Running head: DEFORMING THE HIPPOCAMPAL MAP

Deforming the Hippocampal Map

David S. Touretzky, Wendy E. Weisman, Mark C. Fuhs, William E. Skaggs, Andre A. Fenton, and Robert U. Muller

[1]Center for the Neural Basis of Cognition

[2]Computer Science Department

Carnegie Mellon University

Pittsburgh, PA 15213-3891

Number of words in Abstract: xxx

Number of text pages:

Number of figures:

Number of tables:

Send correspondence to:

David S. Touretzky

phone: 412-268-7561

fax: 412-268-3608

e-mail: dst@cs.cmu.edu

Abstract

To investigate conjoint stimulus control over place cells, Fenton et al. (2000a) recorded while rats foraged in a cylinder with 45° white and black cue cards on the wall. Card centers were 135° apart. In probe trials the cards were rotated together or apart by 25°. Firing field centers shifted during these trials, stretching and shrinking the cognitive map. Fenton et al. (2000b) described this deformation with an ad hoc vector field equation. We present two other models of map deformation. In a maximum likelihood formulation, the rat's location is estimated by a conjoint probability density function. In an attractor neural network model, recurrent connections produce a bump of activity over a 2D array of cells; the bump's position is influenced by landmark features such as distances or bearings. The maximum likelihood and attractor network models yield similar results, supporting previous conjectures (Deneve et al., 2002) that maximum likelihood may be an appropriate framework for describing attractor network behavior.

Deforming the Hippocampal Map

Firing Field Distortions

To investigate conjoint stimulus control over place cells, Fenton, Czismadia, and Muller (2000a) recorded while rats foraged in a cylinder with one white and one black cue card on the wall. The cards each subtended 45° of arc, and their centers were 135° apart, leaving a gap of 90° between the right edge of the white card and the left edge of the black card. In probe trials the cards were rotated to increase or decrease their separation by 25°. Firing field centers shifted systematically during these trials, distorting the cognitive map. In contrast, on trials where one card was removed and the other rotated, the map did not distort, but it did rotate with the remaining card, demonstrating its continuing salience.

Vector Field Model

Fenton, Czismadia, and Muller (2000b) presented a mathematical model of the deformations they had observed in their experiments. The model described how they believe firing field centers move as a result of rotating the cue cards or deleting a card: all cells were controlled by both cards, but to varying degrees based on the distance of the field center to each card. The approach was presented as purely ad hoc: the equations simply describe the desired vector field results, with no claim that the hippocampus actually derived firing fields this way.

The vector field equation had two parts: an angular component that determined how firing field centers rotate around the center of the arena, and a translational component that corrected a problem with the prediction of the rotation equation when the cards were moved together or apart. We begin by presenting this model in a slightly different formulation than Fenton et al. (2000b) for improved clarity and completeness.

Our variant assumes that upon entry into the cylinder during a probe trial, the animal's head direction estimate is reset so that East is the direction defined by the line from the arena center to a reference point on the cylinder wall half way between the closest edges of the two cue cards. If the cards rotate by opposite amounts (e.g., $+10^\circ$ and -10°), their edges move closer together or farther apart, but there is no change in the reference point. If the cards rotate by identical amounts, the reference point rotates as well, so the change is undetectable by the model due to head direction reset. If the cards rotate by unequal amounts, after head direction reset the model will see only their relative motion. Finally, if one card is removed, the reference point is defined to be in "standard" position with respect to the remaining card, meaning 45° counterclockwise from the left edge of the black card or 45° clockwise from the right edge of the white card.

Head direction reset restricts the model to operating in the reference frame defined by the cards. There is no provision for an external reference frame tied to the experimental chamber to influence place cell firing. (And Fenton et al.'s results when the two cards rotate together show no influence of a room frame,) The head direction reset assumption greatly simplifies the vector field equations, ensuring that angular displacements of the cards with respect to the reference point are always equal and opposite. We will therefore adopt the convention that the white card rotates by an angle α and the black card by $-\alpha$.

