Associative Memory in a Continuous Metric Space: A Theoretical Foundation

Memoria Asociativa en un Espacio Métrico Continuo: Fundamentos Teóricos

Enrique Carlos Segura

Department of Computer Science, University of Buenos Aires
Ciudad Universitaria, Pab.I, (1428) Buenos Aires, Argentina
esegura@dc.uba.ar

Article received on December 18, 2007; accepted on August 07, 2008

Abstract
We introduce a formal theoretical background, which includes theorems and their proofs, for a neural network model with associative memory and continuous topology, i.e. its processing units are elements of a continuous metric space and the state space is Euclidean and infinite dimensional. This approach is intended as a generalization of the previous ones due to Little and Hopfield. The main contribution of the present work is to integrate -and to provide a theoretical background that makes this integration consistent- two levels of continuity: i) continuous response processing units and ii) continuous topology of the neural system, obtaining a more biologically plausible model of associative memory. We present our analysis according to the following sequence of steps: general results concerning attractors and stationary solutions, including a variational approach for the derivation of the energy function; focus on the case of orthogonal memories, proving theorems on stability, size of attraction basins and spurious states; considerations on the problem of resolution, analyzing the more general case of memories that are not orthogonal, and with possible modifications to the synaptic operator; getting back to discrete models, i. e. considering new viewpoints arising from the present continuous approach and examine which of the new results are also valid for the discrete models; discussion about the generalization of the non deterministic, finite temperature dynamics.
Keywords: associative memory, continuous metric space, dynamical systems, Hopfield model, stability, Glauber dynamics, continuous topology.

Resumen
Presentamos bases teóricas formales, incluyendo teoremas y sus demostraciones, para un modelo de red neuronal con memoria asociativa y topologia continua, i. e. sus unidades de procesamiento son elementos de un espacio métrico continuo y el espacio de estados es euclidiano y de dimensión infinita. El enfoque es concebido como una generalización de los precedentes debidos a Little y Hopfield. La principal contribución del presente trabajo es integrar -y proveer fundamentos teóricos que den consistencia a tal integración- dos niveles de continuidad: i) unidades de proceso de respuesta continua y ii) topologia continua del sistema neuronal, obteniendo de esta manera un modelo más biológicamente plausible de memoria asociativa. Nuestro análisis es presentado de acuerdo con la siguiente secuencia de pasos: resultados generales sobre atractores y soluciones estacionarias, que incluyen un enfoque variacional para derivar la función de energia; estudio detallado del caso de memorias ortogonales, demostrando teoremas sobre estabilidad, tamaño de cuencas de atracción y estados espurios; consideraciones sobre el problema de la resolución, analizando el caso más general de memorias no ortogonales, y con modificaciones pos-

ibles al operador sináptico; retorno a los modelos discretos, i.e. consideración de nuevos puntos de vista que surgen del presente esquema, y de cuáles de los nuevos resultados son también válidos para los modelos discretos; discusión sobre la generalización de la dinámica no deterministica a temperatura finita.
Palabras clave: memoria asociativa, espacio métrico continuo, sistema dinámico, modelo de Hopfield, dinámica de Glauber, topologia continua, estabilidad.

1 Introduction

In seminal papers, Little [11][12] and Hopfield [8] constructed a content-addressable memory as a dense network of artificial neurons that are represented as elementary bistable processors. Addressability is guaranteed by the dissipative dynamics of the system. It consists of switching each processor from one of its stable configurations to the other as a consequence of the intensity of the local field acting on it. The memories, corresponding to fixed points of the dynamics, are stored in the system in a distributed manner through the matrix of two-body interactions (synaptic efficacies) between the neurons. If this matrix is properly defined, the above dynamics is enough so as to ensure a monotonic decrease of an "energy" function. Thus, starting from an arbitrary configuration the system is led to a local minimum that corresponds to the nearest stored memory.

In a later work, Hopfield [9] aimed at a more realistic model by replacing bistable neurons by graded response devices. In fact, a classical objection to the former model [8][11][12] was that a two-state representation of the neural output is, from a biological point of view, an oversimplification and that it is necessary to describe relevant neural activity by firing rates, rather than merely by the presence or the absence of an individual spike[1]. In either case the retrieval process is again guaranteed by the nature of the matrix of synaptic efficacies. However, in [9] the space of states describing the patterns of activity remained discrete, in the sense that the number of units was, at most, countable. This was an open gap in the plausibility of the model. In fact, since the Little model was formulated to describe the computational ability of an ensemble of simple processing units, it was necessary to reconcile the biological evidence of a true continuum of the neural tissue with the descriptions provided by discrete models inspired in an Ising system. While the empirical evidence always shows patterns of activity or quiescence involving patches with finite sizes, the ferromagnetic approach suggests systems with discrete processing units with no finite dimensions. In spite of this simplification all the discrete models have been remarkably successful in describing emergent processing abilities that correspond to stylized facts concerning basic elementary cognitive processes.

In this paper we introduce a solid theoretical background, including theorems and their proofs, for our neural network model with associative memory and processing units defined as elements of a continuous metric space. This model [15],[16],[17] is intended as a generalization of the previous ones due to Little and Hopfield. Our main purpose is to provide rigorous proofs in the sense that it is actually possible to formulate a system of associative memory with continuous response units and a continuous topological structure on the set of such units. We conceive the network so as to preserve the salient features that made attractive all the discrete models, especially the levels of continuity that the Hopfield model with graded response [9] added to the discrete one [8]: continuous-valued units and continuous scale of time,

[1]However, there is at present an increasing agreement that spiking neurons have some properties for describing certain aspects of neural dynamics not completely covered by rate-coding models.

via the transition from discrete to continuous, differential equation dynamics. In spite of the fact that the corresponding space of configurations is an infinite dimensional functional space, we can define a basic simple dynamics having asymptotic, stationary solutions which can be associated to minima of an energy functional of Lyapunov type and can be taken to represent the memories that are stored in the system.

We have already introduced in a previous article [15] several of the results included here, but none of them was supported on rigorous proof. The present article is devoted to provide a solid theoretical foundation for that sketch. Emphasis is placed on a detailed analysis of orthogonal memories, which constitute a relevant particular case of the general theory (in fact, a deep comprehension of orthogonal memories is essential to understand the general pseudo-orthogonal case).

The paper also presents other more general results. In particular, concerning the biological plausibility of the model, we face the problem of what we call its "unbounded resolution", which unavoidably takes place under the assumptions of Hebbian operator and orthogonal memories. We investigate the possible extensions of the Hebb rule and its robustness against the relaxation of the basic assumption of orthogonality of the memory states. We show that a "finite resolution" can be imposed to the stored memories by limiting the minimum size of the activity patches. Two ways for doing this are presented (Section 7). Moreover, the possibility of modifying the dynamics of the system is discussed, in connection with the desirable (always in terms of biological plausibility) continuity of the solutions, i.e. smoothness of the memory states. As we will see in Section 8, the equation that governs the evolution of the system (1) can be modified, in order for some graded transition to take place between patches with firing neurons and patches without activity.

Some approaches closely related to ours have appeared some time ago [13][14]. In particular, the concept of *field computation*, introduced by MacLennan, has a similar inspiration in the fact that many neural phenomena can be conveniently described as a field, i.e. the distribution of some physical quantity over a continuous space, with some topology associated to it. On the other hand, the big number of neurons per square millimeter that can be found throughout most of the brain cortex, justifies the treatment of neural activity as a field.

All these arguments are related to our approach. However, this is aimed to a different purpose, which is that of formulating an extended model of associative memory and theoretically founding it, including justification of its potential as a tool for modelling cognitive processes of memory and learning. In this sense, we think that these modelling abilities could provide a wide range of possibilities for our understanding of normal as much as pathological brain function.

Moreover, in another previous paper [16], we have already proposed a generalization of the nondeterministic, finite temperature Glauber dynamics [5] to the case of a finite number of graded response neurons (Hopfield'84). We did this by casting the retrieval process of a Hopfield model with continuous-valued units, into the framework of a diffusive process governed by the Fokker-Plank equation. We thus provided a description of the transitional regime that prevails during the retrieval process, which is currently disregarded. In other words, we unified the graded response units model [9] and the stochastic approach, obtaining a description of the retrieval process at both the microscopic, individual neuron level and the macroscopic level of time evolution of the probability density function over the space of all possible activation patterns, i.e. an equation describing, for each possible pattern, how, given an initial probability for the system being in it, this probability changes upon time. For the sake of completeness, we bring here a brief description of the formalisms of this approach.

A not minor issue about the present approach is (as it will be clear, we hope, along the paper and will be discussed more deeply in the Conclusions) how can classical, discrete models and the new ones introduced here be compared. It seems natural to choose some relevant criteria for this comparison and two of them look particularly important: capacity and stability. As it will be seen, the key issue is, concerning capacity, not the finiteness or infiniteness of the storable and retrievable memories, but rather -and this links with the second criterium mentioned above- that of stability of the system as a memory device. In fact, while degradation in the classical Hopfield model is not smooth but certainly subject to a first transition dynamics (i.e. sudden loss of retrievable information beyond a certain critical load value [1]), nothing of this (instability, catastrophic loss of information) happens in neither the continuous topology model nor in that with finite temperature.

It will be also clear along the paper, we think, that such features as stability and biological plausibility are not cost free concerning the complexity of the model and, hence, the time and computer resources required; mathematically speaking, the dynamics of the model implies the numerical integration of nonlinear differential equations instead of a set of discrete difference equations.

Another interesting question that could be stated concerns how general is the work presented in this paper. Besides generalizing the Little and Hopfield models, could it be said that it has any other additional or wider generalization power? In principle, the central concept inspiring our proposal is that of extending the neural dynamics as proposed by those authors to a topologically continuous array of neurons, provided with a metric notion which is not present in the classical discrete models. However, as soon as we introduce variants either on the Hebb-inspired rule for learning the synaptic weights (as we do in Section 7) or directly on the equation that governs the evolution of the system (1), we are going beyond the original framework, defining dynamics that could enclose many different models not identifiable as belongig to the family of the ”Little-Hopfield Models”.

We will also show (see Section 9 for details) that the results to be proven along the paper can also be useful, to a considerable extent, when reconsidering the discrete case in the light of the knowledge of what happens if the state space is continuous. As it will be seen, this ”reverse track” could be useful for both: i) confirming the power of generalization of the proposed model and ii) enhance our understanding of the properties of classical approaches in a more general context.

