A Dynamical System for the Algebraic Approach
to the Genetic Code

M. Magini [1] * and J.E.M. Hornos [2]

[1] *Instituto de Pesquisa e Desenvolvimento, UNIVAP,*
Av. Shishima Hifumi, 2911, 12244-000 São José dos Campos, S.P., Brazil

[2] *Instituto de Física de São Carlos, USP,*
Caixa Postal 369, 13560-970 São Carlos, S.P., Brazil

Explaining codon evolution in the standard genetic code is a remarkable subject in Molecular Biology. There are many works which try to develop a model to represent this evolution, sometimes using a certain amount of mathematical tools. The present work has as its main objective to explain one possible dynamical evolution model, which is based in the algebraic approach proposed in 1993 by Hornos and Hornos. This model made an analogy between the evolution of elementary particles and evolution of codons. As a result, the symmetry group that better adjusts degeneracy of the genetic code is the simplectic group $Sp(6)$. The dynamical model present here is based in the 3 dimensional weight space of $Sp(6)$, called *Codon Space*. This space is invariant under the action of the *Octahedral Group*. A map, $\Re^3 \rightarrow \Re^3$, was constructed using an equivariance principle and its action in the Codon Space was observed. The results were based in the analysis of the attractors and their preserved symmetries. The dynamical system action reproduces the main aspects of the labeling proposed in the algebraic approach. Moreover, the map separates the codons in distinguished sets, coupling codons from the same aminoacid in the same attractor. These dynamical results indicate that the algebraic model proposed in 1993 may be an important contribution to explain codon evolution, both in the algebraic and in the dynamical aspects.

1   Introduction

An evolving code is a widely accepted idea after the discovery of non-standard mithocondrial codes[1, 2]. The deviations from the universal code which characterize mithocondrial and some nuclear species are testimonies of the large scale time evolution from a simpler ancestor code to the present form. Examples of such ideas are the primordial code proposed by Osawa and Jukes[1, 2], the theory of Landweber[3], and Gray et. al.[4].

A code containing only a reduced number of aminoacids ruling the protein synthesis in the beginning of life formation is also a very common assumption. Recently Trifonov [5] listed and compared a family of models designed to select the amino-acid ancestors [6-9].

The algebraic approach to the genetic code [10-13] considers the experimental degeneracy of the standard code as a starting point, and searches for a symmetry group able to generate this degeneracy by a process of symmetry-breaking called the Higgs-Crick mechanism. The best, or unique under some restrictions, symmetry found among Lie groups was the symplectic group $Sp(6)$. From an evolutionary point of view, we can say that the symplectic symmetry have been selected by evolution in the same sense that icosahedral symmetry was advantageous in the formation of virus capsule. However, the analogy is vague: in the case of viruses one deals with the symmetry of an object while in the case of the genetic code it is a dynamical symmetry, manifested in the process of protein translation. This approach is similar to symmetry-based models used in physics, mathematics and recently in biology [14-19]. However, in those theories some kind of dynamical system is naturally associated with the model. The most popular example is Maxwell's theory for electromagnetism or the similar theories for strong and weak interactions. Space-time symmetries and gauge invariance appear as properties of partial differential equations, leading immediately to the possibility of quantitative predictions.

In the case of the algebraic model only the degeneracy is used as an input, and the symmetry-breaking pattern is the main result coming from the theory. A number of predictions about the primordial amino acids, the existence of a freezing process, and some hints about the timing of incorporation of amino acids can be obtained from this approach. However, the absence of a dynamical model limits the quantitative power of that approach.

The definition of a class of dynamical systems appropriate to a given symmetry is a well-defined mathematical problem. One can, for example, look for general first-order differential equations equivariant under the given symmetry, and use that to model the codon evolution and the relation with the amino acids. Indeed, the idea to investigate the properties of codons in some 64th-dimensional structure is older than the algebraic model. It has been used to investigate physical-chemist correlations between codons and amino acids by Giulio [8].

The main point in the algebraic approach is the action of a Lie group on the space of codons. The main goal of this article is to propose a dynamical system for the evolution of codons. The high dimensionality of the codon space and the complexity of the symplectic group turn this problem into a formidable task. Dynamical systems for finite groups have been extensively studied [20, 21, 22], but few examples are available in the literature for continuous groups as complex as $Sp(6)$.

A less ambitious program can be implemented focusing our attention in the weight space of the group [23, 24]. This space, to be defined below, is a tridimensional space instead of a 64th dimensional space as the codon space. They are not equivalent: going down to the weigh space we necessarily loose information. But what is important is that the fundamental features of the dynamical system will appear in this simple model.

