

# Multi-level Synergetic Computation in Brain

Mitja Peruš

*BION Institute, Stegne 21, SI-1000 Ljubljana, Slovenia*

E-mail: mitja.perus@uni-lj.si

(Received 16 November 1999)

Patterns of activities of neurons serve as attractors, since they are those neuronal configurations which correspond to minimal 'free energy' of the whole system. Namely, they realize maximal possible agreement among constitutive neurons and are most-strongly correlated with some environmental pattern. Neuronal patterns-qua-attractors have both a material and a virtual aspect. As neuronal patterns, on the one hand, patterns-qua-attractors are explicit carriers of informational contents. As attractors, on the other hand, patterns-qua-attractors are implicit mental representations which acquire a meaning in contextual relations to other possible patterns.

Recognition of an external pattern is explained as a (re)construction of the pattern which is the most relevant and similar to a given environmental pattern. The identity of the processes of pattern construction, re-construction and Hebbian short-term storage is realized in a net.

Perceptual processes are here modeled using Kohonen's topology-preserving feature mapping onto cortex where further associative processing is continued. To model stratification of associative processing because of influence from higher brain areas, Haken's multi-level synergetic network is found to be appropriate.

The hierarchy of brain processes is of "software"-type, i.e. virtual, as well as it is of "hardware"-type, i.e. physiological. It is shown that synergetic and attractor dynamics can characterize not only neural networks, but also underlying quantum networks. Neural nets are alone not sufficient for consciousness, but interaction with the quantum level might provide effects necessary for consciousness, like, for instance, ultimate binding of perceptual features into an unified experience.

It is mathematically demonstrated that associative neural networks realize information processing analogous to the quantum dynamics. Parallels in the formalism of neural models and quantum theory are listed. Basic elements of the quantum versus neural system (modeled by formal neurons and connections) are very different, but their collective processes obey similar laws. Specifically, it is shown that neuron's weighted spatio-temporal integration of signals corresponds to the Feynman's version of the Schrödinger equation. In the first case weights are synaptic strengths determined by the Hebb or delta correlation rule; in the second case weights are Green functions or density matrices. In both cases encodings of pattern-correlations represent memory. (Re)construction of a neuronal pattern-qua-attractor is analogous to the "wave-function collapse". Transformations of memory (or sub-conscious) representations to a conscious representation is modeled in the same way.

Found mathematical analogies allow translation of the neural-net "algorithm", which in author's simulations works very well, into a quantum one. This indicates how such quantum networks, which might be exploited by the sub-cellular levels of brain, could process information efficiently and also make it conscious.

**Key words:** neural net, quantum, brain, associative, synergetic, perception, consciousness

**PACS numbers:** 87.18.Sn, 87.19.La, 03.67.Lx, 89.70.+c, 87.18.Hf

## Overall introduction

This paper is a review of models of multi-level brain processing with connectionist approach. Parts One

and Two present the topic in a new style, but the third part is a purely original contribution which indicates quantum neural-net-like information processing and their potential relevance for conscious

perception. The aim is that all parts together should present an integrated overview of modeling brain synergetics — at least from a particular (author's) viewpoint, although not exclusively.

In the *first part* a comprehensive introduction to models of associative neural networks is presented as well as their role in large-scale modeling of micro-cognition. Patterns of neuronal activity are interpreted as attractors of the parallel-distributed dynamics of an underlying complex dynamical system.

It will be shown how the system of synaptic connections represents memory. The transmission rates of individual synapses represent correlations between neurons, whereby each neuron may cooperate in global encoding of many patterns as prescribed by the Hebb learning rule. Learning is a process of adjusting synaptic strengths in order to store essential information.

In the *second part* processes of categorization, adaptation, data compression and abstraction, filtering of relevant information and extraction of novel information from stimuli, and perceptual pre-processing of patterns will be presented (now more in detail) using the Kohonen model of topologically-correct feature mapping and its extensions. This model is important for modeling perception and motor action. Instead of free-energy minimization of the previously-presented Hopfield model, in Kohonen's model these tasks are resolved by error minimization, lateral inhibition, and winning of one neuron. This is the cardinal neuron which individually encodes a whole stimulus-pattern in the receptive field. Topology-preserving mapping means that similar input data are projected into adjacent cardinal neurons. This self-organizing processes is responsible for the formation of brain maps with localized encoding of perceptual data. There, even more, correlated patterns are stored next to each other. Methods of vector quantization and principal component analysis will also be presented.

In the *third part* various levels of synergetic computation in the cortex and their relations will be presented. Emergence of a hierarchy of neuronal patterns acting as attractors, and occurrence of thought processes as dynamic pattern sequences or episodes will be discussed. In Haken's synergetic

model, order parameters, which can be embodied by cardinal cells or cardinal domains, act as symbolic codes or representatives of patterns.

Quantum processing will be argued to be necessary for consciousness which remained absent from ordinary neural-net processing. Reasons for this will be listed. Then, neural-net mathematics and quantum mathematics will be systematically compared. *The resulting neuro-quantum mathematical parallelism, which shows an analogy or resemblance, not any rigorous and isomorphic equivalence, indicates that the network metaphor can be used for quantum as well as neural parallel-distributed processing*, in spite of differences of the nature of neural and quantum "units".

If we use these neuro-quantum analogies, effective information processing, similar to neural-net processing, can be searched for in quantum, sub-cellular and other synergetic systems, because the principles of collective information processing are already known from our (including author's) computer simulations of neural nets.

## 1 Introduction to global brain processing: Associative or attractor neural networks

### 1.1 Introduction to neural networks

The term *neural networks* has several meanings. For neurobiologists it denotes webs of neural cells (neurons) in the brain [20, 83]. Neurons are connected through synapses which are essential for encoding of the memory-traces. Information is processed by collective dynamics of neurons and synapses.

Interdisciplinary oriented scientists started to model nets of neurons and synapses mathematically in the forties [6]. Since then, the term neural networks is usually a short name for *neural network models*. Another term for neural network models, which were realized by computer simulations or were implemented in hardware, is *artificial neural networks*. The spectrum of neural network models spreads from biologically-plausible models of real neural processing in living systems on one side to

technical imitations applied for information processing on the other side. The second branch includes all simple artificial brain-like mechanisms or algorithms which establish a sort of model of external data-structures [12, 43, 99]. This is called *empirical (external) modeling* of physical, chemical, technological, economical and other processes.

Here we will be concerned with the first branch of neural network models — the *theoretical (internal) models* of neural network architectures [27, 89]. This kind of analysis has two sub-branches: modeling, as faithful as possible, of real neural webs, and mathematically-oriented research of neural network models. Of course, it is often not possible to make a sharp distinction between these branches and sub-branches of modeling. The mathematical analysis is more or less developed from neurophysiological data. In some cases these *data* are insufficient, but in other cases *models* are insufficient, because they only very roughly obey biological data.

Mathematical approaches to neural network models used for theoretical or internal modeling ("physics of neural networks") can be classified as follows:

1. *Haken's synergetic computers* [40];
2. *statistical mechanics of neural networks* which was developed from the *Hopfield model* and *spin glass theory* [45, 102, 3];
3. *self-organizing feature maps* by Kohonen and others [56, 54].

In this introductory part we examine neural network models, henceforth referred to simply as 'neural networks', which are the most promising and illustrative ones for micro-cognitive modeling [67, 79]. In modeling we usually *neglect whole internal structure* of an individual neuron and an individual synapse. We assume that all information processing is a result of *collective* dynamics of the whole system of neurons *and* synapses. Each neuron is receiving signals from numerous other neurons. This is modeled by a simple addition of such synapse-weighted contributions reflected in an increase of neuron's activity.

Due to their architecture neural network models could be divided into following groups:

1. Fully and bidirectionally connected *associative* or *attractor neural networks* recognize and store patterns without reducing information content. Patterns are stored completely, e.g. as pixel images, representing attractors of the system dynamics. Neurons are symmetrically or homogeneously distributed (figure 3 left).
2. Uni-directionally oriented *networks with "hidden units"* (*cardinal neurons* or *order-parameter-cells*) which are located in the middle layer. This is the second layer of hierarchically ordered population of competitive neurons (figure 6). Such networks are usually *feed-forward nets (multi-layer perceptrons)* [18, 58] which sometimes have feed-back loops executing the so-called back-propagation of errors.

Neural networks are currently the most successful model in cognitive neuroscience [79, 67]. Associative neural network models [12, 54, 56, 40, 72, 27, 36, 89] are a suitable mathematical description and simulation of the processing in associative areas of the cortex [20, 83].

The large-scale dynamics of an associative neural net can be represented in the configuration-energy space (figure 1). Each point on the horizontal axis represents a unique configuration, denoted by vector  $\vec{q}$ . Points on the vertical axis represent the potential or free-energy function  $E$  of each such configuration  $\vec{q}$  [45]. The vector  $\vec{q} = (q_1, q_2, \dots, q_N)$  specifies the state of the entire neural net, i.e., the entire neuronal configuration. Each component  $q_i$  describes exactly the state of each constituent neuron of this configuration.

As can be seen in figure 1, the current neuronal configurations, represented by a black ball, travels in the configuration-energy space of possible configurations in order to find a stable state. Where the ball is stabilized, configuration represents a pattern.

## 1.2 Basic neurophysiological facts and their modeling

The body of a neuron, the soma, acts as an 'adder' in that it receives signals from all other neurons and adds them together [46]. Before being added, a signal must be multiplied in the soma by the strength of the synaptic connection reflecting both its trans-

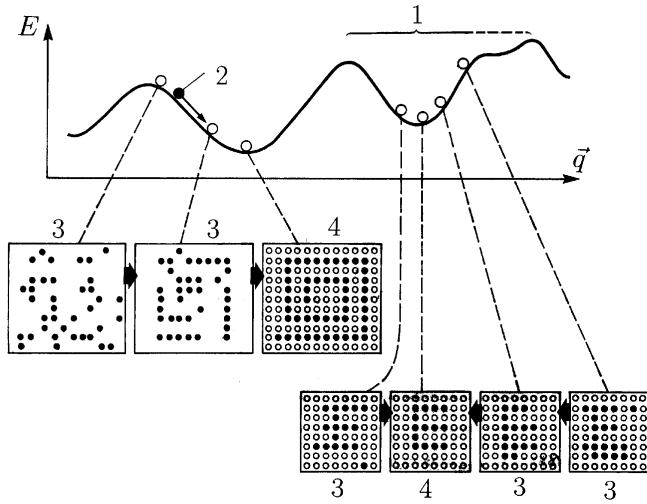


FIG. 1. Configuration-energy space (Note labels: 1 – basin of attraction; 2 – momentary state of the system; 3 – unstable configuration; 4 – pattern, i.e., stable configuration)

missivity and its sign ( $-/+$ ), which can be either inhibitory ( $-$ ), or excitatory ( $+$ ), or even non-existent ( $0$ ). Thus, the nature of a synaptic connection, i.e., its conductivity and 'sign', encodes all relevant information about the influence of a given sending neuron on a given receiving neuron in this model.

Let the variable  $q_i$  denote the state of the  $i^{\text{th}}$  neuron and  $J_{ij}$  denote a synaptic connections  $i - j$ , where  $i$  and  $j$  vary over the set  $(1, 2, \dots, N)$  with  $N$  being the total number of neurons in the net. The state of the  $i$ -th receiving neuron is then expressed as

$$q_i = \text{Sgn}\left(\sum_{j=1}^N J_{ij}q_j\right) \quad (1)$$

where  $q_j$  corresponds to the emitting neuron. This equation is valid for each neuron individually.

Function  $\text{Sgn}$  means that the result  $q_i$  is, in our case, equal to  $+1$  (the neuron is active, or firing) if above sum is greater than a certain threshold, or is equal to  $-1$  (the neuron is quiescent, or inactive) if the above sum is smaller than the neuron's threshold. The threshold is a certain limit value which has to be exceeded in order to fire the neuron. So, the function  $\text{Sgn}$  gives the efferent neuron the sign of the sum of all afferent signals of other neurons.

Letting  $P$  be the integer representing the number of patterns stored in the neural net, then each element of the connection matrix  $\mathbf{J}$  varies over the set of integers  $(-P, -P + 1, \dots, P - 1, P)$ .

Neurons send each other electro-chemical signals [61] according to the principle of 'all for one, and one for all'. One neuron can either promote (through excitatory synaptic connection) or curtail (through inhibitory synaptic connection) the momentary state of another. Specifically, if the sum of signal contributions that a neuron receives from (all) others exceeds a certain threshold of the neuron, then the neuron becomes activated or firing, and it will emit signals.

Individual signals, so-called 'spikes', are arranged in sequences, which are like combs. The frequency of such 'spikes' and the length, or duration, of a 'spike-train' bears quantitatively interpretable informational value. The average frequency is 50 emitted spikes per second. The signal's power is encoded in its frequency modulation. Thus, the more 'dense' is the 'comb' (i.e., the greater the spike frequency), the more 'influential' the signal [40]. The length of a sequence or 'train' of impulses is constrained by the time-duration of the given state of the sending neuron at the time of emission: the longer the sending neuron is active, the longer the 'comb,' and so much the greater will be the influence of that sending neuron on other receiving neurons. The informational content of the frequency of a 'comb' of signals lies in continual re-enforcement of the current state of the sending neuron. Firstly, the neuron sends one signal, and continues to do so if it has remained active in subsequent moments of time. In the absence of sufficient frequency of incoming impulses, the neuron becomes inactive and thereby stops sending its signals.

### 1.3 The system tends spontaneously toward the configuration of minimal energy

Henceforth, we will consider systems of neurons in terms of the Hopfield model [45, 1, 3] which prescribes that all neurons are interconnected bidirectionally and obey the principle 'one for all, and

all for one'. In natural neural nets each neuron is not necessarily interconnected directly with all others. However, the network behaves as though all its neurons were interconnected in this way (mathematico-physical analysis in: [70](pp.45-48) and [72]). A large number of neurons is densely interconnected with a large number of other neurons. In the absence of direct connections, connections are established via a mediator. In either case, the essential property of such systems of neurons is that neurons being closely inter-related via a system of connections one to all others. Therefore various forms of collective organization of neuronal states occur which give rise in turn to the formation of 'attractors' or 'potential wells' in configuration-energy space. Configurations with potential wells or attractors are called 'patterns'. It is exactly such patterns-qua-attractors which are major factors in determining the course of mental processes, not the neurons and synapses alone. Therefore even large-scale lesions don't destroy functioning essentially if attractor structures are preserved.

Let us consider the case of a network of receiving neurons (sensors or receptors) in an eye. The net receives a stimulus from a certain environment. Neurons thus stimulated begin to organize themselves. The self-organizing process is then transferred to centers of higher brain function, firstly to the centers for visual perception in the cerebral cortex. Newly-formed patterns in the cerebral cortex interact associatively with other patterns... The process of neuronal organization consists, in part, in switch of neuronal states on a massive scale. Synaptic connections are being strengthened simultaneously.

The process of changing of synaptic strengths in response to various stimuli is called *learning*. During such processes neuronal configurations and *synaptic connections, which represent memory-components, undergo change*. Such processes occur always in such a way that the energy of the entire system decreases. The bottom of a potential well in such a configuration represents a pattern and attracts, in turn, all other configurations. Therefore, a point at the bottom of a potential well is called an *attractor*: all configurations within the boundaries of a potential well in configuration space converge to

that bottom pattern configuration [3]. This implies that neurons fire or are quiescent in such a way as to form patterns. Once the system 'falls into' such a configuration, all further configuration-changing processes cease until the reception of a new stimulus.

This process is depicted in figure 1. The state of an individual neuron is represented as a pixel of the picture. The transformations from a configuration which resembles nothing recognizable to that of a pattern ('spiral') is readily apparent. Parallel to configurations the corresponding changes to configuration-energy space are drawn. Retrieval of another pattern ('E') from memory corresponds to converging to another attractor. It is critical to know that the energy surface changes in the process! While recognizing the first pattern, the first potential well is especially deepened; while recalling the second pattern, the second attractor becomes dominant and larger. Thus, it is obvious that a neural network is a very dynamic system in which continual change of configuration-energy space, i.e., changing the configurations of the system, enables flexible formation and storage of patterns.

#### 1.4 Recognition and memorizing of patterns

Figure 2 illustrates once again the process of pattern-formation and/or pattern-reconstruction (recognition). We may conceive of configuration-energy space (figure 1) as a 'soft', deformable surface of a mattress with shallow pits and the momentary configuration of the network as a heavy ball resting conformally on this surface. The ball deepens the potential well centered around that portion of configuration-energy space where the ball is located at the given moment. External stimuli are "putting" the ball, which represents the system, in a certain configuration, i.e., they locate the initial point in configuration-energy space. These stimuli force the receptor neurons into a configuration of stimuli (which is none other than the distributed representation of the external stimuli themselves) which propagate, in turn, to regions of higher mental function in the cerebral cortex.

Under the influence of stimuli from the environment a new pattern can be formed (constructed) by the appearance of a new potential well, if it was not there before, of an object that we had never seen before. If, on the other hand, the pattern had already been formed and stored, then the pattern is associatively regenerated. The stimulation following the recurrence of a perceptual event of seeing an object is very similar to the stimulation induced upon the event of the first sighting of this object. If the object is one and the same, it thus stimulates the neurons in the same manner and under circumstances identical to those of the first sighting of the object. The pattern of neuronal activity will be reconstructed as it had been at that time in that previous state of the system.

So, in perceiving an object for the second, tenth, or hundredth time, the same potential well, the same attractor, the same pattern will be reconstructed. It is in this manner that we perceive identity. In the same time that we perceive or recognize a pattern, we also remember it. This implies that the pattern is transferred from short-term storage in the system of neurons to long-term storage in the system of synaptic connections. Thus, the processes of *construction*, *re-construction*, and *short-term storage of a pattern* are identical [40, 39]!

A neural network *transforms a specific external stimulus into a specific internal pattern*. This isomorphic mapping means that an exactly-determined external stimulus always brings about the formation of an exactly-determined potential well and that the neural system stabilizes at the bottom of this particular well. Furthermore, several configurations may converge to (settle into) the bottom of the same potential well if they are sufficiently-similar. For several configurations to be 'sufficiently-similar' means that they are in the neighborhood of convergence of the same attractor. Thus, an attractor acts as a sort of 'funnel' into which similar configurations 'flow together'. So, the system finds the *closest* attractor in configuration space. An example of 'confluence' into the same attractor is given in figure 2 (see also figure 1).

