Complexity and emergence of meaning: toward a semiophysics(*)

F.T. Arecchi
University of Firenze
and
Istituto Nazionale di Ottica Applicata

Outline

1. Introduction

In a debate on "Complexity and emergence", we should first of all provide clear definitions for these terms, in order to ascertain how much of what we say depends on our cultural bias, is an artifact of our linguistic tools, and how much corresponds to hard facts, to our embedding in an open environment, whose features, even though actively elaborated by our semantic memory, can not be taken as sheer "autopoiesis", but are grounded on an ontology.
This inquiry is done from the point of view of a physicist who has been active for decades in investigating the formation of collective or coherent processes out of a large amount of otherwise separate individuals, pointing out the critical appearance (emergence) of new world configurations and the elements of novelty of this emergence, which make this phenomenon complex. By complex we do not mean the trivial fact that the computational cost of their description is high (in such a case I would rather call them complicate) but the fact that available knowledge stored in well established models is not sufficient to predict reliably the emergence, and one must integrate the deductive chains with extra information which introduces a historical flavour into the scientific procedure.

(+) Presented at the Workshop "Complexity and Emergence" organized by the Academie Internationale de Philosophie des Sciences, Bergamo (Italy) on 9-13 May 2001

This presentation is organized as follows.
In a first part we discuss the sources of wonder, what Plato called the origin of science, that is, why among many peculiarities (saliences) we prefer to focus our attention on some ones (pregnancies) (Sec. 2). Then we explore how, as we organize our knowledge into a scientific language, we select the relevant words (names) depending on their relation with an ontology (things) (Sec. 3).

In a second part we try to put order into the debated issue of complexity, introducing a fundamental separation between some purely mental situations without any realistic counterpart (closed systems) and what we in fact come across everyday (open systems) (Secs. 4 and 5).
The third part goes to the very ground of perceptual processes. If we accept – as proper of complex systems – to organize our knowledge over different and mutually irreducible hierarchical levels, each one with its own rules and language, then the most fundamental one in cognitive processes is the physical description of how external stimuli (light, sound, pressure, chemicals) are transformed into sensorial perceptions.
Already at this neurodynamical level, we come across a quantum limitation which forbids the brain operations to be fully simulated by a universal computing machine (Sec. 6). By purpose, I said “brain” since I do not wish to enter the debates on “mind”, “consciousness” etc.
I have called “neurophysics” the combination of neurodynamical facts, whereby neurons are treated as physical objects to be compared to lasers or other nonlinear dynamical systems, and the quantum limitation emerging from the peculiar spike synchronization strategy selected in course of the natural evolution as the optimal strategy to elaborate information into relevant cognitive processes.
As for the Reference list, I have often replaced the specific mention of an article or a book by a Web site, where one can conveniently browse for a more satisfactory answer. I think that time is ripe to consider this reference tool as a standard one.

2. Salience vs pregnancy

The world around us is full of salient features, that is, sharp gradients which denote the transition from one domain to another one. Salience can be captured automatically by a scanning detector equipped to grasp differential features. Saliences have a geometric (space-wise) and dynamic (time-wise) flavor. They correspond to objective features: what Thomas Aquinas called “dispositio rei” and more recently A.Reinach (a follower of Husserl) called “Sachverhalt”[Smith].

We might say that saliences uncover an ontology [Poli], however, in order to classify a set of features and organize them through mutual relations, we need to assign selection criteria.

These descriptive criteria have guided the construction of sectorial ontologies in many AI (Artificial Intellicence) areas [Guarino].

Hence the problem arises: are there individual objects, or instead any world organization is an arbitrary cut that we operate by picking up some saliences and disregarding other ones?

Historically the modern European culture, in line with its Greek – Jewish roots, had chosen the first side of this dilemma; however the contact with Eastern philosophies, through Schopenhauer and Mach, introduced a “conventionalism” or linguistic relativism, whereby one could build different, uncorrelated, ontologies depending on the points of view from where saliences were selected [Feyerabend, Capra].

The recent emphasis on “regional ontologies”, focused on particular saliences and whence on particular classes of objects, is a modern technical limitation. A philosopher of science [Agazzi] would rather say that selecting a point of view gives rise to a particular science, focusing on some truths different from other ones. Yet there is a hard aspect of saliences, that is, they uncover facts having their own existence, and not just dependent on our cultural artifacts.

In line with Gestalt psychology, René Thom has introduced “pregnancy” to denote a subset of saliences which are relevant for the individual observer [Thom 1988].

In the case of animals, pregnancy is related to vital needs (search of food, escaping from predators, sexual appeal). Some of these needs may be genetically imprinted [Lorenz], some others are the result of cultural influences. This latter case is particularly important for human beings. In this regard, it is fundamental the contribution of J. Piaget called “Genetic epistemology”. As one explores the formation of logical structures in children, one realizes that they derive from <u>actions</u> on the objects, not from the objects themselves; in other words, the formation of logical structures is grounded on the coordination of actions, not necessarily on language. In fact, language is one of the possible semiotic functions; the other ones, as gestures, or imitation, or drawing, are forms of expression independent from language, as carefully studied in the case of deaf-mutes [Evans].

Anyway, against relativism, Thom insists on the objective character of the prominent saliences, which he classifies in terms of differential geometry [Thom 1975].

A very convincing dynamical formulation of the emergence of a new feature, or the disappearance of an old one, as a “control parameter” is changed, is given by 1937 Landau’s theory of phase transitions [Landau-Lifshitz]. We present the argument in the updated 1973 formulation called “Synergetic” by Haken [Haken] and initially motivated by a new astonishing phenomenon as the laser threshold, namely, the onset of a collective coherent emission of light out of billion of atoms, which below that threshold instead contribute individual, unrelated (so called spontaneous) emission acts, as it occurs in a conventional light source.

Let me anticipate something I’ll discuss in greater detail in Sec. 3. Assume from the time being that we succeeded in describing the world as a finite set of N features, each one characterized by its own measured value $x_i\,(i = 1 \text{ to } N)$, x_i being a real number, which in principle can take any value in the real domain $(-\infty,\infty)$ even though boundary constraints might confine it to a finite segment L_i.

A complete description of a state of facts (a “dispositio rei”) is given by the N- dimensional vector

$$\underset{\approx}{x} \equiv \left(x_1, x_2, \ldots. x_i, \ldots. x_N\right) \qquad (1)$$

The general evolution of the dynamical system $\underset{\approx}{x}$ is given by a set of N rate equations for all the first time derivatives $\dot{x}_i = dx_i / dt$. We summarize the evolution via the vector equation

$$\dot{\underset{\approx}{x}} = \underset{\approx}{f}(\underset{\approx}{x}, \mu) \tag{2}$$

where the function $\underset{\approx}{f}$ is an N-dimensional vector function depending upon the instantaneous $\underset{\approx}{x}$ values as well as on a set of external (control) parameters μ

Solution of Eq. (2) with suitable initial conditions provides a trajectory $\underset{\approx}{x}(t)$ which describes the time evolution of the system. We consider as ontologically relevant those features which are <u>stable</u>, that is, which persist in time even in presence of perturbations. To explore stability, we perturb each valuable x_i by a small quantity ξ_i, and test whether each perturbation ξ_i tends to disappear or to grow up catastrophically.

However complicated is the nonlinear function $\underset{\approx}{f}$, the local perturbation of (2) provides for ξ_i simple exponential solutions versus time of the type

$$\xi_i(t) = \xi_i(0)e^{-\lambda_i t} . \tag{3}$$

The λ_i can be evaluated from the functional shape of Eq. (2). Each perturbation ξ_i shrinks or grows in course of time depending on whether the corresponding stability exponent λ_i to positive or negative. The λ_i are called the "local Liapunov exponents" .

Now, as we adjust from outside one of the control parameters μ, there may be a critical value μ_c where one of the λ_i crosses zero (goes from + to -) whereas all the other $\lambda_j\,(j \neq i)$ remain positive. We call λ_u the exponent changing sign (*u* stays for "unstable mode") and λ_s all the others (*s* stay for stable) (fig. 1).