2　The algebraic model codon assignment

One important step to understand the present approach is the codon assignment proposed by the algebraic model. The decomposition of the codon space [13] occurs in this way: in the first step six independent groups of codons are generated, corresponding to primitive amino acids. The codon space in this stage has six different codons labeled, and as a result the octahedral symmetry decrease to a symmetry $Z_2 \times D_4$. In the next steps, degeneracy decreases until the standard code is reached, where the symmetry preserved is the *Klein* symmetry. In the final step we have all codons represented by vectors in the *Weight Space*. The assignment of the codons in the Weight Space is showed, separated by their degeneracies, in tab.(1), tab.(2), tab.(3), tab.(4), tab.(5).

In the algebraic model for the genetic code aminoacids with degeneracy six can be found in the weight space of the symplectic group in the plans called *bosonic plans*, which are $z = 0, \pm 2$.

Table 1. The codons and their positions in the Weyl space, degeneracy = 6.

Codons - Aminoacid	Positions
CGU,CGC,CGA,CGG,AGA,AGG - Arg	$(-2,1,0),(2,1,0),(-2,-1,0),(2,-1,0),(0,1,0),(0,-1,0)$
UUA,UUG,CUU,CUC,CUA,CUG - Leu	$(0,1,0),(0,-1,0),(0,1,2),(0,-1,2),(0,1,-2),(0,-1,-2)$
AGU,AGC,UCU,UCC,UCA,UCG - Ser	$(-1,0,0),(1,0,0),(-1,0,2),(1,0,2),(-1,0,-2),(1,0,-2)$

Table 2. The codons and their positions in the Weyl space, degeneracy = 4.

Codons	Aminoacid	Positions
GCU,GCC,GCA,GCG	Ala	$(-1,2,0),(1,2,0),(-1,-2,0),(1,-2,0)$
CCU,CCC,CCG,CCA	Pro	$(-1,1,1),(1,1,1),(1,-1,1),(-1,-1,1)$
GGU,GGC,GGA,GGG	Gly	$(-1,1,-1),(1,1,-1),(-1,-1,-1),(1,-1,-1)$
ACU,ACC,ACG,ACA	Thr	$(-1,1,1),(1,1,1),(1,-1,1),(-1,-1,1)$
GUU,GUC,GUG,GUA	Val	$(-1,1,-1),(1,1,-1),(1,-1,-1),(-1,-1,-1)$

Table 3. The codons and their positions in the Weyl space, degeneracy = 3.

Codons	Aminoacid	Positions
AUU,AUA,AUC	Ile	$(-2,0,1),(0,0,1),(2,0,1)$
UAA,UGA,UAG	Term	$(-2,0,-1),(0,0,-1),(2,0,1)$

Table 4. The codons and their positions in the Weyl space, degeneracy = 2.

Codons	Aminoacid	Positions
UUU,UUC	Phe	$(-1, 0, 0),(1, 0, 0)$
GAU,GAC	Asp	$(-1, 0, 0),(1, 0, 0)$
GAA,GAG	Glu	$(0, 1, 0),(0, -1, 0)$
AAU,AAC	Asn	$(0, 1, 0),(0, -1, 0)$
AAA,AAG	Lys	$(0, 1, 0),(0, -1, 0)$
CAA,CAG	Gln	$(0, 0, 1),(0, 0, -1)$
CAU,CAC	His	$(0, 0, 1),(0, 0, -1)$
UGU,UGC	Cys	$(0, 2, -1),(0, -2, -1)$
UAU,UAC	Tyr	$(0, 2, 1),(0, -2, 1)$

Table 5. The codons and their positions in the Weyl space, degeneracy = 1.

Codons	Aminoacid	Positions
AUG	Met	$(0, 0, -1)$
UGG	Trp	$(0, 0, 1)$

The plan names were defined by the last coordinate in the weight vector. For example a codon with assignment vector type $(a, b, \pm 1)$ has to be in a *fermionic plan* otherwise it will be in a *bosonic plan*.

All aminoacids with degeneracy three are in the fermionic plan[1]. In tab. (4) is showed the aminoacids with degeneracy two.