Insert Figure 1 about here

The next step in our formulation is to calculate displacement vectors, in room coordinates, based on rotations of the individual cue cards. Ignoring the black card for the moment, if the white card rotates by an angle α, then all firing fields influenced by this card should rotate by α. We represent each field by the location of its center. Assume the cylinder is centered at the origin, and let R be its radius. A field on the east edge of the

cylinder, at location $[R, 0]$, would rotate to a new position $[R\cos\alpha, R\sin\alpha]$. Hence, the white card-dependent rotational displacement vector for this specific point would be:

$$W_{R,0} = \begin{bmatrix} R\cos\alpha \\ R\sin\alpha \end{bmatrix} - \begin{bmatrix} R \\ 0 \end{bmatrix} = R \cdot \begin{bmatrix} \cos\alpha - 1 \\ \sin\alpha \end{bmatrix} \tag{1}$$

Any point on the x-axis at a distance r from the origin would have the same rotational displacement vector $W_{R,0}$, scaled by r/R. In the general case of an arbitrary point $[x, y]$ in the cylinder, at a distance $r = \sqrt{x^2 + y^2}$ from the origin and an angular displacement $\theta = \text{atan2}(y, x)$ from the positive x-axis, the white card-dependent rotational displacement vector is given by:

$$W_{x,y} = r \cdot \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \times \begin{bmatrix} \cos\alpha - 1 \\ \sin\alpha \end{bmatrix} \tag{2}$$

The black card-influenced rotational displacement vector $B_{x,y}$ is calculated analogously using $-\alpha$. The overall rotational displacement of a point $[x, y]$ is the average of its white and black card rotational displacement vectors, weighted by relative distance to the two cards, so that the nearer card has proportionally greater influence. If d_w and d_b are the distances from the point $[x, y]$ to the centers of the white and black cards, respectively, then the net rotational displacement is

$$D_{\text{rot}} = \frac{d_w B_{x,y} + d_b W_{x,y}}{d_w + d_b} \tag{3}$$

The result of this equation is shown in Figure 1a. This purely rotational displacement does not accurately capture the distortion of relative field locations observed by Fenton et al. When the two cards move closer together, all field centers are displaced slightly toward the cards; when the cards move apart, field centers are displaced away from the cards. Fields close to a card show less translation, but more rotation, than fields

distant from either card. Fenton et al. added a translational term to their equation to reproduce this effect. Let W_c be the rotational displacement vector denoting the movement of the white card center from its standard position to its current location. Let B_c be the rotational displacement vector for the black card center. The translational displacement of a point at distances d_w, d_b from the white and black cards, respectively, is defined as:

$$D_{\text{trans}} = \frac{W_c + B_c}{c_2 \cdot \left(\frac{1}{d_w} + \frac{1}{d_b}\right)} \tag{4}$$

where $c_2 = 83.4$ cm determines the dropoff of the translational term as distance to the cards decreases. Note that if either d_w or d_b is small, the denominator of the equation will be large, and the translational term will be attenuated relative to the rotational term. But at locations distant from both cards, the translational term is significant. The contribution of this translational displacement term is shown in Figure 1b.

The total displacement of a point is the sum of its rotational and translational displacements, and is shown in Figure 1c:

$$D_{\text{tot}} = D_{\text{rot}} + D_{\text{trans}} \tag{5}$$

Figure 2 shows the rotational, translational, and total displacement vectors when the cards are moved apart by 25°. Once again, the translational component helps to overcome the shortcomings of the rotational component at locations distant from both cue cards.

The Fenton et al. (2000b) model produces a good match to the experimental data of Fenton and Muller on cards moving together or apart by 25°. Furthermore, when one card is removed and the other card rotated, the model correctly predicts that fields will rotate without distortion, i.e., they rotate in the room frame to maintain their "standard" position relative to the remaining card. The original formulation of the model accomplished this by assigning an effectively infinite distance to the missing card, but this

is unnecessary in our revised scheme. Defining the reference point to be in "standard position" whenever a card is missing gives $\alpha = 0$ for Equation 2. The values for d_w and d_b in Equation 4 are thus unimportant, because W_c and B_c are zero vectors. The model is already insensitive to rotations of the cards and cylinder relative to the room frame, so treating a missing cue card as if it were present in standard position does not pose a problem.

Insert Figure 2 about here

Maximum Likelihood Model

Our first alternative to the purely descriptive approach of the vector field equations (3–5) is to look for a theoretically justified account of map deformation. Let us assume the rat uses a probabilistic method to determine its position, with each landmark an independent source of position information. In this framework, each landmark observation generates a probability distribution for the animal's current location. When the cue cards are in the standard configuration, all evidence sources should agree (in the absence of noise.) But when the cards move closer together or farther apart, the evidence becomes inconsistent. One reasonable resolution to this situation is to take all the evidence into account and choose the peak of the combined probability distribution as the animal's most likely location. This maximum likelihood approach offers a probability-theoretic justification for why the firing field map is distorted by cue card movement.

Let the probability distribution generated by a landmark observation be a Gaussian function of distance. If the rat perceives its distance to landmark i to be v_i, then the probability that its actual distance to the landmark is d_i is distributed across a Gaussian annulus with a peak at distance v_i and a variance σ^2 which we assume is proportional to v_i^2:

$$p(d_i|v_i) = \frac{1}{A_i} \exp\left[\frac{-(d_i - v_i)^2}{v_i^2}\right] \tag{6}$$

where A_i is the area under the annulus.