In this manner a neural network realizes *classi-*

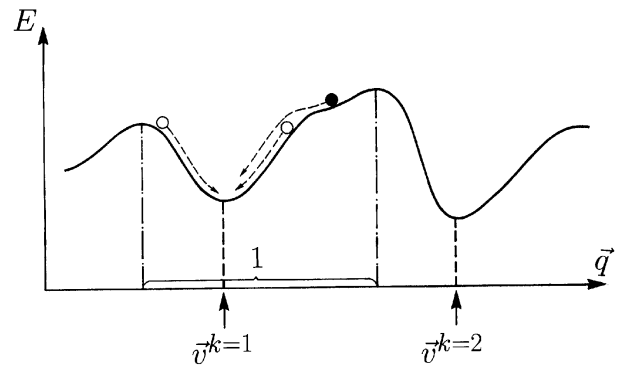


FIG. 2. Similar configurations converge into the same potential well corresponding to a pattern  $\vec{v}$  while recognizing or/and memorizing it. (Note label: 1 – basin of attraction)

*fication*. This enables recognition of an object in slightly different circumstances from those in which one is used to seeing the object in the past. So, we can determine that a purse, which has been rotated through a certain angle or displaced from its usual position, is still the same purse. Repetition of a certain pattern deepens its potential well. Through frequent repetition, the purse configuration becomes stronger and more stable. However, perception of an external pattern under new circumstances is accompanied by changes in corresponding internal configurations. A virtual internal pattern is always formed according to an optimal compromise and some 'melting together' of all possible variations, both of the object and of circumstances, in which the object can occur. Thus, an old pattern of an object (experience or *expectation*) combines associatively and reaches optimal agreement or consistency, respectively, with the new information concerning the new state of this object as acquired through sensory organs. A pattern is synthesized from the outside-world-information, from memory, and from contextual information from other centers. Then such a revised and corrected pattern is stored again. So, indeed, recognition of a pattern is also identical with its formation, reconstruction, and short-term storage in a system of neurons. Long-term storage is achieved through a system of synaptic connections. Short-term storage depends

mostly on bio-electrical processes involving neurons, whereas long-term storage depends on bio-chemical changes in synapses.

### 1.5 Hebbian learning based on search of correlations

*A neural network can store many patterns simultaneously.* Accordingly, each neuron and each synapse can play *multiple roles*, since *each neuron (or each synapse) participates in several patterns*. There can be several potential wells in configuration-potential space, to which the dynamics of the system converges. However, only one pattern  $\vec{v}$  is realized in the system of neurons  $\vec{q}$  at a moment. During the pattern-recall the following equality holds:  $v_i^k = q_i$ .

The strength of synaptic connections are determined by Hebb's 'learning rule' which is a fact of neurophysiology [6]: *If two neurons are both active or both inactive, then the synaptic connection between them is strengthened. Otherwise, if one is active and the other is inactive, then their mutual synaptic connection is weakened.* Therefore, if two neurons corroborate each other, their connection is strengthened. If, however, they contradict or counteract each other, then their connection is weakened.

Let  $v_i^k$  denote a variable ranging over patterns, where  $k$  refers to the  $k^{th}$  pattern and  $i$  to the  $i^{th}$  neuron. So,  $v_i^k$  represents the role which the  $i^{th}$  neuron plays in forming the  $k^{th}$  pattern. Now, let us formulate Hebb's rule mathematically:

If both the  $i^{th}$  and  $j^{th}$  neurons are *either* both active in forming the  $k^{th}$  pattern ( $v_i^k = v_j^k = 1$ ) *or* both are inactive in forming the  $k^{th}$  pattern ( $v_i^k = v_j^k = -1$ ), then their mutual connection  $i - j$  becomes stronger:  $J_{ij}$  is increased. Otherwise, if their activity-rate differs ( $v_i^k = -1$  and  $v_j^k = 1$ ) *or* ( $v_i^k = 1$  and  $v_j^k = -1$ ) then their mutual connection becomes weaker. As can be shown, Hebb's auto-correlation equation, or Hebb's learning rule, consolidates all these facts:

$$J_{ij} = \sum_{k=1}^P v_i^k v_j^k. \quad (2)$$

It describes how synaptic connections encode *corre-*

*lations* between neuronal activities of a single pattern.

Each individual product in equation (2) represents coupling. If an individual product is negative (positive), then connection is weakened (strengthened). The connection strength  $J_{ij}$  is determined by the sum of such coupling of all patterns which are stored simultaneously in the same system. There are  $P$  of stored patterns. Furthermore, if the states of two connected neurons  $v_i^k$  and  $v_j^k$  collaborating in the formation of the  $k^{th}$  neuron are of the same sign, then they will contribute to the stability of the  $k^{th}$  pattern in the neural network. If their signs are different, then they will weaken the pattern.

Patterns stored in synaptic connections are loaded one atop another. Connections are determined locally according to mutual (dis)agreement of pairs of neurons. Globally, in turn, connections determine the (in)stability of all configurations. Thus, neurons locally compete amongst themselves, each seeking to prevail and convert the rest to its state. A neuron succeeds in this if it is in best agreement with the other neurons and if it has the most support from the rest of the system. Then the neuron has excitatory connections.

Globally, neurons mediate competition and coupling among entire patterns. That's similar to the case of two individual sportsmen ("neurons") which mediate in an international sporting event competition of their respective states ("patterns of neurons") which they are representing at the event, thus bringing about a 'meeting of two states'.

In figure 3, on the left, a 'natural' representation of Hopfield's network is depicted and, on the right, a diagram of the functional scheme of the structure of the memory or correlation matrix  $\mathbf{J}$  according to Hebb's equation (2). Individual elements  $J_{ij}$  represent the coupling of the  $i^{th}$  neuron  $v_i^k$  and the  $j^{th}$  neuron  $v_j^k$ , forming parts of a given  $k^{th}$  pattern.

### 1.6 Content-addressable memory in networks

In a neural network information is stored in a parallel-distributed (non-local) manner in the values of connections: patterns are distributed across

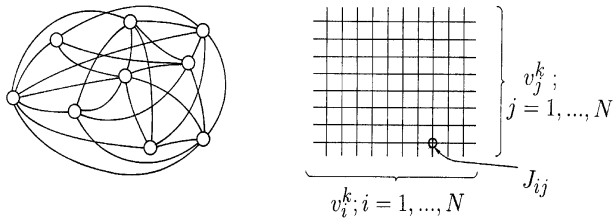


FIG. 3. Diagram of a neural network according to Hopfield's model (left) and diagram of a memory matrix  $\mathbf{J}$  (right).

the entire system of connections. Memorization is *content-addressable* where mutual connection between neurons are essential for associative recall. It differs overall from the memory storage found in von Neumann computers which don't have directly interconnected and directly interacting memory components, and where contents are addressed with special codes. Information content is represented in the state of the network itself. Accordingly, to recall content requires bringing the system in a certain state. Stored patterns co-determine each other. That's similar to the case of a cross-word puzzle ("network") where words ("patterns") determine each other, because they have common letters ("neurons"). These inter-relations enable associative recall.

Neural networks operate *iteratively*. The efferent signal of one neuron is the afferent signal of another neuron. *Cycles* of input—output—new input are densely-interwoven throughout the entire network. So, new-coming patterns get associatively connected with the old memory-patterns and thus *condensely and optimally stored without repetition of common information* (like in a cross-word puzzle where common letters are shared among words).

Memory traces of patterns remain in the connection system even after the system of neurons passes over to the next pattern. There can be only one pattern at one time in the neurons ('in consciousness', or more exactly in the manifest consciousness), whereas in the connections (in memory, or in latent consciousness, or in the subconscious) there can be several patterns simultaneously, albeit it is necessary to recall them from memory. *Through*

*recall a memory is brought from the connection system into the system of neurons*. The prompt for this is typically a similar external stimulus which draws neurons into a 'replication' of the externally-imposed pattern. Such a prompt can come also from some other cerebral sub-networks (and thus internally with respect to all the sub-networks of the brain, yet external to the sub-network under consideration).

To illustrate this let us consider the example of the ball and the pillow from section 1.4. The ball (neuronal configuration) can occupy only one place on the pillow (in potential space), although there can be several pits which it has already indented in the pillow (patterns, 'imprints'). We are 'conscious of' and we are 'experiencing' momentarily only that pattern which is in the neurons, i.e.,  $\vec{v}^{k'} = \vec{q}$ .

In the course of *memorization* information is transferred through the process of learning from 'manifest awareness' (in the neurons – in  $\vec{q}$ ) to 'latent awareness' (stored in the connections – in  $\mathbf{J}$ ). Thus, memorization is a single-valued mapping of some image of an external object into an internal virtual image firstly in the network of neurons themselves (short-term memory, as defined by some authors), whereupon this image is transferred to the synaptic connections (long-term memory).

The pattern is *forgotten* if a certain potential well is erased. This can come about through reception of new stimuli, disturbances, disagreement with other patterns, or disagreement among neurons of the same configuration. Then connection strength diminishes and patterns are weakened. Otherwise, if patterns are frequently reinforced and renewed, then they are strengthened. During learning this comes through repetition. Forgetting can also come about through interference between patterns or 'melting together'. However, the most frequent concomitant to inability to recall a pattern is inability to reconstruct its attractor due to an insufficient 'key', i.e., the partial contents required to prompt the recall of the entire contents.

If a neural network has symmetric connections (the synapse is equally permeable in both directions:  $J_{ij} = J_{ji}$ ), then the system can form stable attractors. If the connections are asymmetric, then



patterns will not be stable and may move around in configuration space. This means that the potential well moves or (what is essentially the same thing) it rolls along configuration space from one pattern into another. One pattern disappears and another appears. The system may describe periodic (after a certain time period the system returns to its original configuration), quasi-periodic, or completely-chaotic paths. Dynamical pattern sequences are the origin of the *flow of thought*.

### 1.7 Explicit/implicit nature of patterns/ attractors; cardinal cells; minimizing free energy

Let us examine how neurons attain the energetically-optimal state. This is the configuration in which they are in *optimal mutual agreement*. Free energy is proportional to the degree of disequilibrium or disagreement among a population of neurons. The optimal state can only be attained through *continual transactions through connections*.

Energy ratios of configurations are carriers of implicit structures – attractors. The *free-energy*-function  $E$  of the system is lower if neurons agree (their activities  $q_i$  and  $q_j$  have the same sign in that particular moment) while having positive connection  $J_{ij}$  (good history of agreement) or disagree while having negative connection (more disagreements in previous times). Furthermore,  $E$  increases if neurons agree while having negative synapse or disagree while having positive synapse. In the same time the energy is inversely-proportional to the agreement of neurons with the neuronal field (“public opinion”). These conditions are encoded in the following equation [36]:

$$E = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N J_{ij} q_i q_j - \frac{1}{2} \sum_{i=1}^N T_i q_i. \quad (3)$$

$T_i$  represents the influence of the neuronal field on the  $i^{th}$  neuron (the influence of “society” on the individual, or the other way around). In another interpretation,  $T_i$  may represent a neuron’s thresh-

old, specifically, its own threshold or its threshold as determined by others.

The Hebb learning rule gives such a value to  $\mathbf{J}$  that  $E$  is *automatically minimized* if stored patterns  $\vec{v}$  enter into the system of neurons  $\vec{q}$ , i.e.,  $\vec{q} = \vec{v}$ . Only in that case the whole network is, with respect to the system of neurons as well as the system of connections, in an optimal overall agreement, and its free energy  $E$  is minimal.

Patterns could not appear on the basis of simple data, which (sensory) neurons receive repeatedly and in fragments, if these patterns would not collaborate and adapt to one-another. On a higher level, a pattern could not collaborate in the formation of complex patterns, if it weren’t in a comparative relation with other patterns. The pattern would be meaningless by itself.

So, a pattern has a role only in *relation to other patterns* and usual configurations. Its place in a set of relations is evident in the energy-configuration space. A pattern has *more meaning* than an ordinary configuration, since it is *more stable*. It is more stable, because it is a more *accurate representation of an actual situation* - in the outside world or on a more abstract internal level. Stated otherwise, a pattern is more stable, since it agrees with outside patterns and with other internal patterns. This is expressed in the fact that the free energy of a system of neurons is lower when that pattern is reconstructed in the system’s configuration. It is because of this that a network tends to this pattern. Let us emphasize again that a map from the environment ‘impresses’ contents into patterns, whereas their meaning emerges from their relations to one another and to the environment. External patterns from the environment have the same role as do internal patterns except that they have greater weight.

In figure 4 two configurations (denoted by a ball) are shown which are in *different ratios* to the other configurations (other points on the horizontal axis). The left one is not stable and the system assumes it only transitionally, since the relief just transformed itself and the system hasn’t yet adapted to the

change. In the right configuration the system equilibrates, because it is especially important (it is a pattern). Its importance/meaning is co-determined by other configurations also.

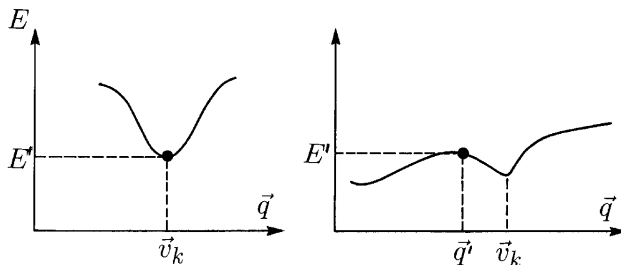


FIG. 4. The same location in the configuration-energy space (the ball) has various meanings in various contexts. On the left it represents a pattern, on the right an ordinary configuration.

Minima of configuration-potential space represent *implicit order* in a neural network and this order determines the formation of further virtual mental structures. In figure 5, on the left, the system is shown in the neighborhood of an attractor and is certainly going to move towards a minimum in configuration space, except if the surface of the relief would suddenly change because of other influences. In this case, the pattern is already latently present in the system of neurons. It is contained *implicitly* in the state vector  $\vec{q}$ , although it is mixed together with traces of other patterns. Then, when the system goes over to the bottom of the potential well (figure 5, on the right), a certain pattern appears clearly and resolutely 'in consciousness'. Then the pattern is manifest *explicitly*. This explicit manifestation is especially obvious if the corresponding *cardinal neuron* is selected and starts to fire. *Switching-on of some most dominant neuron, which then represents the whole neuronal pattern, often accompanies the explication of a pattern from background.* Since the moment of his winning the competition, such a *cardinal neuron encodes and represents its pattern at the next hierarchical level* [40, 73, 75].

When the system is *in the basin of attraction* of an attractor corresponding to a pattern, then

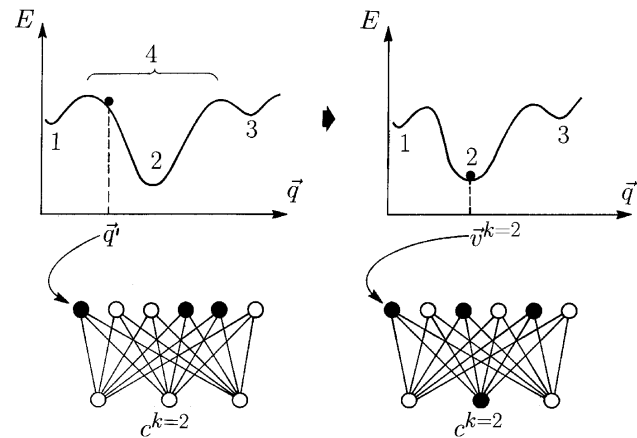


FIG. 5. A pattern (e.g.,  $\vec{v}^{k=2}$ ) is represented implicitly when the system is in its neighborhood of attraction (left: unstable initial situation). When the system reaches the bottom 2 (right: stable final situation), the pattern (black - white - black - white - black - white) becomes explicitly manifested. This may be accompanied by the firing of a cardinal neuron  $c^{k=2}$ . (Note labels: 1, 3 – local spurious attractor; 2 – global attractor; 4 – basin of attraction)

the pattern is represented *implicitly* in the system of neurons, not only in memory (in the system of synaptic connections). When the system of neurons reaches its attractor and, thereby realizing this pattern, in the same time, the value of its cardinal neuron attains a maximum, if it can attain various degrees of strength or intensity. If it can only be active or inactive, then it will be activated at that time. With this, the cardinal neuron *signifies* that there is a corresponding pattern 'in consciousness'. In sending signals throughout the brain, the cardinal neuron *informs* of the presence of its pattern and *represents* it. Therefore a pattern (which is parallel-distributed throughout a network at a lower level of its hierarchical structure) is *encoded implicitly* in the concrete material state of one cell.

The use of cardinal neurons is but a special way of organizing the operation of a neural network. This principle obtains on the *virtual* level as well, on which the cardinal neurons are replaced by *synergetic order parameters* which "act" *implicitly*.

There is a continual transition of configurations

from implicit representation in the system of synaptic connections (memory) to explicit enactment in the system of neurons ('consciousness'), or conversely. Implicitness means the *set of all possibilities*, from which the environment selects that which is actual in that moment and brings these contents into explicit experience. The implicit context determines *the form of the potential well*. So, this implicit background influences those configuration which momentarily represents the 'center of mass' of an attractor with this particular potential well.

The set of all possibilities does not mean literally that we have in the brain, ready in advance, all possible images and situations, in all their details. In memory we have only an *abstract model*, drawn from the most basic features of past experience. In this model each thing is represented only once. Similar things or repeated occurrences of things are encoded automatically with the 'ideal' group of properties of those things (common features), along with additional information about specific features of an actual individual thing.

In this section we have established the interrelation and 'transition' among non-local (parallel-distributed) patterns and local cardinal cells (if the latter exist). If someone asserts the discovery of cells which carry certain information, then this concerns cardinal neurons. However, they cannot process this information alone, rather only through *dynamic collaboration* with their basis patterns. Only these base patterns enable some cardinal cells to exceed their activation threshold.

Obviously, a 'hunger neuron', 'awareness neuron', 'pleasure neuron', etc., (in so far as we may claim that they exist at all) are not hunger, awareness, or pleasure per se, but rather they encode the wider complex of patterns (distributed across the brain) and its influences – just as a signal lamp signifies that a machine is in operation. Cardinal neurons are more clearly-defined at lower levels (integrative neurons), whereas at higher levels, cardinal domains and virtual order parameters perform encoding functions [39, 38].

