

Modeling Attractor Deformation in the Rodent Head-Direction System

JEREMY P. GOODRIDGE¹ AND DAVID S. TOURETZKY^{1,2}

¹Center for the Neural Basis of Cognition and ²Computer Science Department, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213-3891

Goodridge, Jeremy P. and David S. Touretzky. Modeling attractor deformation in the rodent head-direction system. *J Neurophysiol* 83: 3402–3410, 2000. We present a model of the head-direction circuit in the rat that improves on earlier models in several respects. First, it provides an account of some of the unique characteristics of head-direction (HD) cell firing in the lateral mammillary nucleus and the anterior thalamus. Second, the model functions without making physiologically unrealistic assumptions. In particular, it implements attractor dynamics in postsubiculum and lateral mammillary nucleus without directionally tuned inhibitory neurons, which have never been observed *in vivo*, and it integrates angular velocity without the use of multiplicative synapses. The model allows us to examine the relationships among three HD areas and various properties of their representations. A surprising result is that certain combinations of purported HD cell properties are mutually incompatible, suggesting that the lateral mammillary nucleus may not be the primary source of head direction input to anterior thalamic HD cells.

INTRODUCTION

Previous empirical research

Previous research has identified several populations of neurons in the rat brain that fire as a function of the animal's head direction (HD). The firing rate of each HD cell is maximal whenever the animal's head is pointed in one particular direction and tapers off as the animal points its head away from that direction. Different HD cells have different preferred directions so that the entire 360° space is uniformly sampled by the population of HD cells. HD cells have been identified in a number of different brain areas, including the postsubiculum (PoS) (Ranck 1984; Taube et al. 1990), anterior dorsal thalamus (AD) (Taube 1995), and lateral mammillary nucleus (LMN) (Stackman and Taube 1998), although not all cells in these regions fire in a direction-specific manner. Furthermore the particular firing characteristics of HD cells vary from region to region.

PoS head direction cells have simple Gaussian-shaped tuning curves. The firing of these cells is best correlated with the animal's head direction approximately 10 ms in the past, on average (Blair and Sharp 1995; Taube and Muller 1998). In contrast, the activity of the average HD cell in AD is best correlated with the animal's heading approximately 25 ms in the future. In other words, the average AD HD cell has an anticipatory time interval (ATI) of +25 ms. The ATI value for

a single cell appears to be constant over a broad range of angular head velocities.

According to Blair, Lipscomb, and Sharp (1997), the shapes of AD HD cell tuning curves exhibit distortions as a function of the speed and direction of the animal's turning. Furthermore when the animal is still, the curves often have a bimodal shape (two peaks), whereas when the animal is turning, AD HD cell tuning curves are unimodal but skewed in the opposite direction of the turn. The peak firing rate also increases with angular head velocity (Taube 1995).

The extent to which a particular AD HD cell shows this tuning curve distortion depends on how much the cell anticipates the animal's head direction. The greater the ATI, the greater the extent of tuning curve distortion. In addition, the width of an AD HD cell's tuning curve when determined across all velocity levels is positively correlated with ATI (Blair et al. 1997). An accurate model of the HD system must not only capture the average tuning curve differences between AD and PoS, it must also account for the *distribution* of tuning curve characteristics within the AD and PoS populations.

The phenomenon of AD tuning curve distortion has been called into question by Taube and Muller (1998), who did not see this effect in their own experiments. They also did not report bimodality when the animal was still. In this paper, we will be basing our modeling efforts on the Blair et al. (1997) findings. The Taube and Muller findings were based on much shorter recording sessions: only 8 min as compared with 15–30 min in (Blair et al. 1997) and 15–90 min in (Blair and Sharp 1998). Thus because the tuning curve effects reported by Blair et al. are quite subtle, they may very well have been obscured in the 8-min sessions.

Data on LMN head-direction cells have only recently become available and are also somewhat contradictory. Stackman and Taube (1998) describe LMN cells as having Gaussian tuning curves modulated by angular velocity. Cells that prefer clockwise turns increase their firing rate as angular velocity increases in the clockwise direction and decrease their firing rate as angular velocity increases in the counterclockwise direction relative to the baseline rate observed when the animal's head is not turning at all. The observed ATI for LMN cells was approximately 95 ms. But Blair and Sharp (1998) report that LMN HD cells have an ATI of 40 ms, and while their peak firing rate increases with velocity, in addition, the width of the tuning curve contracts as angular velocity increases in the direction of the side containing the cell e.g., counterclockwise turns for left LMN cells or clockwise turns for right LMN cells. The width stays the same for turns in the opposite direction. In

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

this paper, we will be reporting on simulations that address the compatibility of these different LMN findings in a model in which LMN is the primary projection to AD.

Redish, Elga, and Touretzky (1996) developed a model that accounted for some of the differences between AD and PoS HD cell firing. Specifically, their model provided an explanation for why AD HD cell firing was best correlated with the animal's future head direction, whereas PoS HD cell firing was best correlated with current head direction. The Redish et al. model also predicted that AD tuning curves would exhibit distortions as a function of the animal's turn velocity, which was subsequently confirmed by Blair et al. (1997). However, the model did not account for certain other important differences.

In this paper, we present a new model that is the successor to the one described by Redish et al. (1996). Our model accounts for a number of important new results not covered in the previous model, such as the bimodality of AD tuning curves, the constancy of ATI for all angular velocities, the particular distorted shapes exhibited by AD HD cells as velocity increases, and some of the tuning curve changes exhibited by LMN HD cells. Furthermore unlike the model in Redish et al. (1996), our new model does not rely on directionally tuned inhibitory neurons or multiplicative synapses.

Preliminary data concerning some of the results presented in this paper have been published previously (Goodridge et al. 1997).

Previous modeling research

Many findings (Blair and Sharp 1996; McNaughton et al. 1991; Stackman and Taube 1997; Taube and Burton 1995) suggest that HD cells rely on internal sources of information to obtain information about ongoing changes in the animal's heading. Vestibular, proprioceptive, and motor efference copy are all possible sources of self-motion cues. It appears that HD cell activity is updated by integrating angular velocity signals from one or more of these sources.

A number of previous models have been proposed to account for the integration of angular velocity by HD cells (Blair 1996; McNaughton et al. 1991; Redish et al. 1996; Skaggs et al. 1995; Zhang 1996a). Most of these models are based on the *attractor hypothesis* first put forth by Skaggs et al. (1995) that the head-direction system is a circular one-dimensional dynamical system or ring attractor, which integrates angular velocity by moving an activation bump around the ring.

Attractor models postulate the existence of an extensive network of interconnection between HD cells such that cells with similar preferred directions excite more than inhibit one another, whereas cells with dissimilar preferred directions inhibit more than excite one another. This scheme of interconnectivity produces a situation in which a bump or activation hill (the "attractor state") will arise even when the initial activation levels of the cells are random. The hill is then self-maintaining in the absence of external input. The activation hill is made to move around the ring, representing the animal's estimated heading as it turns, by supplying additional input to units on one flank of the hill. The amount of the input supplied to a flank determines the speed at which the hill moves. If the strength of this input corresponds to the animal's angular velocity, then movement of the activation hill will

track the animal's head direction. Simulations showed that this mechanism is sufficient to allow the HD system to accurately track complex head movement profiles taken from a real rat (Redish et al. 1996).