There are nine aminoacids with degeneracy two. Some of them can be found in a fermionic plan and some in the bosonic plan. There are two aminoacids with degeneracy one. The representation of the codons in a three dimensional space leads to a problem: the 64 codons, that were represented in the $Sp(6)$ model by independent vectors, have to fit in a $3d$ space, and thus some information is lost. The arrangement of the codons in the weight space by plans is showed in Fig. 1. The codons position is a fundamental characteristic for a dynamical discrete system behavior, since from these positions was defined the dynamical system action; in other words, they are initial conditions to iterate the discrete map.
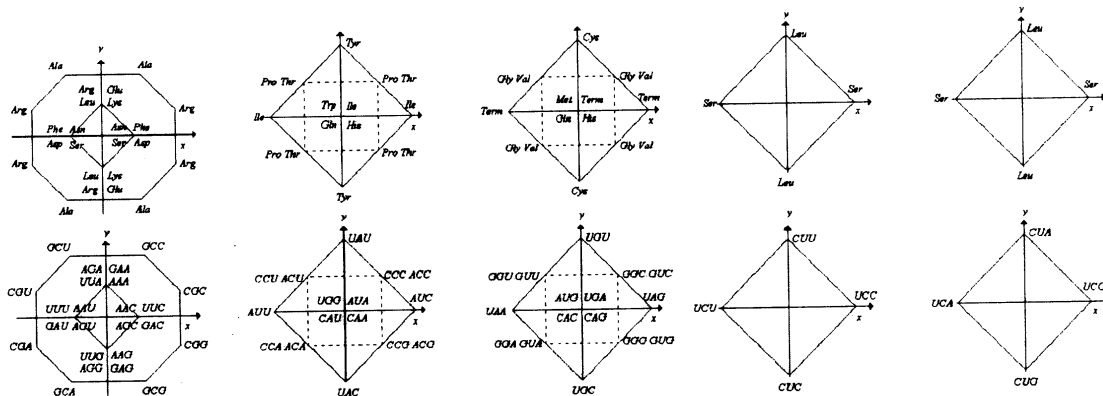

Figure 1. Amino acid and codon assignments for the standard genetic code, according to [13]. From left to right: plane $z = 0$; plane $z = 1$; plane $z = -1$; plane $z = 2$; plane $z = -2$.

3 The octahedral equivariant map

The octahedral equivariant map can be constructed using the generators of the group O_h with dimension three[25]. These generators are matrices 3×3, in total of three. All different combinations of the generators produce the rest of the 45 elements of the O_h group. Each symmetry operation has the corresponding matrix representation. For example: a rotation of 90^0 around the z axis is represented by the matrix

$$M_1 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \qquad (1)$$

This matrix correspond to the permutation element (1234) or a rotation by 90^0 degrees around the z - axis. The other two 3-dimensional generators of O_h are

$$M_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad (2)$$

---

[1]Fermions are particles with spin $\frac{1}{2}$ or half-integer, here the fermionic plan was normalized multiply by two. The same procedure was done with bosonic plans.

and

$$M_3 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \qquad (3)$$

corresponding to the permutation elements $(16)(35)(24)$ and $(13)(24)(56)$, respectively. The last one is the complete reflection. Multiplying these matrices among themselves we generate more 45 matrices, including the identity, which represents the octahedral group.

The dynamical system which has this symmetry as equivariant symmetry can be represented by a set of functions. The equivariance principle is given by: *The function $f(r; \lambda) : M \to N$ is said to be equivariant under the action of the element $\gamma \in \Gamma$, where Γ is a symmetry group if, $f(\gamma r; \lambda) = \gamma f(r; \lambda)$, where the constant λ is the bifurcation parameter* [20]. Through the map iteration it is possible to see the isotropic subgroups when the bifurcation parameter has increased or decreased, depending on the system constructed [25]. The isotropic subgroups are represented by the attractors generated by the iteration process [22]. In our case the map is an application $f(x) : \Re^3 \to \Re^3$, and the attractors are in the same space, but drawing in a plane by perspective.

The first step in the construction of the dynamical system is to apply the equivariance in a generic function $f(x, y, z; \lambda) = \Sigma \Delta_{\tau, \theta, \rho} x^\tau y^\theta z^\rho$, where τ, θ, ρ are variant parameters. Through the equivariance properties by the matrices action it is possible to construct the map. The equivariant map can be reduced in the final step of its construction using the invariant function. The invariance principle is given by *The function $f(r; \lambda) : M \to N$ is said to invariant under the action of the symmetry group Γ if; $f(\gamma r; \lambda) = \gamma f(r; \lambda) = f(r; \lambda)$ where $\gamma \in \Gamma$.* Using the invariance principle was construct the three invariant functions for the octahedral symmetry, as follow.

