0 \; \forall i = 2, ..., n\,, \quad (35)$$

from which we can drawn the same conclusions as in the case where M was diagonalizable: the long–term behaviour is driven by the eigenvalue with largest real part (trivially in this case, because we suppose to have only one eigenvalue with algebraic multiplicity n), while the asymptotic amounts of SRMs are described by the first eigenvector. This result can be straightforwardly generalised as to include the general case where the matrix M has p eigenvalues, λ_i, each one with algebraic multiplicity m_i, and moreover λ_1 is real and has the largest real part of all the remaining eigenvalues. The corresponding treatment is given in the Appendix 4.

4. Conclusions

Let us first comment on a simplification which has been used throughout this work, namely that of assuming that the surface is proportional to a power of the volume. This is certainly the case for a spherical micelle (with exponent 2/3), but in the case of a vesicle it holds (with exponent 1) only in the limit of a very large size.

It can be shown that the finite size effects can be taken into account without modifying our results: synchronization is still obtained. In fact

[d] Let us remark in fact that here we don't need to assume the existence of the limit $\Delta T_k \to \Delta T_\infty$, we only need ΔT_k to be definitely strictly positive.

assuming a generic relation between the volume, and thus the container size, and the surface, $S = f(C)$, for some positive increasing function f, then Eqs. 3 have to be modified into:

$$\frac{d\vec{X}}{dt} = \frac{f(C)}{C} M\vec{X} \quad and \quad \frac{dC}{dt} = \frac{f(C)}{C}\vec{\alpha} \cdot X \,. \tag{36}$$

But then we can observe that the function given by Eq. 13 is still a first integral and thus the same analysis follows. Another explanation of this result is that we can "rescale"the time [e] by the positive function $C/f(C)$ and thus identifying Eqs. 36 and Eq. 13. This result is supported by a dedicated numerical simulation of a linear system with a single self–replicating molecule X (remember that $C = \rho V$) in a "realistic"vesicle with a thick membrane, which is reported in Fig. 5. Thus we can conclude that synchronization is robust with respect to the finite size and the details of the geometry of the protocell.

In the present paper we addresse some relevant questions about the synchronization phenomenon for systems where the kinetic equations are linear, while of course non–linear terms may play a key role. While the analysis of non–linear kinetics lies beyond the scope of the present work, let us briefly mention that there are indeed some cases where unbounded growth of the replicator can be observed, as it may happen (depending upon the values of some parameters) when there are two replicators X and Y whose growth rate is proportional to XY.

We are also considering a model where the growth of each replicator is proportional to its quantity multiplied times a sigmoid function which depends upon the presence of other replicators, i.e. a system of the kind:

$$\begin{cases} \frac{dC}{dt} &= \vec{\alpha} \cdot \vec{X} \\ \frac{d\vec{X}}{dt} &= \vec{X} \cdot \vec{\sigma}(W\vec{X}) \,, \end{cases} \tag{37}$$

where $\sigma_i(W\vec{X}) = \tanh\left(\sum_k W_{ik} X_k\right)$ play the role of an activating function.

In several cases synchronization is achieved but, depending upon the values of the entries of the matrix W, a more intriguing phenomenon can sometimes be observed, where the system seems to approach synchronization, but at a certain point there is a sudden drop of one replicator, with a dramatic increase of the replications time. This is followed by a recovery,

[e] More precisely let us introduce a new non–linear time $\tau = \int^t C^{-1}(s) f(C(s))\, ds$ and let us denote the quantities C and $\vec{X}$ using this new time, respectively by $c(\tau)$ and $\vec{x}(\tau)$, then Eq. 36 is formally equivalent to Eq.3.

which may be followed by a further "crisis", a.s.o. Contrary to what has been observed in some systems with linear kinetics, the crises do not seem to be periodic in time. Further studies are necessary to give a comprehensive account of the behavior of these non–linear systems.