The paper is organized as follows. In Section 2 we give some basic concepts and definitions. Section 3 provides general results concerning attractors and stationary solutions, including a variational approach for the derivation of the energy function. In Section 4, we focus on the orthogonal case, proving theorems on the stability of the memories and of the origin. Section 5 presents a result on the size of basins of attraction and Section 6 analyzes the question of spurious states. In Section 7 we analyze possible extensions of the Hebb rule and its robustness against relaxation of the condition of orthogonality of the memories. The possibility of modifying the dynamics of the system is briefly discussed in Section 8 (although it remains an open question for future investigation). Finally, in Section 9 we get back to discrete models, i.e. discuss new viewpoints arising from the present continuous approach and examine which of the new results are also valid for the discrete models. Additionally, in Section 10 we present a brief description of the finite temperature dynamics (which was already studied in detail in [16]).

2 Basic definitions and first results

Assume that $\mathbf{v}(x,t)$ describes the activity of a point-like neuron located in x at time t. This pattern of activity evolves according to:

$$\frac{\partial \mathbf{v}(x,t)}{\partial t} = -\mathbf{v}(x,t) + g_\sigma\left(\int_K \mathbf{T}(x,y)\mathbf{v}(y,t)dy\right) \tag{1}$$

with $\mathbf{v}(x,t) : K \times R_{\geq 0} \to R$, $K \subset X$. X is a metric space, K a compact domain, g_σ a *sigmoid* function, i.e. $g_\sigma \in C^1(R)$, non decreasing, odd and satisfying $\lim_{x\to\pm\infty} g_\sigma(x) = \pm \mathrm{V_M}$, $\lim_{\sigma\to\infty} g_\sigma(x) = sgn(x)\ \forall x \neq 0$, $|\, g_\sigma(x)\,| < \min\{\mathrm{V_M}, \sigma x\}$ and $g'_\sigma(0) = \sigma$.

Let S be the set of all possible states $\mathbf{v}(x)$ (patterns of activity) of the system. Then a solution $\mathbf{v}(x,t)$ fulfilling (1) is a trajectory in S.

We can assume that $\mathrm{V_M} = 1$. As for $\mathbf{T}$:$K \times K \to R$, we assume it is continuous almost everywhere (a.e.) in order to ensure that the integral is well defined. As a natural extension of the discrete case we introduce the *local field* on (or net input to) the neuron located in x when the state of the system is $\mathbf{v}(y,t)$:

$$h_t^{\mathbf{v}}(x) = \int_K \mathbf{T}(x,y)\mathbf{v}(y,t)dy$$

For $t = 0$ we write $h^{\mathbf{v}}(x) = h_0^{\mathbf{v}}(x)$. Note that $h^{\mathbf{v}}$ is linear in $\mathbf{v}$.

Let $\mathbf{v}_0^\mu(x) = \mathbf{v}^\mu(x,0)$ be an initial condition and $\mathbf{v}(x,t)$ the solution of (1) associated to it. We say that $\mathbf{v}^\mu(x)$ is a *memory* or an *attractor* if and only if:

1) $\mathbf{v}^\mu$ is an equilibrium point, i.e. $\mathbf{v}^\mu(x) = g_\sigma(h_t^{\mathbf{v}^\mu}(x))\quad a.e.$

2) For every $t_0 \geq 0$ and $\mathbf{v}_0$ a different initial condition corresponding to the solution $\mathbf{v}$, there exists $\delta(t_0) > 0$ such that if $\|\, \mathbf{v}^\mu - \mathbf{v}_0 \,\| < \delta$ then $\|\, \mathbf{v}^\mu(\cdot,t) - \mathbf{v}(\cdot,t)\,\| \to 0$ when $t \to \infty$.

Hence, attractors are stationary solutions of (1). Assume that S=$L^2(K)$ and that $|\,K\,| < \infty$ (K has finite Lebesgue measure).

We define the *energy* of the system at time t_0 as the functional:

$$\mathbf{H}\left[\mathbf{v}(\cdot,t_0)\right] = -\frac{1}{2}\int_K\int_K \mathbf{T}(x,y)\mathbf{v}(x,t_0)\mathbf{v}(y,t_0)dxdy + \int_K \int_0^{\mathbf{v}(x,t_0)} g_\sigma^{-1}(s)dsdx \tag{2}$$

where $\mathbf{H}[\mathbf{v}(\cdot,t_0)]$ means that $\mathbf{v}$ is viewed as a function of x. Thus, each $\mathbf{v}$ in $\mathbf{S}$ has an energy $\mathbf{H}(\mathbf{v})$. This is an extension of the energy as defined in [9] for the (discrete) model with graded response functions.

2.1 Attractors and stationary solutions

We begin by proving a useful property:

Lemma 2.1: if $\mathbf{v}(x,t)$ is a general solution (for some initial condition) of (1), then it holds that:

$$\frac{\partial h_t^{\mathbf{v}}(x)}{\partial t} = -h_t^{\mathbf{v}}(x) + \int_K \mathbf{T}(x,y)g_\sigma\left(h_t^{\mathbf{v}}(y)\right)dy$$

Proof: straightforward, noting that

$$\frac{\partial h_t^{\mathbf{v}}(x)}{\partial t} = \frac{\partial}{\partial t}\left[\int_K \mathbf{T}(x,y)\mathbf{v}(y,t)dy\right] = \int_K \mathbf{T}(x,y)\frac{\partial \mathbf{v}(y,t)}{\partial t}dy$$

and using (1).

From now on we assume that $\mathbf{T}$ is symmetric.

Theorem 2.2: $\mathbf{H}$ is monotonically decreasing with t and reaches its minima at $\mathbf{v}_{t_e}(x) = \mathbf{v}(x,t_e)$ such that

$$\left[\frac{\partial \mathbf{v}}{\partial t}(x,t)\right]_{t_e} = 0 \tag{3}$$

a.e. in K (in words, given a solution $\mathbf{v}(x,t)$ corresponding to some initial condition, the minima of $\mathbf{H}$ are equilibrium points of the system). This theorem, as well as its proof, generalizes the classical equivalent result for the discrete Hopfield model with graded response neurons [9] (see for example [2][6]).

Proof: let $\mathbf{v}(x,t)$ be a solution of (1) and assume it has a derivative a.e. First observe that $\frac{\partial}{\partial t}\left[\int_0^{\mathbf{v}(x,t)} g_\sigma^{-1}(s)ds\right] = h_t^{\mathbf{v}}(x)\frac{\partial \mathbf{v}(x,t)}{\partial t}$. Differentiating $\mathbf{H}$ with respect to t (under the integral sign) yields:

$$\frac{\partial \mathbf{H}(\mathbf{v}(x,t))}{\partial t} = -\int_K \frac{\partial \mathbf{v}(x,t)}{\partial t}\left[\int_K \mathbf{T}(x,y)\mathbf{v}(y,t)dy - h_t^{\mathbf{v}}(x)\right]dx \tag{4}$$

Noting that $\frac{\partial \mathbf{v}(x,t)}{\partial t} = g'_\sigma\left(h_t^{\mathbf{v}}(x)\right)\frac{\partial h_t^{\mathbf{v}}(x)}{\partial t}$ and using lemma 2.1 we obtain:

$$\frac{\partial \mathbf{H}(\mathbf{v}(x,t_e))}{\partial t} = -\int_K g'_\sigma\left(h_t^{\mathbf{v}}(x)\right)\left(\frac{\partial h_t^{\mathbf{v}}(x)}{\partial t}\right)^2 dx \le 0$$

Then, $\mathbf{H}$ is monotonic. It always decreases with t, except when it reaches a minimum, i.e. at some t_e:

$$\frac{\partial \mathbf{H}(\mathbf{v}(x,t_e))}{\partial t} = 0 \Longleftrightarrow \frac{\partial h_t^{\mathbf{v}}(x)}{\partial t}\,|_{t_e} = 0 \quad \textit{a.e. in } K$$

since $g'_\sigma > 0$. But

$$\frac{\partial \mathbf{v}}{\partial t}(x,t_e) = g'_\sigma\left(h_{t_e}^{\mathbf{v}}(x)\right)\frac{\partial h_t^{\mathbf{v}}(x)}{\partial t}\,|_{t_e} \Longrightarrow \frac{\partial \mathbf{H}}{\partial t}(\mathbf{v}(x,t_e)) = 0 \Longleftrightarrow \frac{\partial \mathbf{v}}{\partial t}(x,t_e) = 0 \quad \textit{a.e. in } K$$

Memories or attractor states, as defined in this section, satisfy the above conditions. However, the reciprocal implication is not necessarily true: from the previous theorem it does not follow that if a solution $\mathbf{v}(x,t)$ of (1) satisfies condition (3) for some t^*, then $\mathbf{v}(x,t^*)$ is an attractor. For example, the trivial solution

$\mathbf{v}\equiv 0$ satisfies it for all t but, as we soon will see, its stability or instability depends on the slope σ of g_σ at the origin. In general, the possibility to construct nontrivial memories strongly depends on such parameter.

The sigmoid function g_σ plays an important role in determining in which cases the system has nontrivial stationary solutions. A necessary condition is given by the following

Theorem 2.3: (existence and uniqueness of the solution) If $\sigma < \frac{1}{\mathrm{M}|K|}$, being M such that $|\mathbf{T}(x,y)| \leq \mathrm{M}$, then the unique stationary solution of (1) is $\mathbf{v}\equiv 0$.

Proof: by definition of g_σ, σ is a Lipschitz constant for it. Then, assuming that $\mathbf{v}^1$ and $\mathbf{v}^2$ are two fixed points of the operator A defined as

$$A\mathbf{v} = g_\sigma \left[\int_K \mathbf{T}(x,y)\mathbf{v}(y)dy \right]$$

we get (using the L^2 norm):

$$|A\mathbf{v}^1(x) - A\mathbf{v}^2(x)| = |g_\sigma(\int_K \mathbf{T}(x,y)\mathbf{v}^1(y)dy) - g_\sigma(\int_K \mathbf{T}(x,y)\mathbf{v}^2(y)dy)|$$
$$\leq \sigma \left| \int_K \mathbf{T}(x,y)\mathbf{v}^1(y)dy - \int_K \mathbf{T}(x,y)\mathbf{v}^2(y)dy \right| \leq \sigma M |K|^{\frac{1}{2}} \| \mathbf{v}^1 - \mathbf{v}^2 \|$$

being M an upper bound for $|\mathbf{T}(x,y)|$ (which exists since $\mathbf{T}$ is continuous and K is compact). Finally:

$$\| A\mathbf{v}^1 - A\mathbf{v}^2 \| \leq |K|^{\frac{1}{2}} \sup_{x\in K} |A\mathbf{v}^1(x) - A\mathbf{v}^2(x)| < \sigma M |K| \| \mathbf{v}^1 - \mathbf{v}^2 \|$$

Then A is a contraction and has a unique fixed point provided $\sigma < \frac{1}{M|K|}$.

