Chaotic neuron dynamics, synchronization and feature binding

F.T. Arecchi*

Dipartimento di Fisica, Università di Firenze, Largo E. Fermi 6, 50125 Firenze, Italy

Abstract

Neuroscience studies how a large collection of coupled neurons combines external data with internal memories into coherent patterns of meaning. Such a process is called "feature binding", insofar as the coherent patterns combine together features which are extracted separately by specialized cells, but which do not make sense as isolated items. A powerful conjecture, with experimental confirmation, is that feature binding implies the mutual synchronization of axonal spike trains in neurons which can be far away and yet contribute to a well defined perception by sharing the same time code.

Based on recent investigations of homoclinic chaotic systems, and how they mutually synchronize, a novel conjecture on the dynamics of the single neuron is formulated. Homoclinic chaos implies the recurrent return of the dynamical trajectory to a saddle focus, in whose neighbourhood the system susceptibility (response to an external perturbation) is very high and hence it is very easy to lock to an external stimulus. Thus homoclinic chaos appears as the easiest way to encode information by a train of equal spikes occurring at erratic times.

In conventional measurements we read the number indicated by a meter's pointer and assign to the measured object a set position corresponding to that number. On the contrary, a time code requires a decision time $\bar{T}$ sufficiently longer than the minimal interspike separation t_1, so that the total number of different set elements is related in some way to the size $\bar{T}/t_1$. In neuroscience it has been shown that $\bar{T} \simeq 200$ ms while $t_1 \simeq 3$ ms.

In a sensory layer of the brain neocortex an external stimulus spreads over a large assembly of neurons building up a collective state, thus synchronization of trains of different individual neurons is the basis of a coherent perception.

The percept space can be given a metric structure by introducing a distance measure. This distance is conjugate of the duration time in the sense that an uncertainty relation is associated with time limited perceptions. If the synchronized train is truncated at a time $\Delta T < \bar{T}$, then the corresponding identification of a percept P carries an uncertainty cloud ΔP. As two uncertainty clouds overlap, interference occurs; this is a *quantum* behavior. Thus the quantum formalism is not exclusively limited to microscopic phenomena, but here it is conjectured to be the appropriate description of truncated perceptions.

---
* Fax: +39-055-2337755.
  E-mail address: arecchi@ino.it (F.T. Arecchi).

This quantum feature is not related to Planck's action but to the details of the perceptual chain.

1. Homoclinic chaos and its biological relevance

Homoclinic chaos (HC) appears as the optimal strategy for a time code shared by a large crowd of identical coupled objects. Indeed HC provides at each pseudo-period (or inter spike interval = ISI) the alternation of a regular large spike (a) and a small chaotic background (b). (b) is the sensitive region where the activation from neighbors occurs, while the spike (a) provides a suitable signal to activate the coupling. Whence, a chain of weakly coupled objects of this kind will easily synchronize, reaching a state common to all sites.

An easy explanation of such a behavior is offered by the experimental results presented in Fig. 1, showing the alternation of large spikes and a small chaotic background. Rather than the time series (Fig. 1a, b) it is convenient to build the three dimensional (3D) phase space trajectory by an embedding technique (Fig. 1c). For pedagogical reasons, we extract the schematic behavior in Fig. 2. In the course of the orbit, the phase point approaches the saddle focus S with a contraction rate α and then escapes from S with an expansion rate γ. In 3D the escape rate may be complex $\gamma \pm i\omega$ providing oscillations. Whenever $\gamma > \alpha$ we have the so called Shilnikov chaos [1,2] consisting of almost identical peaks P separated by a chaotic background. In that background, the proximity to the saddle focus S implies a large susceptibility, that is, a large sensitivity to a forcing stimulus, thus making HC prone to lock to external signals. Fig. 2 lists some situations in which this locking is useful, they will be discussed in the next section.

Thus the *first key point* of this paper is that an array of weakly coupled HC systems represents the simplest model for a physical realization of feature binding. Various experiments [3,4] prove that a decision is taken after 200 ms of exposure to a sensory stimulus. As we elaborate in Section 5, synchronizing over 200 ms trains of spikes with a minimal interspike separation of 3 ms and an average repetition of 25 ms (the so called gamma band in the EEG jargon) corresponds to a repertoire of about 10^{11} different words. $\bar{T} \simeq 200$ ms seems to be the time necessary to localize an element in the set of 10^{11} virtual sensations and whence activate some action.

How to make use of the synchronization pattern: assume that a sensory layer (visual, auditory, etc.) receives an external stimulus; it is pointless to recur to an hypothetical "*homunculus*" sitting within the brain, reading the synchronization state and acting consequently, activating motor actions as escaping from a predator, catching a prey, or speaking about [5]. We propose a simple model based on available brain components, consisting of coding neurons as described above and a reading system which localizes
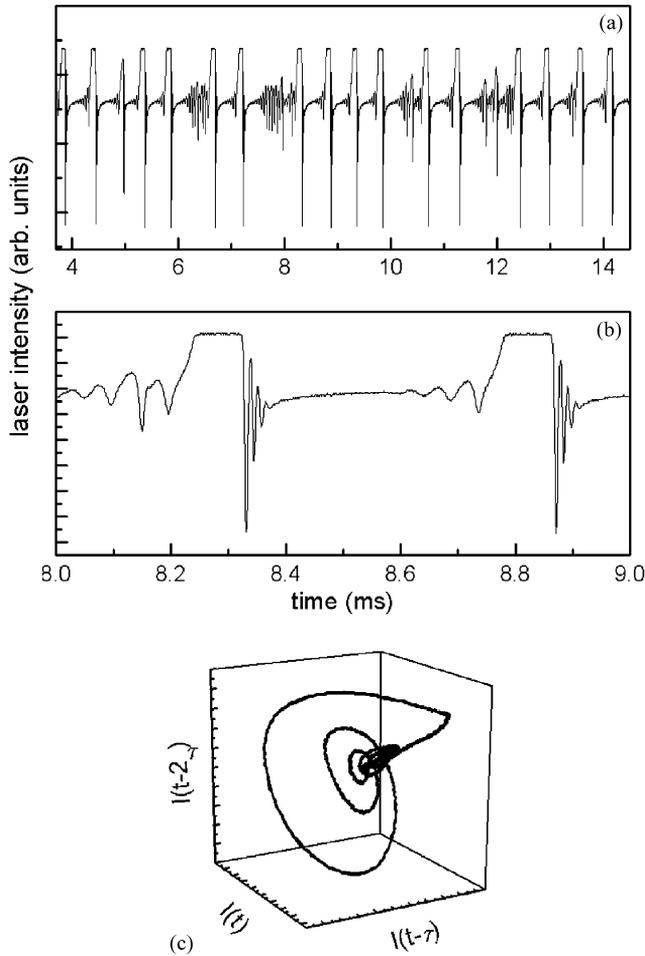
Fig. 1. (a) Experimental time series of the laser intensity for a CO_2 laser with feedback in the regime of homoclinic chaos. (b) Time expansion of a single orbit. (c) Phase space trajectory built by an embedding technique with appropriate delays (from Ref. [8]).

the boundaries of the synchronized domains and decides which is the prominent one, based on a "*majority rule*".

