

# Path Integration and Cognitive Mapping in a Continuous Attractor Neural Network Model

Alexei Samsonovich and Bruce L. McNaughton

Arizona Research Laboratories Division of Neural Systems, Memory and Aging, The University of Arizona, Tucson, Arizona 85749

A minimal synaptic architecture is proposed for how the brain might perform path integration by computing the next internal representation of self-location from the current representation and from the perceived velocity of motion. In the model, a place-cell assembly called a “chart” contains a two-dimensional attractor set called an “attractor map” that can be used to represent coordinates in any arbitrary environment, once associative binding has occurred between chart locations and sensory inputs. In hippocampus, there are different spatial relations among place fields in different environments and behavioral contexts. Thus, the same units may participate in many charts, and it is shown that the number of uncorrelated charts that can be encoded in the same recurrent network is potentially quite large. According to this theory, the firing of a given place cell is primarily a cooperative effect of the activity of its neighbors on the currently active chart. Therefore, it is not particularly useful to think of place cells as encoding any particular external object or event. Because of its recurrent con-

nections, hippocampal field CA3 is proposed as a possible location for this “multichart” architecture; however, other implementations in anatomy would not invalidate the main concepts. The model is implemented numerically both as a network of integrate-and-fire units and as a “macroscopic” (with respect to the space of states) description of the system, based on a continuous approximation defined by a system of stochastic differential equations. It provides an explanation for a number of hitherto perplexing observations on hippocampal place fields, including doubling, vanishing, reshaping in distorted environments, acquiring directionality in a two-goal shuttling task, rapid formation in a novel environment, and slow rotation after disorientation. The model makes several new predictions about the expected properties of hippocampal place cells and other cells of the proposed network.

**Key words:** hippocampus; CA3; place cells; head direction; dead reckoning; path integration; idiothetic; allocentric; spatial learning; cognitive map; attractor; integrate-and-fire

It is known from individual and multiple parallel recordings of single-neuron activity in freely moving rodents that the dynamics of the rodent hippocampus during active locomotion in a planar maze is essentially two-dimensional in its space of states; furthermore, it is a two-dimensional model of the animal’s motion on the maze (O’Keefe and Dostrovsky, 1971; O’Keefe and Nadel, 1978; Wilson and McNaughton, 1993). This statement becomes clear when one considers a *chart*, i.e., an abstract plane, on which all place cells are symbolically represented by *units* (nodes). The fact is that there exists an arrangement of units on a chart such that a typical distribution of neuronal activity over a chart (Fig. 1) is a localized *activity packet* of an invariant shape, the center of which, given a certain fixed mapping from the chart onto the environment, points to the current location of the rat’s head.

As experimental data show, the activity packet has the following dynamical properties. (1) It persists and retains its shape during active locomotion regardless of motion parameters and regardless of the stability and immediate availability of sensory cues, e.g., in complete darkness; (2) under certain conditions, the

whole representation, rather than a fraction of it, can be spontaneously remapped, without distortion of the intrinsic structure of the chart, and this new mapping may subsequently persist; (3) on entering a novel environment, a new chart (i.e., a new spatial code) appears immediately and normally does not undergo subsequent topographical modifications after exploration or changes in environmental stimuli; and (4) under different behavioral conditions, different charts for the same environment are expressed in the hippocampus, showing uncorrelated arrangements of common place cells.

In the present paper, the term “spatial” is used in a restricted sense to refer to location in a plane and does not include yaw. Several proposals have been made regarding possible explanations of the spatially selective firing of hippocampal pyramidal cells (Zipser, 1985; Muller et al., 1991, 1996; Touretzky et al., 1993; Blum and Abbott, 1995; Tsodyks and Sejnowski, 1995; Touretzky and Redish, 1996). None of these explanations, however, is consistent with all of the foregoing facts. To date, there is no satisfactory theory that explains the full range of observed phenomena; however, the scheme proposed by McNaughton et al. (1996) seems to be capable of accounting for most of the existing experimental data. Therefore, the objective of the present work is to test, through numerical simulations, the plausibility of the multichart map-based path integrator (MPI) model proposed by McNaughton et al. (1996).

## MATERIALS AND METHODS

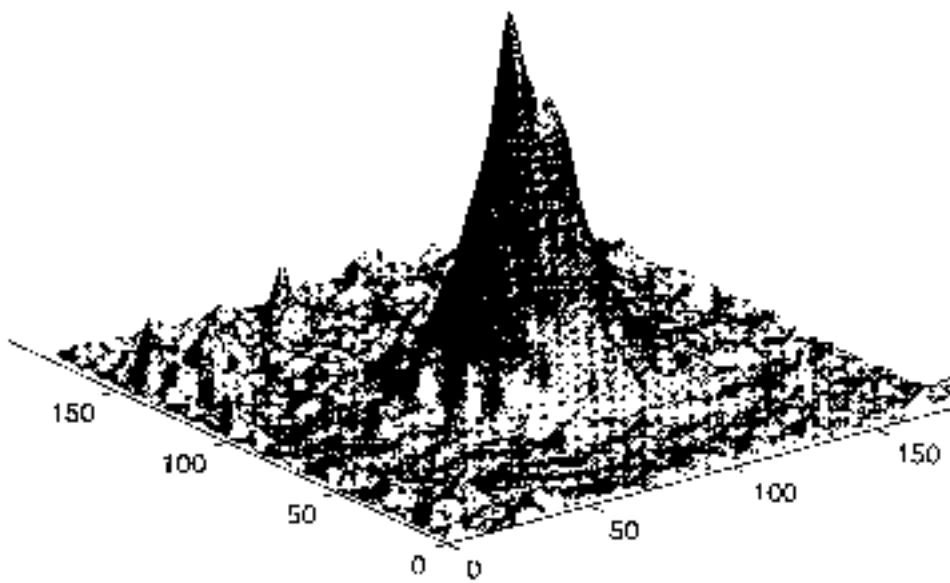
To define the MPI model of the hippocampus, the necessary concepts must be introduced, and then the components of the MPI scheme

Received Feb. 24, 1997; revised May 14, 1997; accepted May 15, 1997.

This work was conducted in partial fulfillment of the requirements for the degree of Doctor of Philosophy (A.S.) and supported by National Institute of Neurological Disorders and Stroke Grant NS20331 and The Office of Naval Research. We are grateful to Drs. C. A. Barnes, K. M. Gothard, J. J. Knierim, W. B. Levy, L. Nadel, J. O’Keefe, W. E. Skaggs, D. S. Touretzky, M. V. Tsodyks, and A. D. Redish for comments on this manuscript and useful discussions.

Correspondence should be addressed to Dr. Bruce L. McNaughton, Room 344, Life Sciences North Building, University of Arizona, Tucson, AZ 85724.

Copyright © 1997 Society for Neuroscience 0270-6474/97/175900-20\$05.00/0



**Figure 1.** Activity packet on a chart constructed from the experimental data of Wilson and McNaughton (1993). The rat was randomly foraging for food in a  $62 \times 62$  cm box. Parallel recordings of  $\sim 100$  hippocampal cells, 36 of which showed robust activity, were taken during the running session. The whole population of recorded place cells is symbolically distributed in the box, each cell being placed at the center of its place field. In the present paper, this planar arrangement of place cells is called a *chart*. A fuzzy snapshot of a momentary firing rate distribution over the chart is taken every 50 msec. The rat's position and orientation are marked on each snapshot, and all of the snapshots are superimposed so that all rat position marks are aligned at the center. The resultant average distribution is shown on the figure; therefore, the plot can be viewed as a typical momentary distribution of the firing rate over a chart (in allocentric coordinates). Units on horizontal axes are centimeters. The animal is located at the center of the square and is moving to the left

and toward the viewer. The total number of P cells (presumably CA1–CA3 pyramidal cells) in a rat's hippocampus is of the order of  $3 \times 10^5$  (Amaral et al., 1990). From empirical studies, for a typical recording environment of  $\sim 1 \text{ m}^2$ , a given P cell has a probability of  $\sim 0.3$  of having a place field. Thus, the density of units on a typical chart can be estimated as  $\sim 10^5 \text{ m}^{-2}$ . The variance of the distribution shown on the figure is  $\sim 0.15 \text{ m}$ , which is consistent with the observation that  $\sim 10^{-2}$  of all P cells fire at a given location. The averaged activity packet seems to have two "subcomponents." In fact, the real activity packet oscillates between these "subcomponents" with the theta frequency (Fig. 9C) (also see Skaggs et al., 1995) and therefore has smaller variance. We performed after-processing of the experimental data as described above with various data selection: right turns only versus left turns only, high velocity versus low velocity, high acceleration versus low acceleration, etc. The results suggest (within the error of measurement) that the shape of the activity packet does not depend on velocity, acceleration, future trajectory of motion, or theta frequency. This result will be presented in more detail elsewhere.

proposed by McNaughton et al. (1996) are identified, the MPI model is defined, and the numerical procedure is described. Finally, on the basis of the results of simulations presented in the Results, a reduced model is defined and its numerical implementation and the simulation procedure are described.

### Basic concepts

**Chart concept.** A chart is defined here as an imaginary arrangement of a population of place cells on an abstract plane, such that when this plane is appropriately mapped onto an environment, each cell appears to be located at the image of the absolute maximum of its firing rate distribution. Therefore, the total activity distribution on the chart appears to be localized around the image of the animal's head. Moreover, whenever there exists a planar arrangement of cells such that the activity distribution appears to be focused at a particular location within it, it is called an *active chart*.

In the models proposed by Muller et al. (1991, 1996) and McNaughton et al. (1996), a chart is associated with a place-cell assembly. As shown below, the distribution of activity on an appropriately "wired" chart is localized regardless of any existing association with an environment. Therefore, the notion of a chart in this case makes sense for an isolated network as well, i.e., in the absence of external input.

In different environments, and even in the same environment under different behavioral paradigms or other conditions, alternative charts may be active in which the spatial relations among place fields of the same place cells may be different. Typically, there are no significant correlations between the alternative charts for the whole population of recorded place cells (O'Keefe and Conway, 1978; O'Keefe and Nadel, 1978; Kubie and Ranck, 1983; Muller and Kubie, 1987; O'Keefe and Speakman, 1987; Bostock et al., 1991; Markus et al., 1994a).

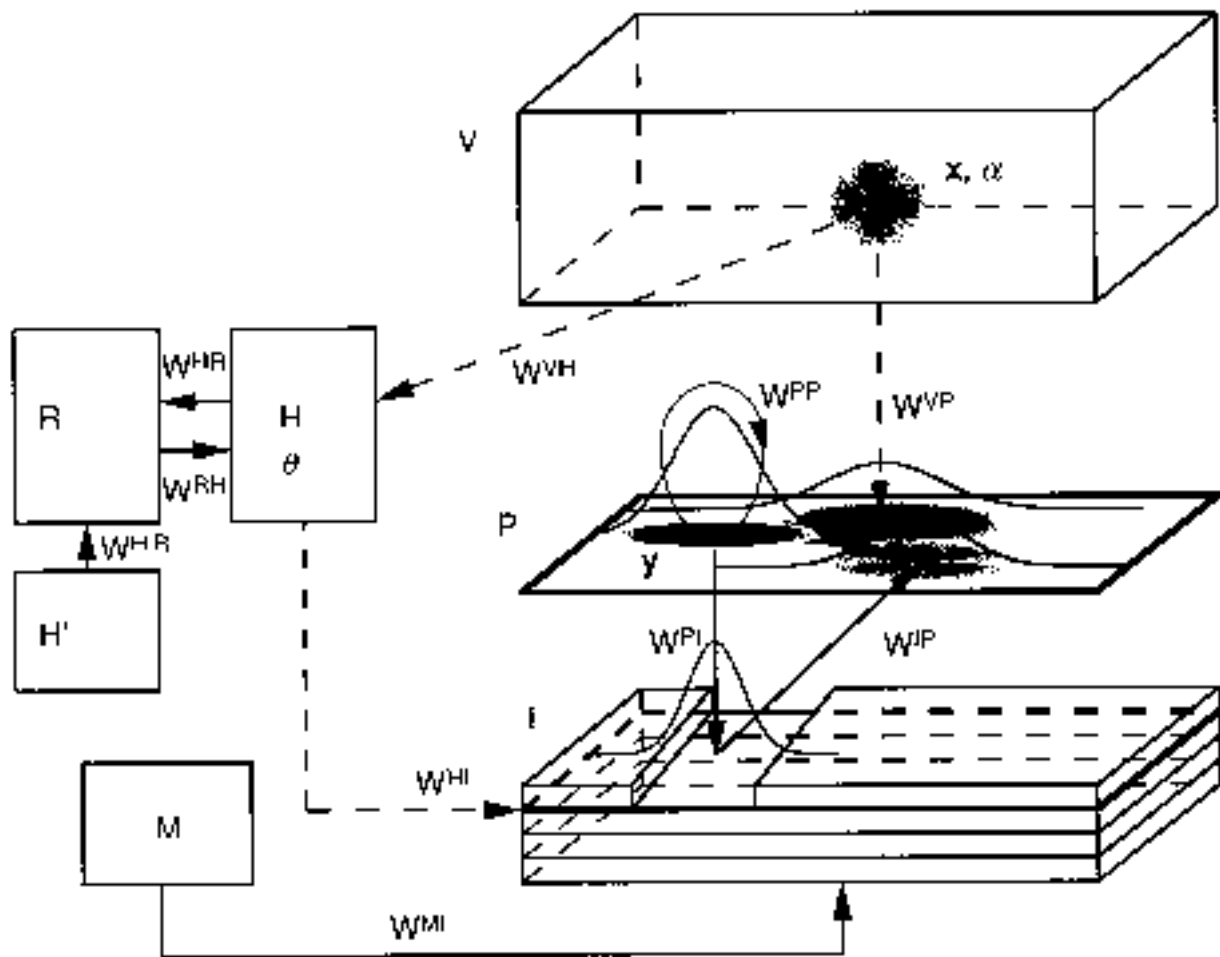
**Attractor map concept.** According to the *cognitive map* concept (O'Keefe and Nadel, 1978), not only do different firing patterns represent different places, but furthermore the spatial relationship between places is encoded by the interconnections between place cells, so that the place cells may fire consistently with each other regardless of the immediate availability of sensory cues, whereas the orientation of the entire map can be changed and subsequently remembered after reorientation and then removal of controlled cues (O'Keefe and Speakman, 1987). Here *consistent firing* means the persistence of the active chart, i.e., the same correlations between individual cell activities, as under previous normal

conditions. In other words, the system refers not to a set of independently stored local views but to a *cognitive map* of the environment, wherein the representation of the location is maintained regardless of external input and updated on the basis of exteroceptive and idiothetic information (here *idiothetic information* means all direct self-motion information, including vestibular signals, motor efference copy, optic flow, and somatosensory feedback). This internal map can be constructed from a set of spatial memory fragments (Worden, 1992) or based on an abstract preconfigured model of space (O'Keefe and Nadel, 1978). Usually only the latter is called a cognitive map.

An attractor map concept (cf. Ranck, 1992; Tsodyks and Sejnowski, 1995; Samsonovich and McNaughton, 1996), which is used in the present paper, is one possible way to introduce a cognitive map mathematically. An *attractor* (Strogatz, 1994) is a minimal closed set  $A$  in the space of states of a dynamic system such that (1) any trajectory that starts in  $A$  stays in  $A$ , and (2)  $A$  attracts all trajectories that start in an open set containing  $A$  (in the present case, this definition makes precise sense for an isolated network without noise). It follows from the definition that there is a finite threshold for an external perturbation to be capable of taking the system out of an attractor.

An *attractor map* can be defined as a two-dimensional, quasicontinuous set of attractors (associated with a particular environment or not), with the following dynamical property: the mobility threshold for transitions between neighboring attractors is negligibly small (tends to zero, with the number of units tending to infinity) as compared with the finite threshold for jumps between distant points or outside of the attractor map. Basically this is a generalization of the one-dimensional *continuous attractor* concept (Amari, 1977; Amit and Tsodyks, 1991a,b; Griniasty et al., 1993; Amit et al., 1994; Cugliandolo and Tsodyks, 1994). It follows from the definition that given a network with an attractor map subjected to subthreshold external perturbations, one may observe an active chart, as defined above. On the other hand, observation of an active chart in a particular network does not necessarily imply the existence of an attractor map in this network: the chart property may result from two-dimensionally organized input.

**Map-based path integration concept.** It is clear from behavioral studies that mammals and many other species possess highly developed path integration capabilities (Mittelstädt and Mittelstädt, 1980; Etienne, 1987; Thinus-Blanc et al., 1987; Müller and Wehner, 1988; Maurer and Séguin-



**Figure 2.** Hippocampal path integration system (according to McNaughton et al., 1996). The main components of the system are sensory array ( $V$ ), array of place cells ( $P$ ), array of integrator cells ( $I$ ), motion cells ( $M$ ), and the head direction system;  $W$  stands for synaptic efficacy. The head direction system ( $H, R, H'$ ) works according to the scheme of Figure 3. In particular, head direction cells are weakly affected by sensory representations in  $V$  that correct the activity packet position. The P–I system involves the two-dimensional array  $P$  of place cells and the three-dimensional array of  $I$  cells. Each “layer” of the latter has asymmetric connections with the  $P$  array, with displacement in the particular direction represented by this layer. The layer associated with the current head direction is selected by the  $H$  array via  $W^{HI}$  connections; the rest of the layers remain silent. Because of local internal connectivity of the  $P$  array, its activity is self-focused into an activity packet. This activity packet excites a small region on the selected  $I$  layer. When projected backward onto the  $P$  layer, the stimulation appears to be displaced in the current direction of the head (which is the direction of the animal’s motion in this model). As a result, the activity packet moves along the chart in this direction. Because total activity of the  $I$  layer is modulated by the  $M$  system, the magnitude of stimulation and therefore the speed of the activity packet motion depend on the speed of the animal’s motion represented by the activity level in  $M$ . In addition to this path integration mechanism, the activity packet location is corrected by sensory information represented by the  $V$  array. The latter can be viewed as a three-dimensional array of cells tuned to different local views (there is a three-dimensional array of local views in a stationary environment). Therefore, the displacement of the activity packet (*horizontal arrow* in the  $P$  layer) is determined by the sum of three “gaussians” that result from activities in  $P, I$ , and  $V$ .

not, 1995; Etienne et al., 1996), and the history of study of path integration goes back more than 100 years (Darwin, 1873).