Besides the condition $\sigma \mathrm{M} |K| \geq 1$, other ones (see next section) have to be fulfilled in order to ensure the actual existence of nontrivial solutions.

2.2 A variational approach

Historically, the function $\mathbf{H}$ as presented in [8] and [9] is probably a late innovation in the formulation of the corresponding dynamics. We may conjecture that the sequence of concepts at the genesis of the Hopfield model was as follows:

i) A stability condition is imposed ($1 \leq i \leq N$):

$$S_i = sgn\left(\sum_j \mathbf{T}_{ij} S_j\right) \quad [8] \quad \text{or} \quad V_i = g_\sigma\left(\sum_j \mathbf{T}_{ij} V_j\right) \quad [9]$$

ii) A dynamics is proposed such that its equilibria are the states satisfying i):

$$S_i(t+1) = sgn\left(\sum_j \mathbf{T}_{ij} S_j(t)\right) \quad [8]) \quad \text{or} \quad \frac{\partial V_i}{\partial t} = -V_i + g_\sigma\left(\sum_j \mathbf{T}_{ij} V_j\right) \quad [9]$$

iii) A "Lyapunov function" is constructed whose minima are the attractors for the dynamics ii):

$$\mathbf{H}[S] = -\frac{1}{2}\sum_{ij}\mathbf{T}_{ij}S_iS_j \quad [8] \quad \text{or} \quad \mathbf{H}[V] = -\frac{1}{2}\sum_{ij}\mathbf{T}_{ij}V_iV_j + \sum_i \int_0^{V_i} g_\sigma^{-1}(V)dV \quad [9]$$

In our case (infinite dimensional state space S), step i) is given by the condition $\mathbf{v}^\mu(x) = g_\sigma\left(h^{\mathbf{v}^\mu}(x)\right)$ a.e. in K while ii) corresponds to equation (1) and iii) to the functional $\mathbf{H}$ as defined by (2).
It is easy to go from i) to ii). As for deriving iii) from ii), it can be suggested that the main source of Hopfield's and Little's works was statistical mechanics. In our case, however, it is possible to obtain the same model through a variational approach. We will prove, by computing the variation of $\mathbf{H}$ (considered as a functional from $L^2(K,R)$ onto R), that every minimum of $\mathbf{H}$ is also an equilibrium for the dynamics.
Let us consider a function $\varphi : K \longrightarrow R$ continuous such that $\varphi \mid_{\partial K} = 0$.

$$\frac{\partial \mathbf{H}(\mathbf{v}+\varepsilon\varphi)}{\partial \varepsilon} \mid_{\varepsilon=0} = -\frac{1}{2}\int_K\int_K \frac{\partial}{\partial \varepsilon}\mathbf{T}(x,y)[\mathbf{v}(x)+\varepsilon\varphi(x)][\mathbf{v}(y)+\varepsilon\varphi(y)]dxdy \mid_{\varepsilon=0}$$

$$+\int_K \frac{\partial}{\partial \varepsilon} \int_0^{\mathbf{v}(x)+\varepsilon\varphi(x)} g_\sigma^{-1}(s)dsdx \mid_{\varepsilon=0}$$

$$= \int_K \left[-\int_K \mathbf{T}(x,y)\mathbf{v}(y)dy + g_\sigma^{-1}(\mathbf{v}(x)) \right] \varphi(x)dx$$

In order for $\mathbf{v}$ being a minimum for $\mathbf{H}$, that quantity must vanish for every φ in the above conditions, which forces the term between brackets to vanish a.e. in K, i.e. $\mathbf{v}(x) = g_\sigma\left(\int_K \mathbf{T}(x,y)\mathbf{v}(y)dy\right)$ a.e. in K. Thus, the following theorem holds:

Theorem 2.4: if $\mathbf{v} \in L^2(K)$ is a local minimum for $\mathbf{H}$ then it is an equilibrium point of (1).

This result is obtained without using the fact that $\mathbf{H}$ is decreasing (theorem 2.2). Adding that property, it is possible to conclude:

Corollary 2.5: $\mathbf{v} \in L^2(K)$ is a local minimum for $\mathbf{H}$ if and only if it is an attractor for (1).

2.3 A general feature of attractor memories

On the base of the previous results, it can be proved that when $\sigma \to \infty$, the attractors approach the asymptotic bounds of g_σ.

Theorem 2.6: if for some $\varepsilon > 0$ it holds that $\mathbf{T}(x,y) \geq 0$ when $\mid x-y \mid < \varepsilon$, then

$$\lim_{\sigma\to\infty} \max_{\mathbf{v}\ attractors,\ x\in K} \min\{1-\mathbf{v}(x), 1+\mathbf{v}(x)\} = 0$$

Proof: consider an attractor $\mathbf{v}(x)$ (keeping in mind that this notation stands for $\mathbf{v}(x,t_e)$ for some solution of (1) at time t_e). Note that, by virtue of theorem 2.4 and corollary 2.5, if $\mathbf{v}(x)$ is an attractor (a memory), then it is a local minimum of $\mathbf{H}$. First remark that

$$\int_K \int_0^{\mathbf{v}(x)} g_\sigma^{-1}(\mathbf{v})d\mathbf{v}dx = \frac{1}{\sigma}\int_K \int_0^{\mathbf{v}(x)} g^{-1}(\mathbf{v})d\mathbf{v}dx \to^{\sigma\to 0} 0 \quad a.e.$$

independently of $\mathbf{v}(x)$. We then need to consider only the quantity

$$\mathbf{H}_1 = -\frac{1}{2}\int_K \int_K \mathbf{T}(x,y)\mathbf{v}(x)\mathbf{v}(y)dxdy$$

Suppose that $\mathbf{v} \in L^2(K)$ is a minimum and there exists an element $x^* \in K$ such that $\mathbf{v}(x^*) \in (-1,1)$. Then, from the piecewise continuity of the solutions, there exists a closed neighborhood U of x^*such that $\mathbf{v}(y) \in (-1,1)\ \forall\ y \in U$ and we may write:

$$\begin{aligned}\mathbf{H}_1 \quad = \quad & -\frac{1}{2}\{\int_{K-U}\int_{K-U} \mathbf{T}(x,y)\mathbf{v}(x)\mathbf{v}(y)dxdy + \int_U \int_{K-U} \mathbf{T}(x,y)\mathbf{v}(x)\mathbf{v}(y)dxdy \\ & + \int_{K-U}\int_U \mathbf{T}(x,y)\mathbf{v}(x)\mathbf{v}(y)dxdy + \int_U\int_U \mathbf{T}(x,y)\mathbf{v}(x)\mathbf{v}(y)dxdy\}\end{aligned}$$

Let us name A, B, C and D the four terms in $\mathbf{H}_1$ from left to right. D is positive since $\mathbf{T}(x,y) \geq 0$ when $|x-y| < \varepsilon$ (choosing U small enough). A does not depend on values in $U(x^*)$. Finally, B and C are equal and reduce to:

$$2\int_U \mathbf{v}(x)\left[\int_{K-U} \mathbf{T}(x,y)\mathbf{v}(y)dy\right]dx = 2\,|\,U\,|\,\mathbf{v}(\xi)\int_{K-U} \mathbf{T}(\xi,y)\mathbf{v}(y)dy$$

for some $\xi \in U$. Then, redefining $\mathbf{v}$ in U as

$$\mathbf{v}(x) = sgn\left(\int_{K-U} \mathbf{T}(x,y)\mathbf{v}(y)dy\right) \quad \text{if} \quad x \in U \tag{5}$$

$B+C$ increases (while A remains unchanged and D also increases). Then $\mathbf{H}_1$ decreases and hence $\mathbf{H}$ decreases as well. This is a contradiction, due to the assumption that $\mathbf{v}$ as defined initially was a minimum.

Remark: the last step of the proof (5) can naturally be regarded from the point of view of a ferromagnetic model (with an infinite number of elements in this case). Every state change that implies the alignment of one (or more) elements according to the local magnetic field acting on it, decreases the total energy of the system.

3 Orthogonal memories, Hebbian synapses

The case we are specially interested in is the storage of orthogonal memories when the matrix of synaptic weights is Hebbian. This can be achieved through a straightforward generalization of the Hebb rule [7]. Let $\{\mathbf{v}^\mu\}$ be an orthogonal set of functions in some space $S(K)$, that is to say $(\mathbf{v}^\mu, \mathbf{v}^\nu) = 0$ if $\mu \neq \nu$. In principle, $S(K)$ may be noncountable and hence we can define in general:

$$\mathbf{T}(x,y) = \frac{1}{|K|} \int_P \mathbf{v}^\rho(x)\mathbf{v}^\rho(y)d\rho$$

for $\rho \in P$ some index set. In the case $\{\mathbf{v}^\mu\}$ is an orthogonal set in $L^2(K)$, it is at most countable (provided the separability of $L^2(K)$). Therefore it is natural to restrict to the case in which P is countable:

$$\mathbf{T}(x,y) = \frac{1}{|K|} \sum_{\mu=1}^{p} \mathbf{v}^\mu(x)\mathbf{v}^\mu(y) \tag{6}$$

Then the following theorem holds:

Theorem 3.1: The system (1), with $\mathbf{T}(x,y)$ defined as in (6), may have any finite number p of orthogonal memories taking values in $\{\mathrm{V}_*, -\mathrm{V}_*\}$, where $g_\sigma(\pm \mathrm{V}_*^3) = \pm \mathrm{V}_*$.