## 2 Perception: Self-organized topology-preserving mapping by Kohonen neural networks

### 2.1 Introduction to perceptual maps with cardinal neurons

The model of self-organizing networks which adapt to perceptual data by evolving brain-maps with preserved input-data relations was established by Teuvo Kohonen in early eighties [53, 55]. It is a relatively biologically-plausible model. It is mostly used for modeling *perception* (using sensor maps) and *motor action* (using motor maps) in living organisms, especially in the brain [72, 82, 27], but it is applied also for pattern-classification tasks in computer science and robotics.

Feed-forward neural networks can perform mappings from sensory inputs (lower layer) to internal representations (middle layer), and further-on from internal representations to motor outputs (upper layer). From now on we will consider only the sensory part of this processing. Decoding of internal representations for triggering motor actions (realized in the upper half of the network in figure 6 right) is anyway following principles quite similar to coding of sensory data  $\vec{x}$  into internal representations  $\vec{w}$  (realized in the lower part of the network in figure 6 right).

In these networks encoding is constructed by *reducing the difference (error) between the external state and the network's internal representation of it*. Both, internal and external states, are presented as activity patterns (figure 6 left) and are mathematically described by vectors. The individual components of each pattern-vector corresponds to an individual element of the pattern. The synaptic connections, where patterns are stored, change according to the rate of this disagreement between the prototype (internal representation)  $\vec{w}$  and ordinary pattern (pattern of environmental stimuli detected by sensory cells)  $\vec{x}$ . Prototype can be prescribed by a "teacher" (supervised learning) or can be established by a self-organizing procedure (unsupervised learning). In the second case, which is

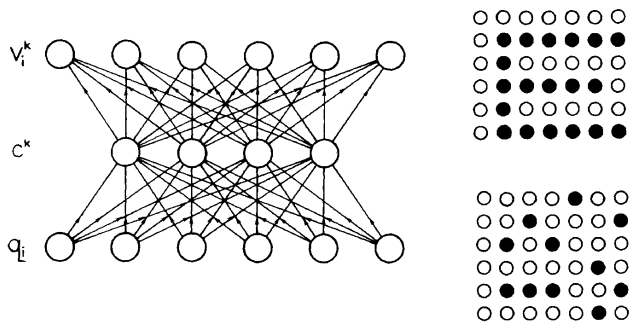


FIG. 6. Right: Feed-forward network (without feedback) with input (sensory) layer, hidden (representation) layer and output (motor) layer. Left: Two examples of a pattern of cell activities as can be formed in the symmetric net as well as in the feed-forward net. Bottom left: input pattern of sensory-cell activities. Upper left: target pattern of synaptic strengths.

biologically more relevant, the prototype is constituted by the *winning neuron*. This is the neuron which *wins the competition between neurons and individually takes the task of coding a specific pattern*. The winning neuron is also called *cardinal neuron*, or *order-parameter-cell* (because it orders the whole layer of neurons after it has "taken all the power"). To be more specific, as long as the cardinal neuron competes with other cardinal neurons we will refer to it as *potentially-cardinal neuron*. In that stage also its competitors are potentially-cardinal neurons. Only one of these potentially-cardinal neurons can win the competition and become the most activated one. In that case it is called *actually-cardinal neuron*. Such a winning neuron located at  $\vec{r}$  represents its corresponding prototype-pattern  $\vec{w}_{\vec{r}}$ . All new-incoming patterns  $\vec{x}$  are compared to this "ideal" (called also *target-pattern*  $\vec{w}$ ) and are classified according to this most similar prototype  $\vec{w}$ . Of course, a neuron can win the competition only if it has support of the actual state in the environment.

Kohonen network is a combination of fully interconnected network, if we look the target layer only, and a feed-forward network if we look vertically through the hierarchy of layers (figure 7: more layers could also be added) [72].

Within the target-layer (output-layer) connec-

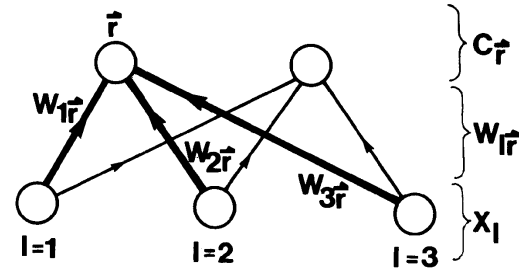


FIG. 7. Kohonen network with the first (input) layer of sensory neurons and the second (target, output) layer of cardinal neurons which constitute maps. Potentially-cardinal neurons inhibit each other.

tions are non-plastic and *inhibitory*, whereas inter-layer connections are plastic. In Kohonen model it is usually taken that neurons have activity-values between 1 (excited) and 0 (quiescent), and that synapses have positive and negative real values including 0.

Such a neural network develops receptive fields [82]. *Receptive field* is a population of sensory cells which are connected to a specific cardinal neuron. This cardinal neuron is specialized to be most sensitive to a specific input-pattern only. A cardinal neuron with its corresponding receptive field classifies and stores a specific input-pattern. Keep in mind that a receptive field is a domain of sensory cells, whereas a map is a domain of neurons.

Each receptive field acts as an *adaptive filter*: it selects those pattern which is the most similar to the right target-pattern. Each target-pattern is represented by a potentially-cardinal neuron. The cardinal neuron whose target-pattern  $\vec{w}$  is the closest to the actual input-pattern  $\vec{x}$  becomes the most activated one.

## 2.2 Topologically-correct feature maps

Neurophysiologists have discovered specific brain regions which process or encode specific kinds of information [68]. The Kohonen model is the most important and biologically-plausible one which can present the process of *localized encoding* of perceptual data (figure 8) [94, 72, 82, 53].

For this localized encoding, topologically-correct

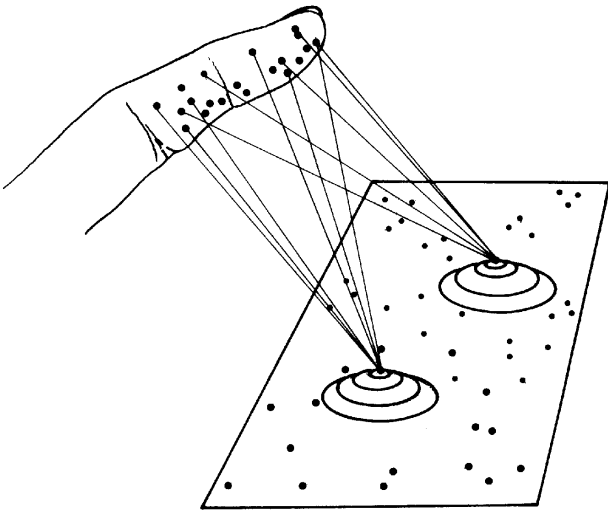


FIG. 8. The network reacts to an input-pattern (in this case represented in sensory cells of the finger's receptive field) by a local activation around a cardinal neuron which is the carrier of encoding of that input-pattern in the brain map.

mapping is essential. In this case, mapping means a transformation of an input-pattern  $\vec{x}$  to a corresponding internal representation or memory  $\vec{w}$  of that input. Topology-preserving mapping means that *adjacent points of an input-pattern are connected to adjacent neurons in the target-layer*. So, topological relations are conserved while the input-pattern  $\vec{x}$  is projected to the target-pattern (output-pattern)  $\vec{w}$ .

*Self-organization* means that the mapping is a result of interactions between neurons and “spontaneous search of collective agreement” [54]. The trigger of this process is a stimulus from environment. This perturbation is enough for forcing the system into search for a new equilibrium. Stability is found by forming a state where the actual input-pattern  $\vec{x}$  and the target-pattern  $\vec{w}$  are as close as possible, i.e.  $\|\vec{w} - \vec{x}\|$  is minimal. Actually, input-pattern  $\vec{x}$  selects those target-pattern  $\vec{w}$  which is the most similar to it. This is done by selective activation of those sensory cells  $x_l$  which “by chance” have at that moment such an activity which matches the stimulus.

Self-organizing mapping is an optimization pro-

cess where the *similarity of input signals is projected to proximity of excited neurons*. So, similar input patterns will be encoded into topologically close neurons. To be “topologically close” means in this case that neurons which are physically located close to each other will react similarly to similar inputs, but neurons located far apart of each other will react differently to similar inputs.

In neural networks dynamics is in general governed by minimizing a co-called *cost function* which is analogous to energy-minimization in physics: the cost-function has to be minimized if the system has to reach equilibrium. In Kohonen model the cost function is proportional to the square of difference between the input-pattern  $\vec{x}$  and the target-pattern  $\vec{w}$ . The Kohonen cost function is analogous to the elasticity energy which is proportional to the dislocation from the equilibrium point (here  $\vec{w}$ ). The network can “spare energy” if neurons with similar tasks communicate over very short connection paths. Therefore the mapping where *similarity relationships among the input-patterns* are transformed into *spatial relationships among the responding neurons* is an “energy”-optimization process [82].

At the same time, the self-organized topology-conserving mapping means *dimensionality reduction of the representation space*. A high-dimensional space of input-patterns is reduced usually to two-dimensional maps which maintain neighborhood relationships. Input-pattern has a dimension equal to the number  $n$  of constituting elements – excited sensory cells  $l$ . Cardinal neurons are determined by  $x$  and  $y$  coordinates belonging to the cortical map. This mapping means also *data compression*. Network learns to discern the most important features of input-patterns. Only these main characteristics of the input-information are stored.

The self-organized map is a result of continuous and non-linear approximative mapping which is defined implicitly by a self-organized process of searching stability through maximal possible agreement of neurons. This process is called *learning*. It is driven by learning examples from the network's evolutionary history.

The two-dimensional maps are mostly located in the primary layer of the neocortex. Examples are somatotopic map (in somatosensory cortex) of the skin surface, retinotopic map (in visual cortex) of the eye's retina, tonotopic map (in auditory cortex) of the ear's cochlea-spiral, etc. In the auditory-cortex-map pitch (frequency) is encoded in one direction, amplitude in the orthogonal direction. In superior colliculus in midbrain direction and amplitude of rapid eye movements (saccades) are spatially encoded [82]. Triggering mechanism for flexion and extension movements of muscles are encoded in the motor cortex (ring map).

In the neocortex there are, vertically cutting its layers, the cortical columns and micro-columns. They can also be understood as a product of self-organized topology-conserving mapping. They are thus *specialized to analyze some particular features* of the input-patterns (orientations, velocities and directions of movements, edges, periodicity, color shades, etc.), or to regulate motor actions (e.g., innervation of a common muscle). In association areas, columns are densely connected to execute collectively complex tasks like recognition of faces, speech comprehension and planning of arm trajectories. Columns provide informational basis for higher brain functions modeled by attractor neural networks.

### 2.3 Learning process by error minimization

Using topology-preserving mapping a large set of  $p$  input-patterns  $\vec{x}^k (k = 1, \dots, p)$  gets encoded into a smaller set of  $p'$  prototypes  $\vec{w}_{\vec{r}}$ . Here  $\vec{r}$  is used as an index. It refers to location of those neuron to which synapses are transmitting signals from sensory cells. The  $k^{th}$  input-pattern is described by a vector  $\vec{x}^k = (x_1^k, x_2^k, \dots, x_n^k)$ . Its components represent individual pixels of the input-pattern which are projected into activities of individual sensory cells on the retina (vision), cochlea (audition), etc.  $n$  is the number of sensory cells.

The prototype is represented by a vector  $\vec{w}_{\vec{r}} = (w_{\vec{r}1}, w_{\vec{r}2}, \dots, w_{\vec{r}n})$ . Its components  $w_{\vec{r}l}$  actu-

ally represent synaptic connections between the potentially-cardinal neuron at location  $\vec{r}$  and the sensory cells with index  $l$ .

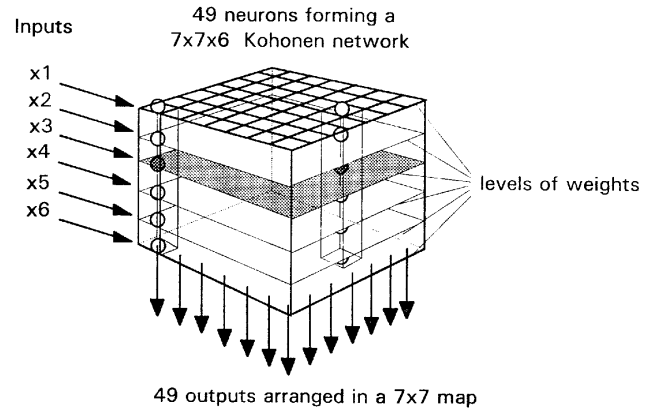


FIG. 9. Kohonen network presented by a block of synaptic strengths or synaptic weights. Each sensory cell  $x_l$  is connected to each neuron in the  $7 \times 7$  output-map. 6-dimensional input  $\vec{x}$  (here presented one only:  $p = 1$ ) is projected into 2-dimensional map of  $7 \times 7$  neurons (as many as there are squares on the top). Each neuron is presented by a vertical column which corresponds to the vector of synapses between this neuron and all sensory cells  $x_l (l = 1, \dots, 6)$ . A “level” of synaptic-weight-values which are given by the  $3^{rd}$  input variable  $x_3$  is shaded. Dark-shaded circles correspond to the  $3^{rd}$  synapse on the neuron (1,1), and the  $3^{rd}$  synapse on the neuron (6,5). In this network-block 49 uncorrelated 6-dimensional input-patterns could simultaneously be stored. If inputs are correlated, memory compression starts... (From Zupan [103], with permission.)

If some prototype  $\vec{w}_{\vec{r}}$  matches the input-pattern  $\vec{x}^k$  more than other prototypes do, then the corresponding potentially-cardinal neuron at location  $\vec{r}$  wins the competition against other potentially-cardinal neurons  $\vec{r}'$  and becomes actually-cardinal neuron – the winner. In this case we say that the network has recognized the input pattern  $\vec{x}^k$  so that it classified this input as an example of the prototype  $\vec{w}_{\vec{r}}$ . By meeting every new input-pattern the prototype changes somewhat: the new updated prototype is a result of a compromise between the old prototype and new input. The prototype is encoded in the synaptic connections  $w_{\vec{r}l}$  and/or in the cardinal neuron at location  $\vec{r}$ . We cannot localize

the encoding since the synaptic strengths  $w_{\vec{r}l}$  and the resulting activity of the cardinal neuron  $\vec{r}$  both contribute to the mapping as a *dynamic process*. Statics does not mean anything to the network; only dynamics has meaning in the sense of realizing specific input-output transformations. Therefore mathematical description which is static is not enough for understanding neural network processing. One needs a lot of intuitive, flexible “mental modeling” also.

All equations written from now on will be valid for all *sequentially occurring adaptation steps following each arrival of an input-pattern  $\vec{x}^k$* , therefore the index  $k$  will always be omitted. The activity of some potentially-cardinal neuron in the second layer (target-layer or map) is given by [82]

$$c_{\vec{r}} = f\left(\sum_{l=1}^n w_{\vec{r}l}x_l + \sum_{\vec{r}'} J_{\vec{r}\vec{r}'}c_{\vec{r}'} - T_{\vec{r}}\right). \quad (4)$$

$f$  is the non-linear (tangens-hyperbolicus-like) activity function of the neuron  $\vec{r}$  called *sigmoid function*.  $T$  is the threshold of the neuron  $\vec{r}$ .  $n$  is the number of sensory cells  $x_l$  in the first layer which are connected with each potentially-cardinal neuron  $\vec{r}'$  in the second layer (map). The neuron's activity  $c_{\vec{r}}$  is updated by a “sigmoid-shaped” sum of three contributions:

1. a sum of all inter-layer synapse-weighted signals<sup>1</sup>  $x_l$  from sensory cells  $l$ ;
2. a sum of all intra-layer synapse-weighted signals  $c_{\vec{r}'}$  from other neurons  $\vec{r}'$  in the map;
3. the threshold  $T$  of the receiving neuron  $\vec{r}$ .

$w_{\vec{r}l}$  is the strength of an inter-layer sensor-neuron synaptic connection.  $J_{\vec{r}\vec{r}'}$  is the strength of an intra-layer neuron-neuron synaptic connection.

Let's assume that the neuron  $\vec{r}'$  is the winner. Then usually,  $J_{\vec{r}\vec{r}'} > 0$  (excitatory synapses)

<sup>1</sup>A signal is propagation of a so-called action potential along a neuron's axon. It is a macroscopic effect of a complex process of selective migration of K and Na ions through the cell's membrane. Because of this migration, a gradient of concentration is formed along the axon. This gradient establishes a dynamic electro-chemical equilibrium: in the previous place the equilibrium is destroyed, in the next segment of the axon it is constituted again. Macroscopically this is manifested as a signal along the axon.

for neighboring neurons, and  $J_{\vec{r}\vec{r}'} < 0$  (inhibitory synapses) for neurons more far away. So, around each neuron there is first a concentric circle of positive synapses (support of similarly-oriented neurons), and then a concentric circle of negative synapses (famous *lateral inhibition*, found in biological perceptual systems). Those neuron who has the most strong support of the nearest neurons (its inside domain), and is at the same time able to inhibit neurons in the neighboring domains (outside its domain), will become the winner. The concentric order of synaptic strengths is illustrated by the “Mexican-hat” function (figure 10).

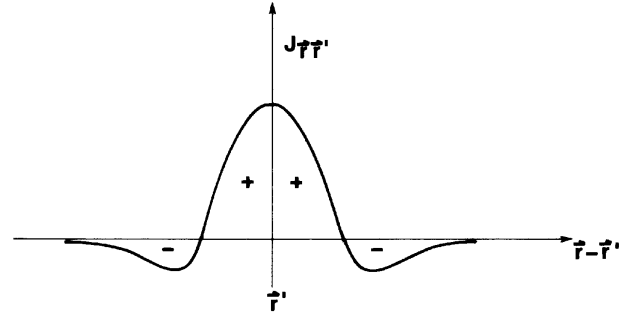


FIG. 10. Mexican-hat function: distribution of excitatory and inhibitory synapses to the neurons encircling the center-point neuron.