In the Redish et al. model, the integration process was implemented by coupling two attractor networks together, one corresponding to AD and the other to PoS. As described in the preceding text, AD contains an HD signal that is best correlated with the animal's future head direction (Blair and Sharp 1995; Blair et al. 1997; Taube and Muller 1998). In contrast, PoS contains an HD signal that best corresponds to the animal's present or recent past head direction. To account for this difference between PoS and AD, Redish et al. (1996) proposed that there are asymmetric connections between AD and PoS. In particular, they proposed that PoS HD cells project to AD HD cells with slightly offset preferred directions, whereas AD HD cells project to PoS HD cells with matching preferred directions. This arrangement is plausible because the anatomical data suggest that AD and PoS are reciprocally connected but do not reveal which cells in PoS project to which cells in AD. In the Redish et al. model, when a left turn was being simulated, PoS projections to AD HD cells with a preferred direction slightly offset to the left ("left offset" projections) were enabled, whereas right offset projections were disabled. The converse was true when right turns were being simulated. As a result of this scheme of interaction, AD contained a hill of activation that slightly anticipated the position of the hill in PoS.

Although the Redish et al. model did account for the temporal difference between AD and PoS HD cell firing, it did not account for the distorted tuning curves in AD. Redish et al. did show that removing the attractor dynamics from AD resulted in distorted tuning curves; however, these distortions were not the same as the ones that AD HD cells actually exhibit. In addition, after removal of attractor dynamics, the amount by which AD HD cell firing anticipated PoS HD cell firing declined with angular velocity, which is not consistent with the known data (Taube and Muller 1998). Nevertheless, as noted by Redish et al., there is good anatomical evidence that the AD does not contain an attractor network. In particular, AD lacks the GABA-containing interneurons (Bentivoglio et al. 1993) that would provide the necessary inhibition for an attractor network. So in our new model, AD does not contain any attractor dynamics, and its connectivity to other structures has been altered.

Blair et al. (1997) proposed their own account of how the distorted tuning curves in AD were generated. They suggested that each AD HD cell received input from two populations of turn-modulated HD cells. One of these populations increased its firing rate when the animal turned clockwise, whereas the other increased its rate for counterclockwise turns. In their scheme, (counter-) clockwise turn-modulated cells projected to AD HD cells with preferred directions slightly (counter-) clockwise to that of the projecting cell. They postulated that AD HD cells that anticipated head direction by a large amount received input from turn-modulated head-direction cells with more highly offset connections and that this scheme would produce the kinds of tuning distortions that AD HD cells exhibited in real rats. Although not intended to account for distortions in AD HD tuning curves, a similar proposal was made by Zhang (1996b). Whereas in the Redish et al. model,

AD tuning curves were a function of only one set of offset connections at a time, in the Blair and Sharp model, AD tuning curves were a function of two sets of simultaneously active offset connections.

In this paper, we present simulations of a model in which the input to AD is as proposed by Blair and Sharp. We show that the model is capable of accurately accounting for AD and PoS tuning curve shapes, that it integrates angular velocity with good accuracy, and that it also produces a constant ATI in AD. The source of input to AD is another attractor module that exhibits the velocity modulated tuning curves described by Stackman and Taube (1998) with respect to LMN. However, our simulations show that the ATI values reported for LMN by Stackman and Taube and by Blair and Sharp (1998) are not compatible with the AD tuning curve properties described by Blair et al. (1997) if it is LMN alone that drives AD.

METHODS

New model of the HD system: overall structure

The organization of our model is shown in Fig. 1. A bump of activation whose location represented the animal's current head direction was maintained in PoS by means of attractor dynamics. This signal was passed to two populations of LMN cells (clockwise: CW and counterclockwise: CCW), where it was modulated by angular velocity information represented in the model by two abstract "angular velocity" units. The two LMN populations projected to AD with opposite offsets, producing bimodal tuning curves. Thus AD activity was purely a function of the input received from LMN, a property which allowed AD to show dramatic shape distortions. AD cells projected to PoS cells with matching preferred directions, thereby updating the position of the PoS attractor bump. This connection scheme is consistent with the known anatomy (Allen and Hopkins 1989; Shibata 1992; van Groen and Wyss 1990, 1995), but it does not include all the connections known to exist for these areas.

Neuronal model

The elements of our model are nonlinear units with continuous-valued outputs in $[0, 1]$. Each unit represents a subpopulation of "real" neurons, which we assume are firing asynchronously. Furthermore the excitatory neurons in the i th subpopulation are assumed to all have preferred direction ϕ_i . (Inhibitory units, representing populations of

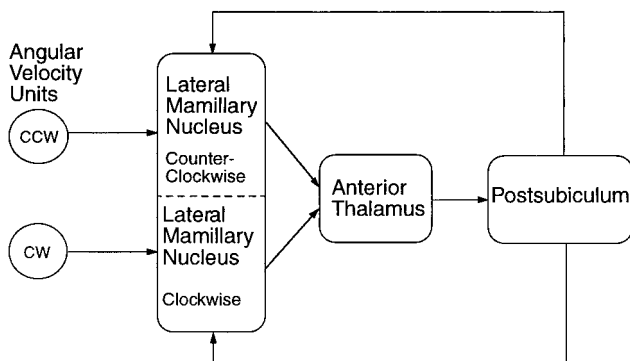


FIG. 1. Connections between structures in the model. The firing rate of the counterclockwise (CCW) angular velocity unit increased monotonically from the baseline (still) value for counterclockwise turns of increasing velocity and decreased monotonically for clockwise turns. The clockwise (CW) angular velocity unit had complementary behavior. The postsubiculum (PoS) module and both pools of the lateral mammillary nucleus module contained recurrent connectivity.