$$\begin{aligned} u &= x^2 y^2 z^2 \\ v &= x^2 + y^2 + z^2 \\ s &= x^2 y^2 + z^2 y^2 + x^2 z^2. \end{aligned} \qquad (4)$$

There are an infinite number of possibilities to produce a dynamical system which preserves the octahedral symmetry. One of these possibilities of dynamics, based in stability concepts and analyzing the fixed points for each value of the bifurcation parameter, is given as follow.

$$F = \begin{pmatrix} f_x = x[\delta s + \sigma v + \lambda + \varphi\{(x^2 - y^2 - z^2) + (y^2 z^2 - x^2 y^2 - z^2 x^2)\}] \\ f_y = y[\delta s + \sigma v + \lambda + \varphi\{(y^2 - x^2 - z^2) + (x^2 z^2 - x^2 y^2 - z^2 y^2)\}] \\ f_z = z[\delta s + \sigma v + \lambda + \varphi\{(z^2 - x^2 - y^2) + (x^2 y^2 - z^2 y^2 - z^2 x^2)\}] \end{pmatrix}. \qquad (5)$$

The parameters δ, σ, φ were determined from parity relations between the matrices and the old general parameters τ, θ, ρ. Their values are determined observing the fixed points possibilities. The role of the parameters is determine the shape of the attractors. Different attractors with O_h can be produced by different values of the fixed parameters δ, σ, φ. This characteristic can be observed in some works were the bidimensional maps with D_m symmetry are analyzed [21]. A fixed point is defined by:

$$f(x_0, y_0, z_0; \lambda) = (x_0, y_0, z_0). \qquad (6)$$

The map is an application $\Re^3 \to \Re^3$ so the fixed points (x_0, y_0, z_0) can be obtained by mathematical relations that preserves the real space. In other words, the fixed points are responsible to define the dynamics behavior in the real space and they have a directly relation $(x_0, y_0, z_0) = \sqrt{f(\delta, \sigma, \varphi; \lambda)}$. Imposed that this square root is a real number for a range of λ values it was achieve the fixed parameters values.

Starting our dynamics with $\delta = 1$, $\sigma = -0.7$, $\varphi = -0.8$ and the bifurcation parameter in $\lambda = -1.440$ produces

an attractor with O_h symmetry. This result is well known in many aspects. However, the interest here is in how this map can change the evolution aspects(labeling) proposed by the algebraic approach. Besides, the map can produce a totally different evolution path when acting in the codon space.

Using the labeling proposed by the algebraic approach described in Fig. 1 it is possible to define what initial conditions are in our perspectives. The λ variation promotes a set of different attractors, where the preserved symmetries are related by the invariant subgroups. For $\lambda = -1.400$ a D_4 symmetry is generated. Through the computational behavior the symmetry relations can be easily determined. Application of the generator matrices in the points resulted by the map iteration leave all points invariant. This fact leads to the D_4 symmetry as invariance group for this value of λ.

4   Results and conclusions

We will now consider the codons positions in the Weight Space as initial conditions to start our dynamics. The sym-

metries that are preserved by different values of λ are: octahedral symmetry in $-1,440$, D_4 symmetry in $-1,400$, *Klein* symmetry in $-1,354$ and Z_2 or *Chiral* symmetry in $-1,332$, which are the isotropic subgroups of O_h. To produce the dynamics attractors the initial condition was fixed in just one value. As an example it is possible to see the octahedral subgroups represented by their attractors in Fig. 2, Fig. 3 acting in the codon space when the values of λ are $-1,440$ and $-1,400$. In this specific case the dynamics does not produce the octahedral symmetry, the D_4 symmetry is preserved. This happens because the aminoacids are in plans which reduce the space in one dimension. In other words, there is no dynamics in a three dimensional space when one is considering the map action in codon space even if $\lambda = -1,440$.
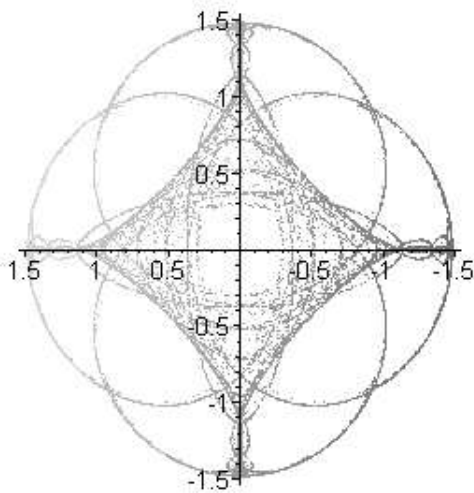

Figure 2. Attractor with D_4 symmetry preserved.