Around μ_c, the perturbation $\xi_u(t) \approx e^{-\lambda_u t} \approx e^o$ tends to be long lived, which means that the variable x_u has rather slow variations with respect to all the others, that we cluster into the subset x_s which varies rapidly. Hence we can split the dynamics (2) into two subdynamics, one 1-dimensional (*u*) and the other (*N-1*) – dimensional (*s*), that is, rewrite Eq. (2) as

$$\begin{aligned} \dot{x}_u &= f_u(x_u, x_s, \mu) \\ \dot{x}_s &= f_s(x_u, x_s, \mu) \end{aligned} \tag{4}$$

The second one being fast, the time derivative $\dot{x}_s$ rapidly goes to zero, and we can consider the algebraic set of equations $f_s = 0$ as a good physical approximation. The solution yields x_s as a function of the slow variable x_u

$$x_s = g(x_u) \tag{5}$$

We say that x_s are “slaved” to x_u. Replacing (5) into the first of (4) we have a closed equation for x_u

$$\dot{x}_u = f_u\big(x_u, g(x_u), \mu\big) \tag{6}$$

First of all, a closed equation means a self consistent description, not depending upon the preliminary assignment of x_s. This gives an ontological robustness to x_u; its slow dependence means that it represents a long lasting feature and its self consistent evolution law Eq. (6) means that we can forget about x_s and speak of x_u alone.

Furthermore as μ crosses μ_c, a previous stable value $x_u^{(1)}$ is destabilized. A growing ξ_u means that eventually the linear perturbation is no longer good, and the nonlinear system saturates at a new value $x_u^{(2)}$ (fig. 2).

Such is the case of the laser going from below to above threshold; such is the case of a thermodynamic equilibrium system going e.g. from gas to liquid or from disordered to ordered as the temperature at which it is set (here represented by μ) is changed.

To summarize, we have isolated from the general dynamics (2) some critical points (bifurcations: see the shape of fig. 2) where new salient features emerge. The local description is rather accessible, even though the general nonlinear dynamics *f* may be rather nasty.

Told in this way, the scientific program seems converging towards firm answers, as compared to the shaky arguments of philosophers. However it was based on a preliminary assumption, that there was a “natural” way of assigning the x_i. In the next Section we explore how to extract the x_i from observations.

3) Names and things

To avoid subjective biases, one should replace definitions in everyday language by quantitative assessments. This is done by isolating something which can be represented in a metrical space as a number, and speaking of a larger or smaller degree of it, of the distance between two values etc., by referring to the corresponding numbers.

In modern science this attitude was consecrated by G. Galilei in his 1610 letter to Marc Welser [Galilei], where he says: “don’t try to grasp the “essence” (i.e. don’t try to define the “nature “ of things) but stick to some quantitative affections”.

For instance, in the case of apples, rather than arguing on the nature of apples, make comparisons among them, based on quantitative measures of “affections” as flavour, colour, shape, size, taste. I have listed five qualities for each of which we know how to introduce a meter and hence set a metrical space. At Galilei’s time, there was a distinction between primary (measurable) and secondary (subjective) qualities. Nowadays, we know how to objectify and measure secondary quality as flavour or taste, thus, that old distinction is no longer relevant.

Two different attitudes may be adopted, namely,

i) Phenomenology: once apples are characterized by a sufficient group of parameters, all apples will be a suitable intersection of the flavour axis, the colour axis etc. in a multidimensional space; such a description is complete (all apples will be included) and unambiguous (two different apples will not overlap in such a multidimensional space, that is, they differ by at least one of their representative numbers).

ii) Reductionism: split the apple into small pieces, and these again into smaller ones, until you reach a set of elementary components (the biomolecules) out of which, with a wise dosage of the elements of the set, one can reconstruct (synthesize) all kinds of apples. This procedure is lengthier than phenomenology, but it is universal; out of a set of components one can also synthesize oranges, dogs etc. Moreover it looks objective; if we come across intelligent beings from elsewhere, we don’t know if our selected affections are relevant for them, but surely they know how to split the apple into components and catch each component's dynamics. When the only known interaction law (2) was Newton's, this approach seemed the ultimate one; thus Newtonianism was considered as the new revolutionary approach upon which to build any world view. An Italian writer of XVIII century,F.Algarotti, wrote in this regard a booklet "Il Newtonianesimo per le dame" ("Newtonianism for ladies") which was the first manifesto of the woman liberation movement, translated into most European languages.

Both approaches can be formalized. A familiar example of a formal theory is Euclid's geometry. Once a set of components has been defined and their mutual relations stated, via a group of axioms, all possible consequences are deducible as theorems, which provide by necessity all explanations as well as predictions on the future behavior.

In phenomenology, we have many sciences, in reductionism, we have a single fundamental science, that of the elementary components, out of which we can extract all relevant levels of organization.

Such an approach has been abundantly criticized [AndersonPW, Arecchi1992, Arecchi1995, Arecchi and Farini]. The main criticism is that the nonlinear dynamics of microscopic components undergoes multiple bifurcations, of the kind of fig. 2, as a control parameter is varied in order to build up a macroscopic object; thus the construction from scratch of a large size system is by no means unique, and the multiple outcomes are a token of complexity, as discussed in Sec. 5.

Since however this essay points at a more fundamental approach to our cognitive acts, for the time being we list current approaches without criticism, just to introduce the technical language and get acquainted with the corresponding problems.

Reductionism does not mean to refer always to Democritus' atoms (nowadays, we would say to leptons and quarks), but to stop at a suitable level where the elementary components are sufficiently characterized. Such are the biomelecules for living beings.

For all practical purposes, the biologists need the descriptive properties of the biomolecules, plus some knowledge of the nearest lower level, that of atomic physics. Think e.g. of the role of Na^+, K^+ and Ca^{2+} ion conductances in neurophysiology or of the devastating effect of some atoms as thallium or plutonium on enzymatic processes, or the balance of hydrogen bonds and van der Waals bonds in stabilizing protein folding.

Thus biochemistry is founded on atomic physics but it does not require nuclear or subnuclear physics. Similarly, atomic physics requires only nuclear physics but not further levels below, and so on.

However, there is no fundamental level which acts as the ultimate explanatory layer.

In fact, recently the problem has been addressed whether a formal description of the state of an elementary particle is sufficient to build a faithful replica of it elsewhere (the so called teleportation problem [Bennett et al. 1993]). A formal description within the current language of quantum mechanics is not sufficient to provide full recovery of the particle. One must add some non formalizable piece of information. It is not the case to expand such a technical part, I just recall that transmitter and receiver must share not only verbal information (the formal description) but also they must be exposed respectively to the two parts of an EPR (Einstein Podolsky Rosen) state, or "entangled" pair. By "entangled we mean a strong quantum correlation which has no classical counterpart, and hence cannot be formalized in the classical physics language [Bennet et al. 1993]. Just like interacting with a baby or somebody of different language; nominal definitions are not enough, the dictionary must be integrated by "ostensive definitions", just putting your finger on the object.

We now discuss how the set (1) of relevant variables and the law of motion (2) are established in the two cases apriori (or reductionistic) and aposteriori, or phenomenological.

3.1. Apriori

This approach started with Newton and has continued up to the present search for a unified theory of all fundamental interactions. It consists in counting the particles in the universe attributing to each one 6 numbers, 3 coordinates in Euclidean space and 3 momenta (or more simply in the non relativistic limits, 3 velocity components).

Quantum mechanics added more specifications for internal degrees of freedom, such as "spin" and electrical charge, both for leptons and quarks, plus "strangeness", "charm" and some other properties for quarks. The numbers corresponding to these internal degrees of freedom do not span over all real, but are confined to a small set of possible values. Most often, they correspond to dichotomous variables with just two values, conventionally denoted as 0 and 1. Anyway, each x_i is a group of 6 real numbers for a classical particle, plus a few other discrete numbers for quantum particles.