Proof: let p be a positive integer number and $\{\mathbf{v}^\mu\}_{\mu=1}^{p} \subset L^2(K)$, $(\mathbf{v}^\mu, \mathbf{v}^\nu) = 0$ if $\mu \neq \nu$, $\mathbf{v}^\mu(x) \in \{-V_*, V_*\}$, $1 \leq \mu \leq p$, $x \in K$, being V_* the number such that $g_\sigma(\pm V_*^3) = \pm V_*$ (which exists and depends on σ). Defining $\mathbf{T}$ according to (6) it holds that, for any μ:

$$g_\sigma(h^\mu(x)) = g_\sigma(\int_K \mathbf{T}(x,y)\mathbf{v}^\mu(y)dy) = g_\sigma(\int_K \frac{1}{|K|} \sum_{\nu=1}^{p} \mathbf{v}^\nu(x)\mathbf{v}^\nu(y)\mathbf{v}^\mu(y)dy)$$

Since the integrals are finite, we can interchange the sum and the integration, thus obtaining:

$$g_\sigma(\frac{1}{|K|} \sum_{\nu=1}^{p} \mathbf{v}^\nu(x)\delta_{\nu\mu} \parallel \mathbf{v}^\mu \parallel^2) = g_\sigma(\frac{\mathbf{v}^\mu(x)\ V_*^2 \,|K|}{|K|}) = \mathbf{v}^\mu(x)$$

Then, $\mathbf{v}^\mu(x,t) = \mathbf{v}^\mu(x) \ \ \forall\, t > 0$ is a fixed point of equation (1).

These solutions $\mathbf{v}^\mu(x)$ look like the example depicted in Figure 1. Activation patterns of this kind agree with the intuitive generalization of the attractors of an Ising-type, spin glass discrete neural network in which patches of full activation alternate randomly with those of full quiescence. They can also be viewed as the vertices of an infinite (noncountable) dimensional hypercube.

A question arising is whether the set of orthogonal fixed points of (1) can be infinite. Note first that it is countable: the elements $\mathbf{v}^\mu$ as they were defined belong to $L^2(K)$, a separable space; hence every orthogonal set in it must be countable. However even an infinite countable number of orthogonal fixed points is not possible while preserving the integrability of $\mathbf{T}$. Observe that if there are k_μ changes of sign in $\mathbf{v}^\mu$ then each term of the form $\mathbf{v}^\mu(x)\mathbf{v}^\mu(y)$ divides the domain $K \times K$ in $(k_\mu + 1)^2$ square regions. Moreover,

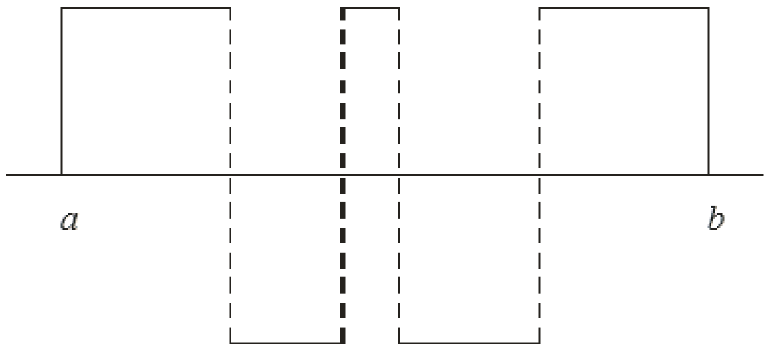

Fig. 1. A memory in the space S=$L^2[a,b]$

each region is separated from the next by discontinuity lines because such term takes the constant values $+\mathrm{V}_*^2$ or $-\mathrm{V}_*^2$. If the set of memories is infinite, the number of terms in **T** that are added is also infinite, therefore those discontinuity lines are dense at least in a neighborhood of some point, and **T** ceases to be piecewise continuous.

Note that theorem 3.1 implies a qualitative difference between discrete and continuous models. Since the memory capacity is now unbounded, there is nothing like a "phase diagram" in which, for a domain K and above some critical number p_c of memories, a transition to a "confusion phase" takes place, implying a rapid degradation of the retrieval ability. While in discrete models the size of the domain is determined by the dimension of the state space ($p_c = \alpha_c N$), in the present case this dimension is infinite and hence there is no p_c.

Discussions on discrete models are mostly made in the thermodynamic limit in which either the number of neurons and the number of memories are taken to tend to infinity while their ratio is kept constant. This possibility is lost in the continuous limit, but this is certainly not a problem as far as the biological plausibility of the model is concerned.

We end this section with the following results that are easy to check.

Lemma 3.2: If the memories are orthogonal, the distance between any two of them is always the same.

Corollary 3.3: Orthogonal memories are never dense in $L^2(K)$.

4 Stability of the solutions

We will now derive conditions for the elements $\mathbf{v}^\mu$, as defined in Section 3, to be stable equilibria (i.e. memories) for equation (1).

Theorem 4.1: The elements $\mathbf{v}^\mu$ are stable fixed points of (1) if and only if $g'_\sigma(\mathrm{V}_*^3) < \frac{1}{\mathrm{V}_*^2}$.

Theorem 4.2 :(Stability of the origin) The solution $\mathbf{v} \equiv 0$ is stable if and only if $g'_\sigma(0) = \sigma < \frac{1}{\mathrm{V}_*^2}$.

Proof (both theorems): the computation of the directional derivatives of $\mathbf{H}(\mathbf{v})$ at an arbitrary point yields:

$$D_w\mathbf{H}(\mathbf{v}) = -\frac{1}{|K|}\sum_{\nu=1}^{p}(\mathbf{v}^\nu,\mathbf{v})(\mathbf{v}^\nu,w) + (g_\sigma^{-1}(\mathbf{v}),w)$$

with $w \in L^2(K)$ and $\| w \| = 1$. Now, if $\mathbf{v} = \mathbf{v}^\mu$, using the condition of orthogonality and noting that $\| \mathbf{v}^\mu \|^2 = \mathrm{V}_*^2 \mid K \mid$, it follows that $D_w\mathbf{H}(\mathbf{v}) = 0$. Similarly, it is easy to check that $D_w\mathbf{H}(\mathbf{v})$ vanishes in general for any element in the set $span\{\mathbf{v}^\mu\}_{\mu=1}^{p}$, i.e. linear combinations of the memories, when those combinations take values on $\{\mathrm{V}_*, -\mathrm{V}_*, 0\}$.

$$D^2_{w^2}\mathbf{H}(\mathbf{v}) \quad = \quad -\frac{1}{|K|}\sum_{\nu=1}^{p}(\mathbf{v}^\nu,w)^2 + (\frac{\partial}{\partial \mathbf{v}}g_\sigma^{-1}(\mathbf{v})w,w)$$

But the $\mathbf{v}^\nu$ are assumed orthogonal. Therefore, the use of Bessel's inequality yields:

$$\sum_{\nu=1}^{p}\frac{(\mathbf{v}^\nu,w)^2}{\| \mathbf{v}^\nu \|^2} \leq \| w \|^2 = 1 \Longleftrightarrow \sum_{\nu=1}^{p}(\mathbf{v}^\nu,w)^2 \leq \mathrm{V}_*^2 \mid K \mid$$

hence

$$D^2_{w^2}\mathbf{H}(\mathbf{v}) \geq (\frac{\partial}{\partial \mathbf{v}}g_\sigma^{-1}(\mathbf{v})w,w) - \mathrm{V}_*^2$$

for any w in S, $\| w \| = 1$. Then, a necessary and sufficient condition for an element $\mathbf{v}$ in S to be a minimum of $\mathbf{H}$ is:

$$(g_\sigma^{-1\prime}(\mathbf{v})w,w) - \mathrm{V}_*^2 > 0 \quad \forall\, w \in \mathrm{S}, \| w \| = 1$$

Theorem 4.2 follows directly from the last equation. Applying this equation to the case in which $\mathbf{v} = \mathbf{v}^\mu$ and keeping in mind that $g_\sigma^{-1\prime}(\mathbf{v}) = (g'_\sigma(g_\sigma^{-1}(\mathbf{v})))^{-1}$
the above condition reduces to

$$(\frac{w}{g'_\sigma(g_\sigma^{-1}(\mathbf{v}^\mu))},w) - \mathrm{V}_*^2 = (\frac{w}{g'_\sigma(\mathrm{V}_*^2\mathbf{v}^\mu)},w) - \mathrm{V}_*^2 = \frac{1}{g'_\sigma(\pm\mathrm{V}_*^3)} - \mathrm{V}_*^2 > 0$$

Since g_σ is odd, this can be rewritten as:

$$g'_\sigma(\mathrm{V}_*^3) < \frac{1}{\mathrm{V}_*^2} \text{ or } g'_\sigma(\mathrm{V}_*^3)\mathrm{V}_*^2 < 1$$

Let us compare the necessary and sufficient condition given by theorem 4.2 for the stability of the origin with the uniqueness condition for the general case (theorem 2.3). When $\mathbf{T}$ is defined according to (6), the $\mathbf{v}^\mu$'s being stationary solutions of (1) and therefore $\mathbf{v}^\mu(x) \in \{\mathrm{V}_*, 0, -\mathrm{V}_*\}$, we have:

$$|\mathbf{T}(x,y)| \leq \frac{p\mathrm{V}_*^2}{|K|} = \mathrm{M}.$$

In this case the condition for the origin to be the only solution is that $\sigma < \frac{1}{\mathrm{M}|K|} = \frac{1}{p\mathrm{V}_*^2}$. This is more restrictive than what follows from theorem 4.2. Therefore, for the case of orthogonal memories there exists an intermediate range for the values of σ ($\sigma \in [\frac{1}{p\mathrm{V}_*^2}, \frac{1}{\mathrm{V}_*^2}]$, which degenerates into a point if $p = 1$) for which the trivial solution $\mathbf{v} \equiv 0$ is an attractor, but not necessarily the only solution of (1). Note, in addition, that the conditions derived in theorems 4.1 y 4.2 are independent of p (number of memories); this is a direct consequence of the orthogonality of the memories.

5 Basins of attraction

On the base of the preceeding results, we can now estimate the size of the basins of attraction.

Theorem 5.1: for $p \geq 2$, the largest sphere contained in the basin of attraction of an orthogonal memory $\mathbf{v}^{\mu}$, $1 \leq \mu \leq p$, has a radius $k = \mathrm{V}_* \sqrt{\frac{|K|}{2}}$.

In other words, whenever $\| \mathbf{v}^{\mu} - \mathbf{v}_0 \| < k$, the distance $\| \mathbf{v}^{\mu}(\cdot,t) - \mathbf{v}(\cdot,t) \| \to 0$ when $t \to \infty$ (being $\mathbf{v}_0$ any initial condition for (1) and $\mathbf{v}$ the corresponding solution).