The appropriate algorithm for extracting the time code is the Wigner function [6] which was introduced in 1932 in quantum mechanics. It compares a signal with a shifted version of itself and sums over all shifts with a phase factor which yields a Fourier transform. The time code is activated on the largest synchronized pattern established on the sensory layer; if there are different stimuli the time coding must be preceded by a space coding which evaluates which one is the winning signal. This space coding is also reasonably performed by a space Wigner function, as discussed later below.
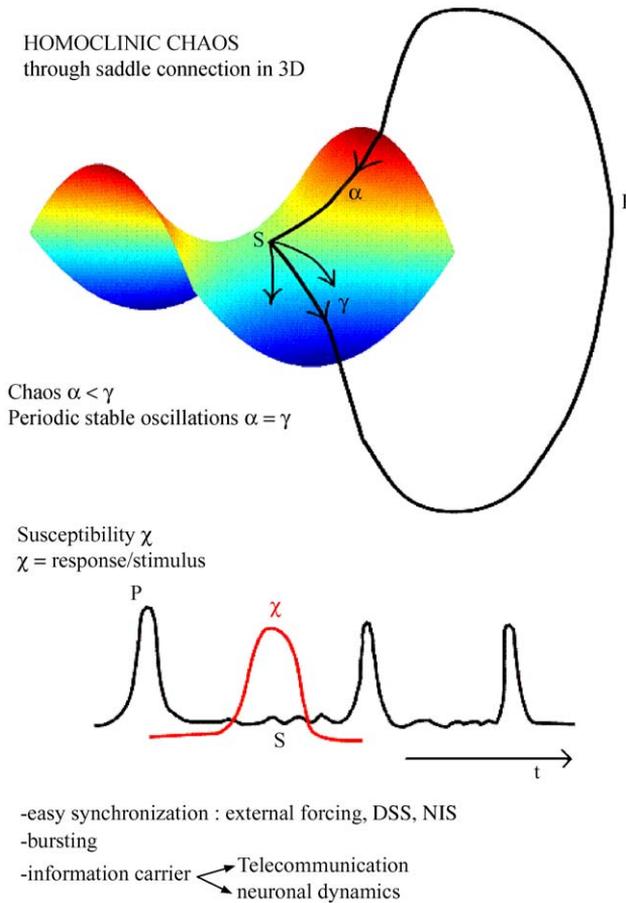
Fig. 2. Schematic view of the phase space trajectory approaching the saddle S and escaping from it. Chaos is due to the shorter or longer permanence around S; from a geometrical point of view most of the orbit P provides a regular spike.

Based on the above considerations, we introduce the *second key point* of percept formation, namely, a quantum limitation in information encoding/decoding through spike trains, whenever the processing session is interrupted. In fact, 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 an uncertainty ΔP emerges in the percept space P. This is related to the finite duration ΔT allotted for the code processing by the uncertainty relation

$$\Delta P \cdot \Delta T \geqslant C \,,$$

where C represents a quantum constraint on the coding. This means 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.

In conclusion, we will elaborate the following points:

(1) Chaos, precisely HC, provides an efficient time code for classifying different percepts.
(2) The code is read via a Wigner distribution; this provides quantum interference between different percepts.
(3) The physiological decision time $\bar{T} \simeq 200$ ms probably corresponds to deleting the interference; it acts as the *decoherence* time in quantum mechanics and a quantum effect can be seen when the perceptual task is truncated to $\Delta T < \bar{T}$.

2. Synchronization of homoclinic chaos

A wide class of sensory neurons demonstrates homoclinic chaotic spiking activity [7].

Precisely, a saddle focus instability separates in parameter space an *excitable* region, where axons are silent (Fig. 4b), from a *periodic* region, where the spike train is periodic (equal interspike intervals (Fig. 4a). If a control parameter is tuned at the saddle focus, the corresponding dynamical behavior (homoclinic chaos) consists of a frequent return to the instability [8]. This manifests as a train of geometrically identical spikes, which however occur at erratic times (chaotic interspike intervals). Around the saddle focus 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 high frequency noise as discussed later.

Such a type of dynamics has been recently dealt with in a series of reports that here I recapitulate as the following chain of linked facts.