The coupling function f of Eq. (2) implies mutual relations. Initially, the single universal one was considered Newton's gravitational interactions. Later, Maxwell electromagnetic theory became the prototype of any field theory. Here, the coupling is no longer between particles but each particle feels forces corresponding to a new entity, the local electromagnetic-field at its position. Viceversa, the fields are generated by moving charged particles. Thus the particle-particle interactions are mediated by the fields; in field dynamics we speak no longer of "action at distance".

In electromagnetic theory one adds a new set of x_i consisting of the 6 components of the electric and magnetic field at each point in space. In this case we have a continuous field problem, since the position is not a discrete set of numbers, but varies with continuity. We write $x(r)$ where *r* denotes the position coordinates in a 3 dimensional space; here *r* is made of three real numbers and we write this as $r \in R^3$ (*r* belongs to the 3 dimensional real space).

The continuum problem has haunted modern physics since its start, and clever devices to deal with it have been produced. However in most cases the continuous fabric of space can be discretized as a lattice of points at finite distances from each other.

I illustrate this trick with reference to a time dependent signal *x(t)* observed over a finite time interval *T*; it depends on all the real values taken by *t* in the segment *T*. Outside *T* the signal is not defined, thus we can arbitrarily assume that it repeats periodically with period *T*, without affecting the values within the observation interval. This means that its information is contained in a discrete Fourier series of pairs of real numbers (A_n, φ_n) sampled at a frequency $n\frac{1}{T}$ which is the *n-th* harmonic of the fundamental repetition frequency *1/T*, that is

$$x(t) = \sum_{n \cong 1}^{\infty} A_n \cos\left(n\frac{2\pi}{T} + \varphi_n\right) \tag{7}$$

Thus, the finite interval *T* limitation has simplified the mathematical description of the signal from continuum to discrete. We do indeed probe with continuity each real *t*, but we synthesize the signal by summing up at each point a discrete set of sinusoids. If furthermore we consider that any detection or signal processing device is a low pass filter with a finite frequency window B (i.e., it responds only to frequencies up to B), then we can truncate the sum (7) up to a maximum value $n_{max}=B^{\cdot}T$ and the signal information of $x(t)$ over *T* is contained in n_{max} sinusoids. Since however for each frequency we have an amplitude A_n and a phase φ_n, the set of numbers which fully specify our signal is twice $B^{\cdot}T$, that is,

$$N = 2BT \tag{8}$$

This important sampling theorem, stated by RADAR investigators during World War II [Shannon], sets the resolution limit for an observation with bandwidth B lasting for a time T. To acquire more information, one must either increase B or T.
In a similar way, a visual system (the eye, or a telecamera) frames a finite two-dimensional domain of sizes L_1, L_2 with bandwidths B_1 and B_2. Thus the number of relevant picture elements (pixels) of a two dimensional image is given by

$$N = 4B_1B_2L_1L_2 \tag{9}$$

The sampling theorem has induced the strong belief that any cognitive process deals with a finite number of elements, and furthermore that the universe is described by a finite, though very large, number of degrees of freedom.
The mathematics of XVIII century physics has been expressed in terms of *ODE's* (ordinary differential equations) for the continuous variation of a variable x_i as a continuous time t flows. If infinitesimally close space points have to be coupled, then we express the co-presence of space derivatives together with time derivatives by *PDE's* (partial differential equations). If time is discretized by sampling it at final distances, then the *ODE's* are replaced by iteration maps, whereby the value of x at the discrete time $n+1$ depends upon the value of x at the previous time n .
If also the space can be discretized as a lattice of disjoint points denoted by discrete indices i,j, than the space derivatives reduce to coupling the iteration maps at different points (*CML*= coupled map lattice)
Eventually, if also the variable x is constrained to assume a finite set of values, in the limit binary values *(0,1)*, than we have a *CAM* (cellular automaton machine) consisting of a network of points each represented by a binary variable which updates at discrete times depending on the values of the neighboring points or "cells" [Wolfram]. We summarize in Table 1 the different types of mathematically modeling the evolution of a physical system.

Table I

State variable	*Time variable*	*Space variable*	
C	*C*	*C*	*PDE*
C	*C*	*D*	*ODE*
C	*D*	*D*	*CML*
D	*D*	*D*	*CAM*

C = continuous, D = discrete

CAM techniques have been very powerful in dealing with model problems, from biology (genetics, population dynamics, immune system) to sociology (traffic problems, econophysics) and meteorology. They have become the basis of a finitistic ideology, whereby the universe can be seen as a large CAM [Toffoli].
However a fundamental limitation to this ideology arises from quantum non commutativity of pairs of complementary observables, as we'll discuss in Sec.6.

3.2. Aposteriori
New classes of phenomena are disclosed by the exploitation of innovative sophisticated systems of investigation, e.g. recording long time series in financial trading or in car traffic, imaging

techniques in brain investigation, automatic machines for sequencing DNA. It is very difficult to fit this new phenomena into a Newtonian frame. A component description is hopeless and one wants to approach the problem directly, without prejudices.
Suppose that, by salience considerations, we have focused our attention on a time dependent quantity *u(t)*. Salience means that *u(t)* displays a patterned behavior, that is, it is strongly correlated with its values at later times. Take *u(t)* as the deviation from an average value, then its time average is zero, that is, *u(t)* looks as a sequence of +/- values. Consider the product of two *u*'s at different times. If they are unrelated, then also the average product of them, called correlation function *C(t,t')*, is zero.

A nonzero *C(t,t')* is a signature of salience. Karhunen (1946) and Loeve (1950) introduced independently the following retrieval method that we call KL [Karhunen, Loeve]. If $\theta_n(t)(n = 1,2...,L)$ are the L most prominent characteristic functions (called eigenfunctions) which retrieve the correlation *C(t,t')* and if *l* is a small number, then we can accurately reconstruct the signal as a weighted sum of L functions as follows:

$$u(t) = \sum_{n=1}^{L} a_n \theta_n(t) \tag{10}$$

If the signal depends on space rather than time, then we grasp the salient features of a given space pattern. Each of these saliences in general is spread over the whole domain. A relevant example in the convective motion of a fluid was given by [Ciliberto et al.], where the three main "modes" of behavior (*L=3*) are distributed over the whole fluid cell.
The opposite limit occurs when saliences are strongly localized. Think e.g. of the face elements (nose, eyes, mouth shape) upon which identikits of criminals are built in police investigation. In such a case, KL would be inconvenient, since it requires a large *L* to converge toward a localized feature. Here the successful phenomenological approach is just the opposite of KL. It consists in reconstructing a pattern, e.g. a face, by a small series of "prototypes". This approach is used in many machine vision programs [Weber et al.].

4. Closed versus open systems

I have discussed elsewhere [Arecchi1995] the failure of what Anderson called the "constructionist" program [AndersonPW]. Trying to build a structured system out of its elementary components does not provide a univocal outcome.
Indeed the components interact via a nonlinear law as Eq.(2) and the emergence of a new stable structure starting from an initial condition P_0 requires the appropriate tuning of the control parameters. Such a tuning provides in general more than one new stable state.
(Fig.3)
The emerging states 2,2' are equivalent, thus, as μ is tuned to μ_{C1}, the system has equal probability of emerging in the state 2 or 2', unless we break the symmetry of the bifurcation by the application of an external field (fig.3b), which makes the two stable states non equivalent, and hence one of the two (the upper one in the figure) chosen with higher probability. The number of equivalent outcomes increases exponentially with the order of the bifurcation: 2 at μ_{C1}, 4 at μ_{C2} and so on.
Hence, a reductionistic program based on the dynamical description of the components does not provide a unique outcome. We must assign some extra information consisting of possible external fields, which specify univocally the final state.
But external fields are beyond the information provided by the dynamical properties of the components.