Proof: the radius of the basin will be the largest number $k > 0$ such that $D_w\mathbf{H}(\mathbf{v}^{\mu} + kw) > 0 \quad \forall w \in S, \| w \| = 1$. We know that

$$D_w\mathbf{H}(\mathbf{v}) = -\frac{1}{|K|} \sum_{\nu=1}^{p} (\mathbf{v}^{\nu}, \mathbf{v})(\mathbf{v}^{\nu}, w) + (g_{\sigma}^{-1}(\mathbf{v}), w)$$

Then:

$$\begin{aligned} D_w\mathbf{H}(\mathbf{v}^{\mu} + kw) &= -\frac{1}{|K|} \sum_{\nu=1}^{p} (\mathbf{v}^{\nu}, \mathbf{v}^{\mu} + kw)(\mathbf{v}^{\nu}, w) + (g_{\sigma}^{-1}(\mathbf{v}^{\mu} + kw), w) \\ &= -\frac{1}{|K|} \left\{ V_*^2 \mid K \mid (\mathbf{v}^{\mu}, w) + k \sum_{\nu=1}^{p} (\mathbf{v}^{\nu}, w)^2 \right\} \\ &\quad + (g_{\sigma}^{-1}(\mathbf{v}^{\mu} + kw), w) \end{aligned}$$

which is positive if and only if

$$(g_{\sigma}^{-1}(\mathbf{v}^{\mu} + kw), w) > \frac{1}{|K|} \left\{ V_*^2 \mid K \mid (\mathbf{v}^{\mu}, w) + k \sum_{\nu=1}^{p} (\mathbf{v}^{\nu}, w)^2 \right\}$$

for every direction w. By virtue of Bessel's inequality:

$$\sum_{\nu=1}^{p} \frac{(\mathbf{v}^{\nu}, w)^2}{\| \mathbf{v}^{\nu} \|^2} \leq \| w \|^2 = 1$$

and, consequently (remembering that $\| \mathbf{v}^{\nu} \|^2 = V_*^2 \mid K \mid$), the condition is satisfied by imposing $(g_{\sigma}^{-1}(\mathbf{v}^{\mu} + kw), w) > (g_{\sigma}^{-1}(\mathbf{v}^{\mu}), w) + kV_*^2$ or equivalently

$$\int_K \frac{g_{\sigma}^{-1}(\mathbf{v}^{\mu}(x) + kw) - (g_{\sigma}^{-1}(\mathbf{v}^{\mu}(x))}{k} w(x)\ dx > V_*^2 \tag{7}$$

In order to prove that inequality (7) holds no matter the direction w, let us take the worst case: w pointing to a different memory, say $\mathbf{v}^{\nu}$, i.e. $w = \frac{\mathbf{v}^{\nu} - \mathbf{v}^{\mu}}{\|\mathbf{v}^{\nu} - \mathbf{v}^{\mu}\|}$. It is easy to check that $\mathbf{v}^{\nu} - \mathbf{v}^{\mu}$ can take only values 0 and $\pm 2V_*$ and that, by virtue of the orthogonality, it is 0 exactly on one half of the domain K and $\pm 2V_*$ on the

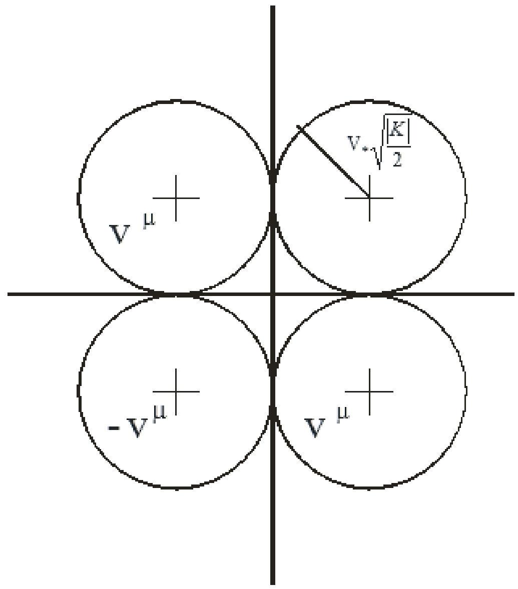

Fig. 2. Two orthogonal memories and their inverses, each one with norm $V_* \sqrt{|K|}$ and a spherical basin of attraction of radius $V_* \sqrt{\frac{|K|}{2}}$

other half. Then w is 0 on a subdomain of size $\frac{|K|}{2}$ and $\sqrt{\frac{2}{|K|}}$ on the remaining subdomain of equal size. Thus, condition (7) may be rewritten as

$$\frac{|K|}{2} \frac{g_\sigma^{-1}(\mathbf{v}^\mu(x) + k\sqrt{\frac{2}{|K|}}) - (g_\sigma^{-1}(\mathbf{v}^\mu(x))}{k\sqrt{\frac{2}{|K|}}} \frac{2}{|K|} > V_*^2$$

(multiplying numerator and denominator by $\sqrt{\frac{2}{|K|}}$), which holds if

$$k\sqrt{\frac{2}{|K|}} < V_* \quad \Longleftrightarrow \quad k < V_* \sqrt{\frac{|K|}{2}}$$

Finally, the largest spherical basin of attraction has a radius equal to $V_* \sqrt{\frac{|K|}{2}}$, since otherwise the basins would not be disjoint, because the distance between any two $\mathbf{v}^\mu$ and $\mathbf{v}^\nu$ is twice that quantity.

Note that this result does not imply that the basins of attraction are spherical. It only limits the radius of *spherical* basins for memories $\mathbf{v}^\mu$ and, consequently, for $-\mathbf{v}^\mu$ as well. Figure 2 shows a simplified bidimensional sketch.

6 Spurious memories

As a consequence of the nonlinearity of the dynamics under consideration, undesired fixed points appear in addition to those purposely stored in the synaptic operator $\mathbf{T}$ with the Hebb prescription. These are called *spurious states* or *spurious memories*.

It is possible to distinguish two types of spurious states: *mixture* and *non-mixture* memories. $\mathbf{v}$ is said to be a mixture state if it can be expressed as a linear combination of the stored memories: $\mathbf{v} = \sum_{i=1}^{q} \alpha_{\mu_i} \mathbf{v}^{\mu_i}$ with $q \leq p$, $\mathbf{v}^{\mu_i} \in \{\mathbf{v}^{\mu}\}$ and α_{μ_i} real constants. A spurious state for which no such $\{\alpha_\mu\}$ exists is called a non-mixture state.

6.1 Mixture spurious states

First note that, just like in the known discrete models, for every memory $\mathbf{v}^\mu$, $-\mathbf{v}^\mu$ is also a memory. In the simple case when $p = 1$, there exist only two spurious states: the origin ($\mathbf{v} \equiv 0$) and the inverse of the (unique) stored memory. Thus, there are no non-mixture states for $p = 1$. If $p \geq 2$, the analysis gets considerably harder.

We have already mentioned the fact that every mixture state is a fixed point if $\mathbf{v}(x) \in \{\mathrm{V}_*, -\mathrm{V}_*, 0\}$ $\forall x \in K$. This may be easily seen either by using the linearity of $h^{\mathbf{v}}$ or from the proof of theorems 4.1 and 4.2. It is also clear that only a small subset of $span\{\mathbf{v}^\mu\}_{\mu=1}^{p}$ contains spurious states. In particular, it follows that if $\mathbf{v}^\mu$ and $\mathbf{v}^\nu$ are memories, then $\pm\frac{1}{2}\mathbf{v}^\mu \pm \frac{1}{2}\mathbf{v}^\nu$ are fixed points of the dynamics. This implies that there exist at least $4\begin{pmatrix} p \\ 2 \end{pmatrix}$ spurious (mixture) states. These are in general unstable, as stated by the following

Theorem 6.1: If $\mathbf{v}$ is a mixture spurious memory and there exists $x \in K$ such that $\mathbf{v}(x) = 0$, then $\mathbf{v}$ is a saddle point of the dynamics.

Proof: let $\mathbf{v} = \sum_{\mu=1}^{q} \alpha_\mu \mathbf{v}^\mu$, $1 \leq q \leq p$ (renaming memories if necessary). $\mathbf{v}$ is piecewise constant (since so are the $\mathbf{v}^\mu$'s, and there is a finite number of them). Therefore, if $\mathbf{v}(x) = 0$ then it vanishes in a neighborhood U of x and it holds that $0 < \sum_{i=1}^{I} \alpha_{\mu_i} \mathbf{v}^{\mu_i} = -\sum_{j=1}^{J} \alpha_{\mu_j} \mathbf{v}^{\mu_j}$ at every point in U, being $\{\mathbf{v}^{\mu_i}\}_{i=1}^{I} \cup \{\mathbf{v}^{\mu_i}\}_{j=1}^{J} = \{\mathbf{v}^{\mu_i}\}_{\mu=1}^{q}$

Let us choose $w = \sum_{i=1}^{I} \alpha_{\mu_i} \mathbf{v}^{\mu_i} - \sum_{j=1}^{J} \alpha_{\mu_j} \mathbf{v}^{\mu_j}$ (we can neglect the normalizing constant). Then $w = 2\sum_{i=1}^{I} \alpha_{\mu_i} \mathbf{v}^{\mu_i}$ in U. Now compute

$$g_\sigma(h^{\mathbf{v}+\varepsilon w}(x)) = g_\sigma(h^{\mathbf{v}}(x) + \varepsilon h^w(x))$$

Keeping in mind that $h^{\mathbf{v}}(x) = V_*^2 \mathbf{v}(x)$ (by hypothesis and by virtue of the linearity of $h^{\mathbf{v}}$) and computing

$$h^w(x) = \sum_{\mu=1}^{q} \alpha_\mu h^{\mathbf{v}^\mu}(x) = V_*^2 \sum_{\mu=1}^{q} \alpha_\mu \mathbf{v}^\mu(x) = 2V_*^2 \sum_{i=1}^{I} \alpha_{\mu_i} \mathbf{v}^{\mu_i}(x)$$

(which holds for every $x \in U$), we obtain

$$g_\sigma(h^{\mathbf{v}+\varepsilon w}(x)) = g_\sigma(V_*^2 \mathbf{v}(x) + 2\varepsilon V_*^2 \sum_{i=1}^{I} \alpha_{\mu_i} \mathbf{v}^{\mu_i}(x))$$