(1) A single spike in a 3D dynamics corresponds to a quasi-homoclinic trajectory around a saddle focus SF; the trajectory leaves the saddle and returns to it (Figs. 1 and 2). We say "quasi-homoclinic" because, in order to stabilize the trajectory away from SF, a second fixed point, namely a saddle node SN, is necessary to assure a heteroclinic connection.
(2) The HC reported in Figs. 1 and 2 has a localized chaotic tangle surrounded by an island of stability (Fig. 3 left). This provides a sensitivity region corresponding to the chaotic tangle and an active region corresponding to the identical spikes. Such a behavior is crucial to couple large arrays of HC systems, as it occurs in the neuron coupling in the brain [9]. On the contrary, standard chaos as Lorenz where the chaotic behavior fills the whole attractor is inconvenient for synchronization purposes (Fig. 3 right).
(3) A train of spikes corresponds to the sequential return to, and escape from, the SF. A control parameter can be fixed at a value B_C for which this return is erratic (chaotic interspike interval) even though there is a finite average frequency. As the control parameter is set above or below B_C, the system moves from *excitable* (single spike triggered by an input signal) to *periodic* (yielding a regular sequence of
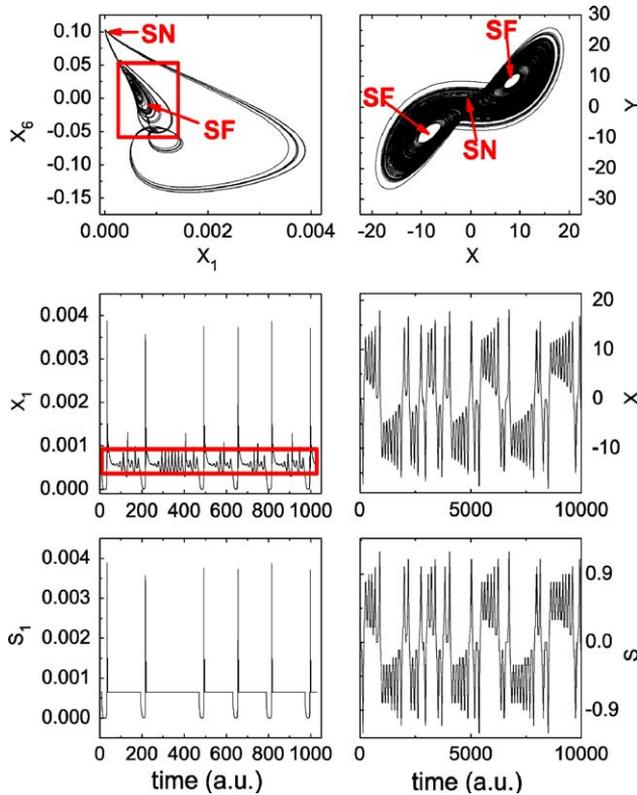
Fig. 3. Comparison between HC (homoclinic chaos) (left column) and Lorenz chaos (right column). The top row is a phase space projection over two dynamical variables; SF denotes the saddle focus and SN the saddle node, on the left the two SF map one onto the other after an inversion ($x \rightarrow -x$, $y \rightarrow -y$) around the origin. The intermediate row shows the time series for variables $x1$ of HC (it represents the laser intensity in the case of the CO_2 laser) and x of Lorenz; in the former case, a suitable threshold cuts on the chaotic background; in the latter case, no convenient region for thresholding can be isolated. Bottom row, left: after threshold, the new variable $S(t)$ alternates spikes with flat regions where the systems has a high sensitivity and short refractive windows where the intensity $S(t)$ goes to zero.

   spikes without need for an input), with a frequency monotonically increasing with
   the separation ΔB from B_C (Fig. 4). A low frequency modulation of B around B_C
   provides alternation of silent intervals with periodic bursts. Such bursting behavior,
   peculiar of neuron on central pattern generators (CPG) as the cardio-respiratory
   system, has been modeled by a laser experiment [10].
(4) Around SF, any tiny disturbance provides a large response. Thus the homoclinic
   spike trains can be synchronized by a periodic sequence of small disturbances
   (Fig. 5). However, each disturbance has to be applied for a minimal time, be-
   low which it is no longer effective; this means that the system is insensitive
   to broadband noise, which is a random collection of fast positive and negative
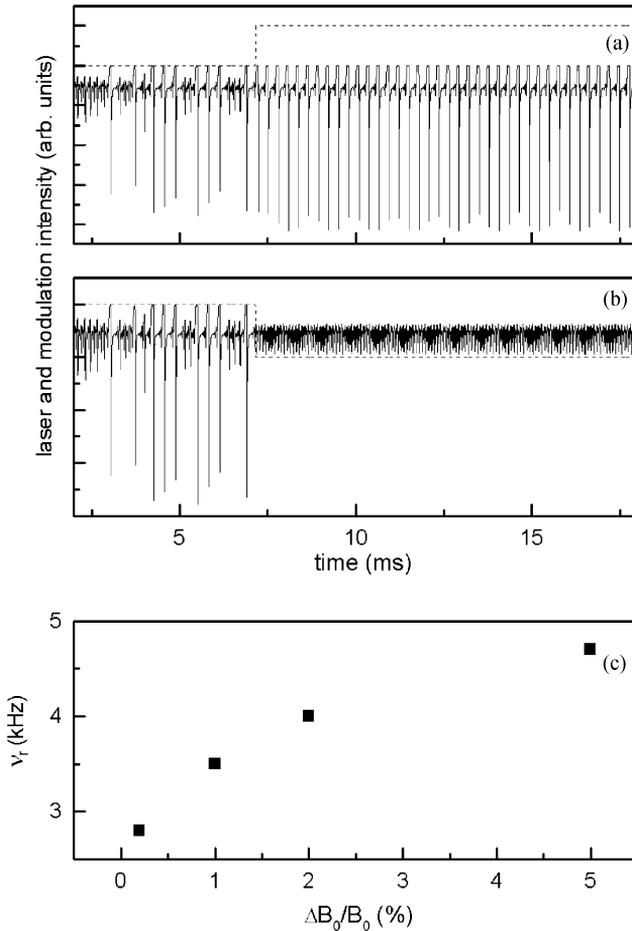   signals [11].

Fig. 4. Stepwise increase (a) and decrease (b) of control parameter $B0$ by $\pm 1\%$ brings the system from homoclinic to periodic or excitable behavior. (c) In case (a) the frequency ν_r of the spikes increases monotonically with $\Delta B0$ (from Ref. [10]).

(5) The above considerations lay the floor for the use of mutual synchronization as the most convenient way to let different neurons respond coherently to the same stimulus, organizing as a space pattern. In the case of a single dynamical system, it can be fed back by its own delayed signal (DSS = delayed self synchronization). As the delay is long enough the system is decorrelated with itself and this is equivalent to feeding an independent system. This process allows to store meaningful sequences of spikes as necessary for a short term memory [12].