We must then distinguish between <u>closed</u> and <u>open</u> systems. The first ones are surrounded by walls which provide precise boundary conditions. Their evolution yields multiple outcomes so that we predict ”potentialities”, never the “actual” system that is observed.
In the case of open systems we must augment our description including the values of the external fields which select a unique outcome at each bifurcation. In general we don’t know how to do that apriori; we can rather proceed backward to an historical reconstruction of external contributions which have obliged our open system to evolve from an initial to a final state. Notice that the setting of μ on the horizontal axis is at our will for a closed system, whereas in an open system it has a proper evolution in time that we do not control. Thus a bifurcation tree as fig. 3 looks as an evolutionary tree, usually plotted in biology as a function of time.
We might think that a metalevel description could treat the overall situation (set of components plus external process) as a closed system. In fact, we would transfer the ambiguities at the metalevel, and to recover uniqueness the metalevel must be affected by its own external forces and so on. In other words, by successive layers we deal with an “onion” structure.
A global description of the whole cosmic onion as a closed system is the dream of theoretical physics. A TOE (theory of everything) would be an equation as (2) where $\underset{\approx}{x}$ are now all the degrees of freedom of the universe and $\underset{\approx}{f}$ is the unified mathematical formulation that one day will be reached among electro-weak, strong and gravitational interactions. In such a situation there would be <u>nothing</u> left out, thus new must itself be a function of $\underset{\approx}{x}$ and hence Eq. (2) of TOE would be <u>closed</u>, with no external control parameters.
In fact, this is in principle impossible. Any foreseeable nominal description (that is, expressed by precise numbers) is incomplete even for a single particle, as discussed in Appendix I. Therefore we must split the vector x_{tot} of all the degrees of freedom of the universe into an observable set x_0 and a complementary set $\bar{x}$ which escapes our description. Our relevant physical equation refers only to the observed part x_0. Thus Eq. (2) must be re-written as

$$\dot{x}_0 = \underset{\approx}{f}(x_0, \mu(x_{0'}, \bar{x})). \tag{11}$$

This way of writing shows that the $\bar{x}$ dependence of μ excludes the above equation from being a closed one.

5. <u>Complication versus complexity</u>

5.1. Complexity of symbolic sequences

In computer science, we define the complexity of a word (symbol sequence) as some indicator of the cost implied in generating that sequence [Hopcroft and Ullman]. There is a "space" cost (length of the instruction stored in the computer memory) and a "time" cost (the CPU time for generating the final result out of some initial instruction).
A space complexity called AIC (Algorithmic Information Complexity) [Kolmogorov, Chaitin] is defined as the length in bit of the minimal instruction which generates the wanted sequence. This indicator is <u>maximum</u> for a random number, since there is no compressed algorithm (that is, shorter than the number itself) to construct a random number.
A time complexity called "logical depth" [Bennett 1987] is defined as the CPU time required to generate the sequence out of the <u>minimal</u> instruction. It is minimal for a random number, indeed, once the instruction has stored all the digits, just command: PRINT IT.
Of course, for simple dynamical systems as a pendulum or the Newtonian two-body problem, both complexities are minimal.

While AIC refers to the process of building a single item, logical depth corresponds to finding the properties of all possible outputs from a known source.

In fact, the exact specification of the final outcome is beyond the ambition of the natural scientist, whose goal is more modest. It may be condensed in the two following items:

i) to transmit some information, coded in a symbol sequence, to a receiver, possibly economizing with respect to the actual string length, that is, making good use of the redundancies (this requires a preliminary study of the language style);

ii)to predict a given span of future, that is to assign with some likelihood a group of forthcoming symbols.

For this second goal, introduction of a probability measure is crucial [Grassberger 1986, Gell-Mann] in order to design a complexity-machine, able to make the best informational use of a given data set.

Such a machine which should mimic the scientific procedure acts as follows [Crutchfield 1992]. Assume that a group of measuring apparatuses have provided the agent A with an information coded as a numerical sequence *s* (for convenience we use a binary code, so that the length $|s|$ of *s* is measured in bits). Agent A has a good understanding of what happens if it can transfer to a received B a compact information *y* upon which B reconstruct the sequence *s'=s*. Of course, *y* has to be shorter than *s* otherwise it would be a tautology, which implies no understanding whatsoever. Thus A is obliged to recur to a class of models built in its memory. Suppose it has chosen a model *m*, then A can simulate the behavior of the observed system and realize that there is an error *e* between the actual measurement *s* and its model reconstruction. If B receives both information *m* and *e*, then B can reconstruct *s'=s*. The bit length of the transmitted information is

$$|y| = |m| + |e| \tag{12}$$

and it has to be minimized for a successful description.

In this case we call complexity of the explanation the compression ratio

$$C(m,s) = \frac{|y|}{|s|} = \frac{|m| + |e|}{|s|} \tag{13}$$

The value of *C* is bounded above by *1*; it depends upon the choice of the model *m*. There are two limit cases for which *C=1* is the worst. When the model is trivial ($|m| = 0$) the entire information is on the error channel ($|e| = |s|$). When the model is tautological *m=s* there is no error $|e| = 0$.

The class of models can be scanned by a Bayes rule [Crutchfield]. This is the case of an “expert system” equipped with a class of models, within which the system formulates the best diagnosis by minimizing *C*.

5.2. A dynamical approach to complexity

As discussed in Sec. 3, the reductionist approach consists of building a hierarchy from large to small and showing how the behaviour of smaller objects should determine that of larger ones. But here a perverse thing occurs. If our words were a global description of the object in <u>any</u> situation (as the philosophical "essences" in Galileo) then, of course, knowledge of elementary particles would be sufficient to make predictions on animals and society. In fact, Galileo's self-limitation to some "affections" is sufficient for a limited description of the event, but only from a narrow point of

view. Even though we believe that humans are made of atoms, the affections that we measure in atomic physics are insufficient to make predictions on human behaviour.
We call complexity the fact that higher levels of organization display features not predictable from the lower ones, as opposed to the previous computer cost of a symbolic task, which we rather call complication .
This way, complexity is not a property of things (like being red or hot) but it is a relation with our status of knowledge, and for modern science it emerges from Galileo's self-limitation.
Reductionism from large to small was accompanied by a logical reduction of the scientific explanation to a deductive task out of a set of axioms.
In this spirit, a scientific theory is considered as a set of primitive concepts (defined by suitable measuring apparatuses) plus their mutual relations. Concepts and relations are the axioms of the theory. The deduction of all possible consequences (theorems) provides predictions which have to be compared with the observations. If the observations falsify the expectations, then one tries with different axioms.
The deductive process is affected by a Gödel undecidability like any formal theory, in the sense that it is possible to build a well formed statement, but the rules of deduction are unable to decide whether that statement is true or false.
Besides that, a second drawback is represented by intractability, that is, by the exponential increase of possible outcomes among which we have to select the final state of a dynamical evolution. As discussed in Sec. 4, during the dynamical evolution of an open system, the control parameters $\{\alpha\}$ may assume different values, hence the cascade of bifurcations provides a large number of final states starting from a unique initial condition.
Thus the reductionist tentative of explaining reality out of its constituents yields an exponentially high number of possible outcomes, when only one is in fact that observed. This means that, while the theory, that is the syntax, would give equal probability to all branches of the tree, in reality we observe an organization process, whereby only one final state has a high probability of occurrence.
Hence, whenever we are in presence of organization, that is of a unique final state, this means that at each bifurcation vertex the symmetry was broken by an external agent which forces a unique outcome, as shown in Fig. 3.
We can thus summarize the logical construction (to rephrase Carnap [Carnap]) of a large system out of its components as follows:

i) A set of control parameters is responsible for successive bifurcations leading to an exponentially high number of final outcomes. If the system is “closed” to outside disturbances, then all outcomes have comparable probabilities and we call complexity the impossibility of predicting which one is the state we are going to observe.

ii) For a system of finite size embedded in an environment, there is a set of external forces applied at each bifurcation point, which break the symmetries, biasing toward a specific choice and eventually leading to a unique final state.

We are in presence of a conflict between (i) “syntax” represented by the set of rules (axioms) and (ii) “semantics” represented by the intervening external agents. The syntax provides many possible outcomes. But if the system is open, then it organizes to a unique final outcome. Once the syntax is known, the final result is therefore an acknowledgement that the set of external events must have occurred, that they have made the evolution meaningful (whence “semantics”).
We define "certitude" the correct application of the rules, and "truth" the adaptation to the reality.
However, due to the freedom we have in formulating theoretical conjectures, the same final outcome would be reached by a different set of rules ,corresponding to a different syntactic tree. In such a case, retracing back the new tree of bifurcations, we would reconstruct a different set of external agents. Thus, it seems that truth, , is language dependent!

