Persistent activity in limbic system neurons: neurophysiological and modeling perspectives

Sidney I. Wiener *, Angelo Arleo

Laboratoire de Physiologie de la Perception et de l’Action, CNRS-College de France, 11 place Marcelin Berthelot, 75231 Paris Cedex 05, France

Abstract

Neural activity persisting for one to hundreds of seconds has been postulated to be a substrate of memory. This review article illustrates examples of such activity in limbic system structures including the hippocampus, postsubiculum, and the anterodorsal thalamus. These neuronal responses include better known correlates with the spatial position as well as with head direction of the animal relative to its environment as well as other lesser known examples. Since head direction responses are greater when the animal is actively moving than when passively rotated, it has been proposed that there might be a general mechanism where the behavioral state of the animal can provide modulatory gating of such persistent signals. This would regulate the relative influence of these signals on downstream structures. Neural network attractor models of the head direction cell system are presented to demonstrate how these responses might originate, as well as the dynamics by which they are updated during movements.

Keywords: Attractor networks; Head direction cells; Place cells; Hippocampus; Memory

1. Introduction

While neurons respond to brief stimuli on the timescale of tens, or even hundreds of milliseconds, we are capable of remembering information for much longer delays. Persistent neuronal activity for periods lasting on the order of seconds has been proposed to carry out this function (see [40] for review) at intermediate time scales prior to consolidation in long term memory. Such a trace would not require those persevering changes in synaptic and neuronal biochemistry and morphology that occur over longer delays. The mechanisms of sustaining persistent activity could involve particular membrane properties as well as dynamic circuit interactions. Many investigators have focused on processing of visual inputs and of discriminative stimuli in delayed response tasks for studying and modeling persistent activity. One particularly useful tool for this has been continuous attractor networks [7,8].

An attractor model is a recurrent neural network in which formal neurons representing similar states (e.g., neighboring locations in two-dimensional physical space) are connected by strong excitatory collaterals, whereas neurons encoding distant states in the state space strongly inhibit each other. The intrinsic dynamics of an attractor network is determined by the interaction between the excitatory and inhibitory signals transmitted between the formal neurons. This allows the system to settle down to stable (self-sustained) attractor states, where subgroups of neurons are active while others remain silent. Extrinsic input signals can then be utilized to shift the state of the network over the continuous attractor state space [1,2,44].

Recently, attention has turned to neurophysiological demonstrations of persistent activity (i.e. sustained discharges) in limbic system neurons. Many of these have been examined principally in terms of spatial orientation and navigation functions, although they could well help to better understand more general cognitive functions attributed to the limbic system, particularly the hippocampus. These discharges are of considerable interest since they are correlated with supramodal, categorical information and could contribute to constructing and reconstructing high-level representations of the ongoing environmental context [42]. Notable examples of this are principal neurons of the hippocampus selective for the position of the body in space (see contributions by
McNaughton and Poucet in this volume) as well as of other limbic neurons selective for the orientation of the head, the ‘head direction cells’ (for review, see [35]; see Fig. 1 for example).

2. Examples of persistent activity in the limbic system

Place responses were first observed in pyramidal neurons of the rat hippocampus [21]. They discharge as the animal passes through a distinct region of the environment. Different neurons are selective for different places and the firing fields for an ensemble of neurons span the environment. Typical reports show the neuronal discharges occurring for durations on the order of 1 s, the time required to cross the firing field. If the animal is then placed in a different environment, the neurons adopt new spatial selective responses (or become silent). There is no apparent topography of the anatomical distribution of these neurons within the hippocampus since cells with adjacent or overlapping fields observed while the animal is in one environment can have totally unrelated responses in another environment. The hippocampal principal neurons are under strong inhibition and at any given moment only a tiny fraction of the neurons will be active [38]. Teîler and DiScenna [37] put forward the hypothesis that this subset of active neurons corresponds to the coactivation of a unique ensemble of cortical neurons appropriate to the situation that the animal is currently experiencing. During sleep these same hippocampal neurons are reactivated [24] and this occurs in the same sequence as when the animal was awake [45]. This process has been proposed to facilitate the consolidation of long term episodic memories by the corresponding extra-hippocampal cortical areas.