(6) A dynamical sketch of a single neuron is presented in Fig. 6 and its HC behavior is discussed, e.g in Ref. [7]. The feature binding conjecture, based on synchronization of different neurons exposed to the same image, is shown in Fig. 7 [13] and the
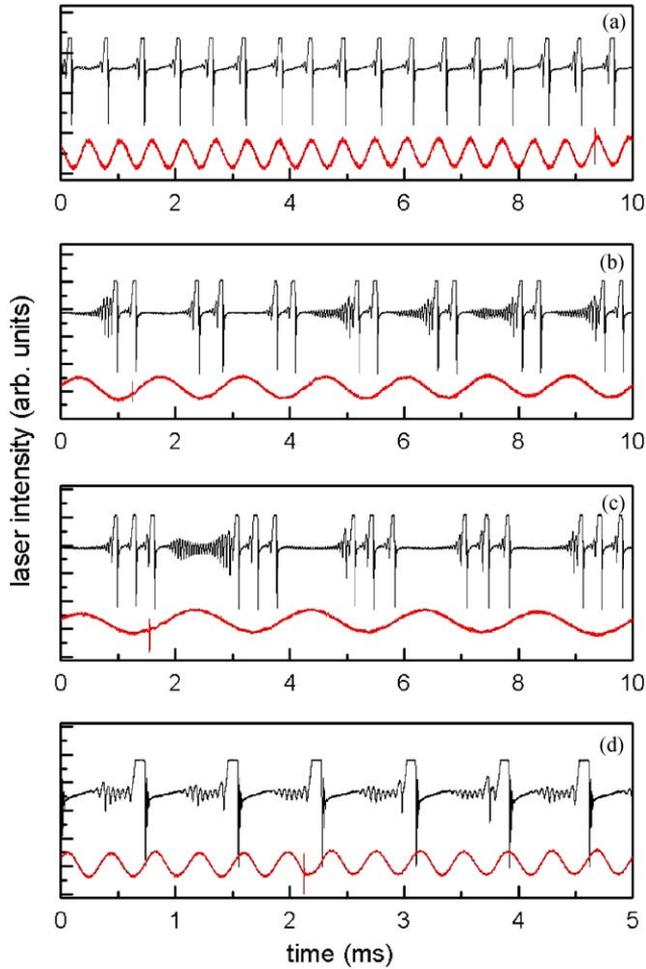
Fig. 5. Experimental time series for different synchronization induced by periodic changes of the control parameter. (a) 1:1 locking, (b) 1:2, (c) 1:3, (d) 2:1 (from Ref. [8]).

role of threshold resetting due to past memories (ART=adaptive resonance theory [14]) is illustrated in Fig. 8.

(7) In the presence of localized stimuli over a few neurons, the corresponding disturbances propagate by inter-neuron coupling (either excitatory or inhibitory); a synchronized pattern is uniquely associated with each stimulus; different patterns compete and we conjecture that the resulting sensory response, which then triggers motor actions, corresponds by a majority rule to that pattern which has extended over the largest cortical domain. An example is discussed for two inputs to a one-dimensional array of coupled HC systems [15].
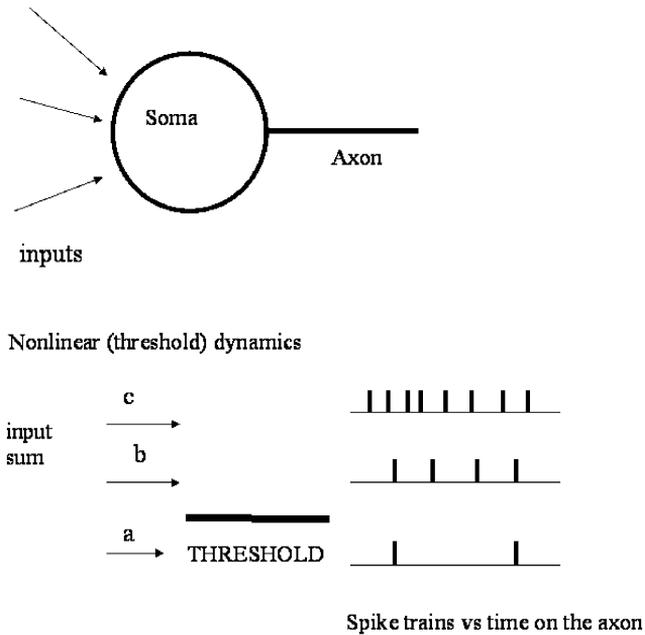
Fig. 6. Dynamical behavior of a single neuron. The sum of the inputs compares with a threshold level; if it is below (a) only a few noisy spikes occur as electrical activity on the axon; if it is just above (b) we have a periodic spike train, whose frequency increases as the input goes high above threshold (c).

These facts have been established experimentally and confirmed by a convenient model in the case of a class B laser [1] with a feedback loop which readjusts the amount of losses depending on the value of the light intensity output [2].

The above listed facts hold in general for any dynamical system which has a 3D sub-manifold separating a region of excitability from a region of periodic oscillations: indeed, this separatrix has to be a saddle focus. If a dynamical system has $N > 3$ dimensions, then a local stability analysis shows that around a saddle focus the systems behaves locally as 3D, since N-3 variation rates are largely negative.

In particular, all biological clocks are of HC type in order to have an adaptive period which can be re-adjusted whenever necessary (think e.g. of the cardiac pace-maker).

---

[1] I recall the classification widely accepted in laser physics. Class A lasers are ruled by a single order parameter, the amplitude of the laser field, which obeys a closed dynamical equation; all the other variables having much faster decay rate, thus adjusting almost instantly to the local field value. Class B lasers are ruled by two order parameters, the laser field and the material energy storage providing gain; the two degrees of freedom having comparable characteristic times and behaving as activator and inhibitor in chemical dynamics [1].
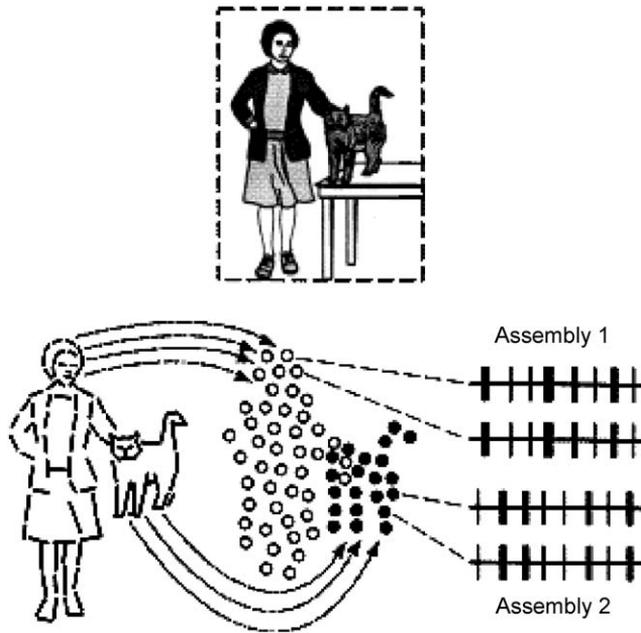
Fig. 7. Feature binding: the lady and the cat are, respectively, represented by the mosaic of empty and filled circles, each one representing the receptive field of a neuron group in the visual cortex. Within each circle the processing refers to a specific detail (e.g. contour orientation). The relations between details are coded by the temporal correlation among neurons, as shown by the same sequences of electrical pulses for two filled circles or two empty circles. Neurons referring to the same individual (e.g. the cat) have synchronous discharges, whereas their spikes are uncorrelated with those referring to another individual (the lady) (from Ref. [13]).