The planar path integration concept involves (1) selecting a physical reference frame, implying the reference location, the reference direction, the metrics, and the clock, and (2) performing integration of the velocity vector over time in this reference frame to update the currently represented, or “perceived,” coordinates. This implies two necessary building blocks: an internal representation of the planar coordinates, maintained independently of immediate exteroceptive stimuli, and a mechanism of its updating based on idiothetic information.

According to McNaughton et al. (1991), the hippocampus (although it could be another brain system connected to it) works as an inertial *path integrator*. In an MPI, the internal representation of coordinates is based on a cognitive (attractor) map, whereas in a “naïve” *path integrator*, coordinates are not based on a map associated with a particular environment, and this implies a universal (i.e., environment-independent) representation of a two-dimensional vector. Because an internally updated

path integration mechanism would be prone to cumulative drift error, visual or other sensory information must be used to correct the representation using previously learned associations between map coordinates and external stimuli (McNaughton et al., 1991). Similar schemes of a planar path integrator were proposed by Droulez and Berthoz (1991) and Zhang (1996) but were not explicitly simulated.

The map-based path integration concept is represented by a possible MPI scheme of the hippocampal spatial representation system (Fig. 2) (McNaughton et al., 1996). It contains the following components:  $P$  (an array of place units implementing an attractor map),  $V$  (an external sensory input array),  $H$  (an array of head-direction units),  $H'$  (an array representing the angular velocity of the head),  $M$  (an array representing the speed of motion),  $R$  [an array sensitive to both horizontal head direction (yaw) and the angular velocity], and  $I$  (an array that receives inputs from  $P, H$ , and  $M$ ). The core of the scheme is the P–I path integrator fragment. The foregoing components and the model are described below. For a simpler treatment of the path integration princi-

ples in the present scheme, the reader is referred to McNaughton et al. (1991) and to an earlier proposal based on learning conditional relationships between locations and movements (McNaughton et al., 1989).

### Identification of components

*V.* The component referred to as the *V network* (Fig. 2) can be identified with sensory association cortex, which provides high-level representations of the local sensory information and sends its output to the hippocampus mainly via the entorhinal cortex and the perforant path. Under normal conditions, a neocortical representation of multimodal sensory stimuli can be thought of as a function of the animal's current location  $x$  and its head direction in the horizontal plane (yaw) given by the angle  $\alpha$ . This function is presumably smooth. For the sake of parsimony, a mnemonic mechanism in the sensory system (e.g., imagery) involved in maintaining the activity of local sensory representations is not assumed. Therefore, in this oversimplified picture the space of states of the *V network* is three-dimensional (two spatial coordinates plus the yaw angle). The term *local view* is used here as a shorthand for the entire sensory representation that is typically specific for a particular combination of location and orientation in the environment. Each local view is associated with a corresponding pattern in the *P network*, because of associative learning in the afferent synapses rather than in the internal P-to-P connections. Different parts of the environment will typically have separate representations in *V*, thus allowing differential binding of the attractor map coordinates to local cues.

*M.* The path integration concept implies that the integrator receives information about self-motion (*M*). Participation of the motor system in the dynamics of hippocampal spatial representations is suggested by the finding that under conditions of movement restraint, both hippocampal place cells (McNaughton et al., 1983; Foster et al., 1989) and thalamic head-direction cells (Knierim et al., 1995, 1996) become virtually silent, even when the animal is passively moved. The pattern of place fields can be reproduced during passive movement, however, provided that the animal retains the possibility of self-motion (Muller et al., 1987; Foster et al., 1989). Thus, for place cells to fire, it is sufficient that the animal is free to move its limbs, even if it does not actually move. Moreover, the firing rates of virtually all hippocampal neurons are modulated by locomotion (Ranck, 1973; Whishaw and Vanderwolf, 1973; McNaughton et al., 1983; Mizumori et al., 1990), which implies that the hippocampal system has access to information about self-motion.

*I.* To perform spatial path integration, it is necessary to know the speed and direction of motion. In general, hippocampal place cells are known to be relatively nondirectional in a two-dimensional environment, if the behavior does not involve the following of specific routes between discrete reinforcement sites (O'Keefe, 1979; Muller et al., 1987; Markus et al., 1994b; Muller et al., 1994); however (and this is very important for the MPI model), some cells in the subiculum, the presubiculum, and the parasubiculum have been found with spatial and directional selectivity at the same time (Sharp and Green, 1994; Taube, 1995b), under conditions in which hippocampal place cells are *nondirectional*. According to McNaughton et al. (1996), they can be considered as candidates for the integrator cells, or *I cells*. A population of such cells would provide a distributed representation of all possible combinations of head orientation and location in an environment. Such a representation could be used to update the activity packet coordinates and must be constructed on the basis of directional information. This population of cells is referred to as the *I network*, which thus could possibly be identified with the subicular complex, although this would require a more complex connective scheme than the simple networks under consideration here.

*H, H', and R.* Directional information is represented in the brain by a population of head-direction cells (*H cells*), which therefore is likely to be the essential *H* component of the spatial path integrating system. An *H cell* fires at a high rate when the rat's head is oriented in a specific absolute direction in the environment, regardless of either the spatial location or the position of the head with respect to the body. *H cells* were first discovered in the dorsal presubiculum (Ranck, 1984; Taube et al., 1990). Later, *H cells* have been found in the anterior nuclei of the thalamus (Mizumori and Williams, 1993; Taube, 1995a; Blair and Sharp, 1995), the retrosplenial (posterior cingulate) neocortex (Chen et al., 1994a,b), the striatum (Wiener, 1993), and the lateral mammillary nuclei (Leonhard et al., 1996). All of these areas are closely connected with the hippocampus proper.

Actually, the path integration mechanisms proposed in the present paper require information about direction of motion rather than head orientation. Head direction will suffice so long as it remains highly

correlated with movement direction. The possibility remains, however, that true "direction-of-motion cells" exist but have not been documented. Such cells could arise from a simple coordinate transformation such as the one that has been suggested by Andersen and his colleagues (Andersen et al., 1985; Zipser and Andersen, 1988) to occur in the primate parietal cortex and could easily have been mistaken for *H cells* in many previous experimental studies.

Knierim et al. (1995) demonstrated experimentally that the origin of the head direction representation is likely to be based on a path integration mechanism rather than on immediate conversion of visual stimuli into *H cell* firing (cf. Blair and Sharp, 1995). A model of such a head-direction path integration mechanism (Fig. 3) has been developed by McNaughton and colleagues (McNaughton et al., 1991; Skaggs et al., 1995). The model includes a circular array of locally interconnected *H cells*, in which an activity bump is stabilized by intrinsic dynamic mechanisms as an attractor state, and a two-dimensional array of angular rotation cells (*R cells*) that force the activity packet to move in a manner consistent with the head angular velocity. These *R cells* are connected to angular velocity cells *H'*, which presumably represent primarily vestibular information. They encode the interaction between *H* and *H'*. Such cells were observed in parietal cortex by McNaughton et al., (1991) and Chen et al. (1994b). More recently, fundamentally similar models, although different in detail, have been proposed (Blair, 1996; Zhang, 1996; Redish et al., 1997).

Slow rotation of the place-field pattern, which was observed by Knierim et al. (1995), can be understood as a result of a weak influence of the symmetry breaking cue card on the head direction system rather than on the *P network*, thus indicating a direct or indirect connection between the *H* and *P* systems.

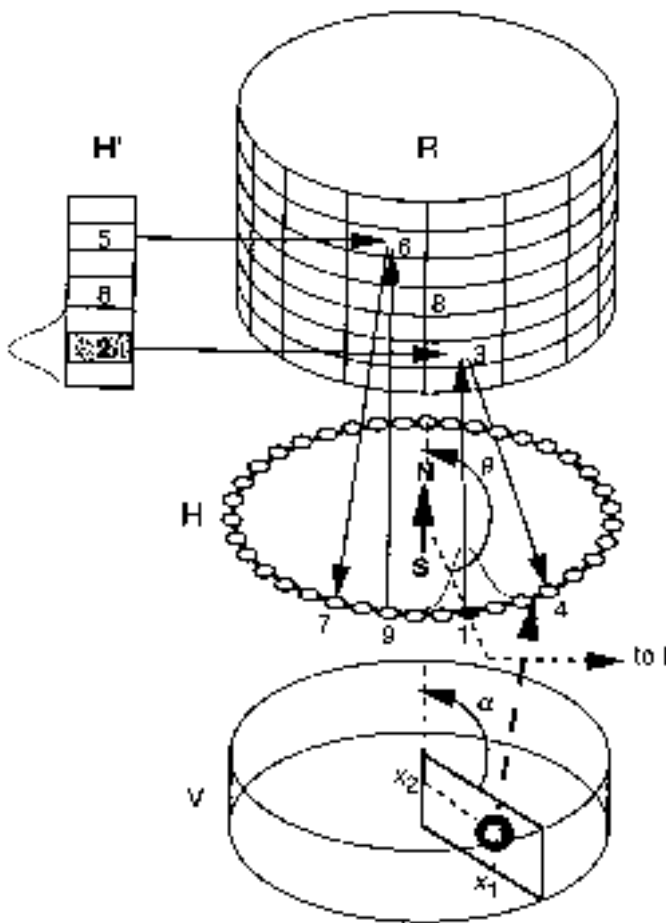
*P.* According to the above considerations, the main candidate for the attractor map as a component of the path integrating system is the *P network*, presumably based on the areas CA3, CA1, and dentate gyrus, with a primary role for CA3 in the origination of the attractor dynamics. Indeed, neuroanatomical data show that CA3 has multiple, long-range excitatory internal connections (Amaral and Witter, 1995) and therefore has a necessary requirement for the *P network*; however, the same architecture could be implemented in other parts of the hippocampal formation, such as the entorhinal cortex, and in general, the present model is not intended to provide a strong argument for any particular anatomical implementation (see Discussion).

It seems natural to extend the principles underlying the head direction path integration model to two dimensions, taking the locally interconnected two-dimensional array of *P cells* as the central element; however, experimental data indicate that this extension cannot be made in a straightforward manner (e.g., as was suggested recently by Zhang, 1996).

The first problem with a parallel between the spatial and the head direction system is that preferred directions of *H cells* typically have unique relations to each other, even though the absolute direction selected by an *H cell* may vary between environments, parts of an environment, and recording sessions (Ranck, 1984; Taube et al., 1990; Knierim et al., 1995; Taube and Burton, 1995). In contrast, in different environments, and even in the same environment under different behavioral paradigms or other conditions, *P cells* may be involved in alternative representations in which spatial relations between their place fields may be different. Typically there are no clear correlations between the alternative place-field patterns for the whole population of recorded place cells (O'Keefe and Conway, 1978; O'Keefe and Nadel, 1978; Kubie and Ranck, 1983; Muller and Kubie, 1987; O'Keefe and Speakman, 1987; Bostock et al., 1991; Markus et al., 1994a). Moreover, in a two-goal shuttling task, switching between representations appears to occur on reaching the goal, making place fields appear directionally tuned (Barnes et al., 1983; O'Keefe and Recce, 1993; Muller et al., 1994; Markus et al., 1995).

A possible solution to this problem is that the *P network* may implement many alternative two-dimensional attractor maps at the same time, and a typical *P cell* participates in a number of these implementations, one of which is selected by the current activity state (McNaughton et al., 1996). Spatial firing properties of a given *P cell* are thus strongly dependent on the current distribution of activity over the whole *P network*, and strictly speaking it is incorrect to use the term "tuning curve" for an individual *P cell*. The problem of multiple place fields (Muller et al., 1987) can be solved by placing *P cells* several times on the same chart, which is similar to placing the same *P cell* on many charts; however, this more complex scheme is not considered here.

Would individual attractor maps not be destroyed by interference? What kind of architecture does this point of view imply? These questions



**Figure 3.** Head direction path integration system (according to McNaughton et al., 1991; Skaggs et al., 1995). The main components of the system are head direction cells ( $H$ ), “tuned” to allocentric head directions; angular velocity cells ( $H'$ ); rotation cells ( $R$ ); and the external sensory representation system ( $V$ ). Because of local connectivity of the circular array,  $H$  unit activity here is self-localized into an activity packet centered at unit 1. Given angular velocity represented by unit 2, unit 3 of the  $R$  array becomes activated by units 1 and 2. This results in stimulation of unit 4 and therefore in displacement of the activity packet in the counterclockwise direction. In another case, when the activity packet (not shown) is centered at unit 9 in the  $H$  system, and the angular velocity is represented by unit 5, activation of units 6 and 7 results in clockwise rotation. Excitation of unit 8 stabilizes the activity packet at its current location. Thus, the activity packet in the  $H$  array points to the current direction of the head. This  $H$  array is coupled to the two-dimensional  $R$  array arranged on the cylinder according to their connections with  $H$  cells, and in another dimension according to connections to the  $H'$  cells. The architecture of the interconnections can be explained as follows: an  $H'$  cell sends equal outputs to the slab of  $R$  associated with it via  $W^{H'R}$ , whereas an  $H$  cell sends equal outputs to its column in  $R$  via  $W^{HR}$ . The nonzero  $W^{RH}$  connections, also all of equal strength, are established with different angular displacement with respect to their counterparts  $W^{HR}$ , depending on the slab of their origin at the  $R$  array; namely, this displacement is proportional to the angular velocity represented by the associated  $H'$  cell. In addition,  $H$  cells are locally interconnected to each other, with the connectivity matrix given by Equation 3. This results in formation of a stable activity packet in the  $H$  network, as described in text. Driven by  $R$  cells, the activity packet in the  $H$  array moves together with head rotation, thus performing angular path integration. The inevitable cumulative error is corrected by representations of visual cues in  $V$ . This presumably requires associative learning between visual representations in  $V$  and activity packet locations in  $H$ , based on modification of the  $V$ -to- $P$  connections. The space of  $V$  representations is three-dimensional: it has two spatial coordinates  $x_1$  and  $x_2$  and the head direction angle  $\alpha$ . Only the latter is distinguished by  $V$ -to- $H$  connections. According to the MPI model (Fig. 2),  $H$  cells send their output to the integrator ( $I$ ) cells.

are examined below. Another problem is that of formation of the attractor map. In particular, a new chart can become active in the dark (Quirk et al., 1990), when path integration is presumably the only source of spatial information for the animal. Briefly, for the present purpose, the prewiring of the multichart attractor map is assumed without consideration for how this occurs.

### Elements and dynamic rules of the MPI model

Elements of the above networks were implemented as model integrate-and-fire units interconnected by synapses. In such a network, dynamic variables are “spikes” ( $S$ ), “EPSPs,” or “voltages” ( $V$ ), and some of the synaptic weights ( $W$ ) that are slow variables. The latter were mostly assumed fixed in the simulations. The network to which a given variable belongs is marked by a superscript.

The discrete time approximation is used here, with a time bin  $\Delta = 6$  msec, which is bigger than the refractory period, close to axonal plus synaptic delays, and yet smaller than the neuronal integration time (for review, see Shepherd, 1990). Therefore,  $S$  can be treated as an array of Boolean variables. The integration time for inhibitory interneurons, however, is smaller than for principal cells (Fox and Ranck, 1981; McNaughton and Morris, 1987; Mizumori et al., 1989; Shepherd, 1990) and is close to the time bin. For this reason, the approximation of fast inhibition is used, assuming that the amount of inhibition (uniformly distributed among all units) is adjusted at every discrete time bin, so that the total number  $M$  of firing units is preserved near the given level, which varies periodically in time according to the theta rhythm.

There are two ways to achieve this: (1) by adjusting the global inhibition  $h$  at each step to match the total number of firing units, and (2) by taking  $h$  as a function of the total number of firing units computed at the previous iteration. The latter is consistent with the idea of hidden inhibitory interneurons and may permit the model to exhibit theta oscillations naturally; however, the time bin is too large for the model to be realistic, and it is not the present goal to study the origin of the theta rhythm. For this reason the first approach is adopted.