Note that these same neurons will again be activated when the animal revisits the place days, or even months later (which we consider as distinct from the persistent

---

Fig. 1. Place cell and head direction cell recordings. (A) Typically the unrestrained rat which has been implanted with chronic electrodes is connected via the headstage with a cable (not shown). The animal then forages for food in an open field (about 76 cm diameter), or performs other tasks. (B) Overhead view of hippocampal place responses recorded in a similar area by Muller et al. [18]. The total number of action potentials is summed for visits to respective parts of the enclosure. The lightest gray zones indicate the absence of cell firing while the shaded zones indicate progressively higher firing rates, usually with a peak rate between 1 and 20 action potentials per second at the center of the firing field, with variations among cells. (C) For head direction cell recordings, two lamps are also fixed to the headstage permitting head direction measurements, shown in this overhead view of recordings in a cylinder (like A) with a video sampling rate of 60 Hz. (D) The directional firing curve is computed by summing together all action potentials recorded for the respective head orientations (here the bin size is 6°) as the rat foraged for food scattered along the cylinder floor. The thick curve corresponds to actual data while the thin curve is a pseudo-Gaussian analytic curve based on these data (see Ref. [34] for details). (A, C, D) provided by M.B. Zugaro. Panel B is adapted with permission from Muller et al. [18], copyright 1987 by the Society for Neuroscience.)
activity discussed here). Furthermore the place responses occur when the animal is returned to the environment even if some of the principal environmental cues have been removed [20]. This has been considered as an example of auto-associative function of the hippocampus (reviewed by Redish [25, pp. 100–107], and Ref. [42]). These models depend upon recurrent connections in the hippocampal CA3 field producing attractor states where a subset of landmark cues triggers activation of a fraction of the subpopulation of neurons corresponding to the representation of a place (for example). The persistent activation of these neurons is required to permit completion of this partial activity pattern. This also permits the neurons to remain active after cues are no longer perceptible, for example when the room lights go off.

The hippocampal pyramidal neurons can also fire continuously if the rat runs in a running wheel placed in an appropriate position and orientation in its environment (Fig. 2; Ref. [14]). These ‘wheel cell’ responses are velocity dependent and drop to zero when the animal is immobile (not shown). This provides a useful model for studying persistent activation of hippocampal neurons, overcoming several difficulties with place responses: lack of response during complete immobility [9] and the need to motivate the animal to pass through the field. Note however that the hippocampal place and wheel neurons tend to discharge in phase with the theta rhythm (about 8 Hz) of the hippocampal EEG, which for the memory storage hypothesis can be an advantage (see Ref. [15]) or a potential complication due to the discontinuity of the signal.

Two other types of persistent activity will be discussed in hippocampal principal neurons. The first involves recordings in rats performing a somewhat complex task. The goal of the experiment was to determine if hippocampal spatial responses had access to non-visual sensory inputs related to body rotations (e.g., vestibular or proprioceptive). The experiments were performed in a high-walled square arena mounted on a rotatable swivel. Water reservoirs were situated at each of the corners and at the center of the floor. The partially-deprived animals were first trained to go to the corner currently in a particular cardinal direction (southeast) for a water reward. Then lights illuminating a cue card were turned off leaving the chamber in total darkness, and the box (still containing the rat) was rotated by multiple of 90°, and the rat had to find the corner now in that same cardinal position. (The experiments were conducted in a Faraday chamber preventing possible use of geomagnetic cues.) Of particular interest here is the unexpected observation that a subgroup of the hippocampal principal neurons discharged continuously during the period that the lights were off and during or after the rotation (see Fig. 3; [16]). This is consistent with the notion of storage of spatial information during the delay until the animal went to the location with the reward. These discharges generally lasted at least several seconds, and went on for as long as on the order of 10 s.