Furthermore, this freedom in choosing the rules (the syntax) means that we can even find a set of axioms which succeeds in predicting the correct final state without external perturbations. This is indeed the pretension of the so called "autopoiesis", or "self-organization" [Krohn].
From a cognitive point of view, a selforganized theory can thought of as a "petitio principi", a tricky formulation tailored for a specific purpose and not applicable to different situations. Rather than explicitly detailing the elements of environment which break the symmetry, the supporter of the selforganized theory has already exploited at a pre-formalized level a detailed knowledge of the process in planning appropriate axioms.
An “ad-hoc” model may fit a specific situation, but in general it lacks sufficient breadth to be considered as a general theory. Think e.g. of the Ptolemaic model of the solar system which holds only for an Earth based observer, as compared to the Newtonian one, which holds also for an observer travelling through the solar system.
However, in describing the adaptive strategy of a living species, or a community etc., “self-organization” may be the most successful description. In other words, once the environmental influences have been known, better to incorporate their knowledge into the model, thus assuring the fast convergence to a given goal.

5.3. Complexity differs from complication

When all the rules of a game and all the partners (components) have been introduced, we are in presence of a definite symbolic system. The corresponding problems can be solved at a cost which may increase more than polynomially with the number of partners, e.g. consider the Travelling Salesman problem or TSP. We prefer to call “complication” the difficulty of solution of a problem within a formal system, and use “complexity” to denote any cognitive task in front of an open system. In such a case, a cognitive machine as an “expert system” is limited to a finite set of models *m*. Furthermore, it is bound to a precise setting of the measuring apparatuses: take for instance the list of the apple properties listed in Sec. 2. As discussed in [Arecchi2001] this limitation of an expert system is overcome by an adaptive strategy, whereby an agent spans not only the class of available models by a Bayesian strategy, as in the computer based model reasoning peculiar of expert systems [Magnani], but also changes in course of time the type of measures performed, thus reaching a meta-level where that cognitive agent is equivalent to a large class of expert systems.
Notice that “large” can be still finite. In Sec. 6 we will discuss a fundamental quantum relation for time dependent processes, which exclude finitism.

6 Neurophysics of perception

6.1 What is neurophysics

It is by now firmly established that a holistic perception emerges, out of separate stimuli entering different receptive fields, by synchronizing the corresponding spike trains of neural action potentials [Von der Malsburg, Singer].
We recall that action potentials play a crucial role for communication between neurons [Izhikevich]. They are steep variations in the electric potential across a cell’s membrane, and they propagate in essentially constant shape from the soma (neuron’s body) along axons toward synaptic connections with other neurons. At the synapses they release an amount of neurotransmitter molecules depending upon the temporal sequences of spikes, thus transforming the electrical into a chemical carrier.
As a fact, neural communication is based on a temporal code whereby different cortical areas which have to contribute to the same percept *P* synchronize their spikes. Spike emission from a nonlinear threshold dynamical system results as a trade off between bottom-up stimuli to the higher cortical regions (arriving through the LGN (lateral geniculate nucleus) from the sensory detectors, video or audio) and threshold modulation due to top-down readjustment.

It is then plausible to hypothesize, as in ART (adaptive resonance theory [Grossberg]) or other computational models of perception [Edelman and Tononi] that a stable cortical pattern is the result of a Darwinian competition among different percepts with different strength. The winning pattern must be confirmed by some matching procedures between bottom-up and top-down signals.
We present two fundamentals aspects of percept formation, namely,

i) The neurodynamics of spike formation

ii) A quantum limitation in information encoding/decoding through spike trains

As for the first aspect, a saddle point instability separates in parameter space an *excitable* region, where axons are silent, from a *periodic* region, where the spike train is periodic (equal interspike intervals). If a control parameter is tuned at the saddle point, the corresponding dynamical behavior (homoclinic chaos) consists of a frequent return to the instability [Allaria et al.]. This manifests as a train of geometrical identical spikes, which however occur at erratic times (chaotic interspike intervals). Around the saddle point the system displays a large susceptibility to an external stimulus, hence it is easily adjustable and prone to respond to an input, provided this is at sufficiently low frequencies; this means that such a system is robust against broadband noise as discussed later.
As for the second aspect, the temporal coding requires a sufficiently long sequence of synchronized spikes, in order to realize a specific percept. If the sequence is interrupted by the arrival of new uncorrelated stimuli, then a fundamental uncertainty ΔP emerges in the percept space P. This is related to the finite duration ΔT allotted for the code processing by a fundamental uncertainty relation

$$\Delta P \cdot \Delta T \geq C$$

where C is a positive dimensional quantity whose non zero value represents a quantum constraint on the coding. This constraint implies that the percepts are not set-theoretical objects, that is, objects belonging to separate domains, but there are overlap regions where it is impossible to discriminate one percept from another. We will discuss later the occurrence of this new class of time dependent perceptual illusions.
We call "neurophysics" the combination of i) and ii), by analogy with "econophysics" which as extracted some general physical from economic phenomena [Mantegna].
Neurophysics is distinct from neurodynamics, which is the investigation of dynamical models of neuron behaviors, as well from neurophysiology, which explores the coupling of different brain areas. Neurophysics is restricted to the two above items, and it is rather model-indipendent, so that it provides a general ground upon which different models can be built and compared.

6.2 Role of duration T in perceptual definition: a quantum aspect
How does a synchronized pattern of neuronal action potentials become a relevant perception? This is an active area of investigation which may be split into many hierarchical levels. At the present level of knowledge, we think that not only the different receptive fields of the visual system, but also other sensory channels as auditory, olfactory, etc. integrate via feature binding into a holistic perception. Its meaning is "decided" in the PFC (pre frontal cortex) which is a kind of arrival station from the sensory areas and departure for signals going to the motor areas. On the basis of the perceived information, actions are started, including linguistic utterances.
Sticking to the neurodynamical level, which is the most fundamental one, and leaving to other sciences, from neurophysiology to psychophysics, the investigation of what goes on at higher levels of organization, we stress here a fundamental temporal limitation.
Taking into account that each spike lasts about 1 msec, that the minimal interspike separation is 3 msec, and that the average decision time at the PFC level is about T=240 msec, we can split T into 240/3 =80 bins of 3 msec duration, which are designated by 1 or 0 depending whether they have a spike or not. Thus the total number of messages which can be transmitted is

$$2^{80} \approx 10^{27}$$

that is, well beyond the information capacity of present computers. Even though this number is large, we are still within a finitistic realm. Provided we have time enough to ascertain which one of the 10^{27} different messages we are dealing with, we can classify it with the accuracy of a digital processor, without residual error.

But suppose we expose the cognitive agent to fast changing scenes, for instance by presenting in sequence unrelated video frames with a time separation less than 240 msec. While small gradual changes induce the sense of motion as in movies, big differences imply completely different subsequent spike trains. Here any spike train gets interrupted after a duration ΔT less than the canonical T. This means that the PFC cannot decide among <u>all</u> perceptions coded by the neural systems and having the same structure up to ΔT, but different afterwards. How many are they: the remaining time is $\tau = T - \Delta T$. To make a numerical example, take a time separation of the video frames $\Delta T = T/2$, then $\tau = T/2$. Thus in spike space an interval ΔP comprising

$$2^{\tau/3} \approx 2^{40} \approx 10^{13}$$

different perceptual patterns is uncertain.

As we increase ΔT, ΔP reduces, thus we have an uncertainty principle

$$\Delta P \cdot \Delta T \geq C$$

The problem faced thus far in the scientific literature, of an abstract comparison of two spike trains without accounting for the available time for such a comparison, is rather unrealistic. A finite available time ΔT places a crucial role in any decision, either if we are trying to identify an object within a fast sequence of different perceptions or if we are scanning trough memorized patterns in order to decide about an action.