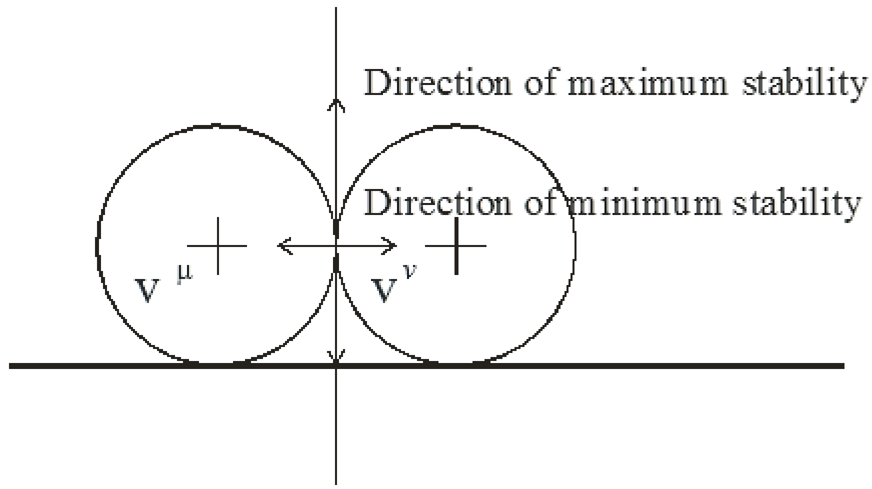

Fig. 3. A mixture spurious state. Dotted lines indicate limits for the spherical basins of the two memories which compose the spurious state

in U. The dynamics at $\mathbf{v}+\varepsilon w$ is (always in $U \subset K$):

$$\frac{\partial(\mathbf{v}+\varepsilon w)}{\partial t} = -2\varepsilon \sum_{i=1}^{I} \alpha_{\mu_i}\mathbf{v}^{\mu_i} + g_\sigma(2\varepsilon V_*^2 \sum_{i=1}^{I} \alpha_{\mu_i}\mathbf{v}^{\mu_i})$$

No matter the sign of $\sum_{i=1}^{I} \alpha_{\mu_i}\mathbf{v}^{\mu_i}$, the stability condition at $\mathbf{v}+\varepsilon w$ is $g_\sigma(2\varepsilon V_*^2 \sum_{i=1}^{I} \alpha_{\mu_i}\mathbf{v}^{\mu_i}) < 2\varepsilon \sum_{i=1}^{I} \alpha_{\mu_i}\mathbf{v}^{\mu_i}$ which is equivalent to $\sigma = g'_\sigma(0) < \frac{1}{V_*^2}$ which is false since, by hypothesis, $g_\sigma(\pm V_*^3) = g_{\sigma V_*^2}(\pm V_*) = \pm V_*$ (otherwise, g_σ would not have any fixed point apart from the origin). Then $\mathbf{v}$ is unstable in the direction w.

Now let us choose $w = \mathbf{v}$. With a similar reasoning, we get $g_\sigma(h^{\mathbf{v}+\varepsilon w}(x)) = g_\sigma(V_*^2\mathbf{v}(1+\varepsilon))$ and the stability condition at $\mathbf{v}+\varepsilon w$ is $g_\sigma(V_*^2\mathbf{v}(1+\varepsilon)) < \mathbf{v}(1+\varepsilon)$. Remembering that $\mathbf{v} = g_\sigma(V_*^2\mathbf{v})$, the condition for the system to be stable at $\mathbf{v}+\varepsilon w$ is $g'_\sigma(V_*^3) < \frac{1}{V_*^2}$, which always holds since the memories $\mathbf{v}^\mu$ are minima of $\mathbf{H}$ (cf. Section 2).

This leads to the useful

Corollary 6.2: The basins of attraction for mixture states have zero radius, in the sense of the $L^2(K)$ norm.

Example: for $q = 2$, all combinations of the form $\mathbf{v} = \pm\frac{1}{2}\mathbf{v}^\mu \pm \frac{1}{2}\mathbf{v}^\nu$ are actual spurious states and the situation can be easily illustrated (Figure 3). The direction of maximum unstability is given by $\mathbf{v}^\mu - \mathbf{v}^\nu$ and that of maximum stability is $\pm\mathbf{v}$ (directly towards or from the origin of coordinates).

6.2 Non-mixture spurious states

Unlike mixture spurious states, which can be calculated analytically, the non-mixture ones are difficult to find. Indeed, in the limit $p \rightarrow \infty$, the following property holds, no matter how **T** is defined.

Lemma 6.3: if $\{\mathbf{v}^{\mu}\}_{\mu=1}^{\infty}$ is complete[2] and $p \rightarrow \infty$, there are no non-mixture spurious memories.

Remark: since S=$L^2(K)$ here, a question about the meaning of lemma 6.3 may arise, i.e. is there some basis of $L^2(K)$ whose elements take on only two values, say $\pm V_*$? The answer is yes. The relevant example for $K \subset R$ (that can be extended to R^n) are the *Haar wavelets*, which form an orthonormal and complete basis in $L^2(-\infty, +\infty)$. They are bi-valued, but since they are normalized, such values change from one function to another. If normality is relaxed, it is possible to force them to take values in $\{V_*, -V_*\}$. If restricted to a bounded interval $K \subset R$, they form an orthogonal (but not orthonormal) and complete set in $L^2(K)$. However, if we construct **T** according to (6), this completeness can be only asymptotical: as we already saw, the number of orthogonal memories can be as large as desired, but it cannot be infinite.

7 The Resolution. Modified Hebb rules

Until now the metrics on X, the space of processing units, has not been specified. This is not necessary, because the contribution of each neuron to the generation of $h_t^{\nu}(x)$ for any $x \in K$ is independent of any distance, just as in the discrete cases [8],[9]. Thus, provided the memories are orthogonal, we have a memory of unbounded "resolution" in the sense that the patches of activation and quiescence can be as small as desired. Any function $v(x) : K \rightarrow \{-V_*, V_*\}$ with a set of discontinuities of zero Lebesgue measure can be memorized by the system. In the case $K \subset R$, this implies that the set of discontinuities must be countable.[3]

It is clear that a realistic model must have some lower bound for such "resolution", at least to take into account the finite size of individual neurons. We will briefly mention two ways to tackle this problem [15]:

i) Introducing a probability distribution on the size of the patches of activity of the memories. This is naturally done by characterizing the memories by the average length λ of their connected regions or "patches" with equal sign and introducing a distribution in which memories with vanishing λ have zero probability. The properties of the associative memory are well preserved, only imposing that the memories v^{μ} be independent, identically distributed and such that $E\left[v^{\mu}(x)\right] = 0 \quad \forall \mu,$ a.e. in K).

Figure 4 shows a computational evaluation of the retrieval capacity (as the mean L^2 distance between the initial state, when this is one of the stored memories, and the final state of the system) as a function on p and λ. The memories $v^{\mu} : K = [a,b] \longrightarrow \{V_*, -V_*\}$ follow a Poisson distribution with parameter λ. Note that stability decreases when p increases, but it only happens for large values of λ. For example, if $\mid K \mid = 8$, a value $\lambda = 1$, implies that each memory has, in average, eight sign changes, i.e., nine zones or patches of activity/inactivity.

[2]The set $\{\mathbf{v}^{\mu}\}$ is said to be *complete* in $L^2(K)$ if and only if the minimal subspace of $L^2(K)$ which contains it is the entire space $L^2(K)$.

[3]For example, the Dirichlet function, valuing 1 -or V_*- on every rational point and 0 -or $-V_*$- on every irrational point, cannot be a memory.

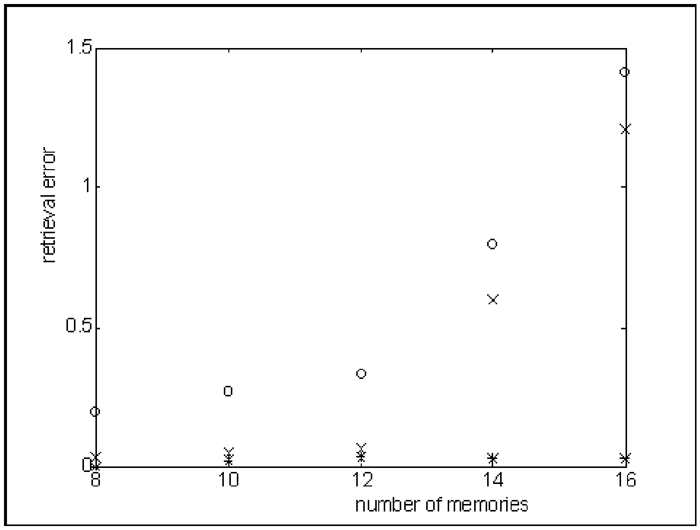

Fig. 4. Error in retrieval (average L^2 distance). Memories follow, for every pair $x \neq y$, a Poisson distribution with parameter λ. Here $X = [0,8]$. λ: 0.25 (dots), 0.5 (crosses), 1 (circles)

ii) Redefining the operator T, by modulating it with a range cut-off function that decreases with the distance between neurons. It is worth mentioning that the properties of the associative memory are preserved only when such cut-off function takes both positive and negative values, as the well known *mexican hat* function that has been used to model lateral inhibition in the brain cortex. This fact has the pleasant feature of being in agreement with the neurophysiological evidences of the modular organization of the cortex [4]. This modulation in which only the interaction between nearest neighbours is preserved causes a degradation of retrieval capacity, increasing the *crosstalk* between memories.
Figure 5 shows a computational evaluation of the retrieval capacity in a similar way as in Fig. 4, with the same λ's and p's. We define in this case:

$$T(x,y) = \frac{1}{|K|} \sum_{\nu=1}^{p} v^{\nu}(x) v^{\nu}(y) m(|y-x|)$$

with

$$m(s) = \{ \begin{array}{ccc} +A & if & 0 \leq s \leq \lambda_1 \\ -B & if & \lambda_1 < s \leq \lambda_2 \\ 0 & otherwise & \end{array}$$

$0 < A < B$ and λ_1,λ_2 appropriately chosen, i.e. the usual discrete approximation for the lateral interaction function. Also in this case stability decreases when p increases, but now this degradation becomes critical for small λ.

We can conclude that these two ways of limiting the resolution are complementary in the sense that, while in the case in which memories are directly affected without changing the Hebbian operator, the stability of the system decreases for large λ, when modulating T with lateral inhibition the effect is the inverse: retrieval gets better when increasing λ.