Within the cylinder, the set of locations at distance d_i from landmark i forms an arc centered on the landmark. A single landmark is thus insufficient to determine location, but a pair of landmarks is usually adequate. We assume that the two edges of each cue card are both utilized as landmarks, so there are normally four landmarks visible. When one card is deleted two landmarks are still available.

If landmarks are independent evidence sources, then they can be combined by multiplying their probability distributions. The two cue cards would be truly independent if cards could overlap. Even when this is prohibited, the cards are nearly independent. But the four cue card edges are not independent, since if we know the location of the left edge of a card, the right edge can only be in one location, determined by the card's fixed width. Nonetheless we will treat the landmarks as independent for purposes of evidence combination, as a naive observer, unaware that a card's two edges must move in unison, might do. The joint probability distribution is thus:

$$p(x, y|v_1, \ldots, v_4) = \prod_{i=1}^{4} p(x, y|v_i) \tag{7}$$

The product of several Gaussian arcs that intersect at a single point is, roughly, a Gaussian bump. Thus, the Gaussian firing rate distribution of a place cell whose field center is at distances v_1 through v_4 from the four landmarks when the cards are in standard position can be viewed as an estimate of $p(x, y)$. Moving the rat (changing x and y) samples the cell's firing rate distribution at different arena locations. Moving the cue cards alters the observed combinations of v_i's and thus changes the entire distribution. We assume the standard deviation σ of each distribution is proportional to the perceived distance to that landmark, so that distributions scale in accordance with Weber's law.

The joint probability distribution $p(x, y|v_1, \ldots, v_4)$ need not be computed directly. It suffices to calculate $\mathcal{L}_{xy}$, the log likelihood, which has the advantage of eliminating the exponential functions and replacing a product with a sum:

$$\begin{aligned} \mathcal{L}_{xy} &= \log p(x, y|v_1, \ldots, v_4) \\ &= \log \textstyle\prod_{i=1}^{4} p(x, y|v_i) \\ &= \textstyle\sum_{i=1}^{4} \log p(x, y|v_i) \\ &= -\textstyle\sum_{i=1}^{4} \log A_i - \sum_{i=1}^{4} \frac{(d_i^{xy} - v_i)^2}{v_i^2} \end{aligned} \tag{8}$$

The maximum of the log likelihood will be found at the same location as that of the underlying probability distribution. To find the location $[x^*, y^*]$ of the peak of the probability distribution we eliminate constant terms and calculate:

$$]x^*, y^*] = \arg\max_{x,y} \sum_{i=1}^{4} \frac{-(d_i^{xy} - v_i)^2}{v_i^2} \tag{9}$$

An interesting consequence of using Weber's law scaling is that nearer landmarks have greater influence on the location of the peak. To see this, consider the one-dimensional case with landmarks at locations l_1 and l_2 on the real line bounding the arena, with $l_1 < l_2$. At any point x within the arena, the distances to these landmarks are $d_1^x = x - l_1$ and $d_2^x = l_2 - x$. Suppose we move the landmarks and try to estimate our position based on observed distances v_1 and v_2:

$$\mathcal{L}_x = -(d_1^x - v_1)^2/v_1^2 - (d_2^x - v_2)^2/v_2^2 \tag{10}$$

The local maximum, attained when $d\mathcal{L}_x/dx = 0$, is

$$x^* = \frac{(l_1 + v_1)v_2^2 + (l_2 - v_2)v_1^2}{v_1^2 + v_2^2} \tag{11}$$

which bears a strong resemblance to Equation 3.

Unlike the vector field model, the maximum likelihood model operates in room coordinates and makes no assumptions about head direction reset. When the two cards rotate by equal amounts, all points rotate around the center by a corresponding amount. When a card is removed, we simply omit the corresponding terms from the sum; there is no inconsistency among the remaining cues so the fields rotate without distortion. But when the cards move closer together or farther apart, the four Gaussian arcs defined by the individual landmarks shift relative to each other, and the map undergoes stretching and shrinking.

We implemented this model using a grid of points spaced 1 cm apart. The 76 cm diameter cylinder contained 4,513 of these points. To plot the vector field with the cards rotated, we calculated v_i' values for the rotated landmarks viewed from all points on the grid. We then selected an evenly spaced subset of points from the interior of the cylinder, and for each point $[x, y]$ representing a firing field center, we found $[x^*, y^*]$, the peak of $p(x, y|v_1', \ldots, v_4')$. This was the location of the cell's firing field center with the cards rotated. We then drew a vector from $[x, y]$ to $[x^*, y^*]$.