TABLE 1. *Individual unit parameters*

| | |
|--|-------|
| Number of excitatory units/pool (N_E) | 101 |
| Number of inhibitory units/pool (N_I) | 1 |
| Integration time step (Δt), ms | 0.1 |
| Excitatory unit time constant (τ_E), ms | 1.0 |
| Inhibitory unit time constant (τ_I), ms | 0.2 |
| PoS excitatory unit tonic inhibition (γ_E^{PoS}) | -1.5 |
| AD excitatory unit tonic inhibition (γ_E^{AD}) | -1.75 |
| LMN excitatory unit tonic inhibition (γ_E^{LMN}) | -3.95 |
| PoS and LMN inhibitory unit tonic inhibition (γ_I) | -7.5 |
| PoS and AD gain ($g_{\text{PoS}}, g_{\text{AD}}$) | 2 |
| LMN gain (g_{LMN}) | 1 |

PoS, postsubiculum; AD, anterior dorsal thalamus; LMN, lateral mammillary nucleus.

inhibitory interneurons, are nondirectional.) The output, or "firing rate," of the i th unit, $F_i(t)$, can be regarded as the fraction of neurons in the subpopulation that were spiking at time t or as the probability that an individual neuron emitted a spike at time t . All structures in the model used the same type of unit, and the preferred directions ϕ_i were uniformly distributed around the circle. The following three equations, modified from Pinto, Brumberg, Simons, and Ermentrout (1996), determined the activity of unit i

$$V_i(t) = \gamma + E_i(t) \quad (1)$$

$$F_i(t) = \frac{1}{1 + \exp[-gV_i(t)]} \quad (2)$$

$$\tau \frac{dS_i(t)}{dt} = -S_i(t) + F_i(t) \quad (3)$$

$V_i(t)$ was the net activation of unit i (or the average membrane voltage of neurons in subpopulation i) at time t . γ was a tonic inhibitory term, and $E_i(t)$ was the weighted input (synaptic drive times coupling strength) received from other units: both external input from other modules and recurrent input from units in the same module. The unit's firing rate, $F_i(t)$, was a sigmoid function of its average membrane voltage. The slope of the sigmoid was determined by the gain parameter, g . The synaptic drive that the unit delivered at time t is denoted $S_i(t)$. Synaptic drive is a measure of the influence of a unit on other units to which it projects. Synaptic drive varied with firing rate, but it changed more slowly and decayed exponentially as governed by the time constant τ in the differential equation. This simulated some of the effects of synaptic delay and temporal integration in real neurons. We kept the excitatory and inhibitory unit time constants short (1.0 and 0.2 ms, respectively) to allow the model to accurately integrate turns at very high speeds, up to 700°/s. We integrated Eq. 3 using a time step Δt of 0.1 ms

$$S_i(t + \Delta t) = S_i(t) + [-S_i(t) + F_i(t)] \frac{\Delta t}{\tau} \quad (4)$$

The units in the LMN, PoS, and AD modules all had identical synaptic drive (S_i) and firing rate (F_i) functions. Within a module, inhibitory and excitatory units used the same gain value. The three modules differed only in their tonic inhibition and gain parameters (γ and g) and their connectivities, which determined their inputs $E_i(t)$. The basic parameters for a unit are shown in Table 1. The equations for the individual voltage functions $V_i(t)$ will be described in the sections that follow.

Model of PoS

Our model of PoS used a modified form of the Wilson-Cowan equations (Wilson and Cowan 1972) to generate an attractor bump (Pinto et al. 1996). The PoS module contained both excitatory and inhibitory units, denoted by superscripts in the following equations.

The recurrent connectivity within PoS was structured to allow a hill of activation to form regardless of each unit's initial activity level. The projection strength between any pair of excitatory units was a Gaussian function of the difference in their preferred directions. There was only one inhibitory unit; it received input from all of the excitatory units and projected back to all of them with equal strength as shown in Fig. 2. In the rat there are many inhibitory interneurons, but if each provides input to a different random subset of the excitatory population independent of their preferred directions, the net effect would be the same as having a single fully-connected inhibitory unit.

The activation equation for the i th excitatory PoS unit was

$$V_i^{\text{E-PoS}}(t) = \gamma_E^{\text{PoS}} + \kappa_{\text{AP}} \cdot S_i^{\text{E-AD}}(t) + \kappa_{\text{IE}} \cdot S^{\text{I-PoS}}(t) + \kappa_{\text{EE}} \sum_j w_{ij} S_j^{\text{E-PoS}}(t) \quad (5)$$

External input was provided by the AD unit with matching preferred direction ϕ_i . κ_{IE} and κ_{EE} were coupling constants determining the strength of input received from the PoS inhibitory unit and from other PoS excitatory units, respectively. $S^{\text{I-PoS}}(t)$ was the synaptic drive of the PoS inhibitory unit, and $S_j^{\text{E-PoS}}(t)$ the drive of the j th PoS excitatory unit. The weight w_{ij} from unit j to unit i was a Gaussian function of the difference in the units' preferred directions. Note that because this difference was a circular variable, its magnitude was bounded between 0 and 180°. The standard deviation σ controlled the width of the Gaussian, which determined (indirectly) the width of the attractor bump

$$w_{ij} = \exp\left(\frac{-[\phi_i - \phi_j]^2}{\sigma^2}\right) \quad (6)$$

The PoS inhibitory unit received no external input. It received inhibitory input from itself, and excitatory input from the excitatory PoS population, governed by the coupling constants κ_{II} and κ_{EI} , respectively

$$V^{\text{I-PoS}}(t) = \gamma_I + \kappa_{\text{II}} \cdot S^{\text{I-PoS}}(t) + \kappa_{\text{EI}} \sum_j S_j^{\text{E-PoS}}(t) \quad (7)$$

The parameters for the PoS attractor module are shown in Table 2. These values were obtained by experimenting with the simulation until an acceptable bump shape was produced. The values are fairly robust in the sense that small changes in parameter values do not produce drastic changes in model behavior.

Model of LMN

Our model of LMN consisted of two pools of units, LMN(cw) and LMN(ccw), with a preference for CW and CCW turns, respectively.

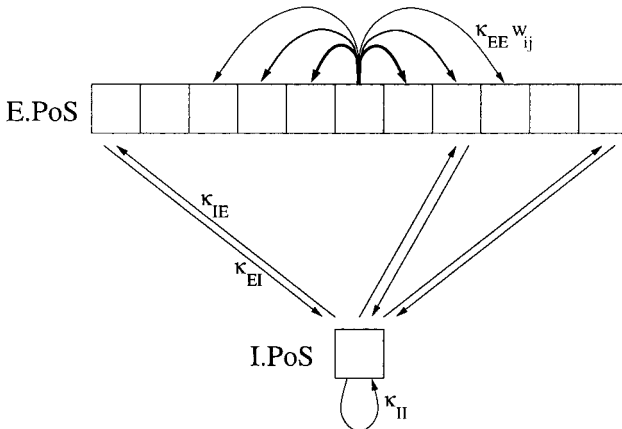


FIG. 2. Connections within the postsubiculum module, illustrating the structure of an attractor network. E.PoS: excitatory units; I.PoS: inhibitory unit.

TABLE 2. Parameters for unit interaction within the attractor modules, PoS and LMN

| | |
|---|------|
| Std. deviation (σ) for recurrent excitation w_{ij} , degrees | 28.5 |
| I \rightarrow E coupling strength (κ_{IE}) | -8.0 |
| E \rightarrow E coupling strength (κ_{EE}) | 6.0 |
| I \rightarrow I coupling strength (κ_{II}) | -4.0 |
| E \rightarrow I coupling strength (κ_{EI}) | 0.5 |

Each pool contained a recurrently connected excitatory population plus a single inhibitory unit. The parameters used for unit interaction within LMN(cw) and LMN(ccw) were identical to those used for the PoS attractor. In addition, the excitatory units in each LMN pool received external input from one of two angular velocity units as shown in Fig. 1. The CW-sensitive angular velocity unit, which projected to units in LMN(cw), fired maximally during a CW high speed turn; it fired minimally during a CCW high speed turn. The CCW-sensitive angular velocity unit, which projected to units in LMN(ccw), responded in the opposite manner.