Therefore, the dynamic equations describing an isolated network of leaky integrate-and-fire units with reciprocal interconnections (e.g.,  $P$  or  $H$  component) can be written as:

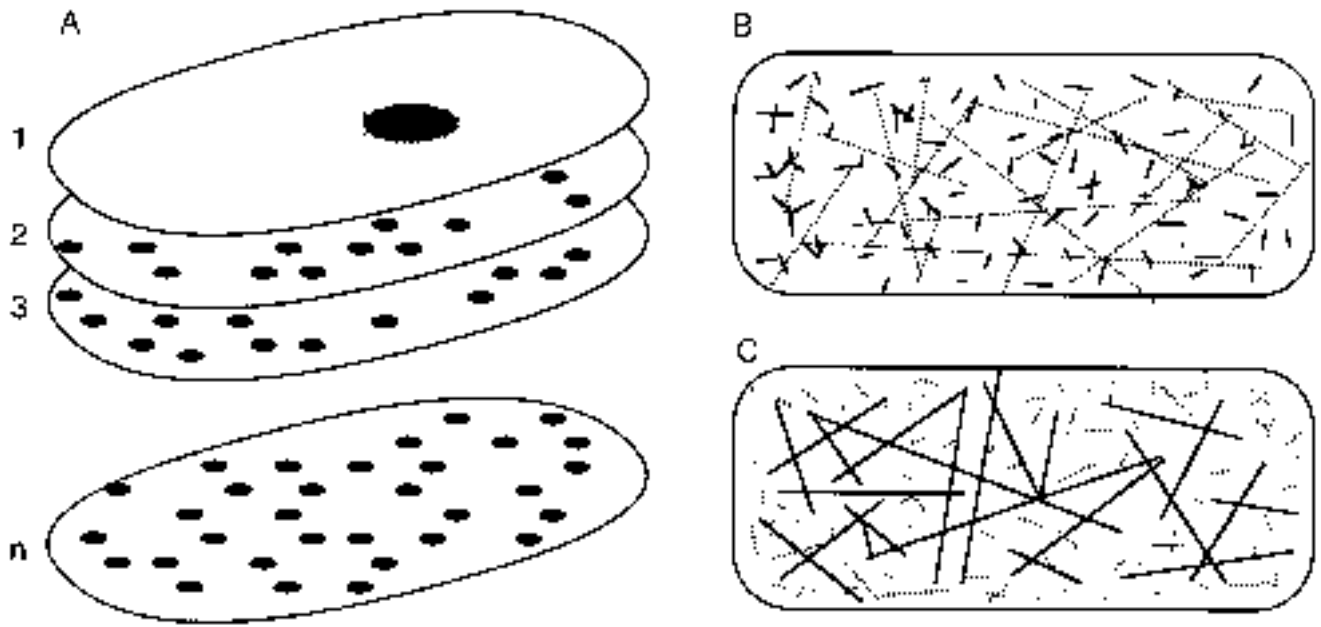
$$V_i^{t+1} = (1 - S_i^t) e^{-(\Delta/\tau)} V_i^t + \sum_{j=1}^N W_{ij} S_j^t, \quad (1)$$

$$S_i^t = \theta(V_i^t - h^t), \quad h^t: \sum_{i=1}^N S_i^t = M. \quad (2)$$

Here  $i$  is the unit number,  $N$  is the total number of units in this subnetwork,  $t$  is the discrete time,  $\Delta$  is the time bin,  $\tau$  is the neuronal integration time,  $W$  is the synaptic matrix, and  $\theta$  is the Heaviside step function. The global threshold  $h^t$  is an implicit function of the set of variables  $\{V\}$  defined by the right Equation 2, where  $M$  is the given total number of active units.

**Multichart architecture of the  $P$  network.** Now the multichart architecture of the  $P$  component is introduced. Consider all  $P$  cells distributed on an abstract plane according to the relative locations of their place-field centers. This arrangement is called a chart; however, there may be multiple such arrangements of the same  $P$  cells that are uncorrelated and can be used to represent different environments. Therefore, consider  $n$  possible arrangements of the same  $P$  cells. In this model they are random permutations of each other. In general, a  $P$  cell may not be found on some charts. In the present model, however, it is assumed that each chart is composed of all model units. To obtain the matrix of internal connections  $W^{PP}$  of the  $P$  network, local interconnections were created on each chart (local in the sense that weights decay rapidly with distance between units on the chart, rather than with the anatomical distance), and then the sum over all charts was taken. The result is a *multichart architecture*. Although charts, being defined in terms of firing rate distributions, make sense without relation to any connectionist model, the structure of the synaptic matrix is an important feature of the MPI model. It is given by the formula:

$$W_{ij}^{PP} = \sum_{k=1}^n \exp\left(-\frac{r_{ijk}^2}{2\sigma^2}\right), \quad (3)$$



**Figure 4.** Multichart architecture. *A*, The set of  $n$  charts, composed of the same P units. Activity that is well localized on one of the charts (1) looks scattered on other charts (2, 3,  $n$ ). *B*, *C*, Two arbitrarily selected charts with some interconnections between units on them. Local interconnections on chart *B* (solid lines) appear to be random (nonlocal) on chart *C*; local interconnections on chart *C* (dashed lines) are random (nonlocal) on chart *B*. Therefore, on a given chart, contributions from other charts to the synaptic matrix can be treated, approximately, as random noise.

where  $r_{ijk}$  is the distance between units  $i$  and  $j$  on the chart  $k$ , or infinity, if at least one of the two units is missing on this chart (also see Muller et al., 1991; Shen and McNaughton, 1996),  $n$  is the number of charts, and  $\sigma$  is a fixed parameter. The resultant multichart architecture is illustrated by Figure 4.

Although this model has all-to-all, symmetric, excitatory connections according to Equation 3, the matrix  $W$  can be called sparse because most of its elements are very close to zero, if  $\sigma \ll L$ , where  $L$  is the chart dimension. Quantities  $r_{ijk}$ ,  $\sigma$ , and  $L$  have the dimensionality of distance. The metrics on the chart, which is assumed for now to be given, is provided by the  $M$  system output, as given by Equations 4 and 8, which specify the relationship between the “perceived” velocity of self-motion and the velocity of the activity packet on the chart.

**Summary of assumptions of the MPI model.** In summary, the proposed MPI model (Fig. 2) is based on the following basic assumptions about the hippocampal formation. (1) The architecture of the P network is preconfigured as a sum of uncorrelated, quasi-two-dimensional architectures. This is most likely to be the architecture of internal CA3 connections; however, other implementations in anatomy would not invalidate the concept. (2) The primary driving mechanism for the activity packet on a chart is based on internal dynamics and is attributable to asymmetry in the connections from I cells to P cells. (3) Activation of I cells is controlled jointly by representations of the speed of motion ( $M$ ) and head direction ( $H$ ) and by return projections from the P network. (4) Learning results in selective strengthening of V to P connections. This enables stimulation of the P layer by the V array to determine on which chart and at which location the activity packet emerges on entry into a familiar environment. (5) Connections between H and I cells are also preconfigured and fixed. In other words, each chart has a built-in compass. Because there is only one chart in the H network, the layered structure of the I network must be the same for all charts.

It follows from the last assumption that although spatial relations between place fields of I cells may be different for different environments, relations between their preferred directions must be the same in all environments and for all representations. This is an untested prediction of the theory.

#### Numerical implementation of the MPI model

The MPI model described above has been implemented numerically on SUN Sparc-20 and Ultra-Spark work stations as a system of networks of integrate-and-fire units. In the simulations of Figure 9 (see Results) each

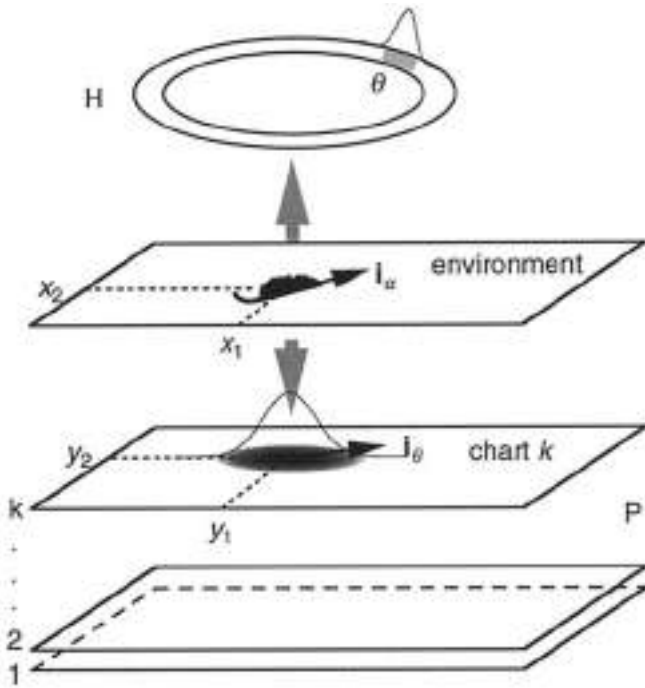
layer of the P–I system was composed of  $n = 256 \times 192 \approx 45,000$  model neuronal units distributed in a square lattice on a torus (i.e., a rectangle with periodic boundary conditions).

In the simulations of Figure 9*A,B,E,F,G*, the external inputs to I from H and M were assumed fixed, meaning that the model rat velocity was constant. In a more general case (see Fig. 9*D*) the I array is modulated by the internal representations of velocity in the M and H systems. The P–I system evolves according to the following system of equations constructed on the basis of Equations 1–3:

$$\begin{cases} V_i^{P,t+1} = (1 - S_i^{P,t})e^{-\frac{\Delta}{\tau}}V_i^{P,t} + \sum_{j=1}^N W_{ij}^{PP}S_j^{P,t} + \sum_{j=1}^N W_{ij}^{IP}S_j^{I,t} \\ \quad + \mu \sum_{k=1}^n \exp\left[-\frac{(\mathbf{x}^t - \mathbf{r}_i^k)^2}{2\epsilon^2}\right]; \\ V_i^{I,t+1} = (1 - S_i^{I,t})e^{-\frac{\Delta}{\tau}}V_i^{I,t} + \sum_{j=1}^N W_{ij}^{PI}S_j^{P,t} + \mathbf{v} \cdot \mathbf{b}_i; \\ S_i^{P,t} = \theta(V_i^{P,t} - h_p^t), \quad h_p^t: \sum S_i^{P,t} = M_t^P; \\ S_i^{I,t} = \theta(V_i^{I,t} - h_i^t), \quad h_i^t: \sum S_i^{I,t} = M_t^I. \end{cases} \quad (4)$$

Here  $t$  is the discrete time with a 6 msec time bin  $\Delta$ ,  $\tau = 10$  msec is the neuronal integration time, and  $\theta$  is the step function. The last term in the first equation of (4) describes the effect of visual input:  $\mathbf{x}^t = \mathbf{x}(t)$  is the given trajectory of the model rat running with a fixed speed of 116 pixels/sec,  $\epsilon = 200$  pixels,  $\mathbf{r}_i^k$  is the fixed coordinate of unit  $i$  on a chart  $k$ , and  $\mu$  is the efficacy of V-to-P connections, which was zero in the simulations of Figure 9*A,B,D,E*, 0.01 in *F*, and 0.1 in *G*. The last term in the second equation of (4) describes modulation of the I array by the H system ( $\mathbf{v}$  is a unit vector pointing in the direction of motion).

The set of constant, random vectors  $\{\mathbf{b}_i\}$  is the same as the set  $\{\mathbf{b}_j\}$  in Equation 8. The whole I component can be viewed as containing cells with a continuous range of gaussian-like directional tuning functions at each chart coordinate. An interval on this range is selected by the “perceived” direction of motion  $\mathbf{v}$ . The M, H, and V systems represented in Equation 4 by  $M_t$ ,  $\mathbf{v}$ , and  $\mathbf{x}$ , respectively, were assumed to be consistent



**Figure 5.** The continuous model. The dynamic state of the whole path integrator system is described by the following variables: the currently selected chart number  $k$ , the activity packet coordinates on this chart ( $y_1$ ,  $y_2$ ), and the “perceived” direction of motion  $i_\alpha$ , dynamically stored in the H array (the symbol  $i_\theta$  here stands for a unit vector pointed at the angle  $\theta$  with respect to north). The actual model rat coordinates in the environment are  $x_1$ ,  $x_2$ , and the actual head direction angle is  $\alpha$ . The direction of motion is assumed to be the same as the head direction; therefore, the model rat moves in the direction  $i_\alpha$ .

with the model rat motion and were not simulated explicitly. The total activities  $M_P$  and  $M_I$  are periodic functions of time:

$$\begin{cases} M'_P = A \cdot \theta \left[ \cos\left(\frac{2\pi\Delta}{T}t\right) \right] \cdot \cos\left(\frac{2\pi\Delta}{T}t\right) + B \\ M'_I = \theta \left[ (C + D) \cdot \cos\left(\frac{2\pi\Delta}{T}t\right) - D \right] \cdot \left[ (C + D) \cdot \cos\left(\frac{2\pi\Delta}{T}t\right) - D \right], \end{cases} \quad (5)$$

with  $A = 1000$ ,  $B = 500$ ,  $C = 250$ ,  $D = 200$ : the relative values  $A/B$  and  $C/D$  were adjusted to fit the experimental data on the population theta rhythm in CA1 and in DG, respectively (Skaggs, 1995). The theta cycle period  $T = 120$  msec. The magnitudes of the parameters  $C$  and  $D$  determine the total activity of the I layer and thus reflect the modulation of the I array by the M system.

All synaptic connections  $W$  in Equation 4 are excitatory; the role of interneurons implicitly present in the model consists of adjustment of the global thresholds  $h$  at each iteration, according to Equation 4. Thus, in the terminology of Amit (1989), the model contains a *hard constraint* for the total activity in each array. The synaptic matrices  $W$  in Equation 4 are defined as:

$$W_{ij}^{PP} = \frac{1 + \eta_{ij}^{PP}}{2\pi\sigma^2} \sum_{k=1}^n \exp\left[-\frac{(\mathbf{r}_{ij}^k)^2}{2\sigma^2}\right], \quad (6)$$

$$W_{ij}^{PI} = \frac{1 + \eta_{ij}^{PI}}{2\pi\sigma^2} \sum_{k=1}^n \exp\left[-\frac{(\mathbf{r}_{ij}^k)^2}{2\sigma^2}\right], \quad (7)$$

$$W_{ij}^{IP} = \frac{1 + \eta_{ij}^{IP}}{2\pi\sigma^2} \sum_{k=1}^n \exp\left[-\frac{(\mathbf{r}_{ij}^k + \mathbf{b}_j)^2}{2\sigma^2}\right]. \quad (8)$$

Here  $\eta$  is the static synaptic noise (centered at 0 and  $<1$  in absolute value) that results from the numerical algorithm (see below) and simulates the anatomical irregularities in the connectivity matrix;  $n$  is the number of charts; the vector  $\mathbf{r}_{ij}^k$  connects the nodes  $i$  and  $j$  on chart  $k$ ; the constant, random gaussian vector  $\mathbf{b}_j$ , which represents the asymmetry in the I–P connections, is centered at zero and has a variance of  $\sigma$ ; the constant gaussian width  $\sigma \approx 10$  pixels. Up to  $n = 6$  charts (see Results) have been implemented simultaneously in the same synaptic matrices  $W^{PP}$ ,  $W^{PI}$ , and  $W^{IP}$  in the simulations of Figure 9, and up to  $n = 200$  charts were implemented in the simulations of Figure 10.

In the simulations of Figure 10 the same model was implemented, but with the following differences in detail. Instead of placing exactly one unit at each node of the square lattice on each chart, as in Figure 9, each unit  $i$  was assigned random coordinates  $\mathbf{r}_i^k$  on each chart  $k$ ; all vectors  $\mathbf{r}_i^k$  were generated independently of each other, with a uniform probability distribution. In addition, an explicit division of the I array into  $L = 6$  head direction layers (Fig. 2) was used, so that the vector  $\mathbf{b}$  in Equations 4 and 8 was the same for all I units in each layer and was rotated by  $60^\circ$  from layer to layer. The fixed absolute value of this vector  $\mathbf{b} = 6$  pixels was the same for all I units. The number  $N$  of P units was the same as the number of I units in each head direction layer; therefore, the total number of I units was six times larger than the number of P units in this implementation. Only one I layer was active in the simulation of Figure 10C. The static synaptic noise  $\eta$  in Equations 6–8 and the parameter  $\mu$  in Equation 4 were set to zero. The values of some other parameters in this implementation were also different:  $\sigma = 3.1$  pixels,  $A = C = 0.01 \times N$ ,  $B = D = 0.002 \times N$ . The values of  $A$  and  $B$  were taken close to the real fraction of place cells that are active at a given location (McNaughton et al., 1996). In simulations of Figure 10A,C,D, each layer consisted of  $n = 300,000$  units (which is approximately the real number of place cells in rat CA3), with  $n = 20$  charts implemented in the synaptic weights. The lattice dimensions in this case were  $256 \times 192$ . In simulations of Figure 10C each layer consisted of  $n = 30,000$  units, with  $n = 100$  (Fig. 10B, *solid line*) or  $n = 200$  (Fig. 10B, *dashed line*); the lattice dimensions in this case were  $96 \times 96$ . The variance of the activity distribution in Figure 10A,B was computed for a torus embedded into a four-dimensional space (to preserve the Euclidean internal metrics of the torus). This provides a measure of the degree of focusing of the activity. In both simulations of Figures 9 and 10 (except for Fig. 9E–G), a network was started from a random configuration (i.e., randomly generated  $S$  and  $V$  variables).

To implement the synaptic matrix numerically, an additional, intermediate array and a two-dimensional filtering procedure were used. At each iteration, for each chart, the output of P and I units was added to the intermediate layer, according to the unit arrangement on this chart. After that, a noisy gaussian filtering procedure was applied, which consisted of redistribution of spike densities among lattice cells in the intermediate layer, and the result was taken as input to P and I units, together with external inputs from V, H, and M according to Equation 4. Two parallel filtering methods were combined: random gaussian filtering that created additional static synaptic noise and the standard gaussian filtering with the kernel (0.09, 0.24, 0.34, 0.24, 0.09). This algorithm substantially reduced the necessary computer memory and enabled the performance of  $O(N \cdot n)$  operations per iteration instead of  $O(N^2)$ , which would be the case for an explicit implementation of synaptic connections. On the other hand, the number of charts that could be simulated was limited by the computation time.

### Reduced continuous model

Because the simulation of complex experimental results using the integrate-and-fire scheme was severely limited by computational constraints, an alternative strategy was to characterize the dynamical state of each component by a set of “macroscopic” variables, instead of a detailed description in terms of variables  $S$  and  $V$ , and to formulate dynamic laws in terms of these “macroscopic variables.” The rationale for this approach is described below.

For the present purpose, the critical results of the integrate-and-fire model of one- and multichart networks (see Results) can be summarized as follows. (1) Under the given conditions, the distribution of activity in a model network possessing an attractor map (P or H network) is described by an activity packet, which retains its shape and moves smoothly, remaining on the same chart. (2) In the two-layer model (I–P fragment), it is possible to control the velocity and direction of motion of the activity packet with reasonable precision via modulation of activity of the I array by the outputs of the M and H arrays. The activity packet in the R–H system can be similarly controlled. (3) The activity packet

position on the chart can be smoothly corrected by an additional stimulation of the substrate array (P or H) by the output of the V array. (4) With the introduction of a strong local stimulation of the substrate array, distant from the current activity packet location, the activity packet may jump, with some probability, to the stimulated area on the same or different chart.