![Fig. 2. Recordings of persistent hippocampal cell activity in rats in a running wheel by Hirase et al. [14]. (A) Recordings of the hippocampal EEG (in theta mode), the activity of a hippocampal pyramidal neuron ('wheel cell spikes') and the head position which remained virtually immobile. (B and C) The firing rate remains constant as the animal runs at a constant speed. (D) The neuron fired selectively when the wheel was placed at a certain position and at a certain orientation within the experimental enclosure (represented similarly to Fig. 1B). (Reproduced with permission from [14].)
A second situation where we have observed persistent activity in hippocampal neurons concerned rats performing on a plus maze where different quantities of water were available at the respective arms [33]. Multiple drops of water were distributed at 1 s intervals. The rats were required to visit the arms with the largest rewards first. They learned new reward distributions in the maze at the beginning and middle of each session. The principal finding was that place responses remained constant after reward values associated with those locations were changed. Of particular interest here was the unexpected discovery of other cells that responded specifically while the rats were immobile waiting for multiple rewards (see Fig. 4). In contrast with place responses which are prominent while the animal is moving, this activity began the instant that the animal was positioned at the water reservoir and continued for as long as 5 s, as the rat remained there waiting for and consuming the water reward. This response occurred only as the rat was at one of the four reward sites, and did not shift location when the reward value at that site was changed. Unfortunately the experimental design was not constructed to help determine what type of information storage this activity might correspond to. It is possible that there is a link between this type of response and those described in the previous paragraph, and further research is required to elucidate a role for these discharges in hippocampal memory functions.

Complementing the hippocampal place responses are the head direction cells, neurons found in a network spanning numerous different structures centered on the brain’s limbic system, ([35]; Fig. 1C and D), including postsubiculum, retrosplenial cortex, anterodorsal thalamic nucleus, lateral mammillary nucleus as well as caudate nucleus. These neurons discharge selectively as the rat, mouse or monkey orients its head in a particular direction. Different cells are each selective for their own particular ‘preferred direction’. Although the responses are primarily anchored on visual cues lying in the background of the environment [47], there is also an influence by self-motion cues [6,34,47,49], and the cells continue to discharge as the animal moves about in darkness [34]. Both place and head direction neurons show a dramatic attenuation of their responses if the animal is tightly restrained [34,36]. However the discharges of head direction cells have no apparent relation to the hippocampal theta rhythm which accompanies certain types of movements—there is no rhythmic
modulation, and the responses occur whether theta is present or not. Thus if the animal remains oriented in a given direction, the corresponding head direction cells will remain active for an indefinite period of time.

Thus these neurons seem to be an excellent example of an ongoing neural trace of the orientation of the head, and could underlie certain memory processes. This information could be exploited by diverse brain structures which have been, respectively, linked with complementary types of spatial orientation strategies. For example, in contrast with the hippocampus, the caudate nucleus has been linked with memory for sequences of actions, or procedures, as well as in simple displacements towards visible (or otherwise perceptible) goals [17,22,23]. Head direction cells have been shown in the medial caudate nucleus [41], and they may well provide orientation references for such processes.

The hippocampus has been implicated in a second type of navigation strategy, involving the use of configurations of landmark cues, often in relation to the contours of the surrounding environment [19]. This would permit navigation toward unmarked goal locations, detours around obstacles and shortcuts. The hippocampus receives a wide spectrum of inputs from diverse cortical and subcortical areas (including the head direction cell network) permitting the elaboration of highly sophisticated activity correlates, including the place responses discussed above. Interestingly, the spatial or behavioral correlate of each cell changes unpredictably whenever the rat starts performing a different behavioral task or goes into a different environment. For example, if the animals are faced with a non-spatial (for example olfactory) discrimination task, the ‘place cells’ can take on task-related, non-spatial responses (e.g., Ref. [43]). Such shifts in the hippocampal representation of the current context is referred to as ‘re-mapping’ when involving place responses. While the mechanism of such shifts remains unknown, one simple proposal is that of mass action—those inputs to the hippocampal system that are the most active will have the greatest influence on what the current representation is. Our recent data from recordings of head direction cells provide a possible basis for how this may come about [48].