These angular velocity units have no specific anatomical correlate in the rat; they are merely a conceptual device for introducing angular velocity information into LMN. While cells tuned to angular velocity have been reported in LMN (a different population than the LMN HD cells), those cells did not discriminate between CW and CCW turns (Stackman and Taube 1998). LMN HD cells might receive angular velocity information from the dorsal tegmental nucleus, which is known to project to LMN and to receive projections from the medial vestibular nucleus and nucleus prepositus hypoglossi (Blair and Sharp 1998; Stackman and Taube 1998). Those nuclei contain angular velocity cells that *are* sensitive to turn direction (Blair and Sharp 1998). Direction-sensitive angular velocity cells have also been reported in PoS (Sharp 1996), which might contribute to the PoS projection to LMN.

In agreement with the observations of Stackman and Taube (1998), the firing rates of our LMN units were modulated by the angular velocity signal, increasing for turns in one direction and decreasing for turns in the opposite direction. To achieve this without distorting the shape of the tuning curve, we relied on a result described by Salinas and Abbott (1996), showing that an additive input supplied to all the units in a one-dimensional attractor bump can produce a multiplicative effect on the bump. The activation equation for the i th excitatory LMN(cw) unit is shown in the following text. A similar equation governed LMN(ccw) units

$$V_i^{\text{E-LMN(cw)}}(t) = \gamma_E^{\text{LMN}} + \kappa_{\text{PL}} \cdot S_i^{\text{E-PoS}}(t) + \rho A^{\text{cw}}(t) + \kappa_{\text{IE}} \cdot S^{\text{I-LMN(cw)}}(t) + \kappa_{\text{EE}} \sum_j w_{ij} S_j^{\text{E-LMN(cw)}}(t) \quad (8)$$

The i th excitatory LMN unit's external input was provided by $S_i^{\text{E-PoS}}$, i.e., the synaptic drive of the PoS unit with matching preferred direction ϕ_i . All LMN excitatory units also received a common input $\rho A^{\text{cw}}(t)$ or $\rho A^{\text{ccw}}(t)$ from the angular velocity unit for their respective pool; ρ is a scale factor. Note that the angular velocity term is unsubscripted because it is a global input to all units in the pool, independent of preferred direction. $A^{\text{cw}}(t)$ ranged from 0 to 1, and $A^{\text{ccw}}(t)$ was always equal to $1 - A^{\text{cw}}(t)$. As a result, the LMN(cw) pool received the complement of the angular velocity signal received by the LMN(ccw) pool. The specific values of this term for different angular velocities were determined experimentally, so that the model could accurately simulate head turns at speeds ranging from 700°/s CCW to 700°/s CW, in 100°/s increments. The tonic inhibition γ_E^{LMN} was set to $\gamma_E^{\text{PoS}} - \rho/2$ to compensate for the mean added excitation from the angular velocity signal, and the gain g_{LMN} was set to half g_{PoS} because LMN receives strong inputs from both its recurrent projections and the PoS projection.

As with PoS, the weight from LMN excitatory unit j to excitatory unit i within the same pool was a Gaussian function of the difference

TABLE 3. *Parameters for module interaction*

| | |
|---|--------------|
| AD → PoS coupling strength (κ_{AP}) | 0.22 |
| PoS → LMN coupling strength (κ_{PL}) | 1.0 |
| LMN → AD coupling strength (κ_{LA}) | 2.5 |
| LMN → AD offset amount (δ), degrees | 27 (nominal) |
| Strength of angular velocity input (ρ) | 4.9 |

in their preferred directions. The same weight matrix w_{ij} was used as for PoS. Also as in PoS, the inhibitory unit in each LMN pool received a projection from each of the excitatory units in its pool, governed by the coupling constant κ_{EI} , and an inhibitory projection from itself governed by the constant κ_{II} . The inhibitory unit received no input from units external to LMN. Parameter values are shown in Tables 2 and 3.

Model of AD

Because AD contained no attractor dynamics, the activity of AD units was entirely dependent on external input. This input consisted of the offset projections from the two LMN pools. The i th unit in AD, with preferred direction ϕ_i , received projections from LMN units with preferred directions $\phi_i \pm \delta$. For example, if the offset amount was 15° , an AD unit with a preferred direction of 90° would receive input from an LMN(cw) unit with a preferred direction of 75° and an LMN(ccw) unit with a preferred direction of 105° . As described in RESULTS, the size of this offset was important in determining both the tuning curve shape exhibited by AD units and the degree of anticipation they exhibit.

AD units were governed by the following equation

$$V_i^{\text{AD}}(t) = \gamma_E^{\text{AD}} + \kappa_{LA} \cdot S_j^{\text{E-LMN(cw)}}(t) + \kappa_{LA} \cdot S_k^{\text{E-LMN(ccw)}}(t) \quad (9)$$

where $j = i - \delta \cdot N_E / 360^\circ$ (modulo N_E), $k = i + \delta \cdot N_E / 360^\circ$ (modulo N_E), δ is the offset angle in degrees, and N_E is the number of excitatory units in an LMN pool. γ_E^{AD} was made slightly larger than γ_E^{PoS} because AD has no inhibitory inputs, but it might be possible to have all three modules use the same γ_E value by adjusting other parameters.

RESULTS

Most of the parameters of the model were kept fixed for all simulations reported here. In particular, the tonic inhibition values γ_E and γ_I and time constants τ_E and τ_I for all units, the weight matrix w_{ij} used in PoS and LMN, the number of units in each structure, and the coupling strengths among the AD, PoS, and LMN modules (κ_{AP} , κ_{PL} , κ_{LA}) were all kept constant. We varied δ , the extent of the offset in the connections from LMN to AD, to fit the tuning curves of several real AD cells. We also varied the strength of the angular velocity input $A^{\text{cw}}(t)$ to simulate turns at various speeds. The effects of varying these parameters are described in the following text.

At the beginning of a simulation, we set the activation levels of units in PoS and LMN to random values. Then at each time step of the simulation, all units updated their activation levels. Soon a stable activation hill formed over the PoS, LMN, and AD populations. This settling process required 10–50 ms of simulation time.

When the activation hills stabilized, we experimentally adjusted the model so that a continuous turn at particular fixed angular velocity values would cause the expected change in the position of the PoS activation hill. We did this by varying the angular velocity input $A^{\text{cw}}(t)$ to LMN until it produced the appropriate speed of activation hill movement. We tuned the

model parameters to accurately integrate angular velocities from 0 to $700^\circ/\text{s}$ in steps of $100^\circ/\text{s}$, both CW and CCW, while maintaining desired tuning curve shapes.