3. Perceptions and feature binding

Let us return to the visual system; the role of elementary feature detectors has been extensively studied in the past decades [16]. By now we know that some neurons are specialized in detecting exclusively vertical or horizontal bars, or a specific luminance contrast, etc. However, the problem arises: how elementary detectors contribute to a holistic (*Gestalt*) perception? A hint is provided by Ref. [13]. Suppose, we are exposed to a visual field containing two separate objects. Both objects are made of the same visual elements, horizontal and vertical contour bars, different degrees of luminance, etc. What are then the neural correlates of the identification of the two objects? We have one million fibers connecting the retina to the visual cortex. Each fiber results from the merging of approximately 100 retinal detectors (rods and cones) and as a result it has its own receptive field. Each receptive field isolates a specific detail of an object, e.g. a vertical bar). We thus split an image into a mosaic of adjacent receptive fields.
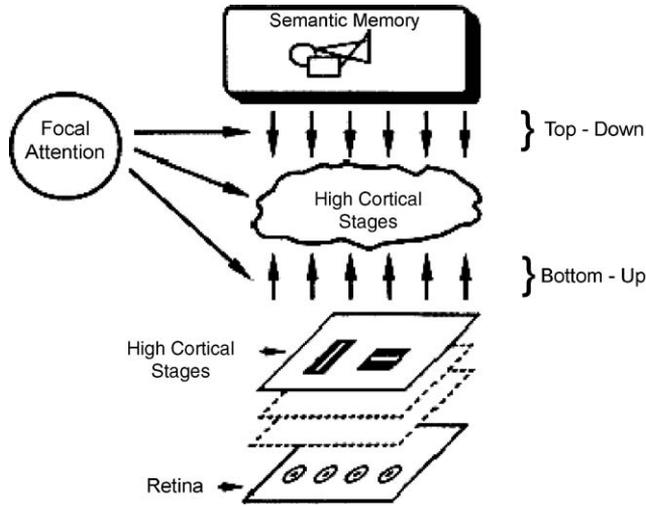
Fig. 8. *ART* = adaptive resonance theory. Role of bottom–up stimuli from the early visual stages an top –down signals due to expectations formulated by the semantic memory. The focal attention assures the matching (resonance) between the two streams (from Ref. [26]).

Now the "feature binding" hypothesis consists of assuming that all the cortical neurons whose receptive fields are pointing to a specific object synchronize the corresponding spikes, and as a consequence the visual cortex organizes into separate neuron groups oscillating on two distinct spike trains for the two objects.

Direct experimental evidence of this synchronization is obtained by insertion of microelectrodes in the cortical tissue of animals just sensing the single neuron (Fig. 6) [13]. Indirect evidence of synchronization has been reached for human beings as well, by processing the EEG (electro-encephalo-gram) data [4].

The advantage of such a temporal coding scheme, as compared to traditional rate based codes, which are sensitive to the average pulse rate over a time interval and which have been exploited in communication engineering, has been discussed in a recent paper [17].

Based on the neurodynamical facts reported above, we can understand how this occurs [14]. The higher cortical stages where synchronization takes place have two inputs. One (bottom–up) comes from the sensory detectors via the early stages which classify elementary features. This single input is insufficient, because it would provide the same signal for, e.g. horizontal bars belonging indifferently to either one of the two objects. However, as we said already, each neuron is a nonlinear system passing close to a saddle point, and the application of a suitable perturbation can stretch or shrink the interval of time spent around the saddle, and thus lengthen or shorten the interspike interval. The perturbation consists of top–down signals corresponding to conjectures made by the semantic memory (Fig. 7).

In other words, the perception process is not like the passive imprinting of a camera film, but it is an active process whereby the external stimuli are interpreted in terms

of past memories. A focal attention mechanism assures that a matching is eventually reached. This matching consists of resonant or coherent behavior between bottom–up and top–down signals. If matching does not occur, different memories are tried, until the matching is realized. In presence of a fully new image without memorized correlates, then the brain has to accept the fact that it is exposed to a new experience.

Notice the advantage of this time-dependent use of neurons, which become available to be active in different perceptions at different times, as compared to the computer paradigm of fixed memory elements which store a specific object and are not available for others (the so called "grandmother neuron" hypothesis).

We have above presented qualitative reasons why the degree of synchronization represents the perceptual salience of an object. Synchronization of neurons located even far away from each other yields a space pattern on the sensory cortex, which can be as wide as a few square millimeters, involving millions of neurons. The winning pattern is determined by dynamic competition (the so-called "winner takes all" dynamics).

This model has an early formulation in ART and has been later substantiated by the synchronization mechanisms. Perceptual knowledge appears as a complex self-organizing process.

4. The role of duration T in perception: 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. Notice however that, due to the above consideration on qualia, the experiments suggested in this section hold for purely a-semantic perceptions; by this we mean the task of detecting rather elementary features which would not trigger our categorization skills.

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 pre-frontal cortex (PFC) 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, motor actions are started, including linguistic utterances [4].

Sticking to the neurodynamical level, and leaving 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 ms, that the minimal interspike separation is 3 ms, and that the decision time at the PCF level is estimated as $\bar{T} \simeq 200$ ms, we can split $\bar{T}$ into $200/3 = 66.6$ bins of 3 ms duration, which are designated by 1 or 0 depending on whether they have a spike or not. Thus the a priori total number of different messages which can be transmitted is

$$2^{66} \simeq 6 \times 10^{19} \,.$$

However, we must account also for the average rate at which spikes proceed in our brain, which is $r = 40$ Hz (so called γ band, average ISI $= 25$ ms). When we account

for this rate we can evaluate a reduction factor $\alpha = S/\bar{T} = 0.54$, where S is an entropy [18–20], thus there are roughly $2^s \simeq 10^{11}$ words with significant probability. Even though this number is large, we are still within a finitistic realm. Provided we have time enough to ascertain which one of the 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 200 ms. 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 $\bar{T}$. This means that the brain cannot decide among *all* coded perceptions having the same structure up to ΔT, but different afterwards.