This experiment was originally designed to determine the role of motor command signals in the generation of...
of the head direction signal. In the recording sessions, first the baseline responses of anterodorsal thalamic head direction cells were determined as the rat searched for food pellets randomly distributed on a circular platform (thus permitting sampling of cell activity for all heading directions at all locations). Next the same cell response was determined as the rat was immobile, drinking water from a spout at the center of the platform. The rat and platform were rotated in order to sample cell activity for different head orientations and hence construct response curves (see Fig. 5). On average the neurons showed a decrease of 30% of the peak firing rate in the passive rotation condition as compared to active movements [48]. In contrast, the directional responses did not shift to new orientations or show changes in resolution or selectivity (as measured by tuning curve width). Thus the amplitude of the directional responses is modulated by the motor state of the animal.

Such reductions of peak firing rate of the head direction neurons during passive movements would be expected to reduce the intensity of the persistent signal transmitted to downstream structures such as the hippocampus. This could bias the hippocampus to be more sensitive to other inputs, perhaps from other non-spatial modalities, eventually leading to remapping or other changes in response correlates. This corresponds intuitively to anecdotal observations in man that while traveling to a new destination, the navigating driver often learns the way better than the passengers do.

This also provides a potential mechanism for the processing and gating of persistent activity signals in brain circuits. Modulation of the amplitude of the responses by other factors (e.g., attention, behavioral or motivational state) could regulate the impact of the neurons on downstream brain structures.

3. Attractor network models of the head direction cells

The functional role of persistent activity of limbic neurons has been the subject of a large body of theoretical work (see Ref. [28]). Most of these studies identify the continuous attractor network model as a useful paradigm to describe the persistence of a pattern of activity over time.

Here we review some theoretical models of the head direction (HD) cells. They all postulate a one-dimensional attractor network as a plausible mechanism underlying the persistent (self-sustained) activity of these limbic neurons, allowing the animal to maintain its internal sense of direction even in the absence of extrinsic sensory inputs.

Skaggs et al. [31] propose a one-dimensional ring attractor to represent a population of cells with the preferred directions evenly distributed in the network (see Fig. 6). Neurons in the network are coupled by recurrent connections so that cells having nearby preferred directions are linked by strong excitatory synapses, whereas cells with distant preferred directions are connected by strong inhibitory projections. Due to the attractor properties of the ring, the network exhibits, at each time, a stable representation consisting of a single cluster of active cells. In the absence of head rotation, this firing pattern is maintained indefinitely, similar to actual head direction cells that remain persistently active when the head of the animal remains oriented in a given orientation. During head rotations, the subset of active neurons shifts along the ring according to the head angular velocity, \( \omega \). The model predicts the existence of ‘rotation cells’, neurons responsive to stimulation of the vestibular system. Two populations of ‘rotation cells’ are employed; one group is responsive for clockwise turns, the other for counterclockwise turns. The firing rate of each rotation cell is proportional to the magnitude of the angular velocity, \( |\omega| \). During clockwise turns, for

![Fig. 6. One-dimensional attractor network representing a population of head direction cells (grey circles) with evenly distributed preferred directions (reproduced with permission from [31]). Each formal unit \( i \) excites its neighbors \( j \) as a function (typically Gaussian) of the distance between their preferred directions, i.e. \( w_{ij} = f(\theta_i - \theta_j) \). In addition, each cell \( i \) inhibits all other units of the network. Two pools of rotation cells (white circles) are responsible for the shift of the packet of activity along the ring. For instance, the firing rate of a clockwise rotation cell \( k \) is non-zero only during clockwise head turnings and is proportional to the magnitude of the angular velocity \( |\omega| \). The rotation cell \( k \) receives an excitatory afferent from the HD cell \( i \) having the same preferred direction, i.e. \( \theta_i = \theta_k \), and projects excitatory efferents to HD cells to the right of cell \( i \). Thus, during clockwise rotations, the subset of active units will shift rightward according to the angular velocity \( \omega \). (Reproduced with permission from Skaggs et al. [31], Adv. NIPS 7, MIT Press.)

Reproduced with permission from [31], Adv. NIPS 7, MIT Press.)
instance, the clockwise rotation neurons will excite the head direction cells to the right of the current peak, thereby shifting the hill of activity rightward.