Insert Figure 3 about here

As shown in Figure 3a, the results of this formulation do not always match that of the Fenton et al. (2000b) model. Specifically, when the two cards move closer together by 25°, most of the vectors point roughly east, but there is a region around the arena center where the vectors are noticeably attenuated. This is because, at the center of the arena, the distance to all four landmarks is equal to the arena radius, and this distance is unchanged by cue card rotation. Another striking feature of the plot is the sharp discontinuity in the eastern portion of the arena, around the vertical line joining the right edge of the white card to the left edge of the black card. The vectors suddenly switch

direction in this region. This is the result of Equation 9 choosing a single location for the maximum value of $\mathcal{L}_{xy}$. Points on the line actually have two maxima, one on either side. As one moves off the line in either direction, the symmetry is broken. These unexpected results are a consequence of the geometry of the arena (a concave, symmetric interior with landmarks located only at the edges) and our assumption that distance from landmarks is the rat's sole source of evidence for its position.

A plausible alternative hypothesis is that rats determine their position based on angles between pairs of landmarks, i.e., the difference in their relative bearings, or equivalently, the retinal angle subtended by a line connecting them. We explored this possibility in a second model. With four visible landmarks there are six possible landmark pairs, but for simplicity we used just the four pairs of circularly adjacent landmarks, since the additional pairs would only provide redundant information. Note that the retinal angle between a card's left and right edge decreases with distance from the card, while the retinal angle between edges of the white card and the black card is a function of both distances to the cards and the card separation angle.

Let u_{ij} be the rat's perceived angle between landmarks i and j, and let a_{ij}^{xy} be the actual bearing difference between these landmarks viewed from location $[x, y]$ with the cards in standard position. We can define a Gaussian probability distribution for position based on the perceived angle between a pair of distinct landmarks as:

$$p(x, y|u_{ij}) = \frac{1}{\sqrt{2\pi\sigma_a^2}} \exp\left[\frac{-(a_{ij}^{xy} - u_{ij})^2}{2\sigma_a^2}\right] \tag{12}$$

Once again, we make the simplifying assumption that evidence sources are independent, so that the overall probability distribution is the product of the probability functions for the four landmark pairs u_{12}, u_{23}, u_{34}, and u_{41}. The result, Figure 3b, shows good behavior in the vicinity of the cue cards but anomalous results in other parts of the cylinder. This plot was made with σ_a proportional to u_{ij}; using a constant value produced

slightly less satisfactory results. In either case, at locations far from either cue card the angles between landmarks change fairly slowly with position, producing a shallow gradient with broad peaks. When the two cards move in a non-rigid fashion, the location of a firing field peak in the western half of the cylinder can shift by a large amount even though the change in magnitude of the probability values is small. Therefore angle-based features are also not ideal for determining position in the cylinder.

The single landmark feature that produced the best firing field deformation pattern was allocentric bearing, i.e., bearing to the landmark with respect to some external coordinate system independent of present heading. The result is shown in Figure 3c.

Neurons in the rodent head direction system have been shown to encode the animal's heading with respect to the environment (Taube, Muller, & Ranck, 1990b). This suggests that rats are capable of computing allocentric bearings of landmarks. Behavioral experiments in gerbils also indicate that rodents can use allocentric bearing information to disambiguate landmarks (Collett, Cartwright, & Smith, 1986). Futhermore, in a familiar environment the alignment of the rodent head direction system is known to be controlled by visual landmarks (Taube, Muller, & Ranck, 1990a). Therefore, to compute allocentric bearings in our maximum likelihood model, we assume that the animal's heading reference (East) is defined as mid-way between the white and black cue cards, just as in our vector field model. Moving the cards together or apart by equal but opposite amounts leaves the reference point unchanged, and thus does not affect the alignment of the head direction system, although individual landmark bearings will of course shift. Rotating the cards by identical amounts rotates the heading reference as well, so the rat does not notice any bearing change in that situation either.

In the standard cue configuration, the reference point is half a card width clockwise from the right edge of the white card, and half a card width counterclockwise from the left edge of the black card. When one card is deleted, we assume that East is in its standard

position relative to the remaining card. Hence, with one card present, firing field centers can rotate but will not deform.