These results allow the introduction of a simple, continuous description of the above components. In this approach the state of the H network is described by one angular variable  $\theta$  characterizing the activity packet position in H, and the state of the multichart network P is characterized by the vector  $\mathbf{y}$ , representing the activity packet location in P, and the chart number  $k$  on which the activity packet is located. Thus, the dynamical variables of the continuous model (see Fig. 5) are  $\mathbf{y} = (y_1, y_2)$ , the animal's perceived location in space, i.e., the coordinates of the activity packet on the chart in the P network;  $\theta$ , the perceived head direction angle, represented by another activity packet on the unique circular chart of the H system; and the current chart number  $k$ . The outputs of M and V subsystems are assumed to be consistent with the model rat's "actual" trajectory given as  $\mathbf{x}(t)$  and with the actual head direction, tangent to  $\mathbf{x}(t)$ , given by the angle  $\alpha(t)$ . In this model the head direction always coincides with the direction of motion. The model is given by a system of stochastic differential equations:

$$\frac{d\mathbf{y}}{dt} = \mu \nabla U(\mathbf{y} - \mathbf{x}, \mathbf{x}) + \xi + \left| \frac{d\mathbf{x}}{dt} \right| \cdot \mathbf{i}_\theta, \quad (9)$$

$$\frac{d\theta}{dt} = \nu \frac{\partial}{\partial \theta} u(\theta - \alpha) + \eta + \frac{d\alpha}{dt}, \quad (10)$$

where  $\mu$  and  $\nu$  are constants that represent activity packet mobility;  $\mathbf{i}_\theta$  is the unit vector in the direction given by  $\theta$ :  $\mathbf{i}_\theta = (\cos \theta, \sin \theta)$ ; the gradient operator  $\nabla$  is acting on  $\mathbf{y}$ ;  $-U$  is the effective attractive potential for the activity packet, and  $-u$  is similarly the attractive potential in the head directional system (when  $U$  and  $u$  are thought of as attractive potentials, they should be taken with the negative sign).  $\xi$  and  $\eta$  are gaussian random variables centered at zero.

In the unchanged familiar environment  $-U$  is a single symmetric well, independent of its second argument, whereas in a stretched or shrunken environment  $-U$  may become a double-well potential, because of the two contributions from the two parts of the environment, displaced with respect to each other. The depth of each half depends on the rat's position  $\mathbf{x}$ , which results in systematic transitions of the activity packet from one well to another with some hysteresis.

It is assumed here that in the unchanged environment,  $U$  is a gaussian, with some variance  $\epsilon$ , as a function of its first argument, and is independent of its second argument. A reasonably simple shape of  $u$  is also assumed. Thus:

$$U(\mathbf{y} - \mathbf{x}, \mathbf{x}) = \exp\left(-\frac{(\mathbf{y} - \mathbf{x})^2}{2\epsilon^2}\right), \quad (11)$$

$$u(\theta - \alpha) = \cos^4\left(\frac{\theta - \alpha}{2}\right), \quad (12)$$

Next, the *differential binding* assumption is made. Behavioral studies indicate independent binding of the cognitive map to individual local parts of the environment. For example, Collett et al. (1986) and Collett (1987) studied goal-directed searching by gerbils when landmarks were displaced. The results indicate binding of internal representations to individual landmarks rather than to the whole configuration. A related study was conducted by Thinus-Blanc and colleagues (1987).

Therefore, it is assumed that in a geometrically changed environment, the gaussian (Eq. 11) is split into several components, each of which is bound to some rigid part of the environment, and their relative strengths depend on the rat's position  $\mathbf{x}$ :

$$U(\mathbf{y} - \mathbf{x}, \mathbf{x}) = \sum_i U_i = \sum_i C_i(\mathbf{x}) \exp\left[-\frac{(\mathbf{y} - \mathbf{x} + \mathbf{a}_i)^2}{2\epsilon^2}\right]. \quad (13)$$

Here  $\mathbf{a}_i$  is the vector of displacement of the  $i$ th part of the environment, and the coefficients  $C_i$  are smooth functions of  $\mathbf{x}$ . They reach their maxima at those environmental locations to which the corresponding terms  $U_i$  are bound.

A similar assumption about the stimulation function was made by O'Keefe and Burgess (1996), although their model does not include

internal P network dynamics or path integration. The partition (Eq. 13) of  $U$  may result from separate representations of the environment in the V system. We do not discuss possible neurophysiological mechanisms of the partition; it is sufficient to know that this assumption (Eq. 13) is consistent with experimental observations, as will be seen below.

To complete the definition of the model, the initial conditions must be specified. It is assumed that when the rat finds itself in a familiar environment (modified or not), the activity packet jumps under the influence of a strong external stimulus to the location on the corresponding chart that is mostly stimulated by V, that is, to the absolute maximum of  $U$  taken over all charts. The same assumption is made for the H system: the first perceived allocentric head direction in the entered environment is determined by the strongest visual cue:

$$\mathbf{y}(0) = \mathbf{y}_{\max}: U(\mathbf{y}_{\max} - \mathbf{x}, \mathbf{x}) = \max_{\mathbf{y}'} \{U(\mathbf{y}' - \mathbf{x}, \mathbf{x})\}, \quad (14)$$

$$\theta(0) = \theta_{\max}: u(\theta_{\max} - \alpha) = \max_{\theta'} \{u(\theta' - \alpha)\}.$$

This "initial condition rule" applies not only at the moment of entry or waking up. It is assumed that under certain circumstances the state of the path integration system can be reset during active running. This means a jump of the activity packet to another location on the same or a different chart. Although there is only one chart in the head direction system, chart switching may happen in the P-H system and appears to do so during shuttling tasks on linear tracks in which the rats follow routes back and forth between goals. In such cases, the charts representing journeys in opposite directions appear to be different, leading to an appearance of directional dependence of place fields (McNaughton et al., 1983). If two charts are associated with the same environment, it is assumed that given  $\mathbf{x}$  and  $\mathbf{y}$ , the probability  $P$  of switching from chart 1 to chart 2 is a function of stimulation on both charts:

$$P(\mathbf{x}, \mathbf{y}) = F(U_1, U_2) = p \cdot \theta(U_2 - U_1) \cdot \theta(U_1 - U_2), \quad (15)$$

where  $\mathbf{y}$  is the current activity packet location on chart 1;  $U_1$  is taken at  $\mathbf{y}$  on chart 1, and  $U_2$  is taken on chart 2 at its absolute maximum (the expected new activity packet location on chart 2 is at the absolute maximum of  $U_2$ );  $p$  and  $U_i$  are constants. The same formula describes possible jumps of the activity packet between two distant points on the same chart, if one takes  $U_1 = U_2 = U$ .

The shape of the activity packet is not a dynamical property in this reduced model; however, it affects place-field shape and dimension and therefore needs to be specified. Thus, a fixed gaussian shape of the activity packet is assumed, with the variance  $\rho$ , and if the activity packet is centered at  $\mathbf{y}$ , then the relative firing rate for a cell located at a coordinate  $\mathbf{z}$  on the currently active chart is:

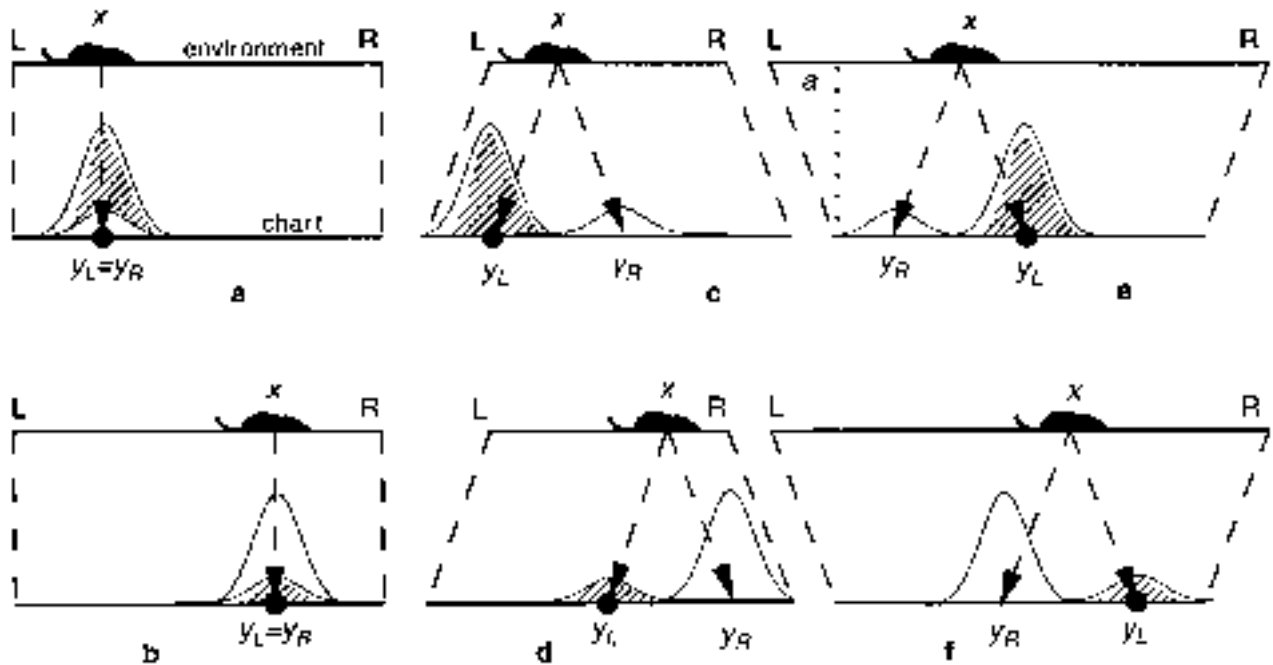
$$\text{rate}(\mathbf{z}) \propto \exp\left[-\frac{(\mathbf{z} - \mathbf{y})^2}{2\rho^2}\right]. \quad (16)$$

Equations 9–16 define the continuous model that was implemented numerically to reproduce the basic experimental results, including the results in unstable environments described below. Learning within this model will be considered afterward.

### Understanding place fields in stretched and shrunken environments

To develop an intuitive understanding, it is useful to consider a further simplification of the model (represented by Eqs. 9–16). The dynamics of the activity packet in this simplified model is described by Equation 9, if the last two terms in it are neglected: the path integration term and the noise  $\xi$ . The case of one-dimensional motion is considered, which is described by  $x = x_1$ , the rat's physical coordinate in the environment, and  $y = y_1$ , the activity packet coordinate on the chart. For now, the approximation of slow motion or strong visual input (however, not so strong that it can change the shape of the activity packet) is taken, in which the activity packet is always located at the local minimum of  $-U$  to which it is trapped, and never jumps to another minimum, unless the current minimum disappears. This may result in a hysteresis loop on the  $x$ - $y$  plane. Thus, suppose that  $U$  is given by:





**Figure 6.** Splitting of the stimulated region on the chart in linearly shrunken and stretched environments. *a, b*, Original environment; *c, d*, shrunken environment; *e, f*, stretched environment. *a, c, e*, The rat is in the first half of the journey; *b, d, f*, the rat is in the second half. The activity packet location is marked by the black dot. The amount of stretching *a* is defined as shown in *e*. On each figure, the top bar represents the environment, and the bottom bar represents the chart. Given two reference points *L* and *R* (walls, landmarks, reward sites, etc.), the current rat location *x* is associated with two locations on the chart,  $y_L$  and  $y_R$ , according to the original map anchored at *L* or *R*, respectively:  $y_L = x + a$  and  $y_R = x - a$ . These locations and surroundings are stimulated by the *V* array. The resultant distribution of stimulation is given by the sum of the two gaussian components,  $U_L$  (hatched) and  $U_R$  (blank), each of which is stronger near its reference point and linearly decays with distance from it (Eq. 17). For some locations *x* the resultant distribution has one maximum and therefore one stable activity packet location for a given *x*; for other locations there are two maxima, in each of which the activity packet may be trapped, and therefore, the activity packet location (and place-cell firing) may depend on the past trajectory, and through it, on the direction of motion. This may be the actual reason for acquired directionality of doubled place fields in a stretched environment, reported by O'Keefe and Burgess (1996). Switching of "host" maxima by the activity packet results in a hysteresis loop (Fig. 7). As the rat moves, the maximum of *U* at which the activity packet is trapped may disappear; in this case the activity packet quickly moves or jumps to another maximum and becomes bound to it. Moreover, with some probability (Eq. 15) the activity packet may jump to another maximum, when the current maximum becomes too weak and the other one is strong enough. In some cases (e.g., consider continuation of motion in *f*) the host maximum may disappear because it exits the part of the chart associated with the original environment.

$$U(y - x, x) = U_L + U_R = \frac{l - 2x}{2l} \exp\left[-\frac{(y - x - a)^2}{2\epsilon^2}\right] + \frac{l + 2x}{2l} \exp\left[-\frac{(y - x + a)^2}{2\epsilon^2}\right], \quad (17)$$

where *l* is the length of the stretched/shrunken environment and  $2a$  is the amount of stretching, which becomes negative in the case of shrinking; if  $l_0$  is the original length, then  $l = (1 + 2a)l_0$ . This situation is depicted in Figure 6.

Then the activity packet coordinate *y*, as a function of *x* and *a*, is given by the transcendent equation:

$$y = x + a \cdot \tanh\left[\frac{a}{\epsilon^2}(y - x) + \frac{1}{2} \log\left(\frac{l - 2x}{l + 2x}\right)\right]. \quad (18)$$

This equation for *y* has one or two stable roots, depending on the values of *x* and *a*. Figure 7 shows the activity packet coordinate *y* as a function of *x*, when *x* changes monotonically from  $-l/2$  to  $l/2$  and back from  $l/2$  to  $-l/2$ , for different values of *a*. Transition to another minimum occurs when the current local minimum disappears, which results in the hysteresis loop. This means doubling (stretched environment) (O'Keefe and Burgess, 1996) or missing (shrunken environment) (Gothard et al., 1996) place fields for the intermediate region on the chart.

Suppose a P cell is located in the middle of the chart ( $y = 0$ ), which is mapped to the middle of the box ( $x = 0$ ). Then, in the original environment (Fig. 7, diagonal line) this cell will fire near one location only:  $x = 0$ . The size of the place field is determined by the size of the activity packet  $\rho$  (measured along the *y* axis) and by the slope of the plot. When the environment is slightly shrunken, the place field shrinks too, remain-

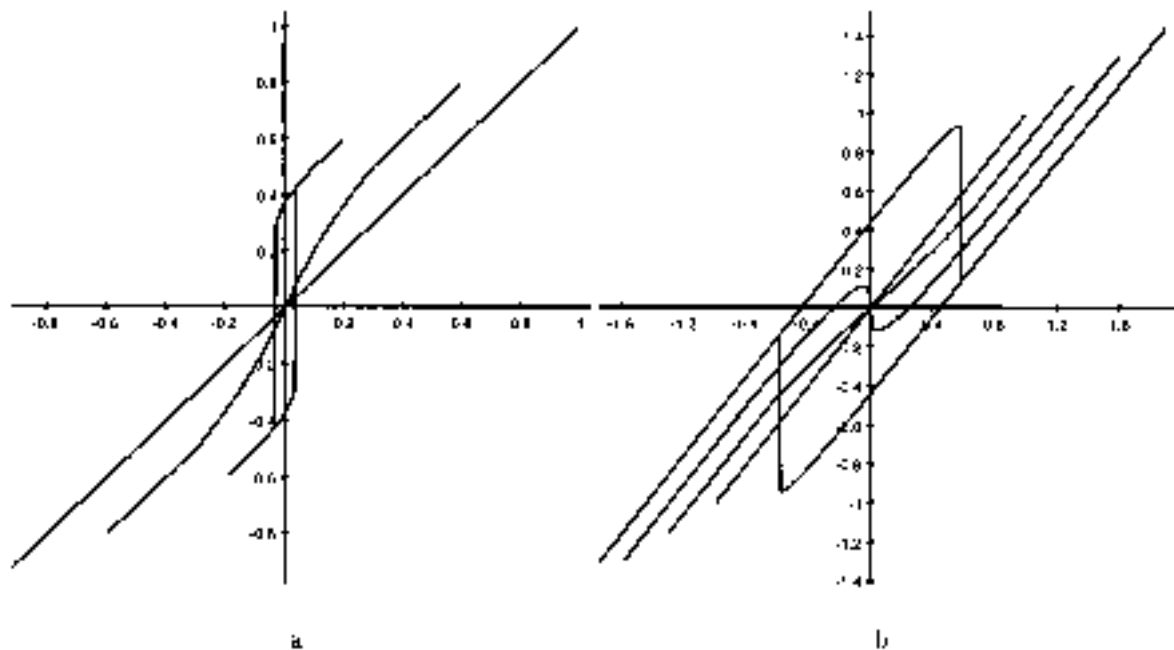
ing in the middle. After a certain degree of shrinking ( $\sim 50\%$  with the gaussian half-width  $\epsilon$  of 0.4) a hysteresis loop appears, which means that this cell will be skipped by the activity packet, and its place field must vanish. This situation is shown in Figure 7a. Given a smaller gaussian half-width, however, the place field may become doubled (not shown). For this particular cell, place-field doubling should always be the case in a stretched environment, if the hysteresis condition is reached (Fig. 7b).

#### Numerical implementation of the continuous model

The model (represented by Eqs. 9–16) was implemented numerically using the simple Euler scheme, which is sufficiently accurate for the qualitative purposes of this model. Typical parameter values were  $\rho = 0.2$  m;  $U_L = 0.6$ ;  $\mu = 0.04$ ;  $\mu = 0.007$ ;  $\nu = 0.0005$ ;  $\langle \xi^2 \rangle^{1/2} = 0.2$ ;  $\langle \zeta^2 \rangle^{1/2} = 0.002$ ; the value of  $\epsilon$  was varied between 0.1 and 2 m; the time step was 6 msec; the model rat's speed was  $\sim 0.2$  m/sec; the original track/box length was 1.5 m. Further details of simulation procedures are given in the next section. Results are represented in Figures 11 and 12.