Zhang [46] puts forth a mathematical framework to describe the intrinsic dynamics of the HD population and considers the distribution of the preferred directions $\theta$ over the entire population as a continuum. He formulates an analytical attractor model accounting for Gaussian-shaped HD tuning curves, stationary self-sustaining activity in the absence of extrinsic signals, and dynamical state shift based upon external inputs. The attractor network is constructed by defining a symmetric weight distribution function $W_1(\theta)$ describing the average synaptic strength between HD cells encoding different directions $\theta$. The shift of the attractor state is produced by adding to the symmetric weight function an antisymmetric component $W_2(\theta)$ (this is referred to as anisotropy in the connection matrix). A key result of the model is that the only analytical solution to rigidly shift the activity profile without deforming its shape consists of taking the antisymmetric weight component as proportional to the derivative of the symmetric component, i.e. $W_2(\theta) = \gamma \cdot W_1(\theta)$ as previously demonstrated by Droulez and Berthoz [10] in other systems. The velocity of the shift is determined by the factor $\gamma$ which defines the magnitude of the antisymmetric component in the weight distribution. It is shown that if $\gamma$ is proportional to the angular head velocity $\omega$, then the network works as a perfect velocity integrator. On the other hand, it is shown that if $\gamma$ is also proportional to the angular head acceleration $\dot{\omega}$, then the directional shift over the population tends to lead the true head direction by a constant time delay. Zhang [46] suggests that this might explain the experimental data showing that during head rotations, directional responses of HD cells in the anterodorsal thalamic nucleus (ADN) and in the lateral mammillary nuclei (LMN) occur prior to the moment that the head is actually oriented in the preferred direction (referred to as anticipatory responses [5]).

Redish et al. [26] propose a thalamocortical circuit, involving the anterior thalamic nucleus (ADN) and the postsubiculum (PoS), to model the head-direction system. Here both ADN and PoS are characterized by an attractor network architecture. Interconnections between ADN and PoS are between their excitatory pools $E_{PoS}$ and $E_{ADN}$, respectively. A set of synapses, termed ‘matching connections’, interlinks neurons in $E_{PoS}$ with those neurons in $E_{ADN}$ having equivalent preferred directions $\theta$. In addition, a set of left and right-offset projections is responsible for the shift of the activity pattern over time (each unit $i \in E_{PoS}$ with preferred direction $\theta_i$ has a left-offset connection to unit $j \in E_{ADN}$ such that $\theta_j = \theta_i - \delta$, and has a right-offset connection to unit $k \in E_{ADN}$ such that $\theta_k = \theta_i + \delta$). All offset synapses are modulated by the head angular velocity $\omega$. Thus, during rightward (clockwise) turns, right-offset connections have a strength proportional to the magnitude of angular velocity $|\omega|$, whereas left-offset connections have strength zero. By contrast, during leftward rotations, left-offset connections are proportional to $|\omega|$ and right-offset connections are zero. When the head is still, both left- and right-offset connections have strength zero, and matching synapses synchronize the PoS and the ADN representations. The model also involves an inhibitory input to ADN coming from the lateral mammillary nucleus. This input is assumed to be proportional to the magnitude of the angular velocity $|\omega|$ and works as a gain control mechanism to compensate for modulated offset connections. This allows the system to maintain the shape of the ADN hill of activity nearly unchanged during rotations. Without this predicted inhibitory modulation, the combined input from the offset and matching connections would distort the ADN tuning curves. However, anatomical findings indicate that the mammillothalamic synapses are excitatory in rats [11,12], requiring another region to replace the mammillary bodies in providing inhibitory inputs.