Lateral inhibition or competition leads to a spatially-localized excitatory response which is the highest at the winning neuron. For simplicity Kohonen often makes approximation by considering only the response of this cardinal neuron  $\vec{r}'$  as if other neurons  $\vec{r}$  would not response at all (“winner takes all”):

$$\sum_{l=1}^n w_{\vec{r}l}x_l = \max_{\vec{r}'} \sum_{l=1}^n w_{\vec{r}l}x_l. \quad (5)$$

If for each neuron  $\|\vec{w}_{\vec{r}}\|$  (that is  $\sqrt{\sum_l w_{\vec{r}l}^2}$  – total synaptic strength per neuron) is constant, and if  $\|\vec{x}\| = 1$  (all input-patterns  $\vec{x}$  have the same “intensity”), then the approximation can be written also as follows:

$$\|\vec{w}_{\vec{r}'} - \vec{x}\| = \min_{\vec{r}} \|\vec{w}_{\vec{r}} - \vec{x}\|. \quad (6)$$

In the Kohonen model it is usual to represent the target-pattern by a vector of strengths of synaptic connections ( $\vec{w}_{\vec{r}l}$ ) from each sensory cell  $l$  ( $l = 1, \dots, n$ ) to a given neuron  $\vec{r}$ . By this definition input-pattern  $\vec{x}$  and target-pattern  $\vec{w}_{\vec{r}}$  have the same dimension ( $n$ ) and can be subtracted. Now we see what is advantage of such kind of mathematical description. Activities of neurons  $c_{\vec{r}}$  are not important in *this* model. Only their *relations* to configurations of sensory-cells are important. These relations are encoded in configurations of synaptic connections ( $\vec{w}_{\vec{r}}$ ) to a given neuron ( $\vec{r}$ ). It is enough if a neuron is denoted by its location  $\vec{r}$  only.

The version (5) chooses the most excited neuron to be the winner. The version (6) chooses those cardinal neuron which represents the target-pattern closest to the input-pattern  $\vec{x}$ . Both ways bring the same effect.

The strengths of synaptic connections are changed according to

$$\begin{aligned} \Delta w_{\vec{r}l} &= \epsilon h_{\vec{r}\vec{r}'}(x_l - w_{\vec{r}l}) \\ \text{or} \quad \Delta \vec{w}_{\vec{r}} &= \epsilon h_{\vec{r}\vec{r}'}(\vec{x} - \vec{w}_{\vec{r}}) \end{aligned} \quad (7)$$

where

$$h_{\vec{r}\vec{r}'} = \exp\left(\frac{-(\vec{r} - \vec{r}')^2}{2\sigma^2}\right). \quad (8)$$

$\sigma$  is the radius of the Gaussian excitation function  $h_{\vec{r}\vec{r}'}$ .  $\sigma$  is fixed in the first approximation, but in the second approximation it can be slowly decreased during learning. In equation (8)  $\vec{r}'$  marks the excitation maximum gained by the actually-cardinal neuron. The “shape” of  $h_{\vec{r}\vec{r}'}$  is fixed with respect to variation of  $\vec{r}$  (locations of the neurons in surroundings). The position of Gaussian function  $h_{\vec{r}\vec{r}'}$  is chosen by an incoming stimulus which enables some specific potentially-cardinal neuron to become the actually-cardinal neuron  $\vec{r}'$ . The response of nearby neurons  $\vec{r}$  is given by  $h_{\vec{r}\vec{r}'}$ . Introduction of this function is necessary in order to achieve localized neuronal excitation as in figure 8.

The prescription (7) originates from neurophysiological research [6, 54]. This so-called learning rule says that the synapses  $w_{\vec{r}l}$  should be changed more if

the difference between the individual input-stimuli  $x_l$  and target-activities  $w_{\vec{r}l}$  is bigger, or that the rates of synaptic transmission  $x_l$  should be changed less if the difference is smaller. So, larger disagreement requires more adaptation. This is not all: the spatial distance between neurons is also important. Its role is encoded in the shape of  $h_{\vec{r}\vec{r}'}$  which controls the size of the neighborhood region. With that the number of neurons affected by a single adaptation step (7) is determined, as well as the rate of affection. The target-patterns  $\vec{w}_{\vec{r}}$  associated with the neighboring potentially-cardinal neurons  $\vec{r}$  are coupled more strongly than target-patterns  $\vec{w}_{\vec{r}}$  corresponding to more distant potentially-cardinal neurons  $\vec{r}$ .

Learning is the process where the adjustments (5) or (6), respectively, and the adjustment (7) are following each other in an *iteration*-procedure [72, 82, 12]. Equations (5), (6) and (7) are equally valid for all adaptation steps marked by  $k$  if we keep in mind that  $\vec{x}^k$  is new in each step. In the beginning of each ( $k^{th}$ ) cycle of iteration the new input-pattern  $\vec{x}^k$  is presented to the network. Then the winning neuron is chosen by (5) or (6), and synapses are changed according to the rate of similarity between the input-pattern and the target-pattern. In a model it can be prescribed whether only one learning cycle or many cycles are following each input. Self-organizing adaptation causes a shift of the input-vector  $\vec{x}^k$  and the target-vector  $\vec{w}_{\vec{r}}$  closer to one another. The adaptation process is finished when  $\vec{w}_{\vec{r}} = \vec{x}^k$ . We say that in this moment the input-pattern was *recognized* and simultaneously it gets *stored* in the new vector of synapses  $\vec{w}_{\vec{r}}$ . In such a way we get a new target-vector which is a new memory representation. If the network is already filled with old memories, we at least get a better memory representation, because it considers also the last input-data.

The learning rate  $\epsilon$  can also decrease gradually with the number ( $k$ ) of learning or adaptation steps. Without proof we shall mention that two important conditions are necessary for  $\epsilon$  in order to achieve



convergent adaptation process [54]:

$$\lim_{t \rightarrow \infty} \epsilon(t) = 0 \quad \text{and} \quad \lim_{t \rightarrow \infty} \int_0^t \epsilon(t') dt' = \infty. \quad (9)$$

Learning time  $t$  flows in parallel with increasing index  $k$  which belongs to sequentially arriving learning examples  $\vec{x}^k$ . The second condition demands that decreasing of  $\epsilon$  towards zero in time  $t$  (as prescribed by the first condition) must not be too fast. At the beginning of the learning process the synapses may be empty or with random transmission rates  $w_{r\ell}$ . The iteration process (6)/(7) using function (8) is actually an alternative description of the signal-exchange determined by (4).

## 2.4 Vector quantization

Two mathematical approaches to description of Kohonen's dynamics will be presented now. In most images, adjacent pixels are significantly correlated. Also if we have to store many similar input-patterns (e.g., many faces) it is useful to *store all correlated pixels close to each other*, to *store all equal data only once* and to *ignore all redundant data*.

Self-organizing mapping helps to solve this task. It executes data storage by nearby-coding of similar information and at the same time it executes *adaptive data compression*. In self-organized maps data are stored in a more compact form, preserving all the most important information, so that the recalled data include as little error as possible comparing to the original information. Usually we introduce a measure of recall-error which should be minimized in order to achieve maximal data compression.

A suitable measure is the sum of all differences between the input-patterns  $\vec{x}^k$  and the target- or prototype-patterns  $\vec{w}_{\vec{r}}$ . By introducing a probability density [27] of input-patterns  $P(\vec{x}^k)$  we can write our measure as an expectation value  $E$  of the square error:

$$E(\vec{w}) = \sum_{k=1}^p \|\vec{x}^k - \vec{w}_{\vec{r}}\|^2 P(\vec{x}^k). \quad (10)$$

We want to minimize this expression in order to compress data. This procedure is called *optimal*

*vector quantization*. The best way to find the minimum of  $E$  is *gradient descent*, i.e. movement in the vector-space along a line of maximal reduction of  $E$  according to change of prototypes  $\vec{w}_{\vec{r}}$ :

$$\begin{aligned} \vec{w}_{\vec{r}}(t+1) &= \vec{w}_{\vec{r}}(t) - \frac{\epsilon}{2} \frac{\partial E}{\partial \vec{w}_{\vec{r}}} = \\ &= \vec{w}_{\vec{r}}(t) + \epsilon \sum_{k=1}^p (\vec{x}^k - \vec{w}_{\vec{r}}(t)) P(\vec{x}^k). \end{aligned} \quad (11)$$

This equation, obtained by inserting (10), is used for every prototype  $\vec{w}_{\vec{r}}$ .

We can simplify equation (11) if the probability distribution  $P(\vec{x}^k)$  is not known explicitly what is often the case:

$$\vec{w}_{\vec{r}}(t+1) = \vec{w}_{\vec{r}}(t) + \epsilon(\vec{x}^k - \vec{w}_{\vec{r}}). \quad (12)$$

Here we have focused our attention only to the adjustment of the prototype  $\vec{w}_{\vec{r}}$  which matches mostly with the *current* input-pattern  $\vec{x}^k$ . Input-patterns are presented successively during learning and in each step (12) is used.

Equation (12) is equivalent to equation (7). In order to get this generalization we would only have to add the neighborhood function  $h_{\vec{r}\vec{r}'}$ . In the most general case which includes the neighborhood function and probability distribution we can write the error function as a functional

$$E' = \sum_{\vec{r}\vec{r}'} h_{\vec{r}\vec{r}'} \sum_{k=1}^p \|\vec{x}^k - \vec{w}_{\vec{r}}\|^2 P(\vec{x}^k). \quad (13)$$

Dimension of vectors  $\vec{x}^k$  and  $\vec{w}_{\vec{r}}$  is  $n$ .  $k$  is the ordinary number of an input-pattern which is to be learned or classified, respectively.

It is a special characteristics of Kohonen maps that it is, say, three times more dense map (three times more neurons) formed where there is three times higher input-probability-distribution  $P(\vec{x}^k)$ . So, we get a better map-resolution corresponding to those receptive fields where there are more stimuli.

## 2.5 Principal component analysis

A very important way for feature extraction is the principal component analysis of the input-data [82,

9]. Input-patterns  $\vec{x}^k$  can be written as a series

$$\vec{x}^k = \vec{w}^0 + \sum_{r=1}^{p'} \vec{w}^r c_r(\vec{x}^k) + \vec{R}(\vec{x}^k). \quad (14)$$

$c_r = \vec{w}^r \vec{x}^k$  are the *principal components* of the input-patterns. Totally there are  $p'$  of them. They could also be treated as the strengths of activity of corresponding cardinal neurons  $\vec{r}$ . Index  $r$  corresponds to location of a cardinal neuron  $\vec{r}$ . At the beginning all  $c_r$  are approximately equal and lower than 1. Then we talk about potentially-cardinal neurons. Later environmental stimulus gives privilege to one pattern and its  $c_r$  increases towards 1 (wins). This means that one neuron becomes actually-cardinal, other neurons get subordinated. We are able to store all inputs  $\vec{x}^k$  completely (so, there is no need for data compression) if  $p = p'$ . In this “ideal” case index  $k$  and index  $r$  are equivalent. If, on the other hand, there is  $p > p'$ , data-compression causes that a higher number ( $p$ ) of inputs  $\vec{x}^k$  is represented by a lower number ( $p'$ ) of cardinal neurons  $c_r$  or/and their set of synapses  $\vec{w}_{\vec{r}}$  (or  $\vec{w}^r$ , respectively).

$\vec{w}^0$  specifies the center of weight of the input distribution  $P(\vec{x}^k)$ . Other  $p'$  prototype-vectors  $\vec{w}^r$  form a *basis of the eigenspace*. These  $p'$  vectors are eigenvectors of the *autocorrelation-matrix*  $\mathbf{C}$  that have the *largest eigenvalues* [82]:

$$\mathbf{C} = \sum_{k=1}^p (\vec{x}^k - \vec{w}^0) \otimes (\vec{x}^k - \vec{w}^0)^T P(\vec{x}^k). \quad (15)$$

$T$  denotes the transpose vector;  $\otimes$  denotes the outer or tensor product:  $(\vec{a} \otimes \vec{b})_{ij} = a_i b_j$ .

Equation (14) defines a hyperplane which passes through the center of weight  $\vec{w}^0$  and is spanned by principal axes along all  $\vec{w}^r$ .  $\vec{R}(\vec{x}^k)$  is a residual vector which represents a non-vanishing distance from the approximating hyperplane perpendicular to it. If  $\vec{R}(\vec{x}^k)$  would be zero, our approximation with principal eigenvectors  $\vec{w}^r$  (prototypes) or, equivalently, with principal components  $c_r$  (cardinal neurons corresponding to prototypes) would be perfect.

To summarize,  $\vec{w}^r$  are eigenvectors of the correlation- or covariance-matrix  $\mathbf{C}$  with the largest

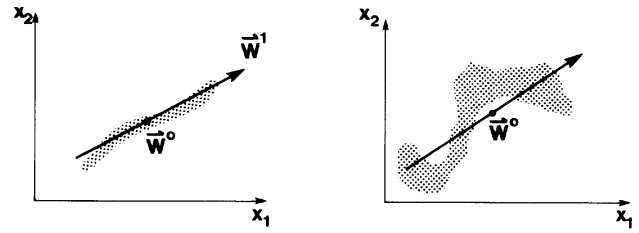


FIG. 11. Approximation of a two-dimensional input-distribution  $P(\vec{x}^k)$ , ( $k = 1, 2$ ) with the center of weight  $\vec{w}^0$  by a principal axis along  $\vec{w}^1$ . Left: linear case. Right: non-linear case where a principal curve would fit the data better than a principal axis.

eigenvalues  $\lambda_r$  [27]:

$$\sum_j C_{ij} w_j^r = \lambda_r w_i^r. \quad (16)$$

$c_r$  are *projections of the input-patterns  $\vec{x}^k$  along the principal axes* defined by the eigenvectors  $\vec{w}^r$ . In each learning cycle, following each input, one potentially-cardinal neuron wins the competition, others are suppressed. This is equivalent to increase of one principal component  $c_{r_0}$  to 1, and decay of other projections  $c_r$  ( $r \neq r_0$ ) onto corresponding eigen-patterns  $\vec{w}^r$  ( $r \neq r_0$ ). After the next input another principal component may get privilege to dominate.

Self-organized topology-preserving maps are actually generalizations of the principal-axes-calculations [82]. Namely, they use (instead of linear principal axes and linear eigenspaces) *principal curves* or *principal surfaces* (figure 11).

## 2.6 Concluding remarks on Kohonen's compressed encoding

We have shown that the self-organizing mapping compresses input information using non-linear projection of input-patterns onto a lower-dimensional prototype-space. This is done in the most efficient way by minimizing the input-target-differences and by transforming similarity-relationships of the input-data into spatial neighborhood-relations of the cardinal neurons which represent prototypes. Principal component analysis was presented as a

special method of feature extraction and data compression.

A special case of mapping is search for an one-dimensional map. This mapping brings approximate solution to the so-called traveling salesman problem – how to connect many points with a shortest trajectory [82, 30]. Kohonen solution of this classic optimization task is a Peano curve, a zig-zag curve which homogeneously covers almost the whole receptive field, and is sometimes fractal. *DNA could be modeled as an one-dimensional map, and its global encoding and decoding<sup>1</sup> processes could be approximated by the Kohonen self-organization.*

All discussed approaches are used for modeling perception and motor action in living organisms. Connection vectors and cardinal neurons are to a very large extent not genetically prespecified [31]. They evolve (like the DNA itself also, but much more gradually) under the selective influence from environment.

Present model is, of course, very important in neurophysics, but one can expect that it will have significant impact on physics in general. It is namely an algorithm for dimension reduction in a complex system, and explains how from self-organizing parallel-distributed processes *localized structures* arise.

### 3 Higher brain functions and consciousness: Synergetic neural networks and neuro-quantum coherence

#### 3.1 Synergetic conglomerates of patterns-qua-attractors: Simultaneous operation at all levels

In the first part we have presented the most important aspects of cognitive processes at a micro-level using symmetric associative neural networks, which

organize themselves into functional and/or virtual hierarchies. Processes such as pattern recognition and storage, associations, adaptation to novel stimuli, extraction of the most relevant informational content, categorization, generalization, etc., can be modeled successfully using Hopfield neural networks (and their generalizations) which were presented in the first part.

In the second part Kohonen-type feed-forward neural network topologies were discussed which realize these functions in another way. Symmetric associative neural nets in the tertiary associative areas of the neocortex could not realize efficient higher-order associative and cognitive information processing without *feature extraction* in the primary cortical areas and *feature integration* in the secondary cortical areas [60]. The preprocessing of input data, which is highly non-linear, is realized by Kohonen-type networks of functional hierarchy.

In a Kohonen neural network neurons interact under the constraint of lateral inhibition in order to achieve *specific responses of individual, localized neurons to specific stimuli*. Such specialization of neurons *emerges* as a consequence of underlying inter-neuronal interaction dynamics from 'competition for an information-representing role'. 'Winners' in this 'competition' are called *cardinal neurons*. Thus, the Kohonen neural network provides a mode of *perceptual preprocessing* as well as triggering motor action [82, 70, 48]. For association areas of the neocortex, our Hopfield-type neural networks provide additional structure, accompanied by virtual attractor-hierarchies, which is more appropriate for higher mental processing once the initial stimulus analysis, primitive categorization and abstraction processes have been realized and carried out by suitable Kohonen nets.

In this part we will extrapolate presented theories, which were to a large extent tested also by author's computer simulations [73] and were compared with experimental (physiological) data, in order to model higher-brain functions and processual basis of consciousness. Various associative, intuitive and even semantic processes could be approximated

<sup>1</sup>During embryonal development DNA's genetic information is decoded in the process of continual external and internal environmental influences.

with models presented here; higher symbolic, linguistic, syntactic, logical, inferential processes, on the other hand, could not be modeled using these models alone, but should be connected with symbolic models of classical artificial intelligence [79] in order to make a hybrid model. Semantic and conceptual networks, language structures, and speech production and recognition will not be presented here.

There are theoretical and experimental indications (the last ones may perhaps be, for some scientists, still controversial) that consciousness is related with quantum systems [23, 14, 76]. Such ideas were theoretically introduced by Wigner, von Neumann, Heisenberg, and later developed by Bohm, Hiley, Stapp [91], Josephson, Penrose [71] and others. The reasons for such a hypothesis, which receives increasing support in recent years (Tucson multi-disciplinary conferences [42]), will be discussed.

We feel that it is convenient to approach to the research of quantum information processing with a method of comparison of quantum and neural network mathematical formalisms. Information processing capabilities of the neural network models were intensively and successfully tested in computer simulations. Here their mathematical formalism will be compared to the quantum formalism in order to find similar, but (hypothetically) additionally conscious, information processing capabilities of quantum systems.