Different initial conditions could produce completely different aspects of our dynamics. For our surprise, the points fixed as codons in the algebraic model weight space gave us one important information: codons that represent the same aminoacid remains in the same attractor until the final step of symmetry breaking in the dynamical model, when the symmetry Z_2 appear as main symmetry, pairing the codons two by two.

The symmetry breaking or codon evolution represented by a dynamical system is unique. As an example of the present approach it will be considered some aminoacids — a complete study is unnecessary because the symmetry is valid for all aminoacid — so the dynamical system acts in the same way when the same position is used as initial condition. Let us present a brief discussion about the action of the discrete map in the aminoacids position labeled represented in the codons space.

When the dynamical system starts in the codon position that represents the aminoacid Gly, it promotes a split

of attractors corresponding to the codons GGC and GGA in a two-points attractor and the other two codons of Gly in another two-points attractor, which are distinct. The same pattern can be viewed when the aminoacid Val is used. A notable fact is this: no codon is mixed with codons from different aminoacids because the different assignment will mix different codons from distinguishable aminoacids. The unexpected pairing GGC and GGA besides GGC and GGU depends on assignment proposed, exclusively, and can be repaired in the algebraic model to make the dynamical model more precise.
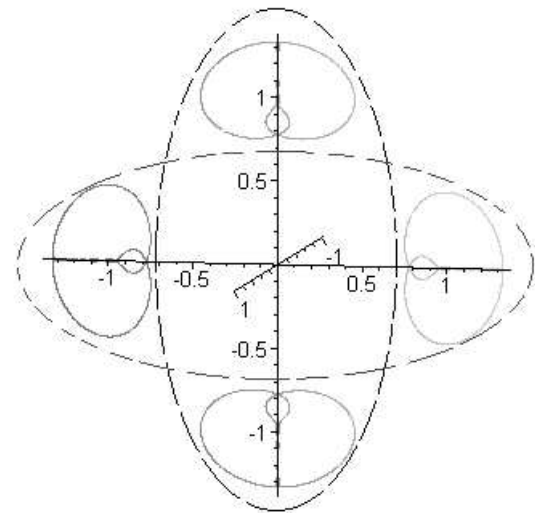

Figure 3. Distinct attractors (divide by traceable lines) with *Klein* symmetry preserved.

The codons of Ala are in a attractor with D_4 symmetry when $\lambda \leq -1.400$ Fig. 2. Together with them, forming in the same orbit, are the Arg codons CGU, CGC, CGA and CGG, see tab.(1). These codons are in the same irreducible representation in the algebraic model, which is an important result. For $-1.400 < \lambda \leq -1.354$ the codons of Ala are separated of the Arg codons in two distinct attractors Fig. 3. Each one has as a setwise group of symmetry, the *Klein* group. The assignment done by the algebraic model and the dynamical system action guarantee a correct separation of codons during their evolution. An undesired situation that not occurred, in this case, will be the mixing of different codons from distinct aminoacids in the same orbit. For lower values of λ there occurs a shift of codons from the same aminoacids forming two attractors for the same aminoacid. Similar result occurred with the Gly.

The codons of Arg, Ser and Leu with degeneracy 6 were separated by the action of the dynamical system in two behaviour families: one family is a set of 4 codons and the another is a set with 2 codons[2] while in the algebraic model

---

[2]This result shows an strong analogy with the family boxes proposed by many biologists[1, 2].

they are in the same representation. The Weyl Group action in the weight space does not connect the six codons of each aminoacid. In this way they evolved to distinct attractors, as discussed.
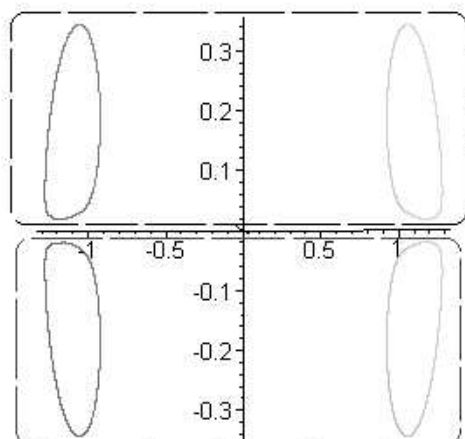

Figure 4. Distinct attractors (divide by traceable lines) with *Quiral* symmetry preserved.