As a result the perceptual space P per se is meaningless. What is relevant for cognition is the joint (P,T) space, since “in vivo” we have always to face a limited time ΔT which may truncate the whole spike sequence upon which a given perception has been coded. Only “in vitro” we allot to each perception all the time necessary to classify it.

A limited ΔT is not only due to the temporal crowding of sequential images, as reported clinically in behavioral disturbances in teenagers exposed to fast video games, but also to sequential conjectures that the semantic memory essays via different top-down signals. Thus in the metrical space (P,T), while the isolated localization of a percept P (however long is T) or of a time T (however spread is the perceptual interval ΔP) have a sense, a joint localization both in percept and time has an ultimate limit when the corresponding domain is less than the quantum area C.

Let us consider the following thought experiment. Take two percepts P1 e P2 which for long processing times appear as the two stable states of a bistable optical illusion, e.g the Necker cube. If we let only a limited observation time ΔT then the two uncertainty areas overlap. The contours drawn in Fig.4 have only a qualitative meaning. The situation is logically equivalent to the non commutative coordinate-momentum space of a single quantum particle. In this case it is well known [Zurek] that the quasiprobability Wigner function has strong non classical oscillations in the overlap region. We cannot split the coordinate-momentum space into two disjoint domains (sets) to which we can apply a Boolean logic or a classical Kolmogorov probability. This is the reason why the Bell inequalities are violated in an experiment dealing with such a situation [Omnès]. The Wigner function formalism derives from a Schroedinger wavefunction treatment for pure state, and corresponding density matrix for mixed states.

In the perceptual (P,T) space no Schroedinger treatment is yet available but we can apply a reverse logical path as follows.

The uncertainty relation $\Delta P \cdot \Delta T \geq C$ forbids a partition of the (P,T) space into sets. Therefore the (P,T) space is non commutative. Thus it must be susceptible of a Wigner function treatment and we can consider the contours of Fig.4 as fully equivalent to isolevel cuts of a Wigner function. Hence we can introduce Schroedinger cat states and violations of Bell inequalities exactly as in quantum physics but with a reverse logical process, as illustrated in Fig.5.

The equivalent of a superposition state should be a bistable situation observed for a time shorter than whole decision time. An experimental test is in preparation in my research group. Such a test should provide an estimation of the C value, which plausibly changes from individual to individual, and for a single one may be age and motivation dependent.
Thus in neurophysics time occurs under two completely different meanings, that is, as the ordering parameter to classify the position of successive events and as the useful duration of a relevant spike sequence, that is the duration of a synchronized train. In the second meaning, time T is a variable conjugate to perception P.
The quantum character has emerged as a necessity from the analysis of an interrupted spike train in a perceptual process. It follows that the (P,T) space cannot be partitioned into disjoint sets to which a Boolean yes/not relation is applicable and hence where ensembles obeying a classical probability can be considered. A set-theoretical partition is the condition to apply the Church-Turing thesis, which establishes the equivalence between recursive functions on a set and operations of an universal computer machine.
The evidence of quantum entanglement of overlapping perception should rule out in principle a finitistic character of the perceptual processes. This should be the negative answer to the Turing 1950 question whether the mental processes can be simulated by a universal computer [Turing].
Among other things, the characterization of the “concept” or “category” as the limit of a recursvive operation on a sequence of individual related perceptions gets rather shaky, since recursive relations imply a set structure.
Quantum limitations were also put forward by Penrose [Penrose] but on a completely different basis. In his proposal, the quantum character was attributed to the physical behavior of the “microtubules” which are microscopic components of the neurons playing a central role in the synaptic activity. However, speaking of quantum coherence in biological processes is very hard to accept, if one accounts for the extreme vulnerability of any quantum system due to “decoherence” processes, which make quantum superposition effects observable only in extremely controlled laboratory situations.

Bibliography

Agazzi, E., 1974, "*Temi e Problemi di Filosofia della Fisica*", Edizioni Abete, Roma.

Allaria E., Arecchi F.T., Di Garbo A., Meucci R. 2001 “Synchronization of homoclinic chaos” Phys. Rev. Lett .86, 791.

Anderson, P.W., 1972, "More is different", *Science* 177: 393-396.

Arecchi F.T. 1992 "Models and metaphors in science", Proceedings of Pontifical Academy of Sciences, Plenary Session on *Emergence of complexity in Mathematics, Physics, Chemistry and Biology,* Rome 26-31 Oct. 1992 (ed. B.Pulmann) Princeton University Press, Princeton NJ.

Arecchi F.T. 1995 "Truth and certitude in the scientific language", International Conference on *"Self-Organization of Complex Structures from Individual to Collective Dynamics*", Berlin 24-28 Sept. 1995, ed. by F. Schweitzer, Gordon and Breach, Amsterdam 1997

Arecchi F.T. and Farini A., 1996, “*Lexicon of Complexity*”, Studio Editoriale Fiorentino, Firenze.

Arecchi F.T. 2000 “Complexity and adaptation: a strategy common to scientific modeling and perception” Cognitive Processing 1, 23.

Arecchi F.T. 2001, “Complexity versus complex system: a new approach to scientific discovery” Nonlin.Dynamics, Psychology, and Life Sciences, 5, 21.

Bennett, C.H., 1987, “Dissipation, information, computational complexity and the definition of organization” in *“Emerging Syntheses in Science”*, D. Pines, ed., Addison Wesley, Redwood City, Ca.

Bennett ,C.H., Brassard, G., Crepeau, C., Jozsa, R., Peres, A., and Wootters, W. 1993,"Teleporting an Unknown Quantum State via Dual Classical and EPR Channels", *Phys. Rev. Lett.* **70**, 1895-1899

Capra F., 1975 *“The Tao of Physics: an exploration of the parallels between modern physics and Eastern mysticism*” The International Bestseller (3d Ed. 1991) Washington D.C.

Carnap, R. 1967 *“The Logical Construction of the World”*: University of California Press, Berkely, CA.

Chaitin, G.J., 1966, "On the length of programs for computing binary sequences", *J. Assoc. Comp. Math*. 13: 547-560.

Churchland P.S. and Sejnowski T.J., 1992, “*The computational Brain*”, MIT Press, Cambridge, Mass.

Ciliberto S., Nicolaenko B., 1991 “Estimating the number of degrees of freedom in spatially extended systems”, Europhys. Lett. 14, 303

Crutchfield J.P. and Young K. 1989,”Inferring statistical complexity”,Phys.Rev.Lett **63**,105

Edelman, G.M., and G. Tononi 1995 “Neural Darwinism: The brain as a selectional system” in “*Nature's Imagination: The frontiers of scientific vision”*, J. Cornwell, ed., pp.78-100, Oxford University Press, New York.

Evans, R.I., 1973,*”Jean Piaget: the Man and His Ideas”*, E.P. Dutton, New York.
Feyerabend P., 1975, *“Against Method”* Verso, London.

Galilei G., Letter to M. Welser "On the sun spots", 1 December 1612, Italian original in: “*Opere di G. Galilei”*, Firenze 1932, Vol. V, pp. 187-188.

Gell-Mann, M., 1994, *“The Quark and the Jaguar”*, W.H. Freeman, New York.

Grassberger, P., 1986, "Toward a quantitative theory of self-generated complexity", *Int. J. Theor. Phys.*, 25: 907-919.

Grossberg S., 1995a “The attentive brain” The American Scientist, 83, 439.

Grossberg S., 1995b “Review of the book by F.Crick: The astonishing hypothesis: The scientific search for a soul” 83 (n.1)

Guarino N.: www.ladseb.pd.cnr.it/infor/Ontology/ontology.html

Haken H.1983, “*Synergetics, An Introduction”*, 3rd edition, Springer Verlag, Berlin

Hopcroft, J.E. and Ullman, J.D., 1979, *"Introduction to Automata Theory, Languages and Computation"*, Addison-Wesley, Reading, MA.

Hubel D.H., 1995 *"Eye, brain and vision"*, Scientific American Library, n. 22, W.H. Freeman, New York.