*Haken's synergetic model* unites associative neural nets and multi-layer neural nets. A multi-level (originally three-level) synergetic network with intra-layer and inter-layer connections is established [40, 39, 38]. Each layer is actually its own associative network which can have functional (*cardinal cells* in the second layer) or virtual interpretation (*order parameters*  $c_k$  – cardinal domains or neuronal oligarchies in the second layer, or patterns  $\vec{v}_k$  in the third layer). In Haken's network order parameters  $c_k$  measure the rate of *overlap* of a pattern  $\vec{v}_k$  with the actual network-state  $\vec{q}$ . So,  $c_k$  is the *projection* of  $\vec{v}_k$  onto  $\vec{q}$ :

$$c_k = \sum_{i=1}^N v_k^i q^i = \langle \vec{v}_k, \vec{q} \rangle.$$

( $i$  is here the index of vector's components and  $k$  the pattern index.) The order parameter  $c_k$  acts as a coefficient in the series  $\vec{q} = \sum_{k=1}^P c_k \vec{v}_k$ .

In a special case of localized feature representation (as presented in the second part) the virtual order parameter  $c_k$  quantitatively represents the strength of a *cardinal neuron* (a “dictator”-neuron) and its corresponding pattern of neuronal population.

Such a synergetic network is probably the most plausible model available for large-scale modeling of (*neo*)*cortical functions* [75]. The neocortex has physiologically relatively uniform structure [28] and his enormous cognitive capabilities arise from collective virtual processes which are still beyond the methods of cognitive neuroscience [32, 19].

In modeling higher brain functions, associative and synergetic neural networks may be used in a generalized way, i.e., so that a generalized interpretation of neurons is given. Generalized “neurons” may be *cortical columns* or mini-columns, patterns-attractors of various orders  ${}^{(n)}\vec{v}$  ( $n^{\text{th}}$  virtual level), etc. Generalized “connections” may be large-scale physiological or virtual connections between cortical areas, etc. [20].

As shown in figure 12, *patterns-qua-attractors are organized into a virtual hierarchy*. Their double nature is shown: they are patterns (figure 12: left “boxes”) as well as attractors (figure 12: right “boxes”). Higher-order patterns contain lower-order patterns. A large hierarchy of patterns is a very flexible, mobile as well as sensitive and continually-changing structure. It consists of a set of levels in two senses: firstly in the sense of various *functional levels* (as in feed-forward nets), and secondly in the sense of various *virtual levels of abstraction*, or meaning, generality, and composition [75]. We usually refer to these virtual levels of structure in the following terms:

- neurons (first biological level)
- patterns (second level – first virtual level, generalized neurons)
- higher-order patterns, schemes, categories, collections of patterns, groupings of patterns; meta-

representations, symbols

- dynamic pattern-sequences: associative chains, episodes, trains of thought
- pattern manifolds: combinations of higher patterns of various types and origins with some common feature; *semantic, symbolic or conceptual networks*
- global attractor conglomerates (personality, ego), consciousness (in multi-level interaction with sub-cellular and quantum systems)

Elements at each level are gestalts of lower-level elements.

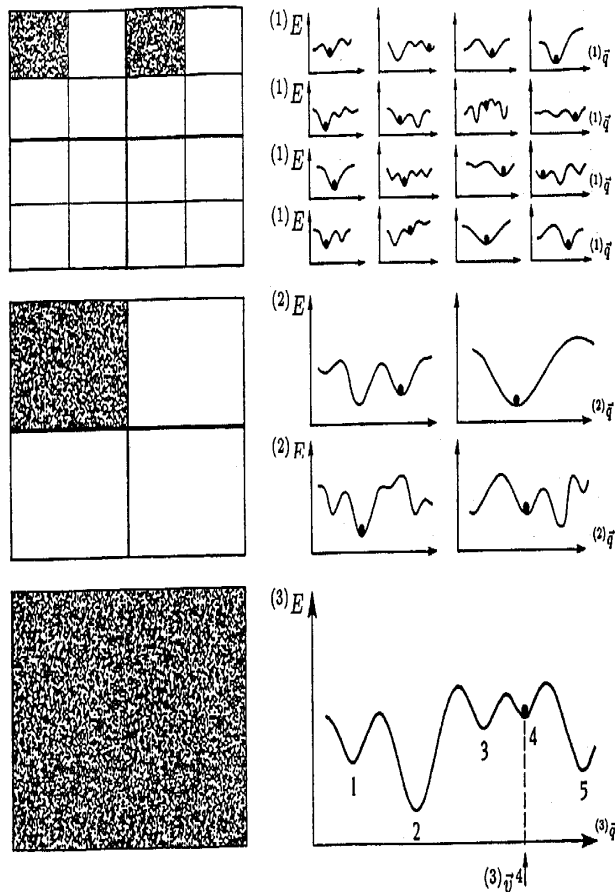


FIG. 12. Virtual hierarchy of patterns-qua-attractors. Left: their configurations; right: their locations in configuration-energy space (in the moment of recognition).

A neural net with such a structure of virtual patterns operates and restructures itself simultaneously

at all levels, so that these structures consolidate optimally with each other in a self-organized and interactive manner.

*Dynamical pattern-sequences* are a special case of higher-order pattern-conglomerate. They arise if neurons (e.g., in Hopfield's model) are interconnected asymmetrically. This means that the influence of one neuron on another is different from the influence of the second neuron on the first, because of different transmission rate in different directions:  $J_{ij} \neq J_{ji}$ . This results in *delays in processing* which gives rise to attractors moving through the configuration space. The system forms at time  $t_1$  a pattern  $\vec{v}_1$ , at time  $t_2$  forms pattern  $\vec{v}_2$ , etc. Potential wells are conserved, but their minima are displaced from configurations which lose stability to other configurations. These sequences of patterns or *episodes* are actually associative chains, since the predecessor pattern is an immediate causal precursor to the successor pattern [70, 72].

Ordered and directed sequencing triggers self-organized formation of a causal process within the framework of a larger parallel-distributed process. A neural net having asymmetric connections causes such inhomogeneities which form a more constant potential gradient, along which the system descends faster from one configuration to another. Such emerging sequential processes have a larger associative context through its connections with other patterns which are within the range of attraction of the observed dynamic pattern. If individual patterns, which are transitive elements of such pattern sequences, are connected with their cardinal cells, or corresponding order parameters, in speech centers (Wernicke area), then such a train of thought is encoded or *symbolized*, and it is possible to verbalize it (Broca area) [60].

We will only mention the last essential way of binding patterns into a complex pattern or an informational unity – coherent oscillations [39, 88]. The frequencies and also phases of oscillation of neuronal activities may get locked: oscillating activities of neurons or whole assemblies are *synchronized* into the *same frequency* and the *same phase*

or at least constant phase difference. Such ordered collective oscillating behaviour is called *coherence*. If patterns of neuronal oscillations are frequency-locked and phase-locked, then they have a *common feature*, or something in common in more abstract sense, and represent an informational unity – a category of higher order.

### 3.2 Synergetic levels in brain: neural, sub-cellular, quantum and virtual processes

In a neural network neurons are constantly interacting, and whole neuronal patterns are interacting. Neuron's activities reflect each other. But it seems that also a very large and complex neural network with its recursive dynamics is not enough for the "real" consciousness as a global and unified self-reflective process [42]. Neuronal patterns merely represent the *objects* of consciousness and take part in information processing concerning these objects, but these patterns do not represent consciousness itself. It is a question how many different perceptions are bound into an unified multi-modal experience where there are no traces of neuronal signaling and other material processes any more. Only a clear, relatively static qualitative experience remains.

There are indications that consciousness is connected with quantum phenomena [71, 91]. The main reasons for this hypothesis are the following:

- Quantum systems are the microscopic basis of all physical processes, including biological or psychophysical processes. All the classical world arises from the overall quantum background.
- Quantum systems transcend even the division of particles and waves, or interactions, or fields [23, 17]. Quantum systems, especially sub-quantum systems, are holistic in nature [16] – they cannot be satisfactorily analyzed into interacting fundamental elements, but act synthetically as indivisible parallel-distributed processes. As such, they are good candidates for correlates of the unity of consciousness.
- Neural networks with their rigid neurons and synapses, in spite of their very subtle virtual pro-

cesses [75], seem to be too mechanistic, too discrete and too deterministic for consciousness and phenomenal qualia, i.e., qualitative perceptual experiences.

On the other hand, *all thought-processes, including consciousness, seem to arise from complex-system-dynamics*. Objects of consciousness and stream of conscious thought seem to be represented in some physical or at least informational (virtual) "medium". That "medium" has to be a complex-system which only is enough flexible, fuzzy, adaptive, and has good self-organizing and recursive abilities. Because the *mathematical formalism of the neural network theory is confined to the collective system-dynamics, it remains to a large extend valid also for complex systems of other basic elements* [74]. So, our formal neurons and formal synaptic connections are not necessary biological neurons and synapses (or, in models, artificial neurons—processors). There are various synergetic biological systems which may be modelled in a neural-network-like way and have relevant roles in micro-cognition processes on various levels:

- *dendritic trees* where neuronal dendrites have similar summation-tasks to that of a neuron [62, 39](Pribram);
- subcellular structures: *cytoskeleton*, especially *microtubules*, and other biomolecular systems of electric or magnetic dipoles which may have a role of an interface between the neural and the quantum level [41, 71];
- webs of quantum particles (e.g., electrons) with their *spins* [92, 93];
- a "continuum" of *sub-quantum "beables"* (Bell) or "*hidden variables*" [16, 44]; etc.

### 3.3 Neuro-quantum coherence needed for consciousness

*Consciousness itself ("pure" consciousness without an object of consciousness)* may be associated with the quantum field, or better, with the "overall sub-quantum sea" [33]. On the other hand, *consciousness having some object as its content can-*

not be associated merely with a specific quantum-informational state. If a specific mental representation is processed under control of consciousness, this specific representation, which is associated with a pattern of neural activity, is coupled or correlated with a specific quantum eigen-wave-function. This quantum eigenstate was made explicit by the "wave-function collapse" which was triggered by corresponding neural processes.

The "wave-function collapse" is a transition of the quantum state from a state described by a linear combination of many quantum eigenstates to a "pure" state which is one eigenstate only. So, a superposition of many "quantum patterns" is transformed into a single "quantum pattern" only.

Collapse of the wave-function means *a selective projection from subconscious memory to the conscious representation* which was explicated from the memory. There are two possible versions of memory and memory-recall: the quantum one (just mentioned), or the classical neural one. In the first case memory may be a parallel-distributed pattern in the system of synaptic connections. In the second case memory could also be realized as a quantum hologram, as a "pattern" in the implicate order of Bohm and Hiley [17], or even as a set of "parallel worlds" according to the many-world interpretation of quantum theory by Everett [26].

Brain-mind is necessarily a multi-level phenomenon, because we cannot totally divide pure consciousness from the object of consciousness which may be an internal virtual image or a real external object. We can then constitute the following overall scheme: Pure consciousness is of quantum nature; virtual representations are associated with neuronal patterns; external objects are of classical nature. Making a synthesis of these levels, *neuro-quantum coherence* is constituted as a basis for consciousness-about-some-object-of-consciousness.

Quantum mechanics governed by the Schrödinger equation doesn't make attractors manifest until the "wave-function collapse" occurs. In that case, because of the interaction of a classical macroscopic system (either measurement apparatus, or environ-

ment, or our sensory apparatus) with the quantum system, the wave-function "collapses" and a specific quantum eigenstate (a "quantum pattern") occurs as an attractor. So, there are quantum virtual structures also, and they cannot be reduced to a quantum eigenstate alone, because they occur only as a result of interaction with a classical system. Thus quantum virtual structures are (re)constructed as a result of so-called quantum measurement where the "measurement apparatus" may be our sensory and associative neural system directly, or a machine which is observed by that neural system. In both alternatives the "wave-function collapse" occurs as a result of a specific interaction with a classical system. The probability of the "collapse" is very much higher if the interaction is *knowledge*-based [74]. That's like in the case of a radio: if we *know* the right frequency, we are able to receive the associated information.

### 3.4 Comparison of the mathematical formalism of associative neural network theory and quantum theory

#### Main functional analogies and their informational significance

New consciousness studies give us reason to investigate parallels between quantum processes and neural-network-processes. Many mathematical analogies of the theory of associative neural networks and the quantum theory can be found. Because we know that neural network simulations perform well, we can search for similar effective cognitive-like information-processing capabilities on the quantum level also. Let us make an overview of neuro-quantum analogies.

#### 1. NEURONAL-STATE-VECTOR $\longleftrightarrow$ QUANTUM WAVE-FUNCTION:

In neural network theory the state of the system of neurons is described by  $q(\vec{r}, t)$  which denotes the activity of an individual neuron (located at  $\vec{r}$ ) at time  $t$ . *Neuronal patterns*  $\vec{v}$  are special neuronal configurations  $\vec{q}$  which represent some

meaningful information. In quantum theory the state of the quantum system at location  $\vec{r}$  and time  $t$  is described by the wave-function  $\Psi(\vec{r}, t)$  [15]. They both represent a state vector describing a parallel-distributed configuration of a complex system.

## 2. NEURONAL STATE IS A SUPERPOSITION OF NEURONAL PATTERNS $\longleftrightarrow$ QUANTUM WAVE-FUNCTION IS A SUPERPOSITION OF QUANTUM EIGENSTATES:

A neuronal configuration  $\vec{q}$  may be described as a linear combination of neuronal patterns  $\vec{v}_k$  ( $k = 1, \dots, p$ ).  $p$  is the number of patterns represented simultaneously in the combination. Similarly, a wave-function  $\Psi$  can be described as a linear combination of eigen-wave-functions  $\psi_k$  ("quantum patterns"):

$$q(\vec{r}, t) = \sum_{k=1}^p c_k(t) v_k(\vec{r}); \quad (17)$$

$$\Psi(\vec{r}, t) = \sum_{k=1}^p C_k(t) \psi_k(\vec{r}). \quad (18)$$

Neuronal patterns and eigen-wave-functions can represent some object of consciousness on different levels. So, as opposed to other configurations, they represent informational states which have a meaning, because they are correlated with some environmental objects.

3. Both sets of vectors,  $\vec{v}_k$  and  $\psi_k$ , usually have the properties of ORTHOGONALITY and NORMALITY. The first property, if satisfied, ensures error-free recall of patterns, because it excludes mixing between patterns or corruption of a pattern because of this. So, the interference between patterns must be destroyed during recall in order to get one of them out properly. However, interference must be eliminated during pattern-recall only; in associative memory interference between patterns is essential. If eigenstates  $\vec{v}_k$  and  $\psi_k$  are not orthogonal completely, recall is not so pure, but associations between patterns may become more

effective.

## 4. COEFFICIENTS OF THE SERIES: SYNERGETIC ORDER PARAMETERS $\longleftrightarrow$ QUANTUM PROBABILITY COEFFICIENTS:

$C_k$  are the quantum probability coefficients, and  $c_k$  are the neural or synergetic order parameters. In the linear combination each pattern is represented by a corresponding coefficient. The coefficients describe how much a specific pattern is represented in the actual state of the system, or how probable it is that the corresponding pattern will be recalled (reconstructed, or explicated from the superposition of many patterns). Thus, the time-dependent coefficients encode quantitatively the *meaning* of their patterns. They describe how strong a role a given pattern has in contextual system-dynamics.

$$c_k(t) = \langle \vec{v}_k, \vec{q} \rangle = \int \int v_k(\vec{r})^* q(\vec{r}, t) d\vec{r} dt; \quad (19)$$

$$C_k(t) = \langle \psi_k, \Psi \rangle = \int \int \psi_k(\vec{r})^* \Psi(\vec{r}, t) d\vec{r} dt. \quad (20)$$

Asterisk denotes the operation of complex conjugation. If variables  $\vec{v}_k$  or  $\psi_k$  are real, we may delete the asterisk. Mathematically, coefficients  $c_k$  and  $C_k$  describe projections (in terms of scalar products  $\langle \cdot, \cdot \rangle$ , in Hilbert space, for example).

## 5. SPATIO-TEMPORAL INTEGRATION OF NEURONAL SIGNALS $\longleftrightarrow$ FEYNMAN'S VERSION OF THE SCHRÖDINGER EQUATION:

The dynamical equation for one neuron includes a spatio-temporal integration of signals from all other neurons which are connected to that neuron. Neurons integrate along spatio-temporal trajectories, not separately in space and time [11]. So, the state of a neuron at position  $\vec{r}_2$  and time  $t_2$  is given by  $J$ -weighted summation of all signals and the whole history of signals<sup>1</sup>:

$$q(\vec{r}_2, t_2) = \int \int J(\vec{r}_1, t_1, \vec{r}_2, t_2) q(\vec{r}_1, t_1) d\vec{r}_1 dt_1 \quad (21)$$

<sup>1</sup> $t_2$  is always greater than  $t_1$  – because of the "arrow of time".



where  $J(\vec{r}_1, t_1, \vec{r}_2, t_2)$  is the strength of an individual synaptic connection.  $J$ 's are the transmissions of real synaptic connections between two neurons (spatially separated, but realized at the same time) or correlations between states of different neurons at different times, represented at least virtually.

In non-relativistic quantum mechanics the dynamical equation is the Schrödinger equation. Here it is written, analogously to the neural equation (21), in Feynman's form [44]:

$$\Psi(\vec{r}_2, t_2) = \int \int G(\vec{r}_1, t_1, \vec{r}_2, t_2) \Psi(\vec{r}_1, t_1) d\vec{r}_1 dt_1. \quad (22)$$

$G(\vec{r}_1, t_1, \vec{r}_2, t_2)$  constitutes the Green function or an element of the quantum propagator [13]. The propagator  $G$  is a matrix which describes a parallel-distributed transformation of the whole system from an initial state  $\Psi(\vec{r}_1, t_1)$  to the final state  $\Psi(\vec{r}_2, t_2)$ . It is essentially a sort of Huygens' principle. The system transforms itself into a new state by exhibiting numerous internal interactions between its constitutive "quantum points" (some mathematical "basic elements" of the system). Informationally, transformation (21) or (22) is *association*. One pattern (concept, idea) is transformed into another pattern (concept, idea) by a collective informational interaction of input data (encoded in the state vector, i.e.  $\Psi$ ) with memory data (encoded in the system of connections or interactions described by the Green matrix  $G$ ).