Finally, the continuous model is extended on the basis of Equations 9 and 10 to incorporate learning mechanisms. To do that, the H circle, the P plane, and the V volume (Fig. 8) were divided into lattice cells to which virtual Boolean synapses are attached. In this sense the model is not continuous anymore but is still referred to here as "continuous" to distinguish it from the original network of integrate-and-fire units. Potentials *U* and *u* are computed as follows. Given the coordinates (*x*,  $\alpha$ ), *y* and  $\theta$ , gaussians centered at these coordinates in V and P with variances  $\epsilon'$  and  $\epsilon$ , respectively, and the bell-shaped curve (Eq. 12) in H were assigned. Then the contribution of each nonzero synapse to *U* (or *u*) is equal to the product of the two bell-shaped curves estimated at the two cells connected by the synapse.

In terms of the original model, the matrices of synapses  $W^{VP}$  and  $W^{VH}$



**Figure 7.** Transformation of the map in distorted environments. These plots show solutions to Equation 18, i.e., stable (with respect to corrections by visual information) activity packet coordinate  $y$  as a function of the rat coordinate  $x$  in an environment of the variable length  $l = (1 + 2a)l_0$  in the attractor map model approximation. *a*, Shrunk environment; *b*, stretched environment. The straight diagonal line represents the map in the unchanged environment ( $a = 0$ ). Values of the parameters are the variance of the “visual gaussian”  $\epsilon = 0.3$ ; the original length of the environment  $l_0 = 2$ . Different curves represent different amounts of stretching:  $a = 0, -0.2, -0.4$  (*a*);  $0, 0.15, 0.3, 0.45$  (*b*). The  $x$  axis represents the environment, with the reference points L and R (Fig. 6) located at  $-1/2$  and  $1/2$  (the two ends of each curve), whereas the  $y$  axis represents the array of place cells (chart). Therefore, horizontal sections of the plot for each  $y$  give environmental locations  $x$  of maximal firing rate for a place cell located at that given  $y$  on the chart. The width of an elementary place field is determined by the size of the activity packet  $\rho$  (measured along the  $y$  axis, not shown) and by the slope of the plot. For small deformations ( $-0.3 < a < 0.25$ ), place fields get shrunk or stretched together with the environment, although not in the right proportion; however, the map preserves topology. For bigger stretching ( $0.25 < a < 0.3$ ) place cells in the middle of the chart develop multiple place fields. When the deformation is too big ( $-0.5 < a < -0.3$ , or  $a > 0.3$ ), the hysteresis loop develops, which implies the appearance of directional place fields for the region within the loop. In the case of shrinking, some place fields disappear: the middle region of the  $y$  axis does not fire at any  $x$ . In the case of stretching, some place cells acquire double place fields, with at least one directional component. Place fields near the reference points, however, retain their compactness and shape. This picture provides intuitive understanding of the experiments in geometrically altered environments (Gothard et al., 1995; O’Keefe and Burgess, 1996).

are implemented as sparse matrices of Boolean values. According to the selected rules, the number of nonzero synapses per unit is limited for all units. In the learning regime, the probabilities of a synapse switching at a given iteration from zero to one ( $P^+$ ) and from one to zero ( $P^-$ ) are:

$$P_{ij}^+ = \beta(1 - \delta_{m_v, m_v^{\max}})(1 - \delta_{m_p, m_p^{\max}})(1 - W_{ij}^{AP})\langle S_i^{P,t} \rangle \langle S_j^{V,t} \rangle; \quad (19)$$

$$P_{ij}^- = \beta W_{ij}^{AP} \left[ \frac{\delta_{m_v, m_v^{\max}} \langle S_j^{V,t} \rangle (1 - \langle S_i^{P,t} \rangle) + \frac{\delta_{m_p, m_p^{\max}} \langle S_i^{P,t} \rangle (1 - \langle S_j^{V,t} \rangle)}{m_p} \right]; \quad (20)$$

where  $\beta$  is the rate of learning;  $m_v$  and  $m_p$  are the total numbers of synapses per unit ( $i, j$ ) in V and in P;  $m_v^{\max}$  and  $m_p^{\max}$  are their maximal allowed values;  $\delta$  is the Kronecker  $\delta$ . In the continuous approximation, the local average firing rates  $\langle S_i \rangle$  and  $\langle S_j \rangle$  in Equations 19 and 20 are approximated as gaussians estimated at locations of the units  $i$  and  $j$ , centered at  $(x, \alpha)$  and at  $y$  (or  $\theta$ ). Thus, learning is accomplished by modifications of the synapses  $[(x, \alpha) \rightarrow y]$  and  $[(x, \alpha) \rightarrow \theta]$ , as illustrated in Figure 8. Further details are given in Results.

## RESULTS

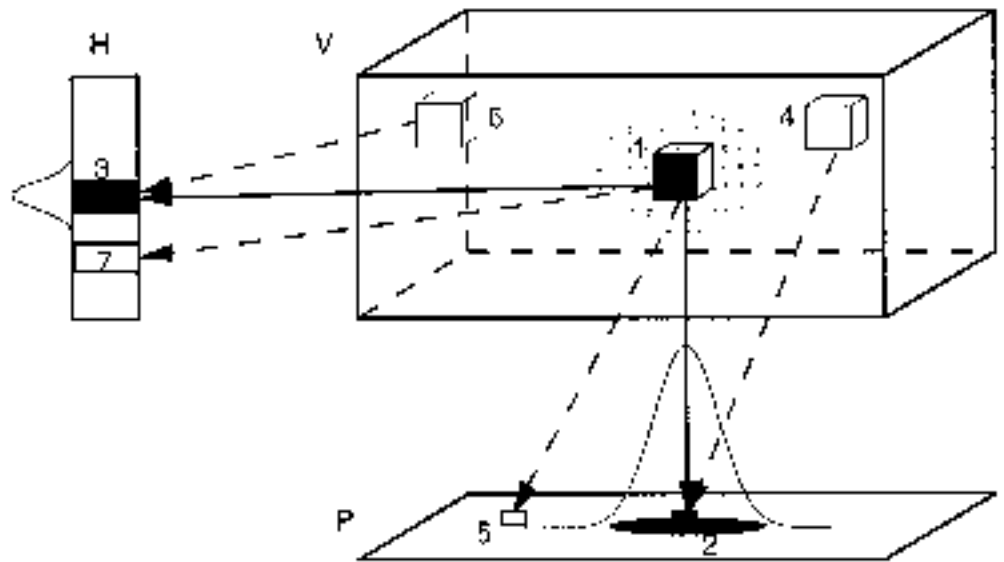
### Simulated dynamics of the integrate-and-fire MPI model

#### Observation of a multichart attractor map

First, the simulation exhibited self-focusing of activity on a chart in the P component (Fig. 9A, 10A,B). In this architecture, within a certain parameter range, the only dynamically stable configurations of cell activity are gaussian-like activity packets on one of

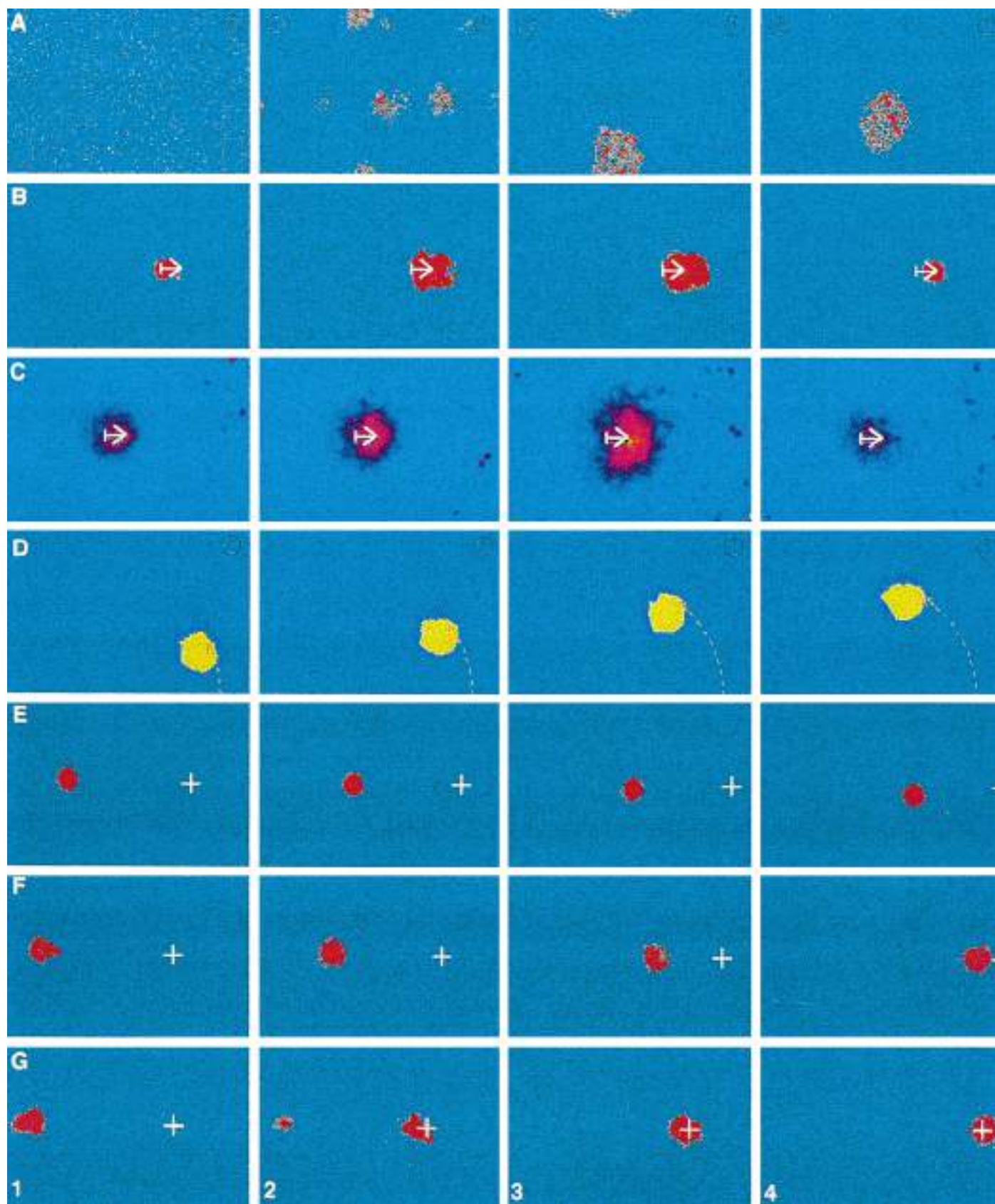
the charts. Such an activity packet was found to be stable in the sense that it maintains its uniqueness and shape, and given constant external input to the P–I system (no input from V) it moves on the chart with a persistent velocity. An activity packet forms spontaneously when the network is started with random activity. The process of self-focusing in the network of 45,000 units in each layer and with six charts is represented on Figure 9A; the variance of the activity distribution as a function of time for a network of 300,000 units in each layer with 20 stored charts is represented on Figure 10A; and the same for a network of 30,000 units with 100 and 200 charts is represented on Figure 10B. Figure 10A shows relatively fast relaxation that lasts less than one theta cycle:  $\sim 40$  msec. It takes 0.8 sec for 100 charts (Fig. 10B, solid line), and it does not happen at all in a network of 30,000 units with 200 charts (Fig. 10B, dashed line). The activity packet remains stable on the same chart while exploring the chart for at least 6 sec (Fig. 10D).

These observations indicate the existence of an attractor map in the P component: the two-dimensional set of fixed-point attractors at zero input to the I component corresponds to the two-dimensional set of possible activity packet locations on a chart. The estimated storage capacity in terms of the maximal number of stable charts stored in the same network is thus  $\sim 0.004 \times N$ . A similar value for a simpler model was obtained analytically by Samsonovich (1997). The latter analytical consid-



Comparison of the two figures (Fig. 9*B,C*; see also Fig. 10*C*) shows that the model reproduces qualitatively the phase precession phenomenon, which can be defined as follows: the firing phase of hippocampal neurons relative to the local theta rhythm advances systematically through almost 360° as the rat passes through the place field of each cell (O'Keefe and Recce, 1993; Skaggs et al., 1996). Stated differently, the activity packet moves ahead of the rat image on a chart during each theta cycle and then jumps backward, at the beginning of the next cycle, to the rat's actual location. The model (Figs. 9*B*, 10*C*) reproduces qualita-

**Figure 9.** Simulation results for the integrate-and-fire model—I. Snapshots in all rows except *B* and *C* are taken at a constant phase of the theta rhythm. *A*, Self-focusing, formation, and propagation of the activity packet in a six-chart network. The network consists of P and I layers. Each pixel on the figure represents a P unit. Each plate consists of  $256 \times 192$  pixels. Boundary conditions are periodic for all charts. The four plates *A1–A4* show the four sequential theta cycles that correspond to different stages of spontaneous self-focusing of activity on the chart 1. Spikes arranged according to this chart are represented by *red*; the background is *blue*. When the same units are arranged according to chart 2 (not shown), their spikes appear almost uniformly scattered and some are grouped into small patches. *B*, Simulated phase precession. The self-focused activity packet propagates to the *right*; the simulated rat location (same on all 4 plates) is shown by the *white arrow*. Only one chart is represented. *B1* through *B4* correspond to the four phases of the same theta cycle (0, 90, 180, and 270°). The center of the distribution clearly oscillates in the direction of motion, which resembles the phenomenon shown in *C*. *C*, Real phase precession of the activity packet in CA1 reconstructed from experimental data. Color on each plate represents an average firing rate distribution on a chart, where the momentary rat location and head direction is shown by the *arrow* in the *center*. High activity is coded by *red*. The *two ends of the arrow* are images of the two infrared light-emitting diodes attached to the rat's head, spaced 0.15 m from each other (for details, see Wilson and McNaughton, 1993). The average firing rate was computed from spikes that occurred within a narrow phase window with respect to the local EEG theta oscillations. Four consecutive phases were selected. Each plot was constructed as described in the caption to Figure 1 (Fig. 1 shows the average of the same data, taken over all phases). These oscillations of the distribution with phase in the direction of motion (from *left* to *right*) are known as the *phase precession* phenomenon (O'Keefe and Recce, 1993). This spatiotemporal structure of the experimentally observed activity packet was independent (within the error level) of the current trajectory configuration (e.g., *left* vs *right* turns), as well as of the velocity and the acceleration of the rat. This observation indicates that the spatiotemporal structure of the activity packet is probably a result of intrinsic dynamics of the hippocampal networks and does not reflect other brain representations, such as future plans or recent memories, goals, or intentions of the animal. (Some of these data can be viewed as movies at <http://www.nsm.aizona.edu/~alexei>.) *D*, The activity packet performs path integration. Four consecutive moments of the activity packet motion are represented. The simulated rat trajectory is a circle (*dashed line*); the simulated rat's position is shown by the *arrow*; the speed is constant. The head direction system was not simulated explicitly, as described in the text; therefore, the direction represented by the active I layer was always consistent with the direction of motion of the simulated rat. After self-focusing at a particular location on the chart, which is taken as the image of the starting point, the activity packet moves around a circle; however, a systematic error in the activity packet position accumulates with time. In this simulation, visual input to the P layer was absent. Spikes are represented by *yellow*. *E*, *F*, *G*, The role of visual input. *E*, The activity packet performs path integration, similarly to *D*, but now the simulated rat trajectory is a straight line. The actual simulated rat's position is shown (*Figure legend continues*)



by the *cross*; stimulation of the P array by V is turned off. *F*, Addition of a gaussian-shaped stimulation to the P layer centered at the *cross* changes the activity packet velocity. The stimulation is relatively weak; the activity packet accelerates following the center of the stimulated area. *G*, The stimulation is strong enough to cause the activity packet to jump to the center of the stimulated area. The jump occurs with a certain probability, when the stimulation magnitude exceeds a certain threshold. The duration of the jump is approximately one to two theta cycles (time scales are different in *F* and *G*).

tively the features specific for a two-dimensional motion (Skaggs et al., 1996), although there are important deviations from the actual phase precession dynamics that will be treated elsewhere.

Thus, the phase precession phenomenon occurs as a byproduct of path integration in this model. This mechanism is quite different from that suggested by Tsodyks et al. (1995), in which the retrograde motion of the activity packet is caused by sensory input. In contrast, according to the present model, the effect should persist regardless of availability of sensory cues.

As can be seen in Figure 10C, the average group velocity of the activity packet remains constant given a constant input to the I layer. In addition, over multiple trials, the average speed of motion appeared to be a monotonic function of the I layer activity.

### *Simulated path integration*

Next, the activity packet was allowed to perform path integration (Figs. 9D, 10D) in a circular motion, in the absence of input from V. The simulated rat moves counterclockwise around a circle. The activity packet repeats this motion with a certain degree of error, which accumulates with time (here, by assumption, the starting position of the activity packet on the chart coincides with the image of the model rat's starting position in the environment). This result shows that the velocity and the direction of motion of the activity packet can be controlled effectively by appropriate stimulation of the I array, and therefore the activity packet position can be accurately updated by idiothetic information; however, after a few seconds, a small correction of the activity packet position becomes necessary.

### *The role of sensory input: accelerated motion and jumps of activity packet*

The next numerical study (Fig. 9E–G) examined the question of correction of the activity packet position by the visual input. Now the simulated rat moves to the right with a constant velocity. The rat's position (projected onto the chart) is marked by the cross. It differs from the activity packet position from the beginning of motion, because of an initially introduced error. The input to the P layer from V representations is absent in Figure 9E, relatively weak in F, and relatively strong in G. Therefore, the V system has no effect on the activity packet motion in E. (See Materials and Methods.)

It was observed that the addition of a gaussian-shaped stimulation to the P layer from V changes the activity packet velocity, thus correcting the activity packet position. Namely, a smooth acceleration of the activity packet occurred at a relatively weak V input (Fig. 9F), and an abrupt jump to the stimulated region occurred when V input was relatively strong (Fig. 9G). Both of these phenomena have recently been observed experimentally (Gothard et al., 1996).