The overall best result was obtained by combining distance with either the angle or bearing based features. The distance-based probability function (Figure 3a) has a stronger gradient in the western half of the cylinder than the angle-based function (Figure 3b). Combining the likelihoods by simple addition allows each to compensate for the other's shortcomings. As shown in Figure 3d, the resulting vector fields are very close in appearance to those of Fenton et al. We set σ_a to unity in this case. The locations of peak firing were estimated by:

$$[x^*, y^*] = \arg\min_{x,y} \left(\sum_{i=1}^{4} \frac{(d_i^{xy} - v_i)^2}{v_i^2} + \sum_{i,j} (a_{ij}^{xy} - u_{ij})^2 \right) \tag{13}$$

Similar results were obtained using distance plus allocentric bearing.

Insert Figure 4 about here

Figure 4 shows the output of various versions of the maximum likelihood model when the cards are moved apart by 25°. Once again, a combination of distance and angle features produces the best results.

Attractor Bumps

Dynamical systems, or "attractor bump" networks, are a popular approach to modeling aspects of hippocampal place cells (Zhang, 1996; Samsonovich & McNaughton, 1997), and have been widely adopted by hippocampal modelers (Redish & Touretzky, 1998; Redish, 1999; Doboli, Minai, & Best, 2000; Káli & Dayan, 2000). One-dimensional attractor models have been used to model the head direction system (Redish, Elga, & Touretzky, 1996; Goodridge & Touretzky, 2000), orientation tuning in visual cortex

(Ben-Yishai, Hansel, & Sompolinsky, 1997), and the oculomotor system (Seung, 1996). Two-dimensional attractor networks have been proposed as models of hippocampus, superior colliculus (Droulez & Berthoz, 1991; Pouget, Deneve, & Duhamel, 2002) and motor cortex (Lukashin, Amirikian, Mozhaev, Wilcox, & Georgopoulos, 1996). Here we examine the ability of an attractor network to function as a deformable map, producing the stretching and shrinking effects observed in the Fenton and Muller two-card experiment.

We begin with a population of place cells arranged as a 2D grid. Let each cell have strong excitatory connections to the cells nearby, weaker excitatory connections to cells somewhat further away, and inhibitory connections to all the remaining cells. With appropriate parameter settings, a network organized this way will have an infinite number of stable states, each consisting of a "bump" of activity localized to some region of the grid. Such a state is analogous to the population activity observed in the hippocampus, because when the rat is at a particular location, the place cell whose field is centered closest to the rat's location will be firing at its maximum rate, while cells whose firing fields just overlap with the rat's location will fire at lesser rates, and cells whose fields are far from the rat's location will be quiescent. As the rat moves through the environment, the activity pattern over the place cell population shifts to reflect this.

If the units comprising an attractor network are initialized with random activity levels, the network will settle into a stable state with a well-formed bump at a random location. However, if a smoothly-varying external input is applied to some region of the grid, the bump will tend to form in the region of maximal external input. The attractor network can thus be regarded as a parallel, distributed mechanism for finding the peak of an input signal projected onto the grid.

To produce visual control of firing fields, the external input may be taken from a collection of visual feature detectors tuned to landmark distances and/or bearings. To

model the two-card experiment, we created a separate set of feature detectors for each landmark. For the ith landmark there was a set of distance detectors $F_{i,j}$ tuned to various distances r_j, and a set of bearing detectors $G_{i,j}$ tuned to various allocentric bearings ϕ_j. The place cell with firing field centered at $[x, y]$ received an excitatory connection from feature detector $F_{i,j}$ or $G_{i,j}$ if the distance from $[x, y]$ to the ith landmark was approximately r_j, or the allocentric bearing of landmark i viewed from $[x, y]$ was approximately ϕ_j.

Once the feature detectors have been wired up to the place cells, consider the rat entering the environment at the start of a trial. Let the rat's perceived distance from its present location to the ith landmark be v_i. Those distance detectors $F_{i,j}$ whose preferred distance value is close to v_i will be active, and will supply excitation to the appropriate subset of place cells. A similar situation holds for bearing detectors $G_{i,j}$. A bump of activity will then form over the place cells with its peak centered at roughly the location receiving the greatest amount of feature detector input. The simulated rat has thereby estimated its position in the arena.

This approximation is exact if the external input to the place cell grid takes the shape of a gaussian bump. However, the projections from individual feature detectors to place cells form arcs, not bumps. For example, a distance-based feature detector $F_{i,j}$ will project to the arc of cells centered at distance r_j from landmark i. If several arcs cross at a single point at roughly equal angles, the resulting pattern of external input will look bump-like. But this condition does not always hold. Angle-based feature detectors produce very broad arcs in the western half of the arena, and the geometry of the arena and cue cards constrains all arcs to be nearly coincident in that region. Bearing-based feature detectors were used instead because they do not suffer this problem. Another problem is that in the eastern half of the arena, when the cards are moved, distance arcs that once overlapped now merely pass close by, producing elongated patterns of external

input that are far from bump-like. These effects were not a problem for the maximum likelihood model because it only looked at the peak of the input distribution; the overall shape of the input was ignored. But the attractor network is sensitive to this shape, and thus requires some refinement of its input features to assure that the peak of the external input is close to the center of the input distribution. We therefore added a feedforward inhibition term from the feature detectors to the place cells that was strong enough to cancel any individual arc or intersection of a few arcs, but not the intersection of many arcs. This "decluttered" the external input signal, producing a stimulus that was more focused and bump-like.