Izhikevich E.M., 2000 "Neural Excitability, Spiking, and Bursting" Int. J. of Bifurcation and Chaos. 10, 1171

Karhunen, K., 1946, "Zur Spektraltheorie Stochasticher Prozess" *Ann. Acad. Sci. Fennicae*, **37**

Kolmogorov, A.N., 1965, "Three approaches to the quantitative definition of information", *Problems of Information Trasmission* 1: 4-20.

Krohn, W., Küppers, G. and Nowotny, H., 1990, *"Self-organization, Portrait of a Scientific Revolution"*, Kluwer Academic Publishers, Dordrecht.

Landau, L.D. & Lifshitz, E.M.1980," *Statistical Physics"* (3rd edition), Pergamon,New York

Loève, M.M., 1955*'Probability Theory'*, VanNostrand, Princeton, N.J.

Magnani L., Nersessian N. J. and Thagard P. (Editors) 1999, "*Model-based Reasoning In Scientific Discovery"* Kluwer, New York.

Mantegna, R.N. and Stanley H.E., 2000 "*An Introduction to Econophysics*" Cambridge University Press, Cambridge UK.

Omnès R. 1994 *"The interpretation of Quantum Mechanics*" Princeton University Press, Princeton NJ

Penrose R. 1994 "*Shadows of the Mind"* Oxford University Press New York.

Poli R.: www.formalontology.it/polir.htm

Rieke, F., Warland, D., de Ruyter van Steveninck, R. and Bialek, W. 1996. "*Spikes: Exploring the neural cod*e", MIT Press, Cambridge Mass..

Rodriguez E., George N., Lachaux J.P., Martinerie J., Renault B.and Varela F. 1999, "Perception's shadow:Long-distance synchronization in the human brain", Nature **397**:340-343.

Shannon, C.E., 1949."Communication in the Presence of Noise," *Proc. IRE*, Vol. 37, No. 1,

Simon, H., 1982, "The architecture of complexity", Proc. Amer. Philos. Soc., 106: 467-485.

Singer W. E Gray C.M., 1995, "Visual feature integration and the temporal correlation hypothesis" Annu.Rev.Neurosci. 18, 555.

Smith B.: http://wings.buffalo.edu/philosophy/faculty/smith/

Thom R., 1975, *“Structural stability and morphogenesis*” Benjamin, Reading,Mass.

Thom R., 1988, *“Esquisse d’une sémiophysique”* InterEditions, Paris

Toffoli,T.1998 “Non -conventional computers”in *Enc.of El. and Electronic Eng*(ed. J Webster)

J.Wiley and Sons

Turing A. 1950 “Computing Machinery and Intelligence” Mind 59, 433

Von der Malsburg C., 1981 “The correlation theory of brain function”, reprinted in E. Domani, J.L. Van Hemmen and K. Schulten (Eds.), *“Models of neural networks II”*, Springer, Berlin.

Weber M., Welling M, Perona P. “Unsupervised Learning of Models for Recognition”, *Proc. 6th Europ. Conf. Comp. Vis., ECCV2000*, Dublin, Ireland, June 2000

Wolfram,S.1984,” Cellular Automata as Models of Complexity “Nature, 311, 419--424

Zurek W.H. 1991 “Decoherence and the transition from quantum to classical” Phys.Today (October issue) p.36

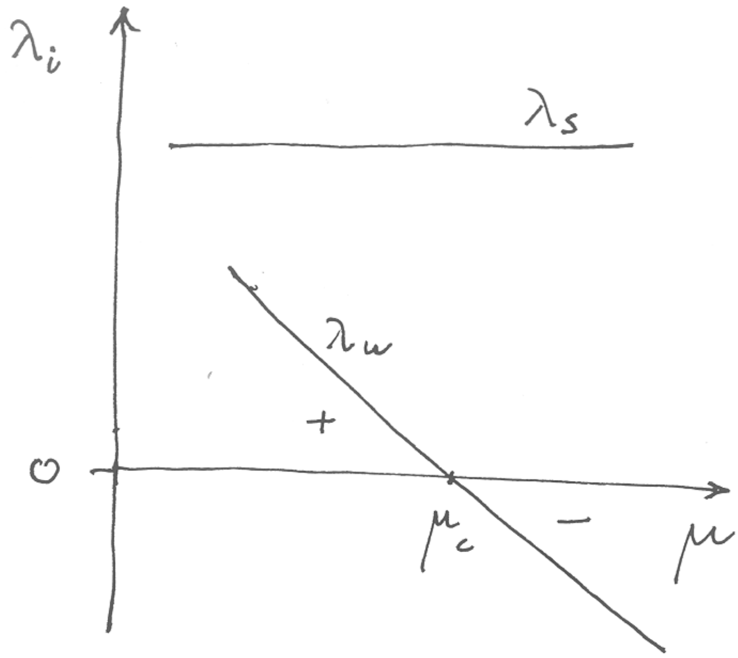

Fig. 1 : When the control parameter crosses the value μ_c, the eigenvalues λ_s remain positive, whereas the eigenvalue λ_u goes from positive to negative.

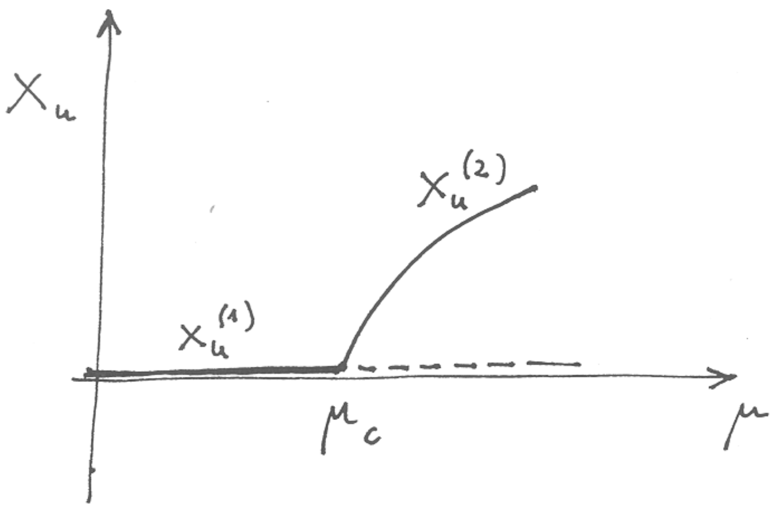

Fig. 2 : The horizontal branch $x_u^{(1)}$ becomes unstable at μ_c and a new stable branch $x_u^{(2)}$ emerges from the bifurcation point.

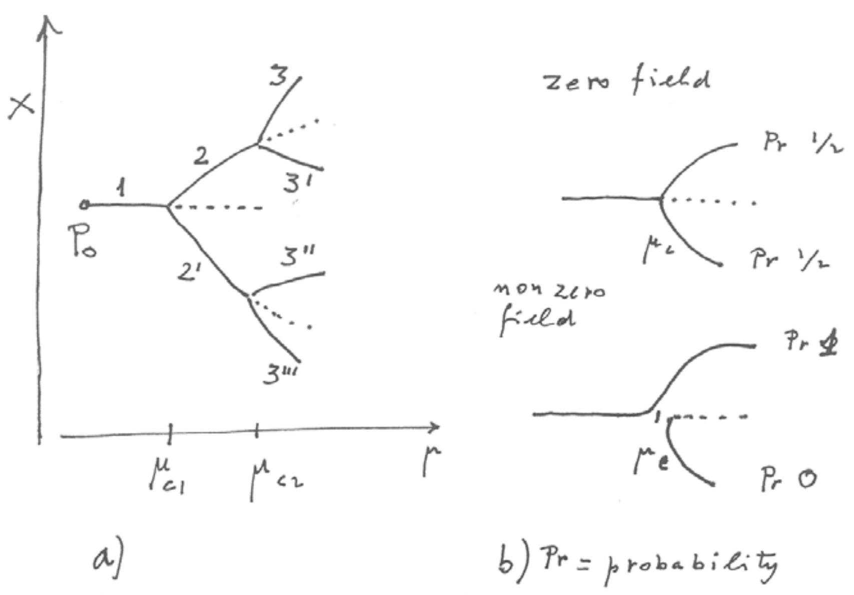

Fig. 3:(a) Example of bifurcation diagram. The dynamical variable x (order parameter) varies vertically, the control parameter μ varies horizontally. Solid (dashed) lines represent stable (unstable) steady states as the control parameter is changed.
(b) Upper: symmetric bifurcation with equal probabilities for the two stable branches.
Lower: asymmetric bifurcation in presence of an external field. If the gap introduced by the field between upper and lower branch is wider than the range of thermal fluctuations at the transition point, then the upper (lower) branch has probability 1 (0).