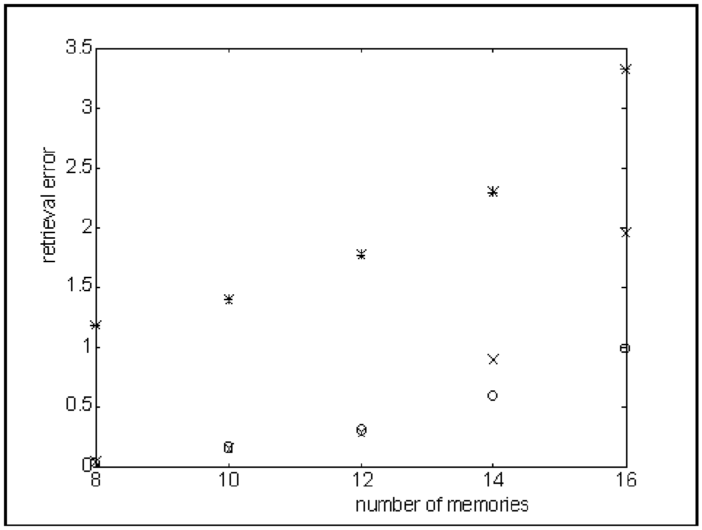

Fig. 5. Error in retrieval (average L^2 distance) for a synaptic operator with lateral inhibition (mexican hat function). Memories follow, for every pair $x \neq y$, a Poisson distribution with parameter λ. Here $X = [0,8]$. λ: 0.25 (dots), 0.5 (crosses), 1 (circles)

8 Beyond the Model: Modifying Dynamics

The purpose of turning associative memory models as much plausible as possible can be aimed also through the equation that governs the evolution of the system (1). In fact, a "landscape" of neuron activity such as depicted in Fig. 1 does not seem quite realistic. It would be more reasonable to expect some graded transition between patches with firing neurons and patches without activity.

We can provide H with some information about the smoothness of its argument, via the gradient of v. Let us redefine

$$H^*[v] = -\frac{1}{2}\int_K\int_K T(x,y)v(x)v(y)dxdy + \int_K\int_0^{v(x)} g_\sigma^{-1}(s)dsdx + \frac{\alpha}{2}\int_K |\nabla v|^2 dx$$

with $\alpha \geq 0$. This function "penalizes" derivatives; it takes the same value as the original (1) if v is a constant. Some computations with the variation of H^* and Green's formula lead to the following candidate equation to replace (1):

$$\frac{\partial v(x,t)}{\partial t} = -v(x,t) + g_\sigma\left(\int_K T(x,y)v(y,t)dy + \alpha\Delta v(x,t)\right)$$

Note that the v^μ taking values in $\{V_*, -V_*\}$ are no longer solutions (except for constants) in a classical sense, as every solution in the above equation has at least two continuous derivatives. However, theorem 2.3 still holds, i.e. when $\sigma \to \infty$, the attractors, if they exist, approach the asymptotic values of g_σ. It would be interesting to investigate whether the original v^μ can still be considered as solutions in some “weak” sense and to study the form of the solutions for the new dynamics and whether it actually provides any advance towards our objectives.

9 Back to the discrete domain

In order for the results of Section 2 to be valid, the only condition on X is to be a metric space (continuous or discrete). Therefore, all theorems of Section 2 hold for the discrete Hopfield model with continuous ranges of activity [9], provided that we replace the $L^2(K)$ norm with the usual euclidean norm and $\mid K \mid$ with N (number of neurons).

As for the results of Section 3, the situation is different. Clearly, theorem 3.1 is no longer true and the same happens, in general, for lemmas and corollaries based on the possibility of memories with an unbounded number of discontinuities, such as corollary 3.3 (which has no meaning for the discrete case). Instead, results concerning stability (Section 4) and size of attraction basins (Section 5) remain valid, with slight changes. The same is true for Section 6 (spurious states), except for theorem 6.3 (non-mixture spurious memories), which has no meaning for the discrete model.

Concerning the Hopfield model with discrete activities [8], first it must be remarked that its metrics based on the Hamming distance is not euclidean. However, being this metrics the discrete version of the L^1 norm, which is equivalent[4] to the quadratic L^2 norm, some results remain qualitatively true. This is the case for theorems 2.2 and 2.4 and corollary 2.5 (as it is well known), but the mathematical tools used here are of no help to obtain them. And, on the other hand, the results of Sections 3 to 6 have no meaning in general (when based on concepts of euclidean distance and directional derivatives, which are of no application in the discrete case).

10 Finite temperature considerations

Until now we where mainly concerned with what is, thermodynamically speaking, a zero temperature dynamics, i.e. a deterministic law of evolution. In the present section we will formulate a finite temperature version of the continuous system that we have introduced, in the same fashion as the stochastic versions of the discrete Hopfield model. In that approach, which is based on the Glauber thermodynamical model, the probability distribution of the state of the i-th processing unit S_i at an instant $n+1$ is given by

$$\mathbf{P}(S_i = \pm 1) = \frac{1}{1+\exp(\mp 2\beta h_i)}$$

As usual $h_i^{\zeta} = \sum_{i=1}^{N} \mathbf{T}_{ij}\zeta_j$ is the local field on site i when all neurons have activities labelled by ζ at time n. The parameter β represents the inverse temperature (we set the Boltzmann constant equal to 1).

[4] Two norms $\|\cdot\|$ and $\|\cdot\|'$ on a vector space V are said to be *equivalent* norms if there exist positive real numbers c and d such that $c \| x \| \leq \| x \|' \leq d \| x \| \; \forall x \in V$.

The activity pattern of the system is therefore defined through a Markov process that can be formulated as a *Master equation* [1]:

$$\mathbf{P}(\xi, n+1) = \mathbf{P}(\xi, n) + \sum_{\zeta \neq \xi} [\mathbf{W}_\beta(\xi \mid \zeta)\mathbf{P}(\zeta, n) - \mathbf{W}_\beta(\zeta \mid \xi)\mathbf{P}(\xi, n)] \tag{8}$$

where

$$\mathbf{W}_\beta(\xi \mid \zeta) = \prod_{i=1}^{N} \frac{1}{\exp\{-\beta \xi_i h_i^\zeta\} + 1} \tag{9}$$

In our case, ξ and ζ are elements in the (normed) state space S and $\mathbf{W}_\beta : \mathrm{S} \times \mathrm{S} \longmapsto [0,1]$ is the transition probability matrix.

Being S continuous infinite dimensional it is no longer possible to construct $\mathbf{W}_\beta(\xi \mid \zeta)$ from point transition probabilities. This problem can be circumvented with a proper definínition of $\mathbf{W}_\beta$. Equation (9) can be formally rewritten in the continuous time limit as the differential form of the Chapman-Kolmogorov equation [10]

$$\frac{\partial \mathbf{P}(\xi, t)}{\partial t} = \int_{\mathrm{S}} \{\mathbf{W}_\beta(\xi \mid \zeta)\mathbf{P}(\zeta, t) - \mathbf{W}_\beta(\zeta \mid \xi)\mathbf{P}(\xi, t)\} d\zeta \tag{10}$$

The integration has to be taken over the whole set S of possible activation patterns. In the present framework in which processing units are assumed to be a continuous metric space this integration has to be expressed as a Feynmann path integral [3]. Assuming for simplicity that $K = [a,b] \subset R$ and $\mathrm{S} = L^2(K)$, we define:

$$\int_{\mathrm{S}} \{\mathbf{W}_\beta(\xi \mid \zeta)\mathbf{P}(\zeta, t) - \mathbf{W}_\beta(\zeta \mid \xi)\mathbf{P}(\xi, t)\} d\zeta \triangleq \int_{\zeta(a)}^{\zeta(b)} \{\mathbf{W}_\beta(\xi \mid \zeta)\mathbf{P}(\zeta, t) - \mathbf{W}_\beta(\zeta \mid \xi)\mathbf{P}(\xi, t)\} D\zeta(\tau)$$

which is a functional integral over all $\zeta \in L^2([a,b])$.

Let us express $\mathbf{W}_\beta$ as a function of the jump r and of the initial state:

$$\mathbf{W}_\beta(\xi \mid \zeta) = \mathbf{W}_\beta(\zeta; r)$$

with $r = \xi - \zeta$. This leads to the following expression for the master equation:

$$\frac{\partial \mathbf{P}(\xi, t)}{\partial t} = \int_{\mathrm{S}} \mathbf{W}_\beta(\xi - r; r)\mathbf{P}(\xi - r, t) Dr - \mathbf{P}(\xi, t) \int_{\mathrm{S}} \mathbf{W}_\beta(\xi; -r) Dr$$

where the integrals are to be also understood as path integrals, now computed over all possible jumps $r \in \mathrm{S}$.

Now assume:

i) Only small values of the jump contribute to $\mathbf{W}_\beta$, i.e. $\mathbf{W}_\beta\ (\zeta; r) \approx 0$ for $\| r \| > \delta > 0$,

ii) $\mathbf{W}_\beta$ depends smoothly on ζ, that is $\mathbf{W}_\beta\ (\zeta + \Delta\zeta; r) \approx \mathbf{W}_\beta\ (\zeta; r)$ for $\| \Delta\zeta \| < \delta$ and

iii) The solution $\mathbf{P}(\xi, t)$ of (10) also varies slowly with ξ.

Then we can expand the first integral of the last equation up to second order in ξ [10]:

$$\begin{aligned}\frac{\partial \mathbf{P}(\xi,t)}{\partial t} &= \int_{S} \mathbf{W}_{\beta}(\xi;r)\mathbf{P}(\xi,t)Dr - \int_{S} \{\mathbf{W}_{\beta}(\xi;r)\mathbf{P}(\xi,t)\}'(r)Dr \\ &+ \frac{1}{2}\int_{S} \{\mathbf{W}_{\beta}(\xi;r)\mathbf{P}(\xi,t)\}''(r,r)Dr - \mathbf{P}(\xi,t)\int_{S} \mathbf{W}_{\beta}(\xi;-r)Dr\end{aligned}$$

where $\{\mathbf{W}_{\beta}(\xi;r)\mathbf{P}(\xi,t)\}^{(\nu)}$ stands for the ν-th order derivative of $\{\mathbf{W}_{\beta}(\xi;r)\mathbf{P}(\xi,t)\}$ at ξ applied to the element $(r,\ldots,r)$ in $\mathrm{S}\times\ldots\times\mathrm{S}=\mathrm{S}^{\nu}$ (being this ν-th order derivative a linear application from S^{ν} onto R). The first and fourth terms of the sum cancel. Then, the equation takes on the form

$$\frac{\partial \mathbf{P}(\xi,t)}{\partial t} = -a_1(\xi,\mathbf{P}) + \frac{1}{2}a_2(\xi,\mathbf{P}) \tag{11}$$

with

$$a_{\nu}(\xi,\mathbf{P}) = \int_{S} \{\mathbf{W}_{\beta}(\xi;r)\mathbf{P}(\xi,t)\}^{(\nu)}(r,\ldots,r)Dr \qquad (r,\ldots,r)\in \mathrm{S}^{\nu}$$

We have obtained an expression for the time evolution of the probability distribution **P** which has the form of a Fokker-Planck equation in an infinite dimensional normed space. In fact, for the case S=R we get the well-known one-dimensional Fokker-Planck equation:

$$\frac{\partial \mathbf{P}(\xi,t)}{\partial t} = -\frac{\partial}{\partial \xi}\{a_1(\xi)\mathbf{P}\} + \frac{1}{2}\frac{\partial^2}{\partial \xi^2}\{a_2(\xi)\mathbf{P}\}$$

with

$$a_{\nu}(\xi) = \int_{S} r^{\nu}\mathbf{W}(\xi;r)dr$$

The interpretation of the Fokker-Plank equation is simple. The (functional) probability density $\mathbf{P}(\xi,t)$ is defined on the space of all possible activation patterns $\xi(x)$ and describes the evolution involved in the "retrieval" process of some particular activation pattern starting from a given initial (arbitrary) configuration. The transition probability matrix $\mathbf{W}_{\beta}(\xi;r)$ governs this evolution through the "moments" involved in the Fokker Plank equation. Upon retrieval, the (stationary) probability density reached when t goes to infinity is expected to be peaked in the neighborhood of the recalled memory. This final activation pattern is however not expected to be δ-like due precisely to the finite temperature fluctuations. A pure deterministic case can only be recovered by letting β go to infinity.