Once this tuning process was complete, the behavior of the model in response to velocity changes could be studied. In particular, we assessed the integration accuracy of the model, the shapes of the tuning curves in all three modules, and the degree of anticipation in AD as a function of angular velocity.

Integrating a real angular velocity profile

To examine whether the network was capable of integrating a real angular velocity profile, we used data obtained from Blair and Sharp (personal communication). This was the same data as were used in Redish et al. (1996). The data consisted of a rat head-direction trajectory lasting 12 s, sampled at 60 Hz. To compute the angular velocity of the animal during each sample, we grouped each sample with the two samples before and two samples after. Then we defined the angular velocity of the sample as the slope of the best fit line through these five points. Because the real data were sampled at 60 Hz, each sample accounted for approximately 16 ms of time. However, the simulations were performed with a time step of 0.1 ms, so we used linear interpolation to generate an angular velocity profile with all the intermediate values required for the model.

We defined the speed at which the model was turning as the rate of change of the location of the PoS activation hill. Location was defined as the weighted mean of the ϕ_i values, where the weights were the firing rates F_i . In other words, we calculated the population vector (Georgopoulos et al. 1983) of the PoS population. Assigning a location to the PoS activation hill is more straightforward than for AD since AD profiles have distorted shapes. The tuning curves in the PoS module did not exhibit the shape distortions that were present in AD even though PoS received direct projections from AD. The primary reason for this was the recurrent connections in PoS, which tended to minimize the effects of distortions in the input. The amplitude of the PoS attractor bump also remained fairly constant, making PoS cells insensitive to angular velocity.

When we exposed the model to the complex angular velocity trajectory taken from a real rat, we found it maintained heading with good accuracy, comparable with the Redish et al. model.

Shapes of AD tuning curves

According to the proposal of Blair et al. (1997), one primary source of differences among AD HD cells with identical preferred directions is the amount of offset in the projections they receive from angular velocity-tuned HD cells. This proposal suggests that for each AD cell there is a particular offset value that determines both the cell's tuning curve shape and the extent to which it anticipates head direction.

To examine the validity of this hypothesis, we chose the HD cell from Blair and Sharp's population that showed the highest degree of tuning curve distortion. Then we varied the level of offset in the projection between LMN and AD until we were able to best fit the tuning curve of that AD HD cell. Figure 3 shows the activation hills produced in our simulated AD in relation to the tuning curves of the actual AD HD cell.

As can be seen, the simulation was able to fit this particular HD cell quite well. The δ value used to generate the simulated

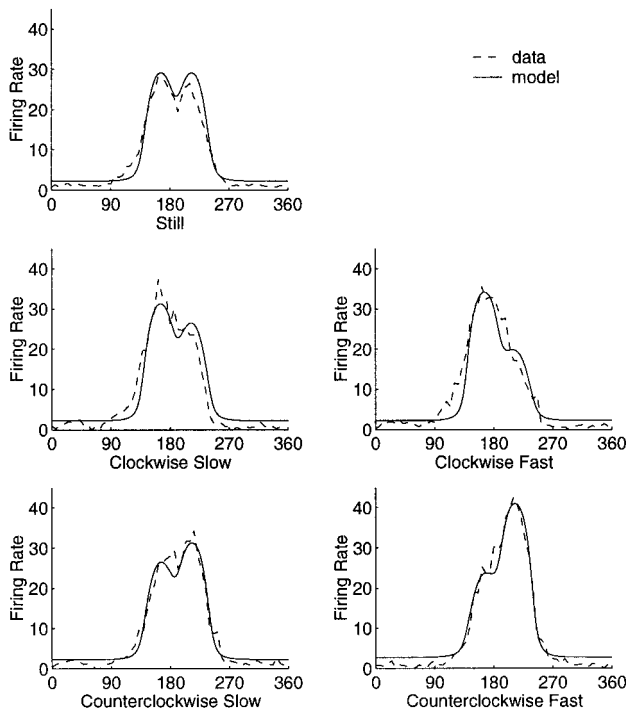


FIG. 3. Real and simulated anterior dorsal thalamus (AD) tuning curves at different angular velocities. For the real curves, Blair and Sharp defined fast turns as exceeding $270^\circ/\text{s}$, slow turns as $30\text{--}270^\circ/\text{s}$, and speeds below $30^\circ/\text{s}$ as “still.” Simulated velocities required to match the real curves were $101^\circ/\text{s}$ for slow turns and $330^\circ/\text{s}$ for fast turns. Model unit “firing rates” (values between 0 and 1) were multiplied by 50 to convert to equivalent spike rates in the graph, except for the fast CCW case where a multiplication factor of 60 gave a better fit.

curves was 27° . It should be noted that this cell shows a very visible degree of tuning curve distortion. We also performed simulations to match another HD cell and found a similar quality of fit. To fit the tuning curves of this new cell, we only had to change the extent of offset δ in the projection from LMN to AD.

To measure anticipatory firing in AD, we arbitrarily designated the position of the attractor hill in the PoS module as the simulated animal’s actual head direction. Thus the ATI for PoS was always 0, and the ATI for AD was measured in relation to this reference. To measure the ATI at any time step in the simulation, we computed the position of the AD activation hill and subtracted the position of the PoS activation hill from this value. This difference represented the actual offset in degrees between the AD and PoS activation hills. The ATI was then computed by dividing this difference by the current angular velocity of the PoS activation hill.

For the simulated cell shown in Fig. 3, the calculated ATI was 28 ms. This value was lower than that of the real cell whose tuning curve we had fit (47 ms), a result that suggests that other mechanisms besides the one proposed here are acting to increase the level of anticipation in AD. In our model, one reason a higher level of anticipation could not be achieved with this offset was the existence of transmission time delays from PoS to LMN and then from LMN to AD. The delays resulted from the nonzero time constants of the individual units. The synaptic drive of a unit (Eq. 3) varied as a function of net input plus the unit’s current drive value at a rate determined by the time constant τ_E . As a result of this delay, the maximum

achievable offset in degrees between AD and PoS hills was less than the value of δ . For the HD cell modeled in Fig. 3, δ was 27° , but simulations showed that even at high velocities, when AD was receiving virtually all its input from the one LMN activation hill projecting ahead in time, the actual offset between AD and PoS hills was only 24° . Thus despite the offset from LMN to AD, the AD activation hill was not as advanced relative to LMN as it would have been in a model with instantaneous-time units, i.e., with τ_E close to zero.

Our simulations showed there was indeed a positive relationship between the offset parameter δ and the extent of anticipation in AD. The result suggests that the degree of offset is a critical factor in explaining tuning curve variability of AD HD cells. Therefore according to this model, the reason that many HD cells do not show a visible degree of distortion is that the δ value is too small. With a very small value for δ , there was almost no visible tuning curve distortion because the two activation hills from LMN overlapped almost completely. Nevertheless even in these cases, there was still a measurable ATI.