Whenever we stop the perceptual task at a time ΔT shorter than the total time T, then the bin stretch $T - \Delta T$ (from now on we measure the times in bin units) is not explored. This means that all stimuli which provide equal spike sequences up to ΔT, and differ afterwards will cover an uncertainty region ΔP whose size is given by

$$\Delta P = 2^{\alpha T} 2^{-\alpha \Delta T} = P_M \exp(-\alpha \Delta T \ln 2) , \tag{1}$$

where $P_M \simeq 10^{11}$ is the maximum perceptual size available with the chosen $\bar{T} \simeq 66.6$ bins per perceptual session and rate $r = 40$ Hz. Relation (1) is very different from the standard uncertainty relation

$$\Delta P \times \Delta T = C , \tag{2}$$

that we would expect in a word-bin space ruled by Fourier transform relations.

Indeed, the transcendental equation (1) is more rapidly converging at short and long ΔT than hyperbola (2). We fit (1) by (2) in the neighborhood of a small uncertainty $\Delta P = 10$ words, which corresponds to $\Delta T = 62$ bins. Around $\Delta T = 62$ bins the local uncertainty (2) yields a quantum constant

$$C = 10 \times 62 = 620 \text{ words} \cdot \text{bins} . \tag{3}$$

To convert C into J_s as Planck's h, consider that:

- 1 bin = 3 ms;
- in order to jump from an attractor corresponding to one perception to a nearby one, a *minimal amount* of energy is needed; we conjecture that this roughly $10k_B T$, that is, neither too small as $k_B T$, otherwise jumps would occur spontaneously at room temperature, nor too high as 1 eV, otherwise our brain would need a large amount of energy per perception. For 10^8 neurons in the visual cortex this makes 4×10^{-12} J. Thus, we estimate

$$C = 6 \times 10^{-12} \ J_s = 10^{22} h . \tag{4}$$

The Wigner function formalism derives from a Schroedinger wavefunction treatment for a pure state, and corresponding density matrix for mixed states.

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

The uncertainty relation $\Delta P \cdot \Delta T \geqslant C$ forbids a partition of the (P,T) space into sets only if 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. 10 as fully equivalent to isolevel cuts of a Wigner function. Hence we can consider interference of distinct states as in quantum physics.

This means that 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 a universal computer machine.

The evidence of interfering perceptions 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 [21].

Among other things, the characterization of the "concept" or "category" as the limit of a recursive 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 [22] 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 at the $\hbar$ level in biological processes is not plausible, 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, and at sub-picosecond time ranges, not relevant for synchronization purposes in the 10–100 ms range.

Our tenet is that the quantum C-level in a living being emerges from the limited time available in order to take vital decisions; it is logically based on a non-commutative set of relevant variables and hence it requires the logical machinery built for the $\hbar$ quantum description of the microscopic world where non-commutativity emerges from use of variables coming from macroscopic experience, as coordinate and momenta, to account for new facts. The recent debate [23] has clearly excluded the need for hidden variables and hence the possibility of considering the quantum formalism as a projection of a more general classical one. Similarly, in our case, the fact that decisions are taken on, say, a 100 ms sequence of neuronal spikes, excludes a cognitive world where longer decision times (say, 300 ms) would disentangle overlapping states. In other words, the universe of perceptual 100 ms events would be correctly described by a C-quantum formalism, since the 300 ms perceptual facts would be considered not accessible in principle.
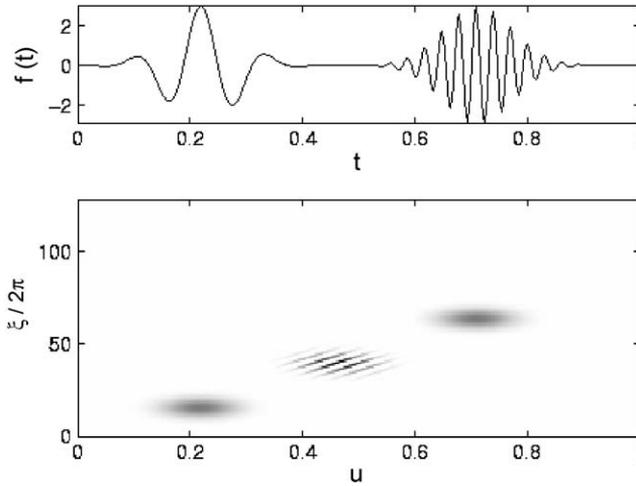
Fig. 9. Wigner distribution of two localized sinusoidal packets shown at the top. The oscillating interference is centered at the middle time–frequency location [24].

The qualitative uncertainty considerations have been taken as "quantum" by postulating an interference within the overlapping region. A more precise consideration consists in classifying the spike trains. Precisely, if we have a sequence of identical spikes of unit area localized at erratic time positions τ_1 then the whole sequence is represented by

$$f(t) = \sum_l \delta(t - \tau_l) \,, \tag{5}$$

where $\{\tau_l\}$ is the set of position of the spikes. A temporal code, based on the mutual position of successive spikes, depends on the moments of the interspike interval distributions

$$(\text{ISI})_l = \{\tau_l - \tau_{l-1}\} \,. \tag{6}$$

Different ISIs encode different sensory information.

A time ordering within the sequence (5) is established by comparing the overlap of two signals as (5) mutually shifted in time. Weighting all shifts with a phase factor and summing up, this amounts to constructing a Wigner function [6,24]

$$W(t, \omega) = \int_{-\infty}^{+\infty} f(t + \tau/2) f(t - \tau/2) \exp(i\omega\tau) \, d\tau \,. \tag{7}$$

If now f is the sum of two packets $f = f_1 + f_2$ as in Fig. 9, the frequency–time plot displays an intermediate interference. Equation (7) would provide interference whatever is the time separation between f_1 and f_2 (see Fig. 10).