### **Continuous model: reproduction of experimental results**

#### *Place fields in stretched and shrunk environments*

O'Keefe and Burgess (1996) examined the question of geometrical determinants of the hippocampal spatial code and showed that a place field can become stretched or even doubled in a linearly stretched rectangular environment. When fields doubled, each half field became directional, with the preferred directions of each half oriented toward each other. When the environment is stretched in both directions, a place field may become tripled or may vanish.

In contrast, Gothard and colleagues (1996) performed a set of experiments involving shuttling between two food sites on a linear track. During the experiment, the distance between reward sites was reduced by varying degrees on a random schedule over trials. As a result, place fields became compressed, but not always in the right proportion, and in some cases disappeared.

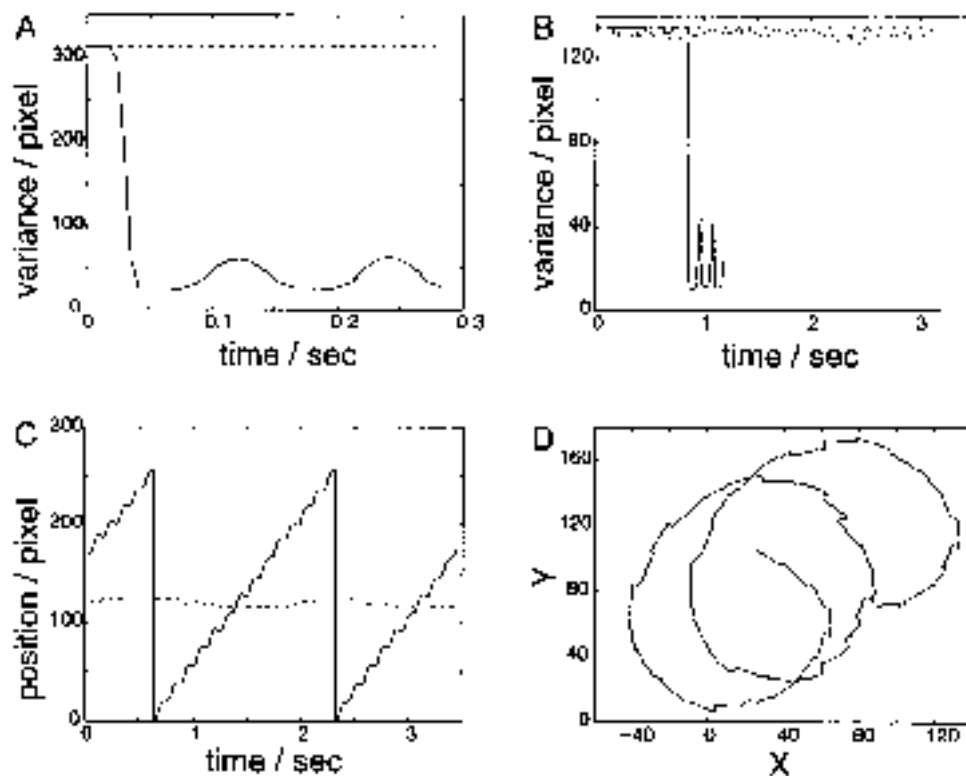
It is remarkable that when two place fields that originally did not overlap became overlapped in the shrunken environment, the two cells actually did not fire simultaneously, as shown by cross-correlation analysis (Gothard et al., 1996). The apparent overlap in the average place fields was thus attributable to trial-to-trial fluctuations in the behavior of the activity packet. This observation supports the attractor map concept, according to which each P cell has its permanent location on the chart, and therefore two cells associated with two distant locations in the original environment cannot be active simultaneously as long as the activity packet remains on the same chart. Nevertheless, they can be activated in sequence during a short time interval, if an activity packet jump occurs.

In both cases (Gothard et al., 1996; O'Keefe and Burgess, 1996), P cells showed binding to a part of the environment that was behind the rat and generally outside of its field of view (the visual field in rat is approximately 300°), although the velocity of the activity packet on the chart in these cases was approximately the same as in an unperturbed environment. This fact makes a local-view explanation (i.e., a direct effect of visual cues) of O'Keefe and Burgess (1996) unlikely (see Discussion). Another straightforward explanation is that after the "position fix" at the beginning of the route, the activity packet is guided primarily by the path integration mechanism, until the rat "bumps" into the opposite wall (McNaughton, 1996). Therefore, as concluded by Gothard et al. (1996), these observations strongly support the path integration concept. The analysis, however, does show that place-field doubling and acquired directionality in this case could result from visual input in the limit of very slow motion (Figs. 6, 7).

First, numerical results for the continuous model (represented by Eqs. 9–16) are presented, with the potential  $U$  given by Equation 17. The predicted place-field modifications computed according to these equations for the cases studied experimentally by O'Keefe and Burgess (1996) and Gothard et al. (1996) are shown in Figure 11a–h. In the case of two-dimensional motion in a box (Fig. 11a–f), the assumptions were that the only source of discrepancy between  $\theta$  and  $\alpha$  is noise (distal directional visual cues remain valid for the original, stretched, and shrunken boxes), and that the rat performs two-dimensional random walks in the box, reflecting from the walls. It was also assumed that the visual stimulus originating from a particular wall decays linearly with the distance from this wall, reaching zero near the opposite wall. The results clearly show the phenomena of doubling, stretching, and vanishing of place fields.

The effect of finite noise is broadening of the place field and later appearance of the hysteresis (not shown). Removal of the path integration term changes nothing in the limit of slow motion, which is the only case when the above simplified consideration (Fig. 6) may be valid. In the opposite limit of fast motion, however, path integration becomes necessary to speed up the activity packet, which because of its finite mobility under a given driving force cannot follow the stimulated area without an additional driving force.

The two components of the doubled place field (Fig. 11c) are highly directional: each component fires when the rat is facing the more distant wall. This directionality results from two mecha-



**Figure 10.** Simulation results for the integrate-and-fire model—II. Networks with 300,000 units in each layer, with 20 charts (*A, C, D*), and with 30,000 units per layer with 100 and 200 charts (*B*) were implemented (see Materials and Methods for details and other parameters). *A*, Self-focusing on one of the 20 charts in a network of 300,000 units. The variance of the distribution on the active chart (*solid line*) and on one of the other charts (*dashed line*) is plotted as a function of time. Both curves correspond to the same simulation epoch. *B*, Self-focusing in the network of 30,000 units implementing 100 charts (*solid line*). The variance of the distribution on the chart that eventually becomes active is plotted as a function of time. The *dashed line* represents the minimal variance over all 200 charts observed in another simulation in the network of 30,000 units. *C*, Linear motion of the activity packet in a network of 300,000 units with 20 charts. Only one layer in the I component is active, which corresponds to the *x* direction. *Solid line*, *x* coordinate of the center; *dashed line*, *y* coordinate of the center. Oscillations of the position of the center in the direction of motion are clearly seen. This somewhat resembles the phase precession phenomenon; however, retrograde motion occurs at particular locations only. *D*, Path integration during circular motion in a network of 300,000 units per layer with 20 charts. The plot represents the trajectory of the activity packet on the active chart; motion is counterclockwise. The duration of the simulated episode is 6 sec. The model rat velocity vector rotates with a constant angular velocity. Cumulative error of path integration is clearly seen as deviation from a circle. A small zigzag at the beginning of the trajectory is attributable to the self-focusing period.

nisms: (1) the hysteresis in activity packet motion on the chart (Fig. 7) and (2) the path integration together with “position fixes” (McNaughton, 1996). The nature of directionality in this case is considered further in Discussion.

In the case of a shrunken environment (Fig. 11*f*), the place field vanishes instead of doubling; however, doubling could occur with slightly different coefficients in Equation 17, as well as according to a pure path integration mechanism, depending on the location of the field (McNaughton, 1996). Directionality of the two components in this case should be opposite to the directionality in the case of Figure 11*c*.

If a cell has its place field closer to one end of the box, then, as can be derived from Figure 7, the place field may be doubled with increased stretching or not doubled at all if its location is very near the wall. In the case of shrinking, such a place field may acquire directionality without doubling.

The question of directionality requires special attention in the case of the Gothard et al. (1996) study and similar experiments using shuttling or route-following paradigms. All place fields observed in this study between the reinforcement sites were 100% directional in the shrunken as well as the original environment. As discussed above, this fact suggests that chart switching occurs when the animal reaches the goal. It is easy to see that chart switching results in a hysteresis loop (of a different origin than that just

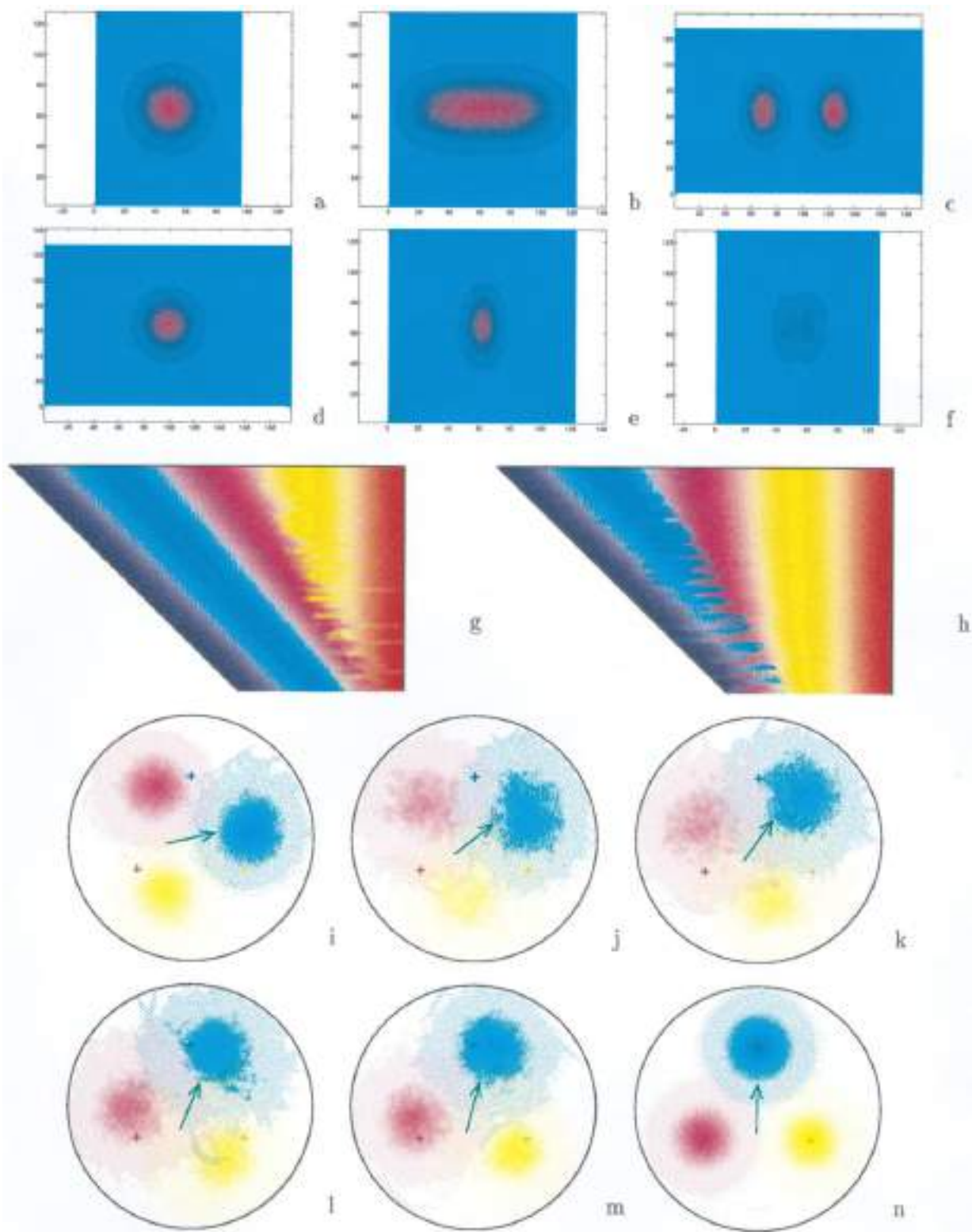
discussed) and therefore in directionality of place fields for cells contained within this loop, i.e., in the middle region of each chart.

Place-field behavior on the shrinking rail (Gothard et al., 1996) was simulated using the model represented by Equations 9–17 at a finite velocity, with finite noise and path integration terms (Fig. 11*g,h*). The assumptions here are the same as before, except that (1) chart switching was assumed to occur each time on reaching a reinforcement site or earlier, according to Equation 15; (2) the angle  $\theta$  was taken fixed (0 or  $\pi$ , depending on the direction of motion); and (3) the moving part of the environment, the box, was assumed to have limited visibility described by an exponentially decaying rather than linear coefficient in Equation 17, with the characteristic length of 0.3*l*.

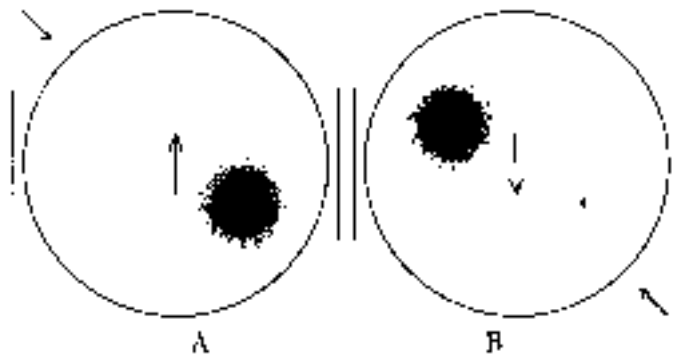
The results (Fig. 11*g,h*) reproduce the principal observations: vanishing and shifting of place fields, including “fractional slope” place fields, i.e., place fields that show smaller displacement than that of the box. The relative influence of the moving box on place field location in this case is different for inbound and outbound journeys, in agreement with the experiment. This difference illustrates the effect of path integration.

Path integration has a clear influence on place fields in this case. It becomes necessary for the activity packet to move when the limited range of visibility of the box behind the rat is taken into account by introducing a finite-range rather than exponential co-





**Figure 11.** Simulation results for the continuous model. *a–f*, Predictions for the place-field modifications in two-dimensional environments computed according to Equations 9, 10, 12–14, 16. The two reference points L and R (Fig. 6) are the left and right walls of the box on *a–f*. In the case of stretching (*a–c*) studied by O'Keefe and Burgess (1996), the original place field (*a*) gets stretched (*b*) and then becomes split into two fields (*c*). In the case of shrinking (*d–f*) studied by Gothard et al. (1996), the original place field (*d*) becomes shrunken (*e*) and eventually disappears (*f*). An intuitive explanation of these phenomena is given by Equation 18 and Figure 7. *g, h*, Place fields on a shrinking rail (*g*) outbound journey; (*h*) inbound journey. The *five colors* represent firing rates of the five selected cells. Each *colored horizontal line* represents a simulated running episode. Therefore, the horizontal axis represents the rat's position on the rail; the vertical axis is the amount of shrinking of the environment (*–a*). "Hatching" results from random jumps of the activity packet according to Equation 15. Similar results were obtained in the Gothard et al. (1996) experiment. *i–n*, Simulated slow rotation of place fields after disorientation in a familiar environment. Starting from a disoriented state (*i*), the system slowly relaxes with time to (Figure legend continues)

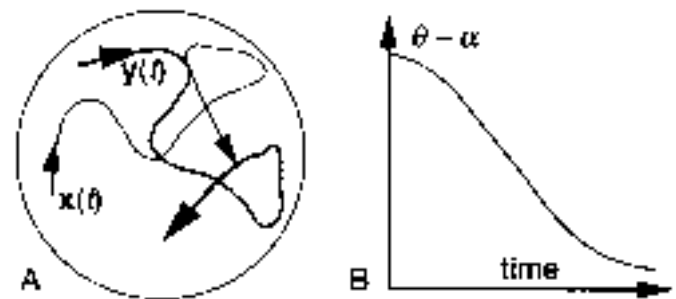


**Figure 12.** Numerical counterpart of the Sharp et al. (1990) experiment. According to the assumption of the model, on entry into a familiar environment, the activity packet appears on the chart that has been associated with this environment, and at the location that is mostly stimulated by V representation of the local view. Thus, if the position of the stimulated domain is determined by a cue card, and the cue card is doubled, then there are two stimulated domains on the chart, and the strongest one is selected by self-focusing of activity. Later place fields remain bound to the selected cue card, if the two maxima of stimulation originating from the two cue cards remain well separated on the chart. This is provided by an appropriate choice of parameters of the potential  $U$ . Therefore, after doubling the cue card, the locations of place fields depend on the entry site, which corresponds exactly to the results observed by Sharp et al. (1990). *A*, The stable place field in the cylinder of one selected model place unit, which is formed on entry to the environment from north-west. *B*, The stable place field of the same unit formed on entry from southwest. The original place field center is shown by the cross. The two cue cards are represented by vertical lines. The arrow outside the circle shows the entry site. The arrow in the center shows the “perceived” north.

efficient into (17), although in this case the activity packet is accelerated by the other stimulated area on the chart, without path integration. It is unlikely, however, that this mechanism of acceleration by chance would result in the correct activity packet velocity.