Goodridge and Touretzky [13] extend the above thalamocortical circuit by taking into account directional responses in the lateral mammillary nuclei (LMN). Attractor networks are employed for the PoS and LMN populations, whereas there is no attractor dynamics in ADN. The model of LMN involves two subnetworks $LMN_{cw}$ and $LMN_{ccw}$, each containing an attractor module. Both LMN pools receive directional information from the PoS by means of matching connections. In addition, $LMN_{cw}$ and $LMN_{ccw}$ activities are modulated by the angular velocity signal $\omega$ during clockwise and counterclockwise turns, respectively. These two LMN populations project to ADN but with opposite offset connections ($\pm \delta$) producing bimodal ADN tuning curves. ADN cells synapse on PoS neurons with matching preferred directions and update the position of the attractor state (the region of activity) according to the current head angular velocity. The authors employ this model to examine the relationships among ADN, PoS, and LMN as well as some properties of their representations. For instance, they study the deformation of the ADN tuning curves during head rotations and replicate the experimental results reported by Blair et al. [4]. However, the model by Goodridge and Touretzky does not fully capture the relation between the anticipatory time intervals observed in LMN and ADN of rats [4,32].

Blair et al. [3] put forth an attractor-integrator network based on the anatomical interaction between the lateral mammillary nuclei (LMN) and the dorsal tegmental nucleus (DTN). They propose that lateral inhibition among LMN HD cells is routed through DTN in which angular-velocity modulated HD cells allow the system to integrate head rotations over time. The properties of these DTN neurons resemble the rotation...
cells proposed by Skaggs et al. [31] except that they are inhibitory rather than excitatory. The authors postulate that this conceptual model (no simulation results are presented) may account for the experimental findings showing that LMN HD cells tend to have narrower tuning curves during contraversive than ipsiversive head rotations [3].

As a generalization of the one-dimensional attractor paradigm employed for HD cells, several authors have proposed a continuous two-dimensional attractor model to study the place responses of hippocampal pyramidal neurons (e.g., [27,29,30,39]). Due to the presence of collaterals in the CA3 region, most of these models identify this area of the hippocampus as a plausible anatomical locus for the recurrent network implementing the two-dimensional attractor map.

4. The ongoing exchange between modeling and neurobiological perspectives

The neural network models described above are based upon the known neuroanatomical and neurophysiological data and subject to continuous modification according to new experimental findings. Conversely, recent neurophysiological experiments have been motivated by an interest in testing for properties predicted by attractor network models. Knierim [15a] studied hippocampal place responses in rats running on a familiar circular track. He predicted that after nearby and distant cues were shifted by different angles, ensembles of place neurons recorded simultaneously should all respond in a similar manner if they were linked in a single strong attractor network. Instead, subsets of the neurons followed either set of cues, and some individual neurons even showed splitting of firing fields to follow the two respective cue sets. He concluded that if these neurons do indeed participate in attractor networks, the latter are weak relative to signals concerning external cues.

In anterodorsal thalamic head direction cells, Zugaro et al. [46a] studied the dynamics of updating of the head direction attractor after a sudden new visual input. The rats were trained to remained stationary facing the preferred direction of a neuron being recorded. Then the lights were turned off, the cue card shifted by 90° along the cylinder wall and the lights turned on again. Since the animal was no longer in the preferred direction the experimenters could measure the time course of the cessation of cell activity. Again the lights were turned off and the card returned to the preferred position, and the timing of the activity onset was measured when the lights were turned on. Two alternate predictions were that the attractor could shift gradually through intermediate states with activation of neurons preferring intermediate positions of the card, or that it could jump directly from one state to the other. Rapid shifts were observed on the order of 80 ms, suggesting a direct jump. Currently neural network models are being studied to confirm this.

5. Summary and perspectives

While attractor networks constitute an elegant hypothesis for the dynamics of neural networks in the limbic system and elsewhere, it remains to be demonstrated that they actually exist. Further neurophysiological work must include experiments designed to test specific predictions provided by modelers. In return, the network architectures should be modified taking into account neuroanatomical and neurophysiological observations. For example the attractor network may be manifested in connections within a large network of interconnected nuclei rather than within only one or two structures. In the head direction cell network, there exist structures that do not have any inhibitory neurons, or have no recurrent collaterals. While the biological infrastructure does not seem to be the simplest or most efficient to perform a single engineering function, it must be recalled that these brain regions may have multiple functions, and must be self-organizing and modulable for the diversity of behavioral and cognitive functions necessary for survival.

References