In fact, we know that the decision time $\bar{T}$ truncates a perceptual task, thus we must introduce a cutoff function $g(\tau) \simeq \exp(-\tau^2/\bar{T}^2)$ which transforms the Wigner
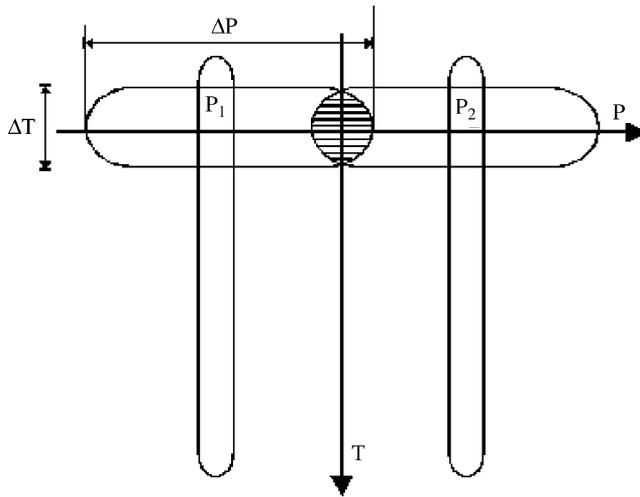
Fig. 10. Uncertainty areas of two perceptions P_1 and P_2 for two different durations of the spike trains. In the case of short ΔT, the overlap region is represented by a Wigner function with strong positive and negative oscillations which go as $\cos((\Delta P/C)T)$ along the T axis; therefore with a frequency given by the ratio of the percept separation $\Delta P = P_2 - P_1$ to the perceptual "Planck's constant" C.

function as

$$W(t,\omega) = \int_{-\infty}^{+\infty} f(t+\tau/2)f(t-\tau/2)\exp(i\omega\tau)g(\tau)\,d\tau\ . \tag{8}$$

In the quantum jargon, the brain physiology breaks the quantum interference for $t > \bar{T}$ (*decoherence* phenomenon).

5. Who reads the information encoded in a synchronized neuron array: the role of the Wigner function in brain operations

We have seen that feature binding in perceptual tasks implies the mutual synchronization of axonal spike trains in neurons which can be even far away and yet contribute to a well defined perception by sharing the same pattern of spike sequence.

The synchronization conjecture, formulated by many people as, e.g. Von der Marlsburg [25] in 1976, was later given direct experimental evidence by inserting several micro-electrodes probing each one single neuron in the cortex of cats and then studying the temporal correlation in response to specific visual inputs [13]. In the human case, indirect evidence is acquired by exposing a subject to transient patterns and reporting the time–frequency plots of the EEG signals [4]. Even though the space resolution is poor, phase relations among EEG signals coming from different cerebral areas at different times provide an indirect evidence of the synchronization mechanism.

Recently the dynamics of HC has been explored with reference to some laboratory phenomena observed in lasers and then explored in its mathematical aspects, which

display strong analogies with the dynamics of many biological clocks, in particular that of a model neuron (Hodgkin and Huxley: see Ref. [7]). HC provides almost equal spikes occurring at variable time positions and presents a region of high sensitivity to external stimuli; perturbations arriving within the sensitivity window induce easily a synchronization, either to an external stimulus or to each other (mutual synchronization) in the case of an array of coupled HC individuals (from now on called neurons).

In view of the above facts, we can model the encoding of external information on a sensory cortical area (e.g. *V1* in the visual case) as a particular spike train assumed by an input neuron directly exposed to the arriving signal and then propagated by coupling through the array. As shown by the series of experiment on feature binding [13], we must transform the local time information provided by Eqs. (5) and (6) into a spatial information which tells the amount of a cortical area which is synchronized.

If many sites have synchronized by mutual coupling, then the read out problem consists in tracking the pattern of values (5), one for each site. Let us take for simplicity a continuous site index x.

In case of two or more different signals applied at different sites a competition starts and we conjecture that the winning information (that is, the one which is then channeled to a decision center to act consequentially) corresponds to a "*majority rule*". Precisely, if the encoding layer is a 1D chain of N coupled sites activated by external stimuli at the two ends ($i = 1$ and N), the majority rule says that the prevailing signal is that which has synchronized more sites.

The crucial question is then: who reads that information in order to decide upon? We can not recur to some homunculus who reads the synchronization state. Indeed, in order to be made of physical components, the homunculus itself should have a reading neuron layer as discussed above, followed by some interpreter which would be a new homunculus, and so on with a "regressio ad infinitum".

On the other hand, it is well known that, as we map the interconnections in the vision system, *V1* exits through the vertical stream and the dorsal stream toward the *inferotemporal cortex* and *parietal cortex*, respectively. The two streams contain a series of intermediate layers characterized by increasing receptive fields; hence they are cascades of layers where each one receives converging signals from two or more neurons of the previous layer. Let us take for the time being this feed-forward architecture as a network enabled to extract relevant information upon which to drive consequent actions.

We show how this cascade of layers can localize the interface between two domains corresponding to different synchronization. It is well known that Hubel and Wiesel simple cells with a center-surround configuration perform a *first and second space derivative*. Suppose this operation was done at certain layer. At the successive one, as the converging process goes on, two signals will converge on a simple cells which then performs *a higher order derivative*, and so on. This way, we build a power series of space derivatives. A translated function as $f(x + \xi)$ is then reconstructed by adding up many layers, as can be checked by a Taylor expansion.

Notice that the alternative of exploring different neighborhoods ξ of x by varying ξ would imply a moving pointer to be set sequentially at different positions, and there is nothing like that in our physiology.

The next step consists in comparing the function $f(x + \xi)$ with a suitable standard, to decide upon its value. Since there are no metrological standards embedded in a living brain, such a comparison must be done by comparing f with a shifted version of itself, something like the product

$$f(x + \xi)f(x - \xi) .$$

Such a product can naturally be performed by converging the two signals $f(x \pm \xi)$ onto the same neuron, exploiting the nonlinear (Hebbian) response characteristic limited to the lowest quadratic nonlinearity, thus taking the square of the sum of the two converging inputs and isolating the double product.

This operation is completed by summing up the different contributions corresponding to different ξ, with a kernel which keeps track of the scanning over different ξ, keeping information on different domain sizes. If this kernel were just a constant, then we would retrieve a trivial average which cancels the ξ information.

Without loosing in generality, we adopt a Fourier kernel $\exp(ik\xi)$ and hence have built the quantity

$$W(x, k) = \int_{-\infty}^{+\infty} f(x - \xi)f(x + \xi) \exp(ik\xi)\,d\xi .\tag{9}$$

$W(x, k)$ is thus a Fourier transform with respect to ξ. It is called Wigner–Ville function. It contains information on both the space position x around which we are exploring the local behavior, as well as the frequency k which is associated with the space resolution.

As well known, it contains the most complete information compatible with the Fourier uncertainty

$$\Delta x \Delta k \geqslant 1 .\tag{10}$$

Notice that building a joint information on locality (x) and resolution (k) by physical measuring operations implies such an intrinsic limitation.