#### 6. MEMORY IN AUTO-CORRELATIONS OF PATTERNS: SYNAPTIC STRENGTHS OBEYING THE HEBB RULE $\longleftrightarrow$ GREEN FUNCTIONS:

The kernels of dynamic equations (21) and (22) are given as a sum of *auto-correlation* functions. The transmission-rate of an individual synaptic connection  $J$  is determined by the Hebb learning rule as a sum of auto-correlations between its two neurons participating in various patterns  $\vec{v}_k$ :

$$J(\vec{r}_1, t_1, \vec{r}_2, t_2) = \sum_{k=1}^p v_k(\vec{r}_1, t_1) v_k(\vec{r}_2, t_2)$$

$$\text{or} \quad J(\vec{r}_1, \vec{r}_2) = \sum_{k=1}^p v_k(\vec{r}_1) v_k(\vec{r}_2). \quad (23)$$

Similarly, the *Green function* [13] is given as a sum of auto-correlations of individual quantum eigenstates  $\psi_k$ :<sup>1</sup>

$$G(\vec{r}_1, t_1, \vec{r}_2, t_2) = \sum_{k=1}^p \psi_k(\vec{r}_1, t_1)^* \psi_k(\vec{r}_2, t_2)$$

$$\text{or} \quad G(\vec{r}_1, \vec{r}_2) = \sum_{k=1}^p \psi_k(\vec{r}_1)^* \psi_k(\vec{r}_2). \quad (24)$$

#### 7. ADAPTIVE MEMORY IN CROSS-CORRELATIONS: SYNAPTIC STRENGTHS OBEYING THE GENERALIZED HEBB RULE OR DELTA RULE $\longleftrightarrow$ DENSITY MATRICES:

The kernels of dynamic equations (21) and (22) can be written also as sums of *cross-correlation* functions. In special cases they can be written as cross-correlation functions of a pattern and a difference between two patterns. In neural network theory we usually use cross-correlation of a *novel pattern*  $\vec{v}_d$  with a *difference or discrepancy of this novel pattern*  $\vec{v}_d$  and a *desired or average pattern*  $\vec{v}_l$  (or  $\vec{v}$ ) where  $\vec{v}_l$  (or  $\vec{v}$ ) represents old knowledge [54, 56]. A Hebb-like learning rule which is based on cross-correlation of a (novel) pattern  $\vec{v}_d$  and a difference  $\vec{v}_d - \vec{v}_l$  is called *delta learning rule*. It is used for novelty filtering: new information, not represented in the old neural pattern  $\vec{v}_l$ , is extracted and stored newly. Therefore the network finally adapts, i.e., optimally condenses and combines components of new information ( $\vec{v}_d - \vec{v}_l$ ) with the whole knowledge collected up to now. The generalized Hebb learning rule using cross-correlations of patterns is given as follows.

$$J(\vec{r}_1, \vec{r}_2, t) = \sum_{k=1}^p \sum_{h=1}^p \lambda_{kh} v_k(\vec{r}_1, t) v_h(\vec{r}_2, t)$$

$$\text{or} \quad J(\vec{r}_1, \vec{r}_2) = \sum_{k=1}^p \sum_{h=1}^p \lambda_{kh} v_k(\vec{r}_1) v_h(\vec{r}_2). \quad (25)$$

<sup>1</sup>Many physicists use another definition of the Green function, and there is an  $i$  ( $i = \sqrt{-1}$ ) in front of integrals of equation (6) in their case. The relation is:  $G_{\text{their-def}} = i G_{\text{our-def}}$ .

This learning rule works also if  $\vec{v}_h$  is a difference pattern  $\vec{v}_d - \vec{v}_l$  or novelty pattern  $\vec{v}_d - \vec{v}$  (delta rule), or even if also  $\vec{v}_k$  is a difference pattern  $\vec{v}_g - \vec{v}_l$  or novelty pattern  $\vec{v}_g - \vec{v}$  (a case similar to equation (15) from section 2.5 where  $\vec{v}_d = \vec{v}_g$ ).

Let us list presented versions of Hebbian correlations once again. In the previous case (no. 6) strengths of individual synaptic connections were determined by a sum of auto-correlations (Hebb rule) of patterns, i.e., activity-correlations of connected neurons which have a role in the *same* patterns. In this case (no. 7) synaptic strengths are given by a sum of cross-correlations (generalized Hebb rule) of patterns, i.e., activity-correlations of neurons which have a role in *different* patterns. If one of these patterns is a difference pattern (difference between the first pattern and a reference or average pattern) then this generalized Hebb rule is called delta rule. So, in the first case (no. 6) a synapse represents correlation of neurons constituting the same pattern, therefore we talk about pattern-auto-correlation, i.e., correlation of parts of the same pattern. In the second case (no. 7) a synapse represents correlation of neurons constituting two different patterns, therefore we speak about cross-correlation or about correlation between parts of two different patterns.

Similarly to the present case (no. 7) of neural cross-correlation in equation (25), the *density matrix*  $\rho$  is given as a sum of cross-correlations of individual quantum eigenstates  $\psi_k$  and  $\psi_h$ :

$$\rho(\vec{r}_1, \vec{r}_2) = \sum_{k=1}^p \sum_{h=1}^p \rho_{kh} \psi_k(\vec{r}_1)^* \psi_h(\vec{r}_2). \quad (26)$$

Comparing equations (25) and (26), and knowing that the generalized Hebb rule for  $J$  realizes efficient associative memory, we may infer about information processing capabilities of  $\rho$  – the quantum probability density matrix or quantum statistical operator, respectively.<sup>1</sup> Just note also that

<sup>1</sup>Of course, the density matrix itself does not describe (or even execute, in simulations) the information-processing input-to-output transformation in such a way as propaga-

the so-called diagonal parts of the density matrix  $\rho(\vec{r}, \vec{r}) = \sum_{k=1}^p \rho_{kk} |\psi_k(\vec{r})|^2$  give a superposition of real-valued and thus measurable quantum probability densities with “weights”  $\rho_{kk}$ .  $\rho_{kh}$  is analogous to  $\lambda_{kh}$  which acts as the so-called learning rate  $\lambda_r$  (see equation (16) and [40]).

Interaction scheme no. 6 realizes auto-association: completion of a neural pattern  $\vec{v}_k$  or quantum pattern  $\psi_k$  from partial information. On the other hand, interaction web no. 7 realizes hetero-association: association from one pattern  $\vec{v}_k$  (or  $\psi_k$ ) to another pattern  $\vec{v}_h$  (or  $\psi_h$ ). Similar quantum information processing using  $\rho$ -decomposition was discussed by Ohya and Matsuoka [69, 65]. They related it to the Ising model.

8. NEURONAL-PATTERN-RECONSTRUCTION  $\longleftrightarrow$  ”WAVE-FUNCTION COLLAPSE”: The most important neuro-quantum analogy is the following. Pattern-reconstruction in a neural network

$$\begin{aligned} q(\vec{r}, t) &= \sum_{k=1}^p c_k(t) v_k(\vec{r}) \\ \implies q(\vec{r}, t_0) &= v_{k_0}(\vec{r}) \end{aligned} \quad (27)$$

is very similar to the ”collapse of the wave-function” in a quantum system

$$\begin{aligned} \Psi(\vec{r}, t) &= \sum_{k=1}^p C_k(t) \psi_k(\vec{r}) \\ \implies \Psi(\vec{r}, t_0) &= \psi_{k_0}(\vec{r}). \end{aligned} \quad (28)$$

For reasons presented in section 3.3 this is a very important feature of cognitive processes. Processes (27) and (28) are both a result of the influence from

tors or Green functions do. They are active-memory descriptors, but the density matrix is a “passive” description (like an “archive”) of statistical and correlational properties of quantum states in the present and history of the system. However, the density matrix may represent an important quantitative description of relations of eigenstates (patterns) which are crucial if a quantum system is harnessed for information processing, or here specifically for cross-correlations or hetero-associations. Authors like Alicki, Ohya and Matsuoka have theoretically described quantum information transmission using density matrices and their tensor products.

the system's *environment*. The environment selects those neural/quantum pattern which is the most similar (or is correlated) to the state of environment.

Neural-pattern-reconstruction and wave-function collapse are results of a transition from the *implicate order* (Bohm's term for latent, implicit, inactive, potential non-Shannon information only) to the *explicate order* (manifest, active, realized Shannon information) [17]. The implicate order represents a combination of very many *possible* states or processes (left-hand side equation in (27) and (28)). It is analogous to a set of so-called "parallel worlds" or parallel sub-branches of the general wave-function proposed by Everett [26]. Explicate order, on the other hand, represents a state or process which is at a moment physically actualized (right-hand side equation in (27) and (28)). It is "chosen" from a set of potential (implicate) states, or is a result of their optimal "compromise". "Taking choice" may be an "act of consciousness" which takes some information out from memory or subconsciousness (implicate order), and makes it explicit and conscious.

## 9. NEURAL DUALITY AND QUANTUM DUALITY:

An interesting neural analogy of the Heisenberg uncertainty principle of quantum mechanics is represented by an *inability of simultaneous determination of patterns in the system of neurons  $\vec{q}$  and of patterns in the system of interactions or connections  $\mathbf{J}$* . We are unable to make a pattern explicit in a system of neurons, and to make a pattern explicit in the system of connections at the same time. Only one pattern, which is at the moment realized in the system of neurons, is made explicit. All the others are present only implicitly in the system of interactions, i.e., in the coupling-rate. In cognition we have a similar situation: we can be aware of one pattern only which has been extracted from memory. Other memory patterns remain unconscious and implicit.

The so-called *position ( $x$ -) representation* of quantum theory can be approximated by the system

of neurons  $\vec{q}$ . The so-called *impulse ( $p$ -) representation* can, on the other hand, be associated with the system of interactions  $\mathbf{J}$  which regulates all transformations of the network-state. Thus, the quantum duality (particles/waves, or particles/interactions, or fermions/bosons, respectively), has a counterpart in neural duality (neurons/synaptic connections, or neurons/interactions).

Sometimes we find that a network where neurons and synapses are exchanged has a similar behaviour as the original one. Thus, in general there is a symmetry of neurons and synapses (or interactions) which represents invariance under neurons-synapses exchange. If we compare, for example, simulations of associative neural networks by this author [73] and simulations by Menneer [66]<sup>1</sup>, we see that explicit processing, which is realized in the system of neurons in this-author's case, is similar to those explicit dynamics which is realized in the system of connections in some Menneer's cases. Menneer in some cases performs "collapse" (pattern selection and reconstruction) in the systems of connections, but this author performs it in the system of neurons. She also emphasizes analogy of this collective dynamics with processes during quantum measurement. We cannot extract information in the system of neurons and in the systems of synaptic connections at the same time – not in both systems. This is a neural version of uncertainty case. We have to make a "collapse" with output information represented *either* in the system of neurons *or* in the system of synapses. Without a "collapse" the interplay of neurons and connections is always an indivisible holistic process.

Like in classical dynamics, for integrable systems, there is a complete equivalence between the Schrödinger equation for  $\Psi$  and the Liouville-von Neumann equation for  $\rho$  [81]. Both kinds of description are complementary on quantum as well as classical (e.g., neural) level.

<sup>1</sup>Comparing with [69, 65] similar symmetry-observations can be made concerning the density matrix  $\rho$  and its decompositions.

# 10. QUANTUM AND NEURAL WAVELET UNCERTAINTY PRINCIPLES:

Some more formalized uncertainty analogies can also be found. In the neural network and early vision theories we have uncertainty principles which are similar to the Heisenberg uncertainty principle. Gabor proofed [62] that the Heisenberg uncertainty principle is a special case of his general frequency/time uncertainty principle:  $\Delta f \Delta t \geq 1$ . Error in measuring frequency  $\Delta f$  is inverse proportional to the error in measuring time  $\Delta t$ . Multiplying both sides of Gabor's inequality by Planck constant  $h$ , and knowing  $E = hf$ , we obtain the well-known Heisenberg's version:  $\Delta E \Delta t \geq h$ . *Gabor uncertainty principle*  $\Delta f \Delta t \geq 1$  is applicable whenever simultaneous measurements of an observable and of its Fourier transform are made [62]. Previously mentioned  $x/p_x$  and  $E/t$  Fourier or uncertainty pairs of conjugate variables are thus similar to the  $f/t$  pair of Gabor's information theory and  $x/u_x$  pair of Daugman's vision theory. In Daugman's theory [22] spatial frequency  $u_x$  determines the wavelet's location in frequency domain so as  $x$  determines location in ordinary spatial domain. It is physiologically supported that Gabor's and Daugman's theories are applicable for representing spatio-temporal visual images and information filtering in the visual cortex [22]. This is a large-scale field-theory-like approach in contrast to usual discrete neural-net-models.

The Gabor inequality changes to equality  $\Delta f \Delta t = 1$  (minimal uncertainty) if *neural wavelets* are modeled by complex oscillatory functions having Gaussian envelope. These are *Gabor functions*  $\psi_{jk}$  which constitute basic information cells represented in space and spatial frequency areas

$$\psi_{jk} = \exp\left(-\frac{\pi(t - j\Delta t)^2}{\alpha^2}\right) \exp(2\pi i k \Delta f(t - j\Delta t)). \quad (29)$$

They are similar to *quantum coherent states* or *wave packets*. An arbitrary function  $\Psi$  can be represented by elementary Gabor functions  $\psi_{jk}$  as their linear superposition  $\Psi = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} C_{jk} \psi_{jk}$ , where  $C_{jk} = \langle \Psi, \psi_{jk} \rangle / \|\psi_{jk}\|$  (if wavelets are orthogonal),

analogously to the neural-net-state expansion (17) with coefficients (19). Gabor wavelets are, however, in general not orthogonal.

Daugman [22] presented psycho-physical analogues of Gabor uncertainty principle for 2 dimensions:  $\Delta x \Delta u_x \geq \frac{1}{4\pi}$  and  $\Delta y \Delta u_y \geq \frac{1}{4\pi}$ , which give together  $\Delta x \Delta y \Delta u_x \Delta u_y \geq \frac{1}{16\pi^2}$ . MacLennan [62] naturally added Gabor uncertainty principle for 3 dimensions:  $\Delta x \Delta u_x \geq \frac{1}{4\pi}$ ,  $\Delta y \Delta u_y \geq \frac{1}{4\pi}$ ,  $\Delta t \Delta f \geq \frac{1}{4\pi}$ , which give together  $\Delta x \Delta y \Delta t \Delta u_x \Delta u_y \Delta f \geq \frac{1}{64\pi^3}$ . So, in the 3-dimensional case each information cell, representing a "quantum of information", is a 6-dimensional hypercube  $(\Delta x, \Delta y, \Delta t, \Delta u_x, \Delta u_y, \Delta f)$  *localized* to spatial area  $\Delta x \Delta y$  and to temporal interval  $\Delta t$ , and furthermore *tuned* to spatial frequency bands of width  $\Delta u_x$  and  $\Delta u_y$ , and to (temporal) frequency band of width  $\Delta f$ .

This overview of uncertainty relations of the theory of processing in the visual cortex showed that the uncertainty relations in general, including Heisenberg's quantum-mechanical version, are connected with informatics of parallel-distributed complex systems in general, even classical ones such as associative neural nets. Uncertainty relations arise as a result of *observational perturbation* of quantum systems, but it may emerge also in the case of classical (neural) complex systems. This observational perturbation has much more evident and exact consequence (such as conjugate pairs of non-commuting observable operators) in quantum complex systems, because of their microscopic scale, than in classical complex systems.

Bohm noted that if we are starting to think about our own thoughts, we change them with this self-observation act. That's uncertainty of conscious self-reflection which Bohm has related to quantum uncertainty [15].

## Discussion of discrepancies in presented analogies

There is a difference between equations (23) and (24): the latter one includes complex-valued variables which are necessary for the quantum case, but

just possible for the classical neural-net case. That prevents us from claiming that quantum systems are a very complex sort of "neural networks". Quantum systems are much more subtle and flexible, because they don't have well-defined fundamental particles and interactions. So, no well defined basic formal "neurons" or formal "connections" can be found – no well-defined "hidden variables" of Bohm [16]. Anyway, in spite of this difficulty, the mathematics of the usual Copenhagen interpretation of quantum mechanics is essentially very close to the mathematics of associative neural networks. So, the limits of our or similar useful [7] analogies are also limits of the orthodox (Copenhagen) quantum theory.

Replacement of complex-valued quantum equations with a system of real-valued equations [15, 17] or *introduction of complex numbers into the neural network theory* (e.g., networks of coupled oscillators [40, 63, 97, 90, 8], quantum neuro-holography [85, 84, 86]) would support the search for further analogies with the neural network theory. Complex-number variables in quantum formalism obviously have a very deep meaning.  $\Psi^*$  as opposed to  $\Psi$  is connected with the *phase-conjugated* or time-reversed dynamics. So, in quantum systems we have a superposition of forward and backward time-evolution. In neural networks this time-flow-superposition is transformed into superposition of associative transformations in the opposite directions in space, or the superposition is completely collapsed into single directed dynamics (time sequence of patterns in the direction of the arrow of time). Observed quantum systems reflect an intrinsic complementary duality which is evident also in the wave/particle double nature or in location and momentum representations. Probably this apparent duality, and uncertainty principles also, are a consequence of an active observer's participation in the observed quantum process.

In spite of the fact that quantum systems are described by a dual complex-valued (real plus imaginary) description, but the usual neural net models only in a real-valued description so far, we must emphasize that there are very interesting functional

and mathematical analogies between quantum and neural levels in the case of oscillatory dynamics of neurons. Specifically, if we have an associative network of coupled *neurons with oscillatory activities* [40, 97, 63] complex-valued mathematical formalism becomes relevant for neural nets also. Our presented neuro-quantum analogies, which in this case contain information about *phase-correlations*, then become much more important. In fact, John G. Sutherland has created a simulation of such a network of neurons with oscillatory activities, called "*Holographic Neural Technology*", which processes very successfully. A condensed presentation by Sutherland, entitled "Holographic model of memory, learning and expression", was published in 1990 in the *International Journal of Neural Systems*, vol. 1, pages 256-267, which is highly recommended.