#### Dependence of place field location on the entry site

By assumption, on entry into a familiar environment, the activity packet appears on the chart that has been associated with this environment, and at the location that is mostly stimulated by the V representation of the local view. Thus, if the position of the stimulated domain is determined by a cue card and the cue card is doubled, then there will be two stimulated domains on the chart. The stimulation originating from the most clearly visible cue card at the moment of entry will be stronger and therefore, according to the assumption of the model, will determine the activity packet location on the chart. Subsequently, place fields will remain bound to the selected cue card. This implies that the positions of the place fields may depend on the entry site and/or the orientation of the rat when first placed into the cylinder. This corresponds to results observed by Sharp et al. (1990). These results are reproduced numerically in Figure 12, using the model represented by Equations 9–17. The Sharp et al. (1990) experiment differs from the case of the O’Keefe and Burgess (1996)



**Figure 13.** Conflict between visual cues and path integrator in the case of disorientation in a familiar environment. *A*, The activity packet and the gaussian (the localized stimulation from the V network) get together at the center of the circle. The activity packet trajectory is  $y(t)$ ; the trajectory of the gaussian (which is the image of the rat trajectory) is  $x(t)$ . The activity packet would be captured immediately in the absence of path integration. Knierim et al. (1995) observed the opposite case: slow rotation of place fields (seconds to minutes), which is evidence for persistent direction of activity packet motion. *B*, Relaxation of the systematic error in the “perceived” head direction  $\theta$  with respect to the “actual” head direction  $\alpha$ , which results from the influence of a symmetry-breaking visual cue representation on the internal head direction representation  $\theta$ ; the solution to Equations 10 and 12.

experiment in that the activity packet remains bound to the originally selected part of the environment at all times, which implies that the corresponding local minimum of  $U$  never disappears during motion in this restricted environment. In the numerical experiment this situation is easily created by appropriate selection of the coefficients  $C_i$  in Equation 13.

#### Slow rotation of place fields

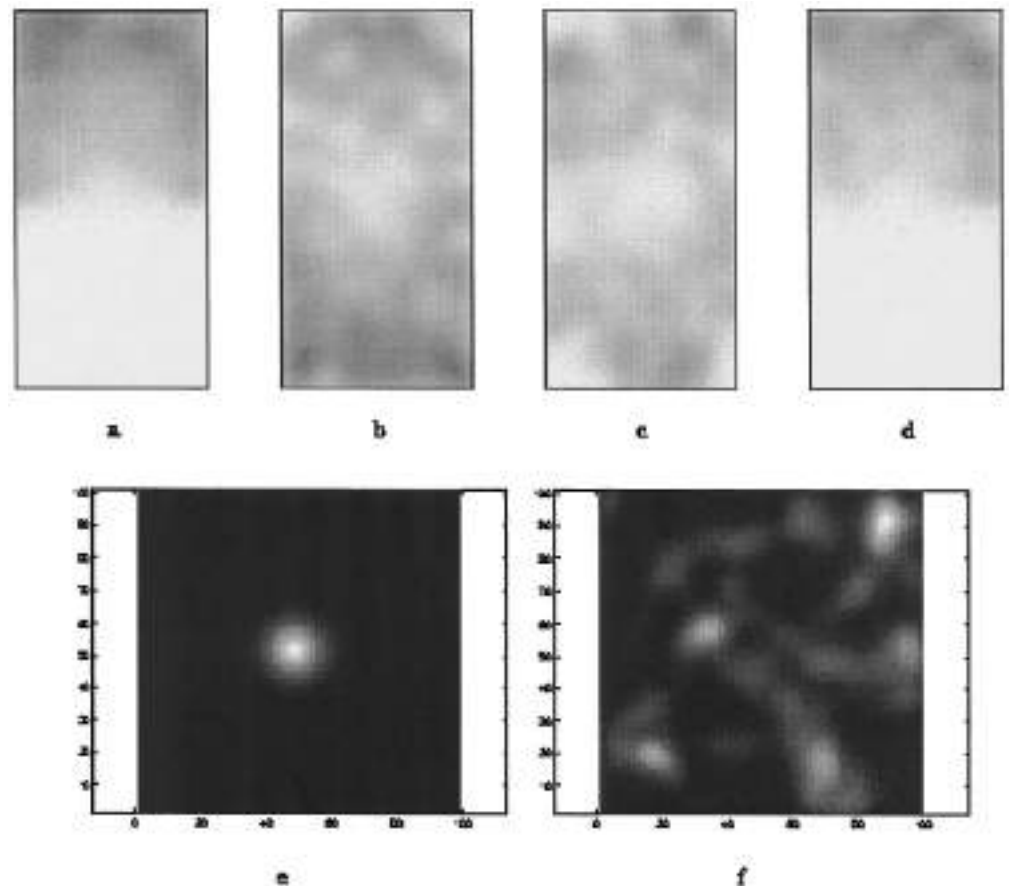
Now consider two-dimensional motion of the model rat in a cylinder according to the model represented by Equations 9–17. Assume here strong binding to tactile cues, i.e., that the radial component of the activity packet coordinate is reset every time the rat reaches the wall of the cylinder and therefore is close to the model rat’s radial coordinate at all times. If the rat is initially disoriented, that is,  $\theta$  substantially differs from  $\alpha$ , then most of the time the activity packet is far enough from the stimulated region and therefore its relaxation to the stimulated region may go relatively slowly; however, at some point, namely, at the center, the activity packet and the stimulated region come together. In the absence of path integration, when the activity packet has no preferred direction of motion, it would follow the stimulated region after the first touch, and thus spatial representations would be reset immediately. Because of path integration, however, the activity packet continues its motion in the wrong direction, missing the stimulated region (Fig. 13). This results in persistent displacement of the place fields from their original locations. Interaction of the V and H systems, however, slowly corrects the perceived yaw angle  $\theta$ , resulting in slow rotation of place fields together with H-cell tuning curves, as was observed by Knierim et al. (1995). The simulation results are shown in Figure 11*i–n*.

←

its stationary state ( $n$ ) determined by the learned  $W^{VH}$  and  $W^{VP}$  connections. The green arrow shows the averaged perceived head direction; the three colors (red, blue, and yellow) represent temporary place fields of the three selected P units. These distributions were obtained by averaging over an ensemble of equivalent model running sessions for five ( $i–m$ ) equal sequential time intervals. The original place fields ( $n$ ) of these units were eventually restored. Their centers are marked by colored crosses on all six figures ( $i–n$ ). Similar results were obtained experimentally by Knierim et al. (1995). Slow relaxation of place fields indicates involvement of a path integration mechanism.



**Figure 14.** Simulation of learning of a novel environment. *a–d*, Numerical counterpart of the Wilson and McNaughton (1993) experiment. The density of each plot represents the distribution of the error in reconstructed rat's position according to the averaged map for this particular session. Error distributions were averaged during equal time intervals. After exploration of the north part (top) of the box (*a*), the rat discovers the south part (*b*). After exploration of this part, the error in it is gradually reduced (*c*), reaching the same level as in the north part. The average error in the north part stays approximately at the same level during all sessions (*a–d*). *e*, Place field formed by learning with path integration; *f*, without path integration. In the absence of the path integration term in Equation 9, the activity packet still moves, driven by initially random  $W^{VP}$  connections (which are modified during this motion). This process, however, did not converge to a one-to-one map from the environment to the chart during the simulated running session. Therefore, a path integration mechanism must be substantially involved in the process of learning.



### Learning a novel environment

Following the scheme described in Materials and Methods, the two synaptic matrices  $W^{VP}$  and  $W^{VH}$  with sizes of  $400,000 \times 90,000$  and  $400,000 \times 10$  were implemented. The maximal number of nonzero synapses per unit was 1 for the V array, 3 for the P array, and 40,000 for the H array; the value of  $\epsilon'$  was 0.05 m. Results are represented in Figure 14.

First, the Wilson and McNaughton (1993) experimental paradigm (Fig. 14*a–d*) was reproduced numerically. A rapid stabilization of place fields in the new environment was observed, with a decrease of the average position error with time. In analogy to how error was defined by Wilson and McNaughton (1993), the error was measured as the discrepancy between the activity packet location on the chart and the simulated rat location in the environment, projected onto the chart according to the average map obtained during this particular simulated epoch. At the same time, persistence of previously formed place fields was observed, in agreement with the results of Wilson and McNaughton (1993).

In agreement with the experimental results, the average error in the new part of the environment was higher than in the familiar part at the beginning of exploration and was reduced to the same level after a certain learning period. At the same time, the average error in the familiar part remained approximately on the same level before, during, and after exploration of the new part.

In addition, the results of learning with (Fig. 14*e*) and without (Fig. 14*f*) the path integration term in Equation 9 were compared, using a "random start" for  $W^{VP}$  and  $W^{VH}$ . Without path integra-

tion, the activity packet moves on the chart, as the animal moves in the environment, being guided by V representations via pre-existing V-to-P connections (which initially were taken as random). This process, which is similar to the Kohonen learning scheme, after a given period of time (which is too short for the Kohonen learning) results in formation of a "map," which is not one-to-one. The results of the simulation show that despite the preexistence of an attractor map in P, appropriate binding of this attractor map to the environment requires a path integration mechanism.

## DISCUSSION

### Principal findings

The principal findings from the integrate-and-fire model concern the formation, shape, stability, oscillations, two-dimensional controllable motion, and controllable jumps of the activity packet in the proposed multichart MPI model. These results are obtained on the basis of numerical simulations of networks of leaky integrate-and-fire model neurons in discrete time, with predetermined total activity.

The principal finding from the continuous model is the reproduction of several basic experimental facts about place fields, including doubling, vanishing, reshaping in distorted environments, acquiring directionality in a two-goal shuttling task, rapid formation in a novel environment, and slow rotation after disorientation. Simulation results strongly suggest that a path integration mechanism is involved in the foregoing processes.

## Significance

The theoretical framework proposed here provides a plausible explanation for how hippocampal place-cell activity may be updated by self-motion cues and how self-motion itself may provide the metrics for representation of spatial relationships. It also provides new insight into how multiple “maps” may be stored within the same synaptic matrix with minimal interference. A multichart architecture implies that the activity of a given place cell has meaning only in the context of the ensemble of other cells that are active with it at a given location on a given chart. In particular, the cooperative interactions that lead to an activity packet suggest that it is misleading and, in general, not useful to think of individual place cells as representing external features at all, in the manner that one thinks of sensory neurons. The multichart architecture easily explains the generally all-or-none “remapping” effects that have been observed in experiments involving cue rearrangements and deletions. Comparison of the simulation results with available experimental data unambiguously supports the underlying general theoretical framework, i.e., the activity packet self-localization property, the multichart property, and the path integration property. The theoretical framework itself provides new insight into how path integration and local view-based mechanisms may interact to update hippocampal spatial representations, and it begins to explain some of the peculiar effects that are observed when path-integration and external sensory inputs are placed in conflict with one another, such as during geometrical distortion of the environment.

The question arises as to how many independent charts can be represented in a globally interconnected network such as CA3 (Amaral and Witter, 1995) while still maintaining stable activity packets that remain on a single chart. No evidence of instability was observed in simulations with up to 100 charts. Analytical calculations of the maximal number of stored uncorrelated spherical attractor maps in a Hopfield-like network (Samsonovich, 1997) give the number  $0.0042 \cdot N$ , where  $N$  is the number of model neuronal units. In conjunction with the numerical simulations, this suggests that a network the size of the rodent CA3 may have the capacity to store substantially more than 100 charts, although the number is likely ultimately to be limited not by the number of neurons but by the number of connections. The possibility remains, however, that the capacity may be improved further by invoking some temporary, experience-dependent plasticity of the intrinsic connections (see below).

The proposed anatomical implementation of the MPI model is not the only one possible. In principle, the P-I path integrator loop of the MPI model could be based on other structures, such as presubiculum and parasubiculum or entorhinal cortex. One difficulty with this interpretation is that these structures appear not to express multiple charts (e.g., Quirk et al., 1992; Sharp, 1997). The possibility remains, however, that dentate gyrus and CA3 may select an active chart that is expressed in these structures and in CA1, whereas a single, “universal” chart is implemented in the subicular complex and entorhinal cortex. This possible variant of the MPI scheme will be discussed elsewhere (Samsonovich and McNaughton, 1997).

## Alternative theoretical approaches and related works

The *trajectory learning concept*, which has not yet been discussed here, is in some sense opposite to the concept of a cognitive map. According to this point of view (Blum and Abbott, 1995), the intrinsic synaptic connections between place cells representing sequentially visited locations are asymmetrically potentiated dur-

ing repeated motion along a specific route. Indeed, the asymmetric expansion of place fields in a direction opposite the rat's motion that has been predicted by such models has been observed experimentally (Mehta et al., 1997), and recent demonstrations of the spontaneous reactivation of sequences of place fields during sleep subsequent to repeated route-following (Skaggs and McNaughton, 1996) suggest that sequence information may be encoded in asymmetric changes in the intrinsic connections of a chart; however, the effects observed to date represent relatively weak changes in the overall structure of existing charts and do not involve any topological rearrangement. Moreover, any asymmetries would tend to cancel out in the case of random two-dimensional foraging, leaving a net simple expansion of the place fields. In general, mild, experience-dependent changes in the intrinsic connectivity of the P layer may be thought of as a higher order effect superimposed on the basic dynamics described here. They may be useful in stabilizing the current or recently visited charts and may play a role in the spontaneous reactivation of recently experienced chart coordinates during sleep (Shen and McNaughton, 1996) but are unlikely to alter the overall dynamics in any major way.

The model of path integration recently proposed by Touretzky and Redish (1996) assumes a rather complicated learning mechanism for P units that store local views in conjunction with internal path integrator coordinates. In addition, this model assumes a rather complicated mechanism of “resets” for internal representations. The path integrator is introduced by Touretzky and Redish (1996) as an abstract element with *ad hoc* functional properties. This model, in its present form, fails to account for multiple uncorrelated charts. In addition, the model does not explain how allocentric features, independent of the head orientation, are extracted by place units from a local view representation.

In another paper (Redish and Touretzky, 1997), the same authors propose a holistic model of the hippocampal spatial code formation based on a concept of a *reference frame*, which includes a reference location, a reference direction, and a set of frame-specific declarative memories, as well as the relationships between these components. A representation of a reference frame is presumably supported by an ensemble of interacting brain structures located inside and outside the hippocampus proper; in particular, this results in selection of an active chart. Thus, the question of “holistic” versus “localized” attractor dynamics underlying chart selection remains open.

O'Keefe and Burgess (1996) offer an alternative explanation of their recent observation of place-field splitting when one or more dimensions of their rectangular recording environment were doubled. Their model is based on the assumption of stimulation of the P cells by visual representations, which resembles the *U* potential of the present model; however, there is no self-localization mechanism (no P-to-P synapses) and no path integration mechanism in their model. Therefore, the model fails to account for the existence of place fields in complete darkness and for slow rotation of place fields together with head direction fields after disorientation. In addition, in this particular case, it is not clear how different “superimposed directional subcomponents” exactly compensate each other's directionality in the original environment, resulting in a compact, omnidirectional place field.

A model called the “cognitive graph,” proposed by Muller et al. (1991, 1996), suggested a possible mechanism for encoding relative distances in the environment based on a synaptic matrix similar to the present P layer. The possibility of storing several

maps in such a network was also discussed. The architecture outlined in the present work, in principle, would also support the computations that motivated the cognitive graph theory; however, there are several fundamental differences between the theories. First, in the cognitive graph theory, the existence of some external, spatially selective input was assumed as the primary source of place-cell activity. The P-to-P matrix itself was learned after coactivation of cells that had been driven by this input and was sensitive to the history of the rat's behavior. The present theory assumes the opposite, namely that the basic attractor map structure in the network is given, and the external inputs become spatially selective only through learning and remain always relatively weak compared with the intrinsic connections. It requires cooperative interactions among a minimum number of cells in the corresponding position on a given chart to generate a stable activity packet. It is this architecture that enables orthogonal representations to be assigned to very similar environments (e.g., Bostock et al., 1991) or sometimes to the same environment under changing behavioral contexts (Markus et al., 1995). Completely orthogonal maps for similar environments cannot be generated easily in a system whose spatial selectivity is initially imposed externally. Similarly, the cognitive graph theory could not account, without *ad hoc* assumptions, for the apparent preexistence of place fields in the absence of any external input (Quirk et al., 1990). Second, although a secondary role of movement information was alluded to, the "cognitive graph" theory as presented actually had no intrinsic metrics, because the synaptic modification process depended on a time interval between locations rather than explicitly on distance, and there was no explicit coupling between self-motion and place-cell activation as in the present model. Thus, the cognitive graph theory could not examine effects that are explicitly dependent on path integration, such as the correct activation of place-cell sequences in the absence of external sensory input or the slow rotation of place fields in cylindrical environments (Knierim et al., 1995). Finally, the cognitive graph theory encountered serious problems with the existence of multiple place fields in a given environment. These do not pose more than a computational inconvenience and a net reduction in capacity in the present theory.

### Predictions

The proposed MPI model predicts (1) the specific, persistent, microscopic architecture of the CA3 network (although, as stated above, the model may have a different implementation in anatomy); (2) that the directional relations between the I cells will be the same for all environments; (3) a possibility of "translational disorientation" in a nontrivial, translationally symmetric environment; (4) "activity packet jumps" occurring at the scale of one to two theta cycles; and (5) coherent behavior of neighboring cells on the active chart during distortions of the environment provided that chart switching does not occur [this prediction distinguishes the present model from that of O'Keefe and Burgess (1996)]. Prediction 4 was recently confirmed experimentally (Gothard et al., 1996). Strictly speaking, statement 1 is an assumption of the model rather than a prediction; however, the agreement between experiments and the simulation results strongly suggests the validity of this assumption.

A corollary of prediction 5 is that in a distorted environment two place cells with overlapping place fields cannot be active simultaneously if their original place fields do not overlap and there is no chart switching. This prediction is consistent with recent observations (Gothard et al., 1996).

Another critical test would consist of an observation of place-cell firing during head turns in a stretched box (O'Keefe and Burgess, 1996) while the rat is located inside a directional component of a split place field. According to the model of the present paper, firing rate in a "directional" component is actually determined by the past trajectory rather than by the current head direction. The prediction of the present model, therefore, is that the place-cell firing should persist in this case for all head directions, as long as the rat stays within the place field. According to the theory of O'Keefe and Burgess (1996), the firing should remain direction-dependent.