In summary, it appears that the Wigner function is the best read-out of a synchronized layer that can be done by exploiting natural machinery, rather than recurring to a homunculus. The local value of the Wigner function represents a decision to be sent to motor areas triggering a suitable action.

Now, let us explore the appropriate function $f(x)$ whose space pattern represents the relevant information stored in a sensitive layer.

The pending problem is how to correlate space and time information. When many sites are coupled, there is a fixed time step t^* between the occurrence of spikes at adjacent sites. Even without a counting for axonal time lags, a delay is due to the fact the input form a previous site is sensed only when the driven system is within the saddle focus region. This implies a site to site delay, whose value depends on whether the coupling is positive or negative. In presence of a single input, the overall synchronization takes a time Nt_1, N being the site number. This correspond to a velocity

$$v = (t^*)^{-1} \text{ sites/s} .\tag{11}$$

This new feature which is peculiar of the Wigner function construction, is what in quantum mechanics is called "the Schroedinger cat paradox": if we have established

that the measurement consists in building up a Wigner function, then we must accept the besides the two patches associated with a live cat and a dead cat, in case of justified laboratory uncertainty there must be also an oscillating interference (Fig. 9). From 1935 to 1995 it had been considered a pseudo-philosophical problem; nowadays such extra term has evidence if the superposition refers to states of microscopic objects as atoms or photons rather than cats. By the way, its existence is crucial for *quantum computation*.

The reason why we do not see interference is "*decoherence*". Let us explain what we mean. If the system under observation is affected by the environment, then the two states of the superposition have a finite lifetime; however, even worse, the interference decays on a much shorter time (the smaller, the bigger is the separation between the two states): so while the separation is still manageable for the two polarization states of a photon, it becomes too big for the two states of a macroscopic quantum system. Precisely, if we call τ_i the intrinsic decay of each one of the two states of the superposition, then the mutual interference in the Wigner function decays with the so-called decoherence time

$$\tau_{dec} = \frac{\tau_i}{D^2} , \tag{12}$$

where D^2 is the square of the separation D in phase space between the centers of the Wigner functions of the two separate states. Notice that, in microscopic physics, D^2 *is measured in units of $\hbar$*. Usually $D^2/\hbar$ is much bigger than unity for a macroscopic system and hence τ_{dec} is so short that any reasonable observation time is too long to detect a coherent superposition. Instead, in neurodynamics, we perform measurement operations which are limited by the intrinsic sizes of space and time resolutions peculiar of brain processes. The associated uncertainty constant C is such that it is very easy to have a relation as (9) with D^2/C comparable to unity, and hence superposition lifetimes comparable to times of standard neural processes. This implies the conjecture that for short times or close cortical domains a massive parallelism typical of quantum computation is possible. The threshold readjustment due to expectations arising from past memory acts as an environmental disturbance, which orthogonalizes different neural states, destroying parallelism.

References

[1] F.T. Arecchi, Instabilities & chaos in single mode homogeneous line lasers, in: F.T. Arecchi, R.G. Harrison (Eds.), Instabilities and Chaos in Quantum Optics, Springer Series Synergetics, Vol. 34, Springer, Berlin, 1987, pp. 9–48.

[2] F.T. Arecchi, R. Meucci, W. Gadomski, Phys. Rev. Lett. 58 (1987) 2205.

[3] B. Libet, E.W. Wright, B. Feinstein, D.K. Pearl, Subjective referral of the timing for a conscious sensory experience, Brain 102 (1979) 193.

[4] E. Rodriguez, N. George, J.P Lachaux, J. Martinerie, B. Renault, F. Varela, Nature 397 (1999) 340–343.

[5] K. MacLeod, A. Backer, G. Laurent, Nature 395 (1998) 693–698.

[6] E.P. Wigner, On the quantum correction for thermodynamic equilibrium, Phys. Rev. 40 (1932) 749.

[7] E.M. Izhikevich, Int. J. Bifurcation Chaos 10 (2000) 1171.

[8] E. Allaria, F.T. Arecchi, A. Di Garbo, R. Meucci, Phys. Rev. Lett. 86 (2001) 791.

[9] F.T. Arecchi, E. Allaria, E., I. Leyva, Phys. Rev. Lett. 91 (2003) 234101.
[10] R. Meucci, A. Di Garbo, E. Allaria, F.T. Arecchi, Phys. Rev. Lett. 88 (2002) 144101.
[11] C.S. Zhou, J. Kurths, E. Allaria, S. Boccaletti, R. Meucci, F.T. Arecchi, Phys. Rev. E 67 (2003) 015205, 066220.
[12] F.T. Arecchi, R. Meucci, E. Allaria, A. Di Garbo, L.S. Tsimring, Phys. Rev. E 65 (2002) 046237.
[13] W. Singer, E.C.M. Gray, Annu. Rev. Neurosci. 18 (1995) 555.
[14] S. Grossberg, The attentive brain, Am. Sci. 83 (1995) 439.
[15] I. Leyva, E. Allaria, S. Boccaletti, F.T. Arecchi, Phys. Rev. E 68 (2003) 066209.
[16] D.H. Hubel, Eye, Brain and Vision, Scientific American Library, No. 22, W.H. Freeman, New York, 1995.
[17] W. Softky, Curr. Opinions Neurobiol. 5 (1995) 239.
[18] F. Rieke, D. Warland, R. de Ruyter van Steveninck, W. Bialek, Spikes: Exploring the Neural Code, MIT Press, Cambridge MA, 1997.
[19] S.P. Strong, et al., Phys. Rev. Lett. 80 (1998) 197.
[20] D. MacKay, W.S. McCulloch, Bull. Math. Biophys. 14 (1952) 127.
[21] A. Turing, Computing machinery and intelligence, Mind 59 (1950) 433.
[22] R. Penrose, Shadows of the Mind, Oxford University Press, New York, 1994.
[23] R. Omnès, The interpretation of Quantum Mechanics, Princeton University Press, Princeton, NJ, 1994.
[24] S. Mallat, A wavelet tour of signal processing, Academic Press, San Diego, CA, 1999.
[25] C. von der Malsburg, The correlation theory of brain function, reprinted in E. Domani, J.L. Van Hemmen, K. Schulten (Eds.), Models of Neural Networks II, Springer, Berlin, 1981.
[26] B. Julesz, Rev. Mod. Phys. 63 (1991) 735–772.