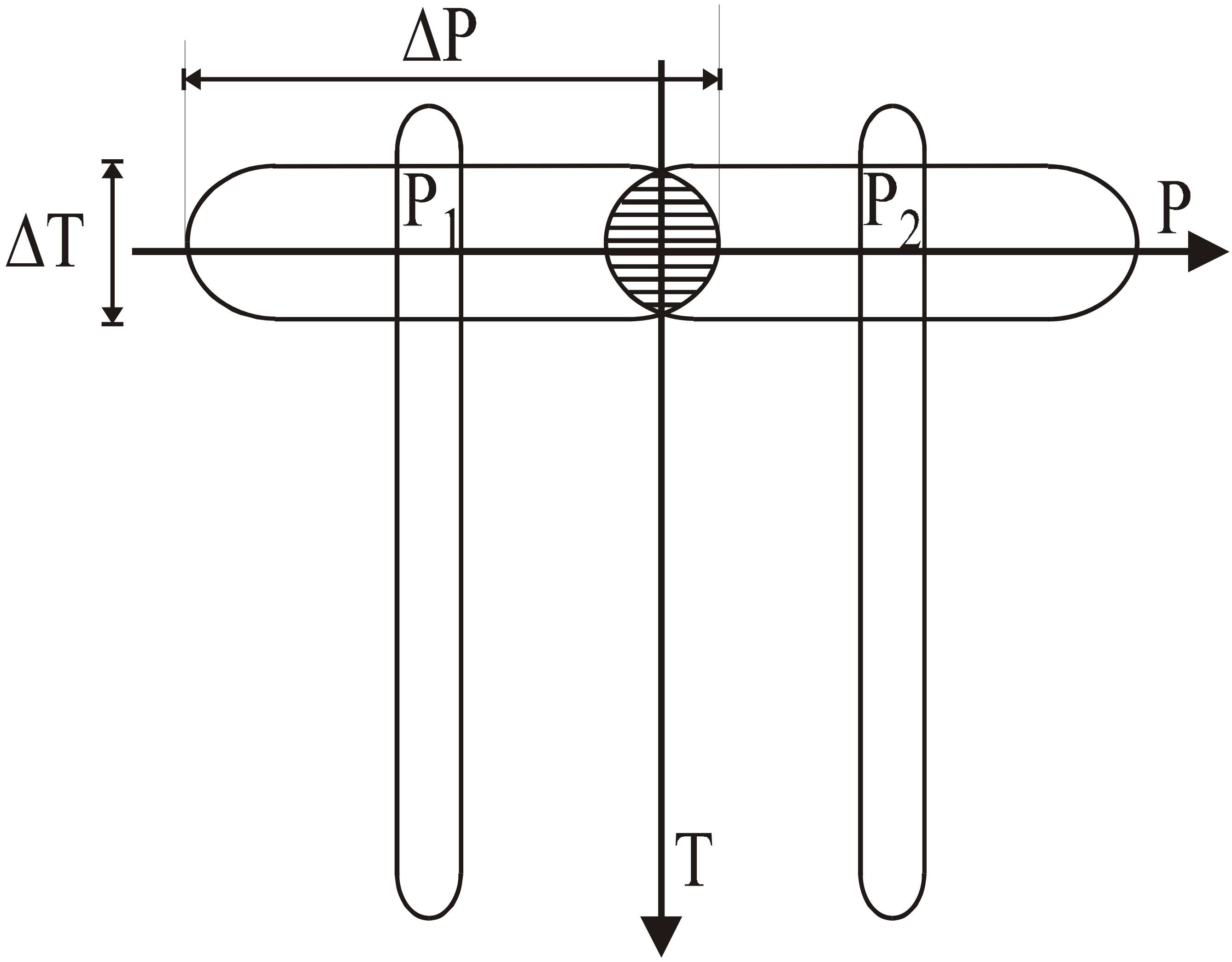

Fig.4 Uncertainty areas of two perceptions P_1 and P_2 for two different durations of the spike trains. In the case of short ΔT, the overlap region is represented by a Wigner function with strong positive and negative oscillations which go as $\cos \frac{\Delta P}{C} T$ along the T axis; therefore with a frequency given by the ratio of the percept separation $\Delta P = P_2 - P_1$ to the perceptual "Planck's constant" C.

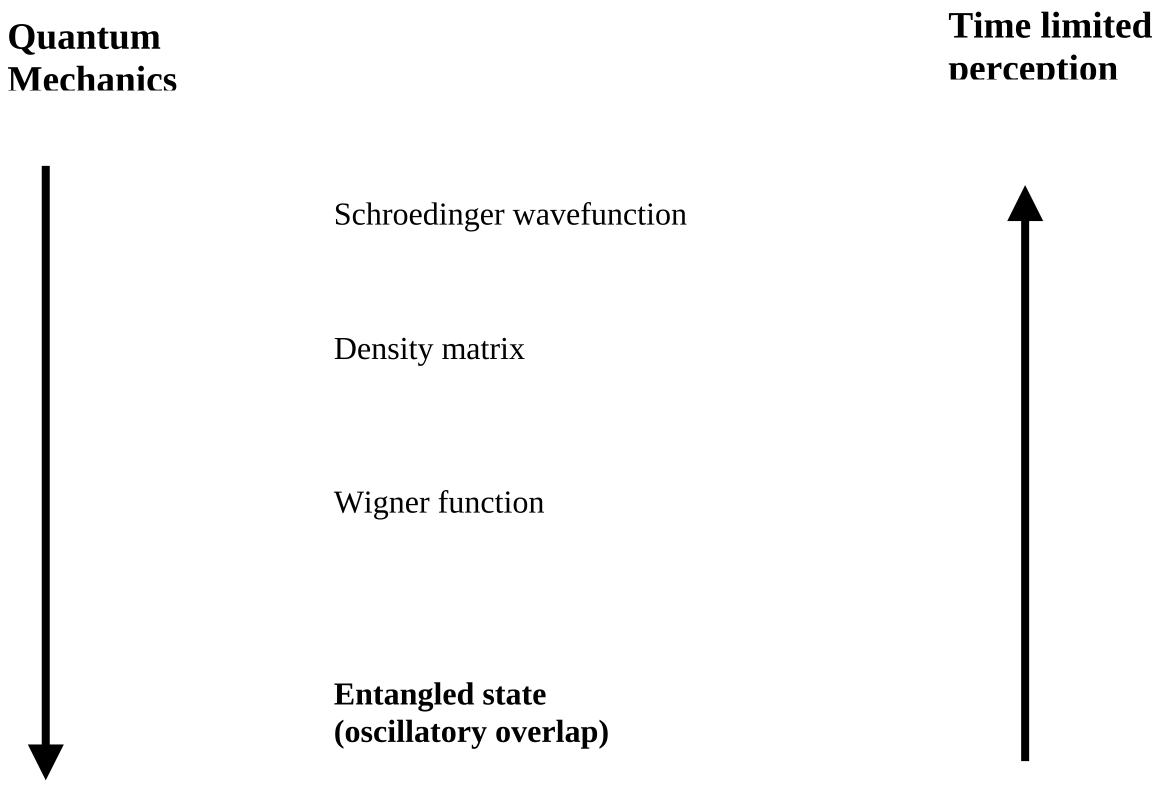

Fig.5 Direction of the logical processes which lead from wavefunction to entangled states or viceversa