Another difficulty arises close to the arena walls. The 2D attractor grid extends beyond the arena boundaries, but since the animal cannot experience the environment beyond the walls, feature detectors were not wired up to place cells lying outside the cylinder. When the simulated animal is at a point along the wall, the attractor bump should be centered on a place cell right at the wall. This cell will excite (and receive excitation from) its nearby neighbors, both those closer to the arena center and those further away (hence outside the wall). However, only the cells within the wall receive external input from the feature detectors; the cells outside the wall do not. Thus there is a danger that the bump may form some distance short of the wall, since that is where the center of mass of the external input lies. To minimize this we use relatively weak weights from the feature detectors to the place cell grid, so that the attractor dynamics dominate, and adjust the feedforward inhibition so that the location of the peak of the input is more important than the overall shape.

Model Details

The attractor network was implemented as a 45×45 grid of cells, toroidally connected to eliminate edge effects. This assures that all cells have the same number of

neighbors, so that in the absence of external input, the attractor bump has a uniform shape everywhere on the grid. The arena was defined as a circular region 38 units in diameter, centered on the grid origin. Each grid unit therefore covered a surface of 4 cm^2. The attractor bump was roughly 17 units in diameter, so a bump located at one edge of the cylinder would have minimal effect on cells at the opposite edge via wrap-around on the torus.

Each place cell's activation $V_i(t)$ was computed as the sum of recurrent excitation from other place cells, a global shunting inhibition term, and the external input received from feature detectors. Shunting inhibition was used because it improves the stability of the network (Káli & Dayan, 2000).

$$V_i(t+1) = \sum_j w_{ij}^{EE} F_j(t) + w^{EI} FI(t) V_i(t) + \sum_k w_{ik}^{EF} FD_k(t) \tag{14}$$

The integral of activation over time is the cell's synaptic drive $S_i(t)$, governed by a time constant τ_E (Pinto, Brumberg, Simons, & Ermentrout, 1996). The cell's firing rate $F_i(t)$ is proportional to the synaptic drive thresholded at zero.

$$\tau_E \frac{dS_i(t)}{dt} = -S_i(t) + V_i(t) \tag{15}$$

$$F_i(t) = [\, S_i(t) \,]_+ \tag{16}$$

A global inhibitory unit receives excitation from all place cells and makes inhibitory projections back to them and to itself:

$$VI(t) = w^{IE} \sum_j F_j(t) + w^{II} FI(t) \tag{17}$$

$$\tau_I \frac{dSI(t)}{dt} = -SI(t) + VI(t) \tag{18}$$

$$FI(t) = [\, SI(t) \,]_+ \tag{19}$$

The strengths of the recurrent connections w_{ij}^{EE} between cells i and j on the torus are a Gaussian function of the distance d_{ij} between them, measured in cm:

$$w_{ij}^{EE} = 0.1125 \ \exp\left(-d_{ij}^2/40.5\right) \tag{20}$$

For each of the four landmarks there were 43 distance-based Gaussian feature detectors, tuned to even distance values from 0 to 84 cm. σ_d was 2 cm, and the distance feature detectors were weighted relative to each other to place more emphasis on detectors tuned to small distances. The weighting function was $24/(24 + d)$, where d is the distance in cm to the landmark, so the relative weights from distance-based feature detectors varied from 1.0 down to 0.22. In addition there were 120 allocentric bearing feature detectors tuned to bearing values from 0° to 357° in increments of 3°. σ_b was 6°, and all the bearing detectors had uniform weights. Figure 5 shows typical activation patterns that distance and bearing-based feature detectors transmit to the place cell grid.

Insert Figure 5 about here

Results from the Attractor Model

Insert Figure 6 about here

Insert Figure 7 about here

Figures 6 and 7 show the map distortion patterns when the cue cards are moved together or apart by 25°. The results are similar to the vector field and maximum likelihood models, though a little less smooth, due in part to the lower resolution of the grid.