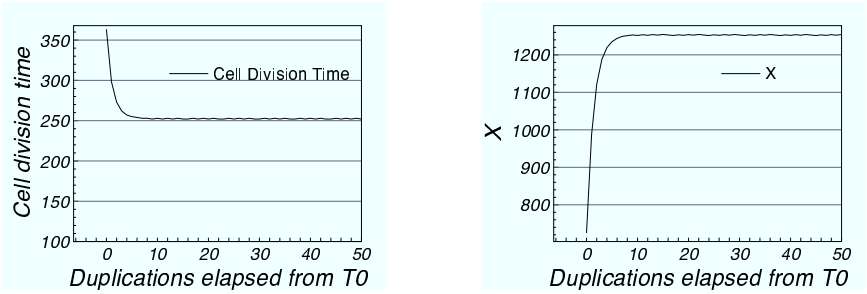

Figure 5. Synchronization in a thick vesicle (we suppose a spherical cell and that the volume of spherical shell increase with the same thickness). On the left the cell division time, while on the right the initial amount of X self replicating molecule in function of the time.

Appendix: Several Jordan blocks

In this last section we briefly show how the synchronization result of the previous sections can be extended as to include the case where the matrix M has p eigenvalues, each one with algebraic multiplicity m_i, and moreover λ_1 is real and has the largest real part of all the remaining eigenvalues.

In fact in this case Eq. 26 can be generalised by stating the existence of a non–singular matrix A such that:

$$AMA^{-1} = M_j = diag\left(J_{m_1}(\lambda_1), ..., J_{m_p}(\lambda_p)\right), \tag{38}$$

where $J_{m_i}(\lambda_i)$ is a standard Jordan matrix $m_i \times m_i$ with eigenvalue λ_i. Then we can introduce once again for all k auxiliary variables, $\vec{Y}_k$, $\vec{Y}_k = A\vec{X}_k$. Hence Eq. 10b can be rewritten as

$$\vec{Y}_{k+1} = \frac{1}{2} A e^{M\Delta T_k} A^{-1} \vec{Y}_k = \frac{1}{2} e^{AMA^{-}1\Delta T_k} \vec{Y}_k = \frac{1}{2} e^{M_j \Delta T_k} \vec{Y}_k \,. \tag{39}$$

The remarkable fact is that one can write a relation similar to Eq. 29 and Eq. 31:

$$\vec{Y}_{k+1} = \frac{1}{2^{k+1}} e^{M_j \sum \Delta T_m} \vec{Y}_0 = \frac{1}{2^{k+1}} diag \left(e^{J_{m_1}(\lambda_1) \sum_{m=0}^{k} \Delta T_{m_1}}, \ldots, e^{J_{m_p}(\lambda_p) \sum_{m=0}^{k} \Delta T_{m_p}} \right) \vec{Y}_0 \,. \tag{40}$$

The algebraic structure is such that the previuos analysis performed on the unique Jordan block, is still applicable to each Jordan blocks, so we can conclude that in the long–time behaviour we have:

$$Y_{k+1}^{(1)} \to Y_{\infty}^{1}, Y_{k+1}^{(m_1+1)} \to Y_{\infty}^{(m_1+1)} \tag{41}$$
$$Y_{k+1}^{(m_1+m_2+1)} \to Y_{\infty}^{(m_1+m_2+1)}, \ldots, Y_{k+1}^{(m_1+\ldots+m_{p-1}+1)} \to Y_{\infty}^{(m_1+\cdots+m_{p-1}+1)},$$

and $Y_{k+1}^{(i)} \to 0$ otherwise. But we can say something more about the remaining components, in fact by Eq. 32 we have:

$$Y_{k+1}^{(1)} = \frac{1}{2^{k+1}} e^{\lambda_1 S_k} \left(Y_0^1 + S_k Y_0^2 + \frac{(S_k)^2}{2} Y_0^3 + \cdots + \frac{(S_k)^{m_1-1}}{(m_1-1)!} Y_0^{m_1} \right) \tag{42}$$
$$Y_{k+1}^{(m_1+1)} = \frac{1}{2^{k+1}} e^{\lambda_1 S_k} \left(Y_0^{m_1+1} + S_k Y_0^{m_1+2} + \frac{(S_k)^2}{2} Y_0^{m_1+3} + \cdots + \frac{(S_k)^{m_1+m_2-1}}{(m_1+m_2-1)!} Y_0^{m_1+m_2} \right),$$

till

$$Y_{k+1}^{(m_1+\ldots+m_{p-1}+1)} = \frac{1}{2^{k+1}} e^{\lambda_1 S_k} \Big(Y_0^{(m_1+\cdots+m_{p-1}+1)} + S_k Y_0^{(m_1+\cdots+m_{p-1}+2)} + \tag{43}$$
$$+ \frac{(S_k)^2}{2} Y_0^{(m_1+\cdots+m_{p-1}+3)} + \cdots + \frac{(S_k)^{m_1+\cdots+m_p-1}}{(m_1+\cdots+m_p-1)!} Y_0^{(m_1+\cdots+m_p)} \Big).$$

By assumption λ_1 has the largest real part, hence recalling that $S_k \to \infty$, we easly obtain:

$$\frac{Y_{k+1}^{(m_1+1)}}{Y_{k+1}^{(1)}} \to 0, \frac{Y_{k+1}^{(m_1+m_2+1)}}{Y_{k+1}^{(1)}} \to 0, \ldots, \frac{Y_{k+1}^{(m_1+\ldots+m_{p-1}+1)}}{Y_{k+1}^{(1)}} \to 0\,, \tag{44}$$

thus in the long–therm behaviour the only positive component is $Y_{k+1}^{(1)} \to Y_{\infty}^{(1)} > 0$ which hence determine the asymptotic amount of SRMs.

Acknowledgments

Support from the EU FET–PACE project within the 6th Framework Program under contract FP6–002035 (Programmable Artificial Cell Evolution) is gratefully acknowledged.

References

1. Alberts B. et al. : *Molecular Biology of the Cell*, Garland, New York, 2002.
2. Rasmussen S. et al. : *Science*, **303**, 2004, pp 963.
3. Szostak D., Bartel P. B. and Luisi P. L. : *Nature*, **409**, 2001, pp 387.
4. Luisi P. L., Ferri F. and Stano P : *Naturwiss.*, **93**, 2006, pp 1.
5. Rasmussen S., Chen L., Stadler B. and Stadler P. : *Origins of Life and Evol. Biosph.*, **34**, 2004, pp 171.

6. Kaneko K. : *Life: an introduction to complex system biology*, Springer-Verlag, Berlin, 2006.
7. Oberholzer T. et al. : *Biochemical and biophysical Research Communications*, **207**, 1, 1995, pp 250.
8. Maynard-Smith J. and Szathmary E. : *Major transitions in evolution*, Oxford University Press, New York, 1997.
9. Serra R., Carletti T. and Poli I. : *Artificial Life*, **13**, 2007, pp 1.
10. Rasmussen S. et al : *Artificial Life*, **9**, 2003, pp 269.
11. Serra R., Carletti T. and Poli I. : *proceedings BIOMAT2006, World Scientific*, 2007.
12. Serra R. et al. : *Accepted proceedings ECCS–07: European Conference on Complex Systems*, 2007.
13. Ganti T. : *Chemoton Theory, Vol. I: Theory of fluyd machineries: Vol. II: Theory of livin system*, Kluwer Academic/Plenum, New York, 2003.
14. Filisetti A. : *M. Sc thesis, Dept. of Social, Cognitive and Quantitative Sciences, Modena and Reggio Emilia University*, 2007.
15. Minc H. : *Nonnegative matrices*, John Wiley, New York, 1988.
16. Lutkepohl H. : *Handbook of matrices*, John Wiley, New York, 1996.