### Generalizations

The above picture of the hippocampal model of space argues for a general concept of an attractor-map-based internal cognitive model. Many cognitive tasks can be represented effectively on the basis of abstract mathematical models such as graphs (e.g., Muller et al., 1996) and manifolds and therefore may require internal "mapping." According to this concept, the underlying attractor map of the cognitive model is presumed to preexist, and representations of particular memory items may become bound to it. For example, in analogy to the multichart architecture for space, it is possible to conceive of attractor map primitives for egocentric space or even for objects such as chairs or people, which exist in a synaptic matrix without yet having been bound to particular exemplars. To navigate around the map, it is necessary to have a sense of possible local transitions. This can be provided by another attractor map, in analogy with the spatial map and the head direction map in the model of this paper. A related concept was proposed by Droulez and Berthoz (1991).

Therefore, attractor maps are likely to be found in various brain areas in addition to the hippocampus; e.g., they may underlie cortical mental rotations (Georgopoulos et al., 1989; 1993) and motion control (Sparkes and Neson, 1987; Sparks et al., 1990; Droulez and Berthoz, 1991; Munoz et al., 1991; Fikes and Townsend, 1995).

### Conclusions

Hippocampal spatial representations can be described efficiently in terms of charts and self-localized activity packets. This representation inspires a new concept of an attractor map, which may have broad applications elsewhere. The attractor map concept together with the path integration concept lead to a plausible connectionist model (MPI) of the hippocampal spatial representation system, which uses previously suggested ideas of multiple wired charts (Muller et al., 1991; McNaughton et al., 1996). Numerical simulations of this theory confirm intuitive assumptions about its dynamics. The proposed point of view, although incomplete, leads to a straightforward explanation of many available experimental facts and currently seems to be the most inclusive among alternative theories of hippocampal place-cell dynamics.

### REFERENCES

- Amaral DG, Witter MP (1995) Hippocampal formation. In: The rat nervous system, 2nd edition (Paxinos G, ed), pp 443–493. San Diego: Academic.
- Amaral DG, Ishizuka N, Claiborne B (1990) Neurons, numbers and the hippocampal network. *Prog Brain Res* 83:1–11.
- Amari S (1977) Dynamics of pattern formation in lateral inhibition type neural fields. *Biol Cybern* 27:77–87.
- Amit DJ (1989) Modeling brain function: the world of attractor neural networks. New York: Cambridge UP.
- Amit DJ, Tsodyks M (1991a) Quantitative study of attractor neural

- network retrieving at low spike rates. I. Substrate-spikes, rates and neuronal gain. *Network: Computat Neural Sys* 2:259–274.
- Amit DJ, Tsodyks M (1991b) Effective neurons and attractor neural networks in critical environments. *Network: J Neurosci* 2:275–294.
- Amit DG, Brunel N, Tsodyks M (1994) Correlations of cortical Hebbian reverberations: theory versus experiment. *J Neurosci* 14:6435–6445.
- Andersen RA, Essick GK, Siegel RM (1985) Encoding of spatial location by posterior parietal neurons. *Science* 230:456–458.
- Barnes CA, McNaughton BL, O'Keefe J (1983) Loss of place specificity in hippocampal complex spike cells of senescent rat. *Neurobiol Aging* 4:113–119.
- Blair HT (1996) Simulation of a thalamocortical circuit for computing directional heading in the rat. In: *Advances in neural information processing systems*, Vol 8 (Touretzky DS, Moser MC, Hasselmo ME, eds), pp 152–158. Cambridge, MA: MIT.
- Blair HT, Sharp PE (1995) Anticipatory head-direction cells in anterior thalamus: evidence for a thalamocortical circuit that integrates angular head velocity to compute head direction. *J Neurosci* 15:6260–6270.
- Blum KI, Abbott LF (1995) A model of spatial map formation in the hippocampus of the rat. *Neural Comput* 8:85–93.
- Bostock E, Muller RU, Kubie JL (1991) Experience-dependent modifications of hippocampal place cell firing. *Hippocampus* 1:193–206.
- Chen LL, Lin LH, Barnes CA, McNaughton BL (1994a) Head direction cells in the rat posterior cortex. II. Contributions of visual and idiothetic information to the directional firing. *Exp Brain Res* 101:24–34.
- Chen LL, Lin LH, Green EJ, Barnes CA, McNaughton BL (1994b) Head direction cells in the rat posterior cortex. I. Anatomical distribution and behavioral modulation. *Exp Brain Res* 101:8–23.
- Collett TS (1987) The use of visual landmarks by gerbils: reaching a goal when landmarks are displaced. *J Comp Physiol [A]* 160:109–113.
- Collett TS, Cartwright BA, Smith BA (1986) Landmark learning and visuo-spatial memories in gerbils. *J Comp Physiol [A]* 158:835–851.
- Cugliandolo LF, Tsodyks MV (1994) Capacity of networks with correlated attractors. *J Physics A* 27:741–756.
- Darwin C (1873) Origin of certain instincts. *Nature* 7:417–418.
- Droulez J, Berthoz A (1991) A neural network model of sensoritopic maps with predictive short-term memory properties. *Proc Natl Acad Sci USA* 88:9653–9655.
- Etienne AS (1987) The control of short-distance homing in the golden hamster. In: *Cognitive processes and spatial orientation in animals and man* (Ellen P, Thinus-Blank C, eds), pp 233–251. Boston: Martinus Nijhoff.
- Etienne AS, Maurer R, Séguinot V (1996) Path integration in mammals and its interaction with visual landmarks. *J Exp Biol* 199:201–209.
- Fikes TG, Townsend JT (1995) Moving models of motion forward: explication and a new concept. *Behav Brain Sci* 18:751–753.
- Foster TC, Castro CA, McNaughton BL (1989) Spatial selectivity of rat hippocampal neurons: dependence on preparedness for movement. *Science* 244:1580–1582.
- Fox SE, Ranck Jr JB (1981) Electrophysiological characteristics of hippocampal complex-spike cells and theta cells. *Exp Brain Res* 41:399–410.
- Georgopoulos AP, Lurito JT, Petrides M, Schwartz AB, Massey JT (1989) Mental rotation of the neuronal population vector. *Science* 260:47–52.
- Georgopoulos AP, Taira M, Lukashin A (1993) Cognitive neurophysiology of the motor cortex. *Science* 260:47–52.
- Gothard KM, Skaggs WE, McNaughton BL (1995) Interaction between multiple spatial reference frames in the rat hippocampus. *Soc Neurosci Abstr* 21:941.
- Gothard KM, Skaggs WE, McNaughton BL (1996) Dynamics of mismatch correction in the hippocampal ensemble code for space: interaction between path integration and environmental cues. *J Neurosci* 16:8027–8040.
- Griniasty M, Tsodyks MV, Amit DJ (1993) Conversion of temporal correlations between stimuli to spatial correlations between attractors in neural networks. *Neural Comput* 5:1–17.
- Knierim JJ, Kudrimoti HS, McNaughton BL (1995) Place cells, head direction cells, and the learning of landmark stability. *J Neurosci* 15:1648–1659.
- Knierim JJ, Kudrimoti HS, Skaggs WE, McNaughton BL (1996) The interaction between vestibular cues and visual landmark learning in spatial navigation. In: *Perception, memory, and emotion: frontiers in neuroscience* (Ono T, McNaughton B, Molotchnikoff S, Rolls E, Nishijo H, eds), pp 343–357. New York: Elsevier Science.
- Kubie JL, Ranck JB (1983) Sensory-behavioral correlates in individual hippocampal neurons in three situations: space and context. In: *Neurobiology of the hippocampus* (Seifert W, ed), pp 433–447. New York: Academic.
- Leonhard CL, Stackman RW, Taube JS (1996) Head direction cells recorded from the lateral mammillary nuclei in rats. *Soc Neurosci Abstr* 22:1873.
- Markus EJ, Barnes CA, McNaughton BL, Gladden VL, Skaggs WE (1994a) Spatial information content and reliability of hippocampal CA1 neurons: effects of visual input. *Hippocampus* 4:410–421.
- Markus EJ, Qin Y, McNaughton BL, Barnes CA (1994b) Place fields are affected by behavioral and spatial constraints. *Cognit Neurosci Soc Abstr* 1:91.
- Markus EJ, Qin YL, Leonard B, Skaggs WE, McNaughton BL, Barnes CA (1995) Interactions between location and task affect the spatial and directional firing of hippocampal neurons. *J Neurosci* 15:7079–7094.
- Maurer R, Séguinot V (1995) What is modeling for? A critical review of the models of path integration. *J Theor Biol* 175:457–475.
- McNaughton BL (1996) Cognitive cartography. *Nature* 381:368–369.
- McNaughton BL, Morris RG (1987) Hippocampal synaptic enhancement and information storage within a distributed memory system. *Trends Neurosci* 10:408–415.
- McNaughton BL, Barnes CA, O'Keefe J (1983) The contribution of position, direction, and velocity to single unit activity in the hippocampus of freely moving rats. *Exp Brain Res* 52:41–49.
- McNaughton BL, Leonard BL, Chen LL (1989) Cortico-hippocampal interactions and cognitive mapping: an hypothesis based on reintegration of the parietal and inferotemporal pathways for visual processing. *Psychobiology* 17:230–235.
- McNaughton BL, Chen LL, Markus EJ (1991) “Dead reckoning”, landmark learning, and the sense of direction: a neurophysiological and computational hypothesis. *J Cognit Neurosci* 3:190–202.
- McNaughton BL, Barnes CA, Gerrard JL, Gothard K, Jung MW, Knierim JJ, Kudrimoti H, Qin Y, Skaggs WE, Suster M, Weaver KL (1996) Deciphering the hippocampal polyglot: the hippocampus as a path integration system. *J Exp Biol* 199:173–185.
- Mehta M, Barnes CA, McNaughton BL (1997) Experience-dependent, asymmetric expansion of hippocampal place fields. *Proc Natl Acad Sci USA*, in press.
- Mittelstädt ML, Mittelstädt H (1980) Homing by path integration in a mammal. *Naturwissenschaften* 67:566–567.
- Mizumori SJY, Williams JD (1993) Directionally selective mnemonic properties of neurons in the lateral dorsal nucleus of the thalamus of rats. *J Neurosci* 13:4015–4028.
- Mizumori SJY, McNaughton BL, Barnes CA (1989) A comparison of supramammillary and medial septal influences on hippocampal field potentials and single-unit activity. *J Neurophysiol* 61:15–31.
- Mizumori SJY, Barnes CA, McNaughton BL (1990) Behavioral correlates of theta-on and theta-off cells recorded from hippocampal formation of mature young and aged rats. *Exp Brain Res* 80:365–373.
- Müller M, Wehner R (1988) Path integration in desert ants, *Cataglyphis fortis*. *Proc Natl Acad Sci USA* 85:5287–5290.
- Muller RU, Kubie JL (1987) The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. *J Neurosci* 7:1951–1968.
- Muller RU, Kubie JL, Ranck JB (1987) Spatial firing patterns of hippocampal complex-spike cells in a fixed environment. *J Neurosci* 7:1935–1950.
- Muller RU, Kubie JL, Saypoff R (1991) The hippocampus as a cognitive graph: abridged version. *Hippocampus* 1:243–246.
- Muller RU, Bostok E, Taube JS, Kubie JL (1994) On the directional firing properties of hippocampal place cells. *J Neurosci* 14:7235–7251.
- Muller RU, Stead M, Pach J (1996) The hippocampus as a cognitive graph. *J Gen Physiol* 107:663–694.
- Munoz DP, Pelisson D, Guitton D (1991) Movement of neural activity on the superior colliculus motor map during gaze shifts. *Science* 251:1358–1360.
- O'Keefe J (1979) A review of the hippocampal place cells. *Prog Neurobiol* 13:419–439.
- O'Keefe J, Burgess N (1996) Geometric determinants of the place fields of hippocampal neurones. *Nature* 381:425–428.
- O'Keefe J, Conway DH (1978) Hippocampal place units in the freely moving rat: why they fire where they fire. *Exp Brain Res* 31:573–590.
- O'Keefe J, Dostrovsky J (1971) The hippocampus as a spatial map:

- preliminary evidence from unit activity in the freely moving rat. *Brain Res* 34:171–175.
- O'Keefe J, Nadel L (1978) *The hippocampus as a cognitive map*. New York: Clarendon.
- O'Keefe J, Recce ML (1993) Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus* 3:317–330.
- O'Keefe J, Speakman A (1987) Single unit activity in the rat hippocampus during a spatial memory task. *Exp Brain Res* 68:1–27.
- Quirk GJ, Muller RU, Kubie JL (1990) The firing of the hippocampal place cells in the dark depends on the rat's recent experience. *J Neurosci* 10:2008–2017.
- Quirk GJ, Muller RU, Kubie JL, Ranck JB (1992) The positional firing properties of medial entorhinal neurons: description and comparison with hippocampal place cells. *J Neurosci* 12:1945–1963.
- Ranck Jr JB (1973) Studies of single neurons in dorsal hippocampal formation and septum in unrestrained rats. I. Behavioral correlates and firing repertoires. *Exp Neurol* 41:461–531.
- Ranck Jr JB (1984) Head direction cells in the deep cell layer of dorsal presubiculum in freely moving rats. *Soc Neurosci Abstr* 10:599.
- Ranck Jr JB (1992) Neuronal firing as differentiable manifolds embedded in neuronal firing rate space. *Soc Neurosci Abstr* 18:321.
- Redish AD, Touretzky DS (1997) Cognitive maps beyond the hippocampus. *Hippocampus* 7:15–35.
- Redish AD, Elga AN, Touretzky DS (1997) A coupled attractor model of the rodent head direction system. *Network* 7:671–686.
- Samsonovich A (1997) *Attractor-map theory of the hippocampal representation of space*. PhD thesis, The University of Arizona.
- Samsonovich A, McNaughton BL (1996) Attractor-map-based path integration model of the hippocampus reproduces the phase precession phenomenon. *Soc Neurosci Abstr* 22:1872.
- Samsonovich A, McNaughton BL (1997) A model of formation of the hippocampal spatial code. *Soc Neurosci Abstr*, in press.
- Sharp PE (1997) Subicular cells generate similar spatial firing patterns in two geometrically and visually distinctive environments: comparison with hippocampal place cells. *Behav Brain Res* 85:71–92.
- Sharp PE, Green C (1994) Spatial correlates of firing patterns of single cells in the subiculum of freely moving rat. *J Neurosci* 14:2339–2356.
- Sharp PE, Kubie JL, Muller RU (1990) Firing properties of hippocampal neurons in a visually symmetrical environment: contributions of multiple sensory cues and mnemonic processes. *J Neurosci* 10:3093–3105.
- Shen B, McNaughton BL (1996) Modeling the spontaneous reactivation of experience specific hippocampal cell assemblies during sleep. *Hippocampus* 6:685–692.
- Shepherd GM (1990) *The synaptic organization of the brain*, 3rd edition. New York: Oxford UP.
- Skaggs WE (1995) Relations between the theta rhythm and activity patterns of hippocampal neurons. PhD thesis, The University of Arizona.
- Skaggs WE, McNaughton BL (1996) Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience. *Science* 271:1870–1873.
- Skaggs WE, Knierim JJ, Kudrimoti HS, McNaughton BL (1995) A model of the neural basis of the rat's sense of direction. In: *Advances in neural information processing systems* (Tesauro G, Touretzky D, Leen T, eds), pp 130–180. Cambridge: MIT.
- Skaggs WE, McNaughton BL, Wilson MA, Barnes CA (1996) Theta phase precession in hippocampal neuronal populations and the compression of temporal sequences. *Hippocampus* 6:149–172.
- Sparks DL, Nelson JS (1987) Sensory and motor maps in the mammalian superior colliculus. *Trends Neurosci* 10:312–314.
- Sparks DL, Lee C, Rohrer WH (1990) Population coding of the direction, amplitude and velocity of saccadic eye movements by neurons in the superior colliculus. *Cold Spring Harbor Symposium on Quantitative Biology*, Vol LV, pp 805–811. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Strogatz SH (1994) *Nonlinear dynamics and chaos*. Reading, MA: Addison-Wesley.
- Taube JS (1995a) Head direction cells recorded in the interior thalamic nuclei of freely moving rats. *J Neurosci* 15:70–86.
- Taube JS (1995b) Place cells recorded in the parasubiculum of freely moving rats. *Hippocampus* 5:569–583.
- Taube JS, Burton HL (1995) Head direction cell activity monitored in a novel environment and during a cue conflict situation. *J Neurophysiol* 74:1953–1971.
- Taube JS, Muller RU, Ranck JB (1990) Head direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis. *J Neurosci* 10:420–435.
- Thinus-Blank C, Bouzouba L, Chaix K, Chapuis N, Poucet B (1987) A study of spatial parameters encoded during exploration in hamsters. *J Exp Psychol: Animal Behav Proc* 13:418–427.
- Touretzky DS, Redish AD (1996) A theory of rodent navigation based on interacting representations of space. *Hippocampus* 6:247–270.
- Touretzky DS, Redish AD, Wan HS (1993) Neural representation of space using sinusoidal arrays. *Neural Comp* 5:869–884.
- Tsodyks M, Sejnowski T (1995) Associative memory and hippocampal place cells. *Int J Neural Sys [Suppl]* 6:81–86.
- Tsodyks M, Skaggs WE, Sejnowski TJ, McNaughton BL (1995) On the phase relation between excitatory and inhibitory populations during hippocampal theta rhythm. *Soc Neurosci Abstr* 21:941.
- Whishaw IQ, Vanderwolf CH (1973) Hippocampal EEG and behavior: changes in amplitude and frequency of RSA (theta rhythm) associated with spontaneous and learned movement patterns in rats and cats. *Behav Biol* 8:461–484.
- Wiener SI (1993) Spatial and behavioral correlates of striatal neurons in rats performing a self-initiated navigation task. *J Neurosci* 13:3802–3817.
- Wilson MA, McNaughton BL (1993) Dynamics of the hippocampal ensemble code for space. *Science* 261:1055–1058.
- Worden R (1992) Navigation by fragment fitting: a theory of hippocampal function. *Hippocampus* 2:165–188.
- Zhang K (1996) Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. *J Neurosci* 16:2112–2126.
- Zipser D (1985) A computational model of hippocampal place fields. *Behav Neurosci* 99:1006–1018.
- Zipser D, Andersen RA (1988) A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Nature* 331:679–684.