Insert Table 1 about here

The attractor network model also reproduces the reduction in peak firing rates observed by Fenton et al. when the cards were rotated together or apart, as shown in Table 1. Peak activation levels in dimensionless units for three cells are shown in Figure 8. The mean peak activation (over all 1129 cells) declined by 10% in the cards apart condition, and by 12% in the cards together condition. Fenton et al. report reductions in the mean centroid firing rate, i.e., the firing rate of a cell in the centroid pixel of its firing field, of 36% (apart) and 35% (together), and reduction in the mean in-field firing rate of 21% (apart) and 15% (together).

The reduction in peak activation in the model is a consequence of the feature detector input becoming defocused when the cards are moved out of standard position. The Gaussian arcs from the various feature detectors no longer intersect perfectly at a single point, and so the place cell at the center of the stimulus bump receives less total input than before.

Insert Figure 8 about here

Figure 8 also shows that the shapes of firing fields distort along the direction of cue card motion. The least distortion is observed at the center of the arena, where only the

card bearings change, not the distances. **[Bob had a comment here: "Refer to DATA." What should go here?]**

Discussion

Comparison of Models

In the vector field model, a cue card's influence in the rotational component (Equation 3) is weighted inversely by distance in order to achieve the desired map distortion effect (Figures 1a and 2a.) A similar effect is obtained in the maximum likelihood model by giving distance-based landmark features a standard deviation σ_d proportional to the perceived distance v_i, on the assumption that such perceptual measurements should obey Weber's law. The scaled variance gives the evidence from a closer landmark a steeper gradient than that from more distant landmarks, hence the position estimate shows greater influence by the closer card.

In the model of Káli and Dayan (2000), EC cells are tuned to both distance and egocentric bearing to walls. The distance tuning is sharper for closer walls, and also sharper for walls behind the animal versus those ahead of it. Kali and Dayan justify this by assuming that if the animal is headed away from a wall, then it has been close to the wall recently and is likely to have a more accurate distance estimate based on path integration.

The boundary vector cells (BVCs) of Hartley, Burgess, Lever, Cacucci, and O'Keefe (2000) use a product of two Gaussians, one tuned to distance and one to allocentric bearing. Our maximum likelihood model and our attractor network model compute the sum of distance-based and bearing-based tuning functions. The log of a product being equal to the sum of the logs, the two approaches appear similar. However, Hartley et al.'s place cells compute a thresholded linear combination of two or more boundary vector cells, an extra layer of processing our model lacks. This appears to be necessary to derive

spatially compact place fields, because BVC's use entire walls as landmarks, whereas our models use only point landmarks. Furthermore, the attractor network model utilizes recurrent connections and network dynamics to produce roughly gaussian shaped firing fields from inputs with more varied shapes, while the Hartley et al. model produces these shapes in a strictly feed-forward manner. One drawback of the attractor approach is that the network does not produce crescent-shaped fields along cylinder walls, as the Hartley et al. model does. Such fields have been reported by Muller, Kubie, and Ranck (1987). Their existence suggests that place cells – at least those associated with boundaries – should be tightly sensory-bound. Place cells not associated with a boundary are presumably driven by distal landmarks and recurrent excitation, making them robust against landmark deletion.

Unlike in our maximum likelihood model, distance-based feature detectors in the attractor network model have a fixed variance, but the strength of their projections to place cells is weighted inversely by distance. This was necessary because active feature detectors with large variances would supply excitation to broad swaths of place cells, which deforms the input governing the location and shape of the attractor bump. The maximum likelihood model was unaffected by broad excitation because it simply picked the single point of maximum input as the animal's most likely location. But the attractor model relies on network dynamics to settle the activity bump over approximately the peak of the input, and is therefore less tolerant of inputs that depart significantly from a compact, roughly symmetric, unimodal shape. The feed-forward inhibition mechanism discussed previously helps to cut off the tails of an elongated input pattern, but is not sufficient in itself. The effect of the distance weighting function is similar to that of increasing the variance in the maximum likelihood model, in that it reduces the feature detector's contribution to the gradient of the activation.

The good agreement between the maximum likelihood and attractor network models

supports previous suggestions that attractor networks can function as maximum likelihood reasoners (Deneve, Latham, & Pouget, 2002), integrating evidence from multiple sources and cleaning up noise in the input.

Strength of Recurrent Connections

The strength of the recurrent connections relative to feature detector afferents determines how the attractor network model responds to inconsistent cues. Recurrent connections provide both excitation and inhibition, making firing fields robust against changes in total input and largely preserving their shapes when the cards are rotated.. If the recurrent connections are too weak, so that the activity bump with the cards in standard configuration is mainly a result of feature detector input, the reduction in input when a cue card is removed could cause the attractor bump to collapse. And the change in the input distribution when cards are moved together or apart by a significant amount could cause gross distortions in firing field shapes.