Another feature of AD tuning curve shapes is that their peak amplitude rises with increasing velocity (Taube 1995). That result is also consistent with the AD tuning curves generated by this model.

Constancy of anticipation in AD

Taube and Muller (1998) reported that the extent to which the signal in AD anticipates the signal in PoS does not vary with angular velocity. We were interested in determining whether our model exhibited similar behavior. To assess this in the current model, we conducted simulations at a range of different angular velocities. For each angular velocity, we computed the ATI in AD relative to PoS. Figure 4 plots the relationship between angular velocity and temporal difference between AD and PoS for real HD cells and for the model.

As can be seen from the graph, the model exhibited a nearly constant ATI as a function of angular velocity. However, it is also important to note that there is a great deal of variability in the real HD data, which makes it difficult to determine how constant the anticipation of real AD HD cells actually is.

Constancy of anticipation means that the difference in position of AD versus PoS activation hills varies linearly with the velocity at which the activation hills are moving. In order for a linear relationship to be preserved, we found it is critical that

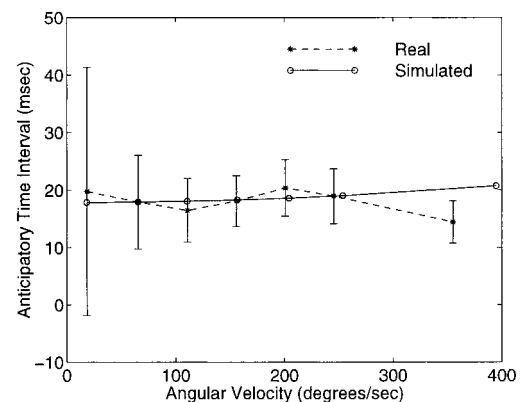


FIG. 4. Comparison of the anticipatory time interval for model anterior thalamic cells vs. real anterior thalamic cells (real data courtesy of Jeffrey S. Taube).

there be conservation of strength in the projections from LMN to AD. If during a turn the increase in total activity of CW LMN units is not exactly balanced by a decrease in activity for CCW LMN units, then AD units will receive more total external input during high-speed turns than during low-speed turns. Depending on how this extra input is distributed, the change in the shape of the AD activation hill will produce an effect on the speed of the PoS hill *above and beyond* the effect produced by the difference in position between AD and PoS activation hills. As a result, there will not be a linear relationship between the difference in position of AD versus PoS activation hills and the velocity at which the activation hills are moving, which means ATI will not be constant. Thus the model suggests that equal and opposite changes in the two LMN populations' projections to AD are critical for the stability of AD anticipation across different angular velocities.

Does LMN drive AD?

All of our results for AD were obtained by driving AD with nonanticipatory HD signals offset by a constant amount $\pm \delta$. In the model, AD was driven by a velocity-modulated signal from LMN, which was in turn driven by PoS, giving LMN an ATI value of -3 ms relative to PoS due to transmission delay. But observed ATI values for LMN range from 40 to 95 ms (Blair and Sharp 1998; Stackman and Taube 1998). This raises the question of whether the data on AD and LMN response properties are compatible with the hypothesis that LMN drives AD. Although bilateral LMN lesions eliminate the AD head direction signal (Blair et al. 1999; Tullman and Taube 1998), this does not prove that AD activity is principally derived directly from the LMN projection.

To investigate this hypothesis, we modified the model to force LMN to anticipate PoS by a modest amount by shifting the PoS input to LMN CW or CCW as a function of angular velocity. This produced ATI values of 31–46 ms in LMN. (The variation is due to the fact that the PoS activity vector was always shifted by a whole number of units, i.e., in steps of 3.56° since there were 101 units per module.) We found that LMN anticipation combined additively with the effect of offset connections from LMN to AD, producing an ATI of 60–80 ms in AD. Thus AD anticipated LMN, when the reverse should be true. In addition, the shapes of the AD tuning curves no longer fit the data of Blair and Sharp, as will be discussed in the following text.

Why does PoS in the model have a smaller ATI than AD, which drives it, but AD does not have a smaller ATI than LMN? The AD projection to PoS does not involve offset connections. But more importantly, PoS is an attractor module: input from AD is only one influence on PoS unit activity. Recurrent connections in PoS have significant influence by stabilizing and maintaining the bump that the AD input perturbs. In AD, which lacks recurrent connections, unit activity is determined solely by the offset LMN projections, with the offset amount δ determined by the distance between the two peaks in the bimodal still tuning curve. We produced anticipation in AD by increasing the height of one of the peaks during turns while decreasing the height of the other. The *peak* locations do not actually shift with velocity (Fig. 5A), but since their relative amplitudes change, the *recovered direction* in AD, calculated using the population vector, shifts toward the

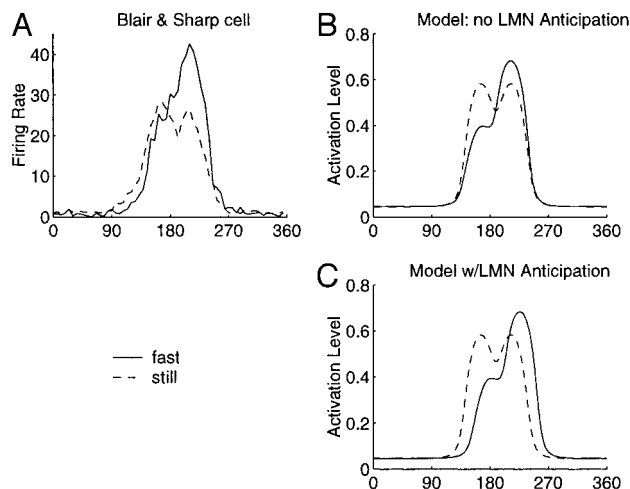


FIG. 5. Effect of anticipation on tuning curve shape. *A*: comparison of fast CCW curve and still curve from the same cell as Fig. 3: the fast peak is at the same location as the right peak of the bimodal still curve, and the flanks are also similar. *B*: comparison of model's fast and still tuning curves without lateral mammillary nucleus (LMN) anticipation shows that both the peak and the edges are aligned. *C*: with anticipation in LMN, the fast AD curve shifts relative to the still curve; peaks and edges fail to align.

rising peak, producing an anticipatory effect. For the curves in Fig. 3, the recovered directions were 173° (fast CW), 176° (slow CW), 184° (still), 192° (slow CCW), and 198° (fast CCW).

The anticipatory LMN simulation did not produce correct AD tuning curves because the fast tuning curve was shifted in position, not just height, relative to the still curve. Figure 5A compares fast and still tuning curves for the cell shown in Fig. 3. Note that the peak of the fast CCW curve is at the same location as the right peak of the still curve. Figure 5B shows that the model correctly reproduces this relationship when there is no LMN anticipation. But if LMN anticipates PoS, the AD tuning curve shifts in the direction opposite the turn, as shown in Fig. 5C. Hence the AD curve is not correct since the fast peak does not coincide with the still peak.