Note that this derivation of the finite temperature dynamics for the continuous system is a natural extension of the master equation (8) based on the transition probability matrix (9). Therefore, it could have been applied to the analog Hopfield model [8] in a direct way, since that case would have required only a simple multivariate formulation of (11) in the form

$$\frac{\partial \mathbf{P}(\xi,t)}{\partial t} = -\sum_{i=1}^{N}\frac{\partial}{\partial \xi_i}\{a_1^i(\xi)\mathbf{P}\} + \frac{1}{2}\sum_{i=1,j=1}^{N}\frac{\partial^2}{\partial \xi_i \partial \xi_j}\{a_2^{ij}(\xi)\mathbf{P}\}$$

being now a_1^i the usual first order moment for the i-th coordinate, a_2^{ij} the element (i,j) of the covariance matrix and N the number of processing units.

11 Conclusions

We introduced a formal theoretical background, including theorems and their proofs, for our neural network model with associative memory in which processing units are elements of a continuous metric space. This approach is intended as a generalization of the previous ones due to Little and Hopfield. Our main purpose was to provide a mathematical foundation in the sense that it is actually possible to formulate a system of associative memory with graded response units and a continuous topological structure on the set of such units, obtaining a more biologically plausible model of associative memory.

On the other hand, our approach preserves the salient features that made attractive all the discrete models, especially the levels of continuity that the second Hopfield model [9] added to the discrete one [8]: graded response of the activation functions and continuous scale of time, via the transition from discrete to continuous, differential equation dynamics.

Firstly (Section 2) general results were proved assuming only a symmetric weight matrix **T** with non-negative diagonal elements. These results are generalizations of well known properties of discrete, Ising-type models.

Then (Sections 3 to 6) we analyzed the case when the memories are orthogonal and the synaptic operator is constructed following the autocorrelation (Hebb) rule. We proved:

- *Hebb rule*: it can be naturally extended to the infinite dimensional case.
- *Capacity*: any finite set of orthogonal memories can be stored and retrieved. However there are, concerning capacity, some differences with regard to discrete approaches.
- *Stability*: there exist necessary and sufficient conditions for the memories and the origin to be stable, in terms of the relation between parameters of the transfer function g_σ.
- *Basins of attraction*: they have the same radius for all memories, positive in the L^2 norm.
- *Spurious memories*: they exist. If a spurious state vanishes at some point, then its basin of attraction has zero radius ($\| L^2 \|$) and it is a saddle point of the dynamics.

When considering some "weak" aspects of the model from the point of view of its biological plausibility, possible variations were discussed:

- It is possible to impose a "finite resolution" to the stored memories by limiting the minimum size of the activity patches. We have briefly indicated two possible alternatives for doing this (Section 7).
- It is possible to modify the equation that governs the evolution of the system (1), in order for some graded transition to take place between patches with firing neurons and patches without activity (Section 8).

The validity of our results was discussed when applied to the classical, discrete models of associative memory [8][9]. Such application looks more natural for the Hopfield model with graded response [9], since in this case the concepts of euclidean distance and directional derivatives remain valid, while in the discrete case [8] only some general, well known properties (concerning stability and convergence to attractors) are preserved, maintaining anyhow the qualitative similarity with the infinite dimensional system.

We also showed (briefly reviewing our previous approach [16]) that a continuous model can be extended to the case in which the evolution law is non deterministic, in the same way as the Hopfield, discrete model is extended to include finite temperature effects through the Glauber dynamics. We have shown that this leads to a Fokker-Planck equation (in an infinite dimensional normed space) that governs the time evolution of the probability density distribution defined in the space of functions describing all possible activity patterns.

In view of this multiplicity of possible alternative formulations, the question arises about how and on the base of which parameters might all of them be compared. It seems natural to include among these criteria at least those concerning capacity and stability. Regarding the first one, it is clear that the key issue is not the finiteness or infiniteness of the storable and retrievable memories (since we have seen how to incorporate control parameters in order to bound the memory capacity for the sake of biological and also technological plausibility). The key issue is, we think, -and this leads to the second criterium mentioned above- rather the fact that, in any case, degradation in the classical Hopfield model is not smooth but certainly subject to a first transition dynamics implying a sudden, catastrophic loss of recoverable information as soon as a certain critical load value is reached [1]. Nothing like this (instability, sudden loss of information) happens in neither the continuous topology model nor (and even less) in that with finite temperature.

The above observations are valid even regardless of the questions concerning comparative costs in time and computing resources: clearly, the price for such features as stability and biological plausibility is, in part, a qualitative increase of the complexity of the model; mathematically speaking, it implies the numerical integration of nonlinear differential equations instead of a set of discrete difference equations.

We think that this approach can be useful for modelling in biology and neurophysiology (bringing a wide range of possibilities of application in technology, robotics and even in our understanding of cognitive processes as much as normal and pathological brain function). It retains all the stylized facts that have made attractive the Hopfield neural network model and its modifications, yet giving the possibility of modelling the brain cortex as a continuous space. In other words, it integrates two levels of continuity:

- Continuous response units, which was already present in [9] and permits description of relevant neural activity by firing rates, rather than merely by the presence or the absence of an individual spike.

- Continuous topology of the neural system, obtaining a model of associative memory that reconciles biological evidence of a continuum of the neural tissue with descriptions provided by discrete models inspired in Ising systems.

In addition, the results proved here can also be useful, with the limitations pointed out in Section 9, when performing the reverse track of what we have done, namely when reconsidering the discrete case in the light of the knowledge of what happens if the state space is continuous. This ”reverse track” could have the double utility of confirming the power of generalization of the proposed model and, at the same time, enhance our understanding of the properties of classical approaches in a more general context.

Acknowledgements: this work was partially supported by grants X166 from the University of Buenos Aires and 26001 from the National Agency for Scientific Research, Argentina.

References

1. **Amit, D. J.**, *Modeling Brain Function.* Cambridge: Cambridge University Press, 1989.
2. **Dreyfus, G.**, *Neural Networks. Methodology and Applications*. Berlin: Springer, 2005.
3. **Feynman, R. P.** and **Hibbs, A. R.**, *Quantum Mechanics and Path Integrals.* New York: McGraw-Hill, 1965.
4. **Fodor, J. A.**, *The Modularity of Mind.* Cambridge: MIT Press, 1983.
5. **Glauber R. J.**, ”Time-dependent Statistics of the Ising Model”. *Journal of Math. Phys.*, 4:294-307 (1963).
6. **Hagan, M. T.**, **Demuth, H. B.** and **Beale, M. H.**, *Neural Network Design*. Boston: PWS Publishing Company, 2002.
7. **Hebb, D. O.**, *The Organization of Behavior: A Neuropsychological Theory.* New York: Wiley, 1949)
8. **Hopfield, J. J.**, ”Neural networks and physical systems with emergent collective computational abilities”. *Proc. Natl. Acad. Sci.*, 79:2554-2558 (1982).
9. **Hopfield, J. J.**, ”Neurons with graded response have collective computational properties like those of two-state neurons”. *Proc. Natl. Acad. Sci.*, 81:3088-3092 (1984).
10. **Kampen, N. G. van**, *Stochastic Processes in Physics and Chemistry.* Amsterdam: Elsevier, 1997.
11. **Little, W. A.**, ”The Existence of Persistent States in the Brain”. *Mathematical Biosciences*, 19:101-120 (1974).
12. **Little, W. A.**, ”Analytic Study of the Memory Storage Capacity of a Neural Network”. *Mathematical Biosciences*, 39:281-290 (1978).
13. **MacLennan, B. J.**, ”Field Computation in Natural and Artificial Intelligence”. *Information Sciences*, 119:73-89 (1999).
14. **MacLennan, B. J.**, ”Continuous Information Representation and Processing in Natural and Artificial Neural Networks”. *JCIS 2002*, Research Triangle Park, North Carolina, USA, 561-565 (2002).
15. **Segura, E. C.** and **Perazzo, R. P. J.**, ”Associative memories in infinite dimensional spaces”. *Neural Processing Letters*, 12(2):129-144 (2000).
16. **Segura, E. C.** and **Perazzo, R. P. J.**, ”Biologically Plausible Associative Memory: Continuous Unit Response + Stochastic Dynamics”. *Neural Processing Letters*, 16(3):243-257 (2002).
17. **Segura, E. C.**, ”Infinite Dimensional Associative Memories: Foundations (with Theorems)”. *Res. Comp. Science*, 28(3):15-28 (2007).

Enrique Carlos Segura is a Professor and Researcher at the University of Buenos Aires. He obtained M.Sc. degrees in Computer Sciences and in Mathematics (1988), and a Ph.D. in Mathematics (1999), all from the University of Buenos Aires. Previously he had a Fellowship from the CONICeT (National Research Council, Argentina), 1988-1992. He also received a grant from the Thalmann Foundation for research at the Polytechnical University of Catalonia, Spain (1996). From 2001 to 2003 he was a Research Fellow at South Bank University, London. His main areas of interest are Neural Networks (theory and applications) and, more generally, cognitive models of learning and memory.