The map distortion effects visualized in Figures 6 and 7 as shrinking and stretching of the grid of firing field centers do entail some change in firing field shapes. To see this, consider two place cells i and j whose firing field centers c_i and c_j move closer together, to locations c_i' and c_j', when the cards are rotated. When the rat is at location c_i', by definition cell i is firing at its maximum rate with the cards rotated, and c_i' is closer to cell j's firing field center (now at c_j') than when the rat was at c_i with the cards in standard position. Hence cell j will be firing at a rate that is closer to its peak rate rate than when the cards are in standard position. Due to recurrent excitation and inhibition, a change in one cell's firing rate affects the behavior of other cells with overlapping firing fields. The effects will be seen as changes in peak firing rates and in the steepness of the gradient of firing rate change with distance from the center. Hence, the shapes of firing fields of cells in a recurrent network must change somewhat when the locations of their centers move

relative to each other.

When recurrent connections are too strong relative to the external input, that input can only influence where on the grid the activity bump appears, but not its shape. Place field centers could shift uniformly in response to card manipulations, but map distortions caused by relative motion would not be possible.

Threshold for Remapping

Knierim (2002) argues that recurrent connections must be weak relative to external inputs because in a double cue rotation task where local and distal cues rotated in opposite directions, some cells followed the local cues, some followed the distal cues, some developed split firing fields, and some remapped entirely.

The cue card manipulations studied here are sufficiently subtle that they do not trigger remapping. But moving the cards by greater amounts must eventually exceed some threshold beyond which the vector field transformation of Fenton et al. (2000b) no longer applies. Note that rotating the cards apart by 90° is equivalent to a mirror image reflection of the arena, where the white and black cards swap places. Unlike in the double cue rotation experiments (Tanila, Shapiro, & Eichenbaum, 1997; Knierim, 2002; Brown & Skaggs, 2002), where the discordance is between local vs. distal cues, here the two discordant cues are of the same type and presumably equal in salience. We therefore predict that at some critical amount of card rotation between 25 and 90 degrees, either a complete remapping will occur, or one card will lose its influence over firing fields, which will rotate with the other card without deforming.

Our current attractor-based model produces neither of these effects. It cannot undergo remapping because there is only one map stored in the recurrent connections. Neither can it reject the influence of one cue card when the inputs from the two cards are so far apart that they do not overlap. In attractor models with strong recurrent

connections, a spatially localized stimulus applied to one flank of the bump will cause the bump to shift in that direction, but the same input applied far from the bump will be suppressed due to recurrent inhibition (Redish, 1999). However, the feature detectors in our attractor model of the cylinder have broad projections (see Figure 5), so rotating the cue cards still produces inputs that intersect with and influence the bump location. Thus, for the rat to ignore one card when the separation is increased to 90 degrees, some attentional mechanism not considered here would have to come into play.

Acknowledgments

This work was funded by National Institutes of Health award MH59932. Wendy Weisman was funded by a National Science Foundation REU supplement to award IIS-9978403.

References

Ben-Yishai, R., Hansel, D., & Sompolinsky, H. (1997). Traveling waves and the processing of weakly tuned inputs in a cortical network module. Journal of Computational Neuroscience, 4, 57-77.

Brown, J. E., & Skaggs, W. E. (2002). Concordant and discordant coding of spatial location in populations of hippocampal CA1 pyramidal cells. Journal of Neurophysiology, 88(4), 1605–1613.

Collett, T., Cartwright, B. A., & Smith, B. A. (1986). Landmark learning and visuo-spatial memories in gerbils. Journal of Comparative Physiology A, 158, 835-851.

Deneve, S., Latham, P., & Pouget, A. (2002). Efficient computation and cue integration with noisy population codes. Nature Neuroscience, 4, 826–831.

Doboli, S., Minai, A. A., & Best, P. J. (2000). Latent attractors: A model for context-dependent place representations in the hippocampus. Neural Computation, 12(5), 1003–1037.

Droulez, J., & Berthoz, A. (1991). A neural network model of sensoritopic maps with predictive short-term memory properties. Proceedings of the National Academy of Sciences, USA, 88, 9653-9657.

Fenton, A. A., Czismadia, G., & Muller, R. U. (2000a). Conjoint control of hippocampal place cell firing by two visual stimuli: I. The effects of moving the stimuli on firing field positions. Journal of General Physiology, 116(2), 191–209.

Fenton, A. A., Czismadia, G., & Muller, R. U. (2000b). Conjoint control of hippocampal place cell firing by two visual stimuli: II. A vector-field theory that predicts modifications of the representation of the envrionment. Journal of General Physiology, 116(2