### Some further analogies

#### 11. QUANTUM NEURAL HOLOGRAPHY: WAVELET-MIXING OR CONVOLUTION:

In neural networks memory consists of *(auto)correlations between patterns*. In quantum mechanics *phase differences between different parts of the wave-function* may have a similar role [15]. They control the time evolution of probability distribution involving *interference* of the contributions of different stationary eigen-wave-functions. In quantum neuro-holography we have interference patterns of coherent neural wavelets. Here *phase shift*  $\varphi$  encodes the degree of correlation between components of coherent wavelets. On the other hand, their amplitude encodes the significance of synchronized wavelet components [85]. Indeed, *wavelet mixing*  $\psi_k(t)dt \otimes \psi_h(t')dt'$ , (mapped onto) spatio-temporal *cross-correlation function*  $\int e^{2\pi i \nu \varphi t} \psi_k(t + \Delta t) \psi_h(t) dt^1$  [64], or *convolution*, etc., are examples of *bilinear forms* similar to the *generalized Hebb learning rule* used in neural networks. So, quantum holography and theory of associative neural nets (in particular, if neurons have oscillatory activities) have significant mathematical analogies which are a consequence

<sup>1</sup> $\nu$  denotes a fixed spectral parameter.

of similar collective dynamics in both complex systems.

In generalized neural network theory we have “matrices of (synaptic) connections” which encode correlations of two neural patterns. In holography, on the other hand, we have holograms which encode correlations of two quantum-optical signals, i.e., their interference. In both theories we use similar mathematics for similar information processing as long as we can neglect details in the structure of basic elements of the network.

Elements of the system are, of course, different again. This brings differences between quantum holography and associative neural nets, but it seems that brain realizes quantum-holographic-like and neural-network-like processing on different scales. So, in spite of differences in “anatomical” details, we may speak about *fractal-like multi-level synergetic processing in brain*. This is caused by similar “sociology” in networks on various biophysical levels, i.e., *quantum, sub-cellular, neural, virtual or conceptual* networks or “holograms”.

## 12. PHASE RELATIONS IN QUANTUM HOLOGRAMS $\longleftrightarrow$ STRENGTHS OF FORMAL SYNAPSES:

*Changing the phase relations between eigen-wave-functions* (e.g., quantum-optical signals) is analogous to the *learning*-process in neural networks where new pattern correlations are added to the synaptic correlation-matrix. As mentioned, conventional neural network theory does not incorporate neurons with oscillating activities, although this is observed in brain. Only some more advanced neural theories introduce oscillating neuronal activities and put relevance to phase-relations or phase-correlations between neurons with oscillating firing rates. Phase-correlations are an essential ingredient of (quantum) holography, therefore associative networks with oscillating neuronal activities are certainly more similar to holographic devices than conventional associative neural networks. If we study networks with oscillating neuronal activities, we find similar associative memory capabilities

as in conventional associative networks [40] and similar phase-coupling or frequency-coupling (coherent behaviour) as with wavelets in holography.

## 13. “COLLAPSED” QUANTUM SUPERPOSITION $\longleftrightarrow$ FORMAL NEURAL NETWORK:

Quantum holography, however, introduces something more. It gives opportunity for introduction of *complex conjugate* form of a wavelet. Complex conjugate wavelet  $\psi^*$  is a time mirrored version of the original wavelet  $\psi$ .

The well-known quantum bilinear form  $\psi\psi^*$  represents a combination of a state vector (wave function)  $\psi$  with its phase-conjugated (time-mirrored) form  $\psi^*$ . Together they give real-valued probability density  $|\psi|^2 = \rho$ , in the case of a pure state. For “mixtures” a weighted sum of  $\psi_k\psi_k^*$  has to be written [81] (in general this gives expression (26)). Thus, in quantum world we have superpositions of many possible networks, processes and “arrows of time”.

Only one network, process and “arrow of time” is, however, realized explicitly during measurement process, and this makes situation similar to the case of neural nets. To summarize roughly, a generalized (mathematical) neural network emerges as a result of “collapse” of a quantum superposition of implicit “networks”. To say as simple and short as possible: mathematical neural network (network of mathematical points or formal “neurons” which interact intensively) is a “collapsed” quantum system. Biological neural network evolves as a macroscopic replica of these fundamental dynamical principles.

14. Fourier-like CORRESPONDENCE RULES between the above equations can be shown first for the neural case (30), and then for the quantum case (31), using the following calculation.

$$\begin{aligned} & q(\vec{r}_2, t_2) \\ &= \int \int J(\vec{r}_1, t_1, \vec{r}_2, t_2) q(\vec{r}_1, t_1) d\vec{r}_1 dt_1 \\ &= \int \int \left[ \sum_{k=1}^p v_k(\vec{r}_1, t_1) v_k(\vec{r}_2, t_2) \right] q(\vec{r}_1, t_1) d\vec{r}_1 dt_1 \end{aligned}$$

$$\begin{aligned}
&= \sum_{k=1}^p v_k(\vec{r}_2, t_2) \left[ \int \int v_k(\vec{r}_1, t_1) q(\vec{r}_1, t_1) d\vec{r}_1 dt_1 \right] \\
&\Rightarrow q(\vec{r}, t) = \sum_{k=1}^p c_k(t) v_k(\vec{r}, t) \\
&\text{or } q(\vec{r}, t) = \sum_{k=1}^p c_k(t) v_k(\vec{r}) \quad (30)
\end{aligned}$$

In the last step we have used definition (19) for time-dependent patterns. For the stationary case we delete the time integral.

$$\begin{aligned}
&\Psi(\vec{r}_2, t_2) \\
&= \int \int G(\vec{r}_1, t_1, \vec{r}_2, t_2) \Psi(\vec{r}_1, t_1) d\vec{r}_1 dt_1 \\
&= \int \int \left[ \sum_{k=1}^p \psi_k(\vec{r}_1, t_1)^* \psi_k(\vec{r}_2, t_2) \right] \Psi(\vec{r}_1, t_1) d\vec{r}_1 dt_1 \\
&= \sum_{k=1}^p \psi_k(\vec{r}_2, t_2) \left[ \int \int \psi_k(\vec{r}_1, t_1)^* \Psi(\vec{r}_1, t_1) d\vec{r}_1 dt_1 \right] \\
&\Rightarrow \Psi(\vec{r}, t) = \sum_{k=1}^p C_k(t) \psi_k(\vec{r}, t) \\
&\text{or } \Psi(\vec{r}, t) = \sum_{k=1}^p C_k(t) \psi_k(\vec{r}) \quad (31)
\end{aligned}$$

15. The following version of dynamic equations can be derived from neuro-synergetic equations (21) and (23) which are here somewhat generalized. The so-called *attention parameter*  $\lambda_k$  is added [40]. Attention parameter  $\lambda_k$  of the pattern  $\vec{v}_k$  is an eigenvalue of the matrix  $\mathbf{J}$  with eigenvectors  $\vec{v}_k$ :  $\mathbf{J} \vec{v}_k = \lambda_k \vec{v}_k$ . (In quantum case, analogously,  $\psi_k$  are eigenstates of the density operator  $\rho$  with eigenvalues  $\rho_k$  [65]). We insert a generalized version of the Hebb rule (23) into the time-dependent equation (21). In the second step we use definition (19) for time-dependent patterns. For stationary case we can delete the time integral.

$$\begin{aligned}
&\dot{q}(\vec{r}_2, t_2) \\
&= \int \int \left[ \sum_{k=1}^p \lambda_k v_k(\vec{r}_1, t_1) v_k(\vec{r}_2, t_2) \right] q(\vec{r}_1, t_1) d\vec{r}_1 dt_1
\end{aligned}$$

$$\begin{aligned}
&= \sum_{k=1}^p \lambda_k v_k(\vec{r}_2, t_2) \left[ \int \int v_k(\vec{r}_1, t_1) q(\vec{r}_1, t_1) d\vec{r}_1 dt_1 \right] \\
&\Rightarrow \dot{q}(\vec{r}, t) = \sum_{k=1}^p \lambda_k c_k v_k(\vec{r}, t) \\
&\text{or } \dot{q}(\vec{r}, t) = \sum_{k=1}^p \lambda_k c_k v_k(\vec{r}). \quad (32)
\end{aligned}$$

In quantum case we can make the following calculation starting from the time-dependent Schrödinger equation with the Hamiltonian operator  $\hat{H}$ . In the last step we use the stationary Schrödinger equation  $\hat{H} \psi_k(\vec{r}) = E_k \psi_k(\vec{r})$  where  $E_k$  is the energy-eigenvalue.

$$\begin{aligned}
i\hbar \dot{\Psi}(\vec{r}, t) &= \hat{H} \Psi(\vec{r}, t) \\
&= \hat{H} \left[ \sum_k C_k \psi_k(\vec{r}) \right] \\
&= \sum_k C_k \hat{H} \psi_k(\vec{r}) = \sum_k C_k E_k \psi_k(\vec{r}). \quad (33)
\end{aligned}$$

We thus obtain

$$\begin{aligned}
\dot{\Psi}(\vec{r}, t) &= -\frac{i}{\hbar} \sum_{k=1}^p E_k C_k \psi_k(\vec{r}) \\
\text{or } \dot{\Psi}(\vec{r}, t) &= -\frac{i}{\hbar} \sum_{k=1}^p E_k C_k \Psi_k(\vec{r}, t) \quad (34)
\end{aligned}$$

Here  $E_k$  (an eigenvalue which represents result of a measurement of an *observable* quantity, i.e., energy) has a similar role as  $\lambda_k$ , the neuro-synergetic *attention* parameter. Note the interdependence of the process of attention and observation which is here, probably not completely by chance, replicated in the mathematical formalism.

16. The analogy no. 6 can be extended to relativistic domain also. In the relativistic case, where  $\psi_k$  are 4-dimensional vectors, the role of the Green function  $G$  is realized by the *S-matrix* (scattering matrix) [13]. Its elements are:

$$S(\vec{r}_1, t_1, \vec{r}_2, t_2) = - \sum_{k=1}^p \sum_{j=1}^2 \psi_k^j(\vec{r}_1, t_1)^* \psi_k^j(\vec{r}_2, t_2)$$

$$\text{if } t_2 > t_1; \quad (35)$$

$$S(\vec{r}_1, t_1, \vec{r}_2, t_2) = \sum_{k=1}^p \sum_{j=3}^4 \psi_k^j(\vec{r}_1, t_1)^* \psi_k^j(\vec{r}_2, t_2)$$

$$\text{if } t_2 < t_1. \quad (36)$$

The same Green function which propagates a solution of the Schrödinger equation forward in time propagates its complex conjugate backward in time. It contains *all* the solutions of the Schrödinger equation, including the bound states as required in the completeness theorem

$$\sum_{k=1}^P \psi_k(\vec{r}_1, t)^* \psi_k(\vec{r}_2, t) = \delta^3(\vec{r}_1 - \vec{r}_2),$$

with equal weight. In the relativistic S-matrix one half of components ( $j = 1, 2$ ) corresponds to the positive-energy waves and the second half ( $j = 3, 4$ ) to the negative energy-waves (antiparticles, i.e., particles propagating into the past:  $t_2 < t_1$ ) which were not present in the non-relativistic case [13].

17. Let us only mention two additional cases of the “learning matrix”, similar to those of analogy no. 6, which an interested reader may find useful for further exploration. In his historical papers on the theory of positrons [29] Feynman introduced the quantum propagator (for  $t_2 > t_1$ )

$$G(\vec{r}_1, t_1, \vec{r}_2, t_2) = \sum_k \psi_k(\vec{r}_1) \psi_k(\vec{r}_2) e^{-iE_k(t_1 - t_2)}$$

which is inserted into (22)-like dynamical equation. This version of the Green function explicitly encodes the “phase”-differences in the exponents.

A similar expression from statistical physics is the Green function of the *Fokker-Planck equation* [37]<sup>1</sup>

$$G(\vec{r}_1, \vec{r}_2, t) = f(\vec{r}_1, \vec{r}_2) \times \sum_{h=0}^{\infty} \sum_{k=-\infty}^{\infty} \psi_{kh}(\vec{r}_1) \psi_{kh}(\vec{r}_2) e^{ik(\varphi_1 - \varphi_2) - \lambda_{kh} \tilde{t}}$$

which is inserted into (22)-like equation for statistical distribution.  $f(\vec{r}_1, \vec{r}_2)$  is an exponential function of new variables incorporating  $r_1$  or  $r_2$ .

These short examples are only to remind the reader that there are numerous cases in physics where “neural-network”-like collective processes take place which could be given informational interpretation. Indeed they could, at least in

principle, realize information processing similar (or more advanced) than those described in the first part.

18. L.I. Gould presented *analogies of Pribram’s holonomic brain theory and Bohm’s quantum interpretation* [35]. He emphasized the importance of non-local processing in two different synergetic levels in brain, which these two theories describe, using quantum and neural versions of Hamilton-Jacobi equation (his general theory see in [34]). Furthermore, he showed that the non-local Schrödinger equation

$$i\hbar \frac{\partial \Psi}{\partial t}$$

$$= -\frac{\hbar^2}{2m} \nabla^2 \Psi + \int V(\vec{r}_1, \vec{r}_2, t) \Psi(\vec{r}_2, t) d\vec{r}_2 \quad (37)$$

is quite similar to the non-local neural wave equation (derived after Pribram)

$$i\nu \frac{\partial \Psi}{\partial t}$$

$$= \frac{\nu^2}{2} \nabla^2 \Psi + \int U(\vec{r}_1, \vec{r}_2, t) \Psi(\vec{r}_2, t) d\vec{r}_2. \quad (38)$$

Integrals in both integro-differential equations take the whole system’s spatial volume. The potential per unit volume is denoted by  $V(\vec{r}_1, \vec{r}_2, t)$ .<sup>1</sup>  $U(\vec{r}_1, \vec{r}_2, t)$  is the potential per unit volume at the classical biomolecular level resulting from its many-body interactions.  $\nu$  has the dimensions of  $(length)^2/(time)$ . It combines the information about the flow-velocity of ionic bioplasma to the spatial frequency of isophase contours. For numerous physiological in biophysical details (e.g., concerning the ionic bioplasma in surroundings of dendritic membranes) we must here invite the reader to see the works of Pribram [80], Umezawa, and Yasue and Jibu [47].

If  $\Psi = A \exp(i\frac{S}{\hbar})$  ( $S$  has here the role of the phase) is inserted into the non-local Schrödinger

<sup>1</sup>For a detailed context please see section 10.3. “Computer Solution of a Fokker-Planck Equation for a Complex Order Parameter” in Haken’s book. ( $\lambda_{kh}$  is an eigenvalue.)

<sup>1</sup>It is not just potential as in usual Schrödinger equation.



equation (37), the following “non-local probability continuity equation” is obtained:

$$\frac{\partial \rho}{\partial t} + \text{div} \left( \frac{\rho \nabla S}{m} \right) = 2 \text{Im} \left( \int V(\vec{r}_1, \vec{r}_2, t) \Psi(\vec{r}_1, t)^* \Psi(\vec{r}_2, t) d\vec{r}_1 d\vec{r}_2 \right). \quad (39)$$

Here  $\rho$  is the probability density:  $\rho = |A|^2$ . The right-hand side contains the generalized-Hebb-rule-like expression, as one can see after the integration is made (see analogy no. 7), which might manifest a “learning”-process as defined in the neural network theory. Actually, because there is no  $\sum_k$  (sum over patterns) in expression (39), it could represent gradual learning with iterative presentation of new patterns - one after another.

### 3.5 Discussion

Although only some basic mathematical analogies were presented here, numerous other parallels can be found between the neural and the quantum processing. They show that there may be a subtle “division of labour” and an isomorphism-like relation and cooperation between the neural and the quantum levels. The relation is not “isomorphic” strictly in a mathematical sense, but in the sense of resemblance of levels and their dynamics. These levels may be in a sort of *fractal*-relationship.

Here the question arises, what is the difference between ordinary quantum processes and mental or even conscious information processes. The first main difference between the physical and psychophysical processes is that *a synergetic (quantum or neural) system itself is not intentional* (does not carry any mental information), but *mind-brain is intentional* (carries specific mental contents). Only consciousness is those entity which gives mental interpretation to some synergetic processes, and makes them, by this very act, informational in a psychological sense (i.e., they start to represent information for humans).

The second difference is that *the quantum system itself does not have a relatively independent environment, but brain does*. Therefore the brain models its

macroscopic environment in a specific and flexible manner by *using the biological neural network as a macro-micro-interface and a (subconscious) pre-processor for an unified conscious experience which involves neuro-quantum coherence*.

## 4 Overall discussion and conclusions

We presented an introductory review of synergetic models describing multi-level system-processing backgrounds of perception, cognition and consciousness. Starting with neural-net models and using mathematical analysis we tried to provide, by including novel quantum approaches, an unified picture of brain-mind and its cooperating levels.

We have argued that mind-like processing needs neural or/and quantum virtual structures (attractors). Indeed, every collective state of a complex system may constitute a specific gestalt (a specific virtual unity) which cannot be reduced to the state of constitutive elements of the system alone. The problem is the formation of a specific isomorphic (e.g., fractal) multi-level coherence. The practice of computer simulations of neural nets shows that we can govern the artificial virtual level implicitly by regulating the artificial neuronal level explicitly (“bottom-up influence”). If our *dynamic equations for neurons and synapses regulate the patterns only, the attractors always accompany this dynamics implicitly!* The opposite can also be done: in physical modeling by regulating system’s energy, in brain-mind by “top-down influence” of consciousness.

Neuronal dynamical equations (used in computer programs) are differential equations (with local range of validity), but attractor-structures may be mathematically described by variational calculus (with global range of validity). We do not necessarily need both mathematical descriptions – one is sufficient. Thus, we may organize one level and the others will follow automatically. This is the reason, why the *reductionist approach usually works for all practical purposes*, and this is a chance for ad-

vanced system-based artificial intelligence in form of neural, quantum [59, 25, 21], and neuro-quantum [49, 50, 51] or other hybrid parallel (synergetic) computers.

Although there is no strict mathematical isomorphism in neural and quantum formalisms, section 3.4 lists remarkable parallels which enable us to use neural-net processing as a metaphor for quantum network processing. These neuro-quantum analogies are an original contribution which enable development of quantum neural-net-like information-processing "algorithms". Indeed, in [78] we presented our new *quantum associative network* model.<sup>1</sup>

## Acknowledgements

TEMPUS grant (JEP-4310-93) for my research in London, OSI grant, and discussions with Prof. K.H. Pribram, Prof. B.J. Hiley, Prof. L.I. Gould, Prof. A.O. Župančič, Dr. J. Šlechta, Dr. T. Matsuoka and T. Menneer are gratefully acknowledged. Many thanks also to P. Ečimovič for his significant help in preparation of the first part, and to Prof. A. Linkevich, Prof. A. Ezhov and Dr. V. Samuilov.