The dynamical system differs the termination codons UAA and UAG, see tab.(3), from the codon UGA. This distinction can be observed in non-standard codes when UGA starts to codify Ile. Surprisingly these two codons are in the same orbit in Fig. 4. The same pattern occurs with Met and Trp.

The algebraic approach [10] was proposed to explain codon degeneracy in the standard genetic code. After some time a series of works and tools were created to produce stronger conclusions about evolution. Most of them were a success in the sense that they prove that ideas suggested by the model are correct, in biological and mathematical context. Here was presented one important approach, which only uses as a starting point the Weyl Group of $Sp(6)$. Note that the possibilities of distinction using the dynamical model are more than 1. In other words, from the action of the map in the Codon Space it could be expected a set of symmetry breaking chains. Actually the possible number of chains are bigger than 20. More than this, the points in the map are unstable and hyperbolics so the behavior of the map in unpredictable. However, it was observed that even in the dynamical system the codon evolution produces practically the same path proposed by the algebraic approach, with one difference: the aminoacids with degeneracy 6 are not well represented by the dynamics, this is a weak point in this approach. Nevertheless, aminoacids with degeneracy 6 were a doubt problem when the algebraic approach was constructed, and are considered by many researchers as "problematic" aminoacids when the evolution context is analyzed. In a general context the dynamics action in Codon Space is responsible to shift different aminoacids leaving codons from a same aminoacids in a same dynamic symmetry. These results generate families of codons in a evolution sense provided by variation of λ.

References

[1] S. Osawa, T.H. Jukes, K. Watanabe, and A. Muto, Microb. Rev. **56** 229 (1992).

[2] S. Osawa, *Evolution of the Genetic Code*, (Oxford University Press, Oxford, 1995).

[3] R.D. Knight and L.F. Landweber, Cell vol. **101**, 569 (2000).

[4] M.W. Gray, G. Burger, and B. F. Lang, Science vol. **283**, 1476 (1999).

[5] E.N. Trifonov, Gene **261**, 139 (2000).

[6] M. Eigen and R. Winkler- Oswatitsch, Die Naturwissenshaften **68**, 217 (1981).

[7] M. Eigen and R. Winkler- Oswatitsch, Die Naturwissenshaften **68**, 282 (1981).

[8] M. Di-Giulio, Orig. Life Evol. Biosph. **26**, 589 (1996).

[9] K. Nieselt-Struwe and P.R. Wills, J. Theor. Biol. **187**, 1 (1997).

[10] J.E.M. Hornos and Y.M.M. Hornos, Phys. Rev. Lett. **71**, 4401 (1993).

[11] J.E.M. Hornos and Y.M.M. Hornos, J. Biol. Phys. **20**, 289 (1994).

[12] M. Forger, J.E.M. Hornos, and Y.M.M. Hornos, Phys. Rev. E **56**, 7078 (1997).

[13] J.E.M. Hornos, Y.M.M. Hornos, and M. Forger, Int. J. Mod. Phys. B **13**, 2795 (1999).

[14] A. Arima and F. Iachello, *The Interacting Boson Model, Cambridge Monographs on Mathematical Physics*, (Cambridge University Press, Cambridge, 1987).

[15] M. Eigen, B.F. Lindermann, M. Tietze, R. Winkler-Oswatitsch, A. Dress, and A. von Haeseler, Science **244**, 673 (1989).

[16] J. Chela-Flores, Acta Biotheretica **37**, 267 (1988).

[17] J. Maddox, Nature, **367**, 111 (1994).

[18] I. Stewart, New Scientist **141**, 16 (1994).

[19] M.A. Jiménez-Montaño, Biosyst. **64**, 47 (1999).

[20] P. Chossat and M. Golubitsky, Siam J. Math. Anal. **19** n.6, 1259 (1988).

[21] P. Chossat and M. Golubitsky, Phys. D **32**, 423 (1998).

[22] E. Barany, M. Dellnitz, and M. Golubitsky, Phys. D **67**, 66 (1993).

[23] M. Hammermesh, *Group Theory and It is Application to Physical Problems*, (Dover Books, New York, 1989).

[24] A.O. Barut and R. Raczka, *Theory of Group Representations and Applications*, (World Scientific Publishing, Singapore, 1986).

[25] M. Golubitsky, I. Stewart e D. Schaeffer, *Singularities and Groups in Bifurcation Theory, vol. II*, (Springer - Verlag, New York 1989).