To further investigate the effect of LMN anticipation on AD, we constructed a separate linear network model to determine whether there was *any* pattern of weights between LMN and AD that would allow LMN to be anticipatory and AD to exhibit the desired tuning curve shapes. This model was a two-layer network of linear units, where the input layer's 202 units represented the two LMN pools, and the output layer's 101 units represented AD. The training set consisted of a set of LMN activity patterns corresponding to bumps centered at each of 101 positions around the circle, for five different angular velocities from fast CCW (approximately $-300^\circ/\text{s}$) through still to fast CW. Thus there were a total of 505 training patterns. The amplitudes of the activity patterns in the two LMN pools were modulated by angular velocity in the appropriate way. In simulations where LMN was supposed to be anticipatory, we also shifted the bumps in the two LMN pools by appropriate amounts corresponding to an ATI of roughly 46 ms. When simulating a nonanticipatory LMN, the bumps were not shifted, only their amplitudes varied.

Given a set of LMN input patterns X and a set of desired AD output patterns Y , we sought a weight matrix W that could satisfy $WX = Y$. The weights can be derived by gradient

descent learning using the LMS (least mean squares) algorithm (Hertz et al. 1991) or they can be computed directly as $W = YX^{-1}$, where X^{-1} denotes the pseudo-inverse of the 202×505 matrix X . The latter solution is faster to compute but numerically unstable; it produces more erratic weight vectors than gradient descent.

Results for these simulations are shown in Fig. 6. The network was easily able to associate nonanticipatory activity profiles in LMN with the proper AD tuning curve shapes (Fig. 6A). It was unable to do this for anticipatory activity profiles in LMN (Fig. 6, B and C). The matrix inverse solution produced the same unsatisfactory result as LMN. The bimodality of AD tuning curves during zero-turn cases was especially hard to reproduce.

Because the LMS algorithm with a two-layer network is known to converge to a solution if one exists (Hertz et al. 1991), these additional simulations suggest that there is no linear mapping between an anticipatory LMN and an AD with the tuning curve profiles described by Blair et al. (1997). Therefore although velocity-driven amplitude modulation in LMN provides a mechanism that can account for tuning curve distortions in AD, it would seem that the anticipatory nature of real LMN cells precludes their providing a full account of this phenomenon.

PoS has a reciprocal projection to AD. PoS lesions do not eliminate HD cell firing in the AD, but they do alter the shapes of AD tuning curves, making them slightly broader and increasing the ATI (Goodridge and Taube 1997). On the other hand, bilateral lesions of LMN abolish head direction cell firing in AD (Blair et al. 1999; Tullman and Taube 1998). To investigate the possibility that LMN and PoS inputs together could produce the observed tuning curves in AD, we added a PoS input to the two-layer linear network model, expanding the input layer to 303 units. With this addition, the network was able to reproduce the desired AD response, as shown in Fig.

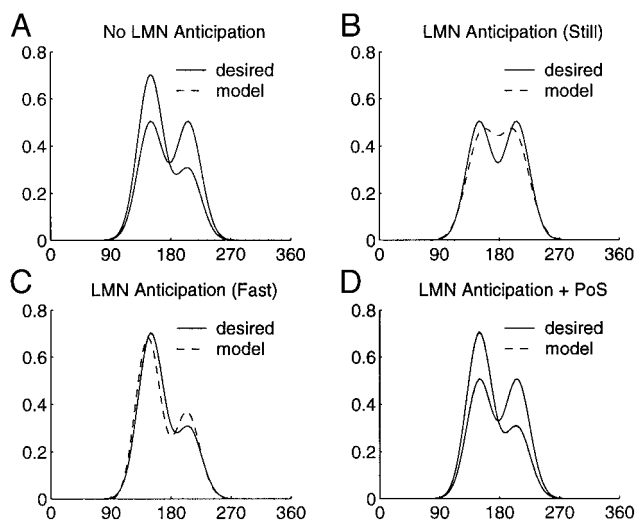


FIG. 6. Results from training a linear network to produce curves with shapes similar to what is exhibited in the anterior thalamus, given either anticipatory or nonanticipatory LMN inputs. *A*: with no LMN anticipation, the network reproduces the desired AD tuning curve shapes perfectly. — and ··· overlap so completely that ··· is not visible. *B*: when LMN patterns exhibit substantial anticipation, the network reproduces the AD still tuning curve poorly. *C*: the AD fast tuning curve also cannot be fit perfectly with an anticipatory AD. *D*: when LMN anticipates, but a PoS input is also provided, the network can match AD tuning curves exactly.

6D. Our intuition about why the PoS input helped is that PoS is somewhat “out of phase” with an anticipatory LMN, so the PoS input vector contributes a new set of basis functions on which the linear approximator can draw. The idea is analogous to adding more terms to a Fourier series.

Besides PoS, another area that could be affecting AD tuning curve shapes is retrosplenial cortex, which like LMN receives projections from PoS and provides input to AD (van Groen and Wyss 1990, 1992). In addition, evidence suggests that some cells in retrosplenial cortex fire as a function of the animal’s movements (Chen et al. 1994), implying there is velocity modulation there.

DISCUSSION

The contributions of this modeling effort to our understanding of the rodent head-direction system are 1) a confirmation of the Blair and Sharp hypothesis that a single offset angle parameter (δ) can account for both degree of tuning curve distortion and degree of anticipation in AD; 2) a prediction that to achieve a constant ATI over a range of angular velocities, the head-direction system must be balancing CW and CCW offset inputs so as to maintain an overall constant input to AD; 3) a demonstration that an additive angular velocity signal can produce suitable amplitude modulation (a multiplicative effect) in an attractor network model of LMN; 4) a demonstration that LMN anticipation of AD is incompatible with LMN being the sole input responsible for AD tuning curves, if AD curves are bimodal and distort with velocity as described by Blair and Sharp. However, 5) a combination of anticipatory, velocity-modulated LMN input and nonanticipatory, velocity-independent PoS input does suffice to reproduce AD tuning curves.

Several issues remain. We derived a workable combination of LMN and PoS to AD weights using a simple linear model, but we have not yet tried to retrofit these weights into the more complex nonlinear model. Also a more recent report by Blair and Sharp (1998) describes LMN cells as being width-modulated as well as showing velocity modulation; the tuning curves narrow with increasing velocity in the cell’s preferred turning direction. This effect was not reported by Stackman and Taube (1998). We await the results of further experiments that could resolve the issue. In the mean time, a model for producing width-modulated LMN tuning curves would be useful so that the implications for AD tuning curve shapes may be investigated.

We are grateful to T. Blair and P. Sharp of Yale University and B. Stackman and J. Taube of Dartmouth College for sharing head-direction cell data with us. We thank all of the preceding people and D. Redish and K. Zhang, for helpful discussions.

This work was funded by National Science Foundation Grant IBN-9631336.

Present address of J. P. Goodridge: 336 Central Park West, Studio 6, New York, NY 10025.

Address for reprint requests: D. S. Touretzky, Computer Science Department, Carnegie Mellon University, Pittsburgh, PA 15213-3891.

Received 4 January 1999; accepted in final form 16 February 2000.

REFERENCES

ALLEN GV AND HOPKINS DA. Mammillary body in the rat: topography and synaptology of projections from the subicular complex, prefrontal cortex, and midbrain tegmentum. *J Comp Neurol* 286: 311–336, 1989.

- BENTIVOGLIO M, KULTAS-ILINSKY K, AND ILLINSKY I. Limbic thalamus: structure, intrinsic organisation, and connections. In: *Neurobiology of Cingulate Cortex and Limbic Thalamus: A Comprehensive Handbook*, edited by Vogt BA and Gabriel M. Boston: Birkhauser, 1993, p. 71–122.
- BLAIR HT. A thalamocortical circuit for computing directional heading in the rat. *Adv Neural Inform Process Syst* 8: 152–158, 1996.
- BLAIR HT, CHO J, AND SHARP PE. The anterior thalamic head-direction signal is abolished by bilateral but not unilateral lesions of the lateral mammillary nucleus. *J Neurosci* 19: 6673–6683, 1999.
- BLAIR HT, LIPSCOMB BW, AND SHARP PE. Anticipatory time intervals of head-direction cells in the anterior thalamus of the rat, implications for path integration in the head-direction circuit. *J Neurophysiol* 78: 145–159, 1997.
- BLAIR HT AND SHARP PE. Anticipatory head direction signals in anterior thalamus: evidence for a thalamocortical circuit that integrates angular head motion to compute head direction. *J Neurosci* 15: 6260–6270, 1995.
- BLAIR HT AND SHARP PE. Visual and vestibular influences on head-direction cells in the anterior thalamus of the rat. *Behav Neurosci* 110: 643–660, 1996.
- BLAIR HT AND SHARP PE. Angular velocity-modulated head direction cells are segregated by hemisphere in the lateral mammillary nucleus of the rat. *Neuron* 21: 1387–1397, 1998.
- CHEN LL, LIN LH, GREEN EJ, BARNES CA, AND MCNAUGHTON BL. Head-direction cells in the rat posterior cortex. I. Anatomical distribution and behavioral modulation. *Exp Brain Res* 101: 8–23, 1994.
- GEORGOPOULOS AP, CAMINITI R, KALASKA JF, AND MASSEY JT. Spatial coding of movement: a hypothesis concerning the coding of movement direction by motor cortical populations. *Exp Brain Res Suppl* 7: 327–336, 1983.
- GOODRIDGE JP, REDISH AD, TOURETZKY DS, BLAIR HT, AND SHARP PE. Lateral mammillary input explains distortions in tuning curve shapes of anterior thalamic head direction cells. *Soc Neurosci Abstr* 23: 503, 1997.
- GOODRIDGE JP AND TAUBE JS. Interaction between the postsubiculum and anterior thalamus in the generation of head direction cell activity. *J Neurosci* 17: 9315–9330, 1997.
- HERTZ J, KROGH A, AND PALMER PG. *Introduction to the Theory of Neural Computation*. Reading, MA: Addison-Wesley, 1991.
- MCNAUGHTON BL, CHEN LL, AND MARKUS EJ. Dead reckoning landmark learning, and the sense of direction: a neurophysiological and computational hypothesis. *J Cognit Neurol* 3: 190–202, 1991.
- PINTO DJ, BRUMBERG JC, SIMONS DJ, AND ERMENTROUT GB. A quantitative population model of whisker barrels: re-examining the Wilson-Cowan equations. *J Comput Neurosci* 3: 247–264, 1996.
- RANCK JB JR. Head-direction cells in the deep cell layers of dorsal presubiculum in freely moving rats. *Soc Neurosci Abstr* 10: 599, 1984.
- REDISH AD, ELGA AN, AND TOURETZKY DS. A coupled attractor model of the rodent head direction system. *Network Comput Neural Syst* 7: 671–685, 1996.
- SALINAS E AND ABBOTT LF. A model of multiplicative neural responses in parietal cortex. *Proc Natl Acad Sci USA* 93: 11956–11961, 1996.
- SHARP PE. Multiple spatial/behavioral correlates for cells in the rat postsubiculum: multiple regression analysis and comparison to other hippocampal areas. *Cereb Cortex* 6: 238–259, 1996.
- SHIBATA H. Topographic organization of subcortical projections to the anterior thalamic nuclei in the rat. *J Comp Neurol* 332: 117–127, 1992.
- SKAGGS WE, KNIERIM JJ, KUDRIMOTI HS, AND MCNAUGHTON BL. A model of the neural basis of the rat's sense of direction. *Adv Neural Inform Process Syst* 7: 173–180, 1995.
- STACKMAN RW AND TAUBE JS. Firing properties of head direction cells in the rat anterior thalamic nucleus: dependence upon vestibular input. *J Neurosci* 17: 4349–4358, 1997.
- STACKMAN RW AND TAUBE JS. Firing properties of rat lateral mammillary single units: head direction, head pitch, and angular head velocity. *J Neurosci* 18: 9020–9037, 1998.
- TAUBE JS. Head direction cells recorded in the anterior thalamic nuclei of freely moving rats. *J Neurosci* 15: 1953–1971, 1995.
- TAUBE JS AND BURTON HL. Head direction cell activity monitored in a novel environment and during a cue conflict situation. *J Neurophysiol* 74: 1953–1971, 1995.
- TAUBE JS, MULLER RI, AND RANCK JB JR. Head direction cells recorded from the postsubiculum in freely moving rats. Description and quantitative analysis. *J Neurosci* 10: 420–435, 1990.
- TAUBE JS AND MULLER RI. Comparisons of head direction cell activity in the postsubiculum and anterior thalamus of freely moving rats. *Hippocrates* 8: 87–108, 1998.
- TULLMAN ML AND TAUBE JS. Lesions of the lateral mammillary nuclei abolish head direction cell activity in the anterior dorsal thalamus. *Soc Neurosci Abstr* 24: 1912, 1998.
- VAN GROEN T AND WYSS JM. The postsubicular cortex in the rat: characterization of the fourth region of the subicular cortex and its connections. *Brain Res* 529: 165–177, 1990.
- VAN GROEN T AND WYSS JM. Connections of the retrosplenial dysgranular cortex in the rat. *J Comp Neurol* 315: 200–216, 1992.
- VAN GROEN T AND WYSS JM. Projection from the anterodorsal and anteroventral nucleus of the thalamus to the limbic cortex in the rat. *J Comp Neurol* 358: 584–604, 1995.
- WILSON HR AND COWAN JD. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys J* 12: 1–24, 1972.
- ZHANG K. Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. *J Neurosci* 16: 2112–2126, 1996a.
- ZHANG K. Spatial orientation and the Papez circuit: dynamics of the interacting cell groups. Presented at CNS*96, *The Fifth Annual Computation in Neural Systems Meeting*, 1996b.