## References

- [1] S.V.B. Aiyer, M. Niranjana, F. Fallside. Theoretical investigation into the performance of the Hopfield model. *IEEE Transact. on Neural Netw.* **1**, 204 (1990).
- [2] S.-I. Amari, K. Maginu. Statistical Neurodynamics of Associative Memory. *Neural Netw.* **1**, 63-73 (1988).
- [3] D. Amit. *Modeling Brain Functions (The world of attractor neural nets)*. (Cambridge Univ. Press, Cambridge, 1989).
- [4] R. Amoroso. The Extracellular Containment of Natural Intelligence. *Informatica* **19**, 585-590 (1995).
- [5] R. Amoroso. Modeling the Heisenberg Matrix: Quantum Coherence and Thought at the Holographic Manifold and Deeper Complementarity. In: *Scale in Conscious Experience: Is the Brain too Important to Be Left to Biologists to Study?* Edited by K. Pribram. (Lawrence Erlbaum, Hillsdale (NJ), 1995).
- [6] *Neurocomputing - Foundations of Research*. Edited by J.A. Anderson, E. Rosenfeld. (MIT Press, Cambridge (MA), 1988).
- [7] J.B. Anderson. A random-walk simulation of the Schrödinger equation:  $H_3^+$ . *J. Chem. Phys.* **63**, no. 4, 1499-1503 (1975).
- [8] B. Baird. Bifurcation and category learning in network models of oscillating cortex. *Physica D* **42**, 365-384 (1990).
- [9] P. Baldi, K. Hornik. Neural Networks and Principal Component Analysis. *Neural Netw.* **2**, 53 (1989).
- [10] W. Banzhaf, H. Haken. An energy function for specialization. *Physica D* **42**, 257-264 (1990).
- [11] H. Barlow. Intraneuronal information processing, directional selectivity and memory for spatiotemporal sequences. *Network: Comp. in Neural Sys.* **7**, 251-259 (1996).
- [12] R. Beale, T. Jackson. *Neural Computing (An Introduction)*. (Adam Hilger, Bristol, 1990).
- [13] J.D. Bjorken, S.D. Drell. *Relativistic Quantum Fields*. Chapter 6: Propagator Theory. (McGraw-Hill, New York, 1965).
- [14] G.J.F. Blommesteijn. A Theory on the Relation between Quantum Mechanical Reduction Process and Consciousness. Presented at the 2<sup>nd</sup> Conf. on Consciousness, Tucson (AZ), 1996 (preprint).
- [15] D. Bohm. *Quantum Theory*. (Constable and Co., London, 1954).
- [16] D. Bohm. *Wholeness and Implicate Order*. (Routledge and Paul Kegan, London, 1980).
- [17] D. Bohm, B.J. Hiley. *The Undivided Universe (An ontological interpretation of quantum theory)*. (Routledge, London, 1993).
- [18] H. Bourlard, Y. Kamp. Auto-Association by Multilayer Perceptrons and Singular Value Decomposition. *Biol. Cybern.* **59**, 291 (1988).

<sup>1</sup>It triggered, since 1998, several sub-conferences on quantum neural-like nets as parts of multi-conferences in USA (IJC Information Sciences '98 [101], '99, and NIPS 2000), Russia (Neuroinformatika '99, 2000) and Belgium (CASYS '99, 2000).

- [19] F. Brentano. *Psychology from an Empirical Standpoint*. (Routledge and Kegan Paul, London, 1973).
- [20] Y. Burnod. *An Adaptive Neural Network: the Cerebral Cortex*. (Prentice Hall, London, 1990).
- [21] R.L. Chrisley. Quantum Learning. In: *New Directions in Cognitive Science*, Proc. Int. Conf. in Lapland. Edited by P. Pylkkänen, P. Pylkkö. (Finnish AI Soc., Helsinki, 1995, pp. 77-89).
- [22] J.G. Daugman. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by 2-D visual cortical filters. *J. Opt. Soc. Amer. A* **2**, no. 7, 1160 (1985).
- [23] P.C.W. Davies, J.R. Brown. *The Ghost in the Atom*. (Cambridge Univ. Press, Cambridge, 1986).
- [24] J.S. Denker. Neural network models of learning and adaptation. *Physica D* **22**, 216 (1986).
- [25] D. Deutsch. Quantum computation. *Phys. World*, pp. 57-61 (June 1992).
- [26] *The Many-Worlds Interpretation of Quantum Mechanics*. Edited by B.S. DeWitt, H. Graham. (Princeton Univ. Press, Princeton, 1973.)
- [27] *Models of Neural Networks* (Series "Physics of NN"). Edited by E. Domany, J.L. van Hemmen, K. Schulten. (Springer, Berlin).
- [28] M. Ebdon. Is the Cerebral Neocortex an Uniform Cognitive Architecture? *Mind & Language* **8**, 368 (1993).
- [29] R. Feynman. The Theory of Positrons. *Phys. Rev.* **76**, no. 6, 749 (1949).
- [30] J.C. Fort. Solving a Combinatorial Problem via Self-Organizing Process: An Application of the Kohonen Algorithm to the Traveling Salesman Problem. *Biol. Cybern.* **59**, 33 (1988).
- [31] A. Gierer. Spatial Organization and Genetic Information in Brain Development. *Biol. Cybern.* **59**, 13 (1988).
- [32] G. Globus: Towards a Noncomputational Cognitive Neuroscience. *J. Cognit. Neurosci.* **4**, 299 (1992).
- [33] A. Goswami. Consciousness in Quantum Physics and the Mind-Body Problem. *J. Mind & Behav.* **11**, no. 1, 75 (1990).
- [34] L.I. Gould. Nonlocal Conserved Quantities, Balance Laws, and Equations of Motion. *Int. J. Theor. Phys.* **28**, no. 3, 335-364 (1989).
- [35] L.I. Gould. Reflections on the Relevance of Non-locality to Cognitive science and the Philosophy of Mind. In: *New Directions in Cognitive Science*, Proc. Int. Conf. in Lapland. Edited by P. Pylkkänen, P. Pylkkö. (Finnish AI Soc., Helsinki, 1995, pp. 104-114).
- [36] H. Gutfreund. From Statistical Mechanics to Neural Networks and Back. *Physica A* **163**, 373-385 (1990).
- [37] H. Haken. *Advanced Synergetics (Instability Hierarchies of Self-Organizing Systems and Devices)*. (Springer, Berlin, 1987).
- [38] *Neural and Synergetic Computers*. Edited by H. Haken. (Springer, Berlin, 1988).
- [39] *Synergetics of Cognition*. Edited by H. Haken, M. Stadler. (Springer, Berlin, 1989).
- [40] H. Haken. *Synergetic Computers and Cognition (A Top-Down Approach to Neural Nets)*. (Springer, Berlin, 1991).
- [41] S.R. Hameroff. Quantum coherence in microtubules: a neural basis for emergent consciousness? *J. Consciousn. Stud.* **1**, no. 1, 91-118 (1994).
- [42] Towards a Science of Consciousness, Tucson I. Edited by S.R. Hameroff, A.W. Kaszniak, A.C. Scott. (MIT Press, Cambridge (MA), 1996.)
- [43] R. Hecht-Nielsen. *Neurocomputing*. (Addison-Wesley, Reading, 1990).
- [44] B.J. Hiley, F.D. Peat. *Quantum Implications (Essays in Honour of David Bohm)*. (Routledge, London, 1987).
- [45] J.J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceed. Natl. Acad. Sci. USA* **79**, 2554-2558 (1982).
- [46] J.J. Hopfield. Neurons with graded response have collective computational properties like those of two-state neurons. *Proceed. Natl. Acad. Sci. USA* **81**, 3088 (1984).
- [47] M. Jibu, K. Yasue. *Quantum Brain Dynamics and Consciousness*. (John Benjamins, Amsterdam / Philadelphia, 1995).
- [48] M. Kaipainen, P. Toivainen, J. Louhivuori. A Self-Organizing Map that Recognizes and Generates Melodies. In: *New Directions in Cognitive Science*, Proc. Int. Conf. in Lapland. Edited by P. Pylkkänen, P. Pylkkö. (Finnish AI Soc., Helsinki, 1995, pp. 286-315).

- [49] S.C. Kak. Information, Physics, and Computation. *Foundat. Phys.* **26**, no. 1, 127-137 (1996).
- [50] S.C. Kak. On quantum neural computing. *Inform. Sci.* **83**, 143-160 (1995).
- [51] S.C. Kak. Quantum Neural Computing. *Advances in Imaging & Electron Phys.* **94**, 259-313 (1995).
- [52] I. Kanter. Potts-glass models of neural networks. *Physical Rev. A* **37**, no. 7, 2739-2742 (1988).
- [53] T. Kohonen. Self-organized formation of topologically correct feature maps. *Biol. Cybern.* **43**, 59 (1982).
- [54] T. Kohonen. *Self-Organization and Associative Memory*. (Springer, Berlin, 1984).
- [55] T. Kohonen. Adaptive, Associative, and Self-Organizing Functions in Neural Computing. *Appl. Optics* **26**, 23 (1987).
- [56] T. Kohonen. An Introduction to Neural Computing. *Neural Netw.* **1**, 3 (1988).
- [57] A.D. Linkevich. Anticipation, perception, language, mind and nonlinear dynamics. *Informatica* **21**, 435-464 (1997).
- [58] R. Linsker. Self-Organization in a Perceptual Network. *Computer*, pp. 105-117 (March 1988).
- [59] S. Lloyd. Quantum-Mechanical Computers. *Sci. Amer.*, pp. 44-50 (October 1995).
- [60] A.R. Lurija. Foundations of Neuropsychology. (Russ. orig.: *Osnovy Neyropsihologii*. Izdat. moskov. univ., Moscow, 1973).
- [61] R.J. MacGregor. *Neural and Brain Modeling*. (Academic Press, San Diego, 1987).
- [62] B. MacLennan. Information Processing in the Dendritic Net. Chapter 6 in: *Rethinking Neural Networks (Quantum Fields and Biological Data*. Edited by K. Pribram. (Lawrence Erlbaum Assoc., Hillsdale (NJ), 1993).
- [63] C.L.T. Mannion, J.G. Taylor. Information Processing by Oscillating Neurons. In: *Coupled Oscillating Neurons*. Edited by J.G. Taylor, C.L.T. Mannion. (Springer, London, 1992).
- [64] P. Marcer, W. Schempp. The brain as a conscious systems. *J. General Sys.* **27**, 231-248 (1998).
- [65] T. Matsuoka, M. Ohya. Fractal Dimensions of States and Application to Ising model. *Reports Math. Phys.* **36**, no. 2/3, 365-379 (1995).
- [66] T. Menneer. Quantum-inspired two-layer neural networks. Tech. rep., Univ. of Exeter, Comp. Sci. Dept., 1996 & Tucson II conf.
- [67] J.L. McClelland, R. Rumelhart, PDP group: Parallel Distributed Processing – Explorations in the Microstructure of Cognition: vols. 1, 2. (MIT Press, Cambridge (MA), 1986).
- [68] M.E. Nelson, J.M. Bower. Brain maps and parallel computers. *Trends in Neurosci.* **13**, no. 10, 403 (1990).
- [69] M. Ohya. Fractal Dimensions of States. In: *Quantum Probability and Related Topics IV*. (World Scientific, Singapore, 1991, pp. 359-369).
- [70] R.G. Palmer, J. Hertz, A. Krogh. *Introduction to the Theory of Neural Computation*. (Addison-Wesley, Redwood City, 1991).
- [71] R. Penrose. *Shadows of the Mind (A Search for the Missing Science of Consciousness)*. (Oxford Univ. Press, Oxford, 1994).
- [72] P. Peretto. *An Introduction to the Modeling of Neural Networks*. (Cambridge Univ. Press, Cambridge, 1992).
- [73] M. Peruš. *All in One, One in All (Brain and Mind in Analysis and Synthesis)*. (DZS, Ljubljana, 1995). M. Peruš. *Bio-Networks, Cognition and Consciousness*. (DZS & Satjam, Ljubljana & Maribor, 2000). (Both books are in Slovene.) [Quantitative details in: Neurophysical Foundations of Cognitive Science; unpubl. manuscript.]
- [74] M. Peruš. Analogies of neural and quantum processing – Consequences for cognitive science. In: *New Directions in Cognitive Science*, Proc. Int. Conf. Edited by P. Pyllkänen, P. Pyllkö. (Finnish AI Soc., Helsinki, 1995, pp. 115-123).
- [75] M. Peruš. Synergetic Approach to Cognition-Modeling with Neural Networks. In: *Bilder im Geiste*. Edited by K. Sachs-Hombach. (Rodopi, Amsterdam & Atlanta, 1995, pp. 183-194).
- [76] M. Peruš. System-processual backgrounds of consciousness. *Informatica* **21**, 491-506 (1997). [Alternatively, see in Slovene: Concepts of Holistic "Models" of Consciousness (Bohm's Quantum Implications, Holograms and Neural Networks). *J. Critics of Sci.* **174**, 11-22 (1995).]
- [77] M. Peruš. Neuro-Quantum Parallelism in Mind-Brain and Computers. *Informatica* **20**, 173-183 (1996).
- [78] M. Peruš. From neural to quantum associative networks: A new quantum "algorithm". AIP Conf. Proceed. of CASYS'99 (Computing Anticipatory Systems – 3rd Int. Conf., Liege, Belgium), vol. **517**. Edited by D. Dubois. (American Institute of Physics, Melville (NY), 2000,

- pp. 289-295). [Alternatively: M. Peruš: Proceed. JCIS'98, vol. 2, RTP, NC, USA, pp. 197-200. Neural Netw. World **10** (no. 6), in press (2000). Appl. Math. Lett. **13**, no. 8, 31-36 (2000) (with S.K. Dey).]
- [79] *Foundations of Cognitive Science*. Edited by M.I. Posner. (MIT Press, Cambridge (MA), 1989).
  - [80] K.H. Pribram. *Brain and Perception*. (Lawrence Erlbaum Assoc., Hillsdale (NJ), 1991).
  - [81] I. Prigogine. Why Irreversibility? The Formulation of Classical and Quantum Mechanics for Nonintegrable Systems. Int. J. Quantum Chem. **53**, 105-118 (1995).
  - [82] H. Ritter, T. Martinetz, K. Schulten. *Neural Computation and Self-Organizing Maps*. (Addison-Wesley, Redwood City, 1992).
  - [83] *Neurobiology of Higher Cognitive Functions*. Edited by A.B. Scheibel, A.F. Wechsler. (Guilford Press, London, 1990).
  - [84] W. Schempp. Bohr's Indeterminacy Principle in Quantum Holography, Self-adaptive Neural Network Architectures, Cortical Self-Organization, Molecular Computers, Magnetic Resonance Imaging and Solitonic Nanotechnology. Nanobiol. **2**, 109 (1993).
  - [85] W. Schempp. Analog VLSI network models, cortical linking neural network models, and quantum holographic neural technology. In: *Wavelets and Their Applications*. Edited by J.S. Byrnes et al. (Kluwer, Amsterdam, 1994, pp. 213-260).
  - [86] W. Schempp. Phase coherent wavelets, Fourier transform magnetic resonance imaging, and synchronized time-domain neural networks. Proceed. Steklov Inst. Math. **3**, 323-351 (1995).
  - [87] G.I. Shipov. A Theory of Physical Vacuum (chapters 2.4 and 2.5, especially). (book manuscript, Moscow, 1997).
  - [88] L. Shastri, V. Ajjanagadde. From simple associations to systematic reasoning: A connectionist representation of rules, variables and dynamic bindings using temporal synchrony. Behav. & Brain Sci. **16**, 417 (1993).
  - [89] H. Sompolinsky. Statistical Mechanics of Neural Networks. Phys. Today, pp. 70 (December 1988).
  - [90] H. Sompolinsky, M. Tsodyks. Segmentation by a Network of Oscillators with Stored Memories. Neural Computat. **6**, 642-657 (1994).
  - [91] H.P. Stapp. Quantum Propensities and the Brain-Mind Connection. Foundat. Phys. **21**, no. 12, 1451 (1991).
  - [92] A. Stern. *Matrix Logic and Mind (A probe into the unified theory of mind and matter)*. (North-Holland / Elsevier, Amsterdam, 1992).
  - [93] A. Stern. *The Quantum Brain*. (North-Holland / Elsevier, Amsterdam, 1994).
  - [94] N.V. Swindale. The development of topography in the visual cortex: a review of models. Network: Comp. in Neural Sys. **7**, 161-247 (1996).
  - [95] J. Šlechtá. On a Quantum-Statistical Theory of Pair Interaction Between Memory Traces in the Brain. Informatica **17**, 109-115 (1993) (and other ref. within).
  - [96] J. Šlechtá. A Study of Structurally Disordered Single-Element Magnetic Materials beyond the Homogeneous Molecular Field Approximation. Phys. stat. sol. (b) **67**, 595-536 (1975).
  - [97] J.G. Taylor, F.N. Alavi. Mathematical Analysis of a Competitive Network for Attention. In: *Mathematical Approaches to Neural Networks*. Edited by J.G. Taylor. (Elsevier, Amsterdam, 1993).
  - [98] V. Tryba, K. Goser. A Modified Algorithm for Self-Organizing Maps Based on Schrödinger Equation. In: *Artificial Neural Networks*, Proceed. Int. Workshop IWANN, 1991.
  - [99] M. Tušar, J. Zupan, J. Gasteiger. Neural networks modelling in chemistry. J. Chem. Phys. **89**, 1517 (1992).
  - [100] G. Vitiello. Coherence and Electromagnetic Fields in Living Matter. Nanobiol. **1**, 221-228 (1992).
  - [101] *Proceedings IJC "Inform. Sci. '98" – vol. 2: Proceed. Int. Conf. "Comput. Intellig. & Neurosci. '98"*. Edited by P.P. Wang et al. (Assoc. Intellig. Mach., RTP (NC, USA), 1998, pp. 167-224) (papers by Behrman, Vitiello, Ventura, Ohya, Linkevich, MacLennan, Gould, Matsuoka, Beck, Kak, Peruš, et al.).
  - [102] A. Zippelius. Statistical Mechanics of Neural Networks. Physica A **194**, 471-481 (1993).
  - [103] J. Zupan. Introduction to Artificial Neural Network (ANN) Methods: What They Are and How to Use Them. Acta Chimica Slov. **41**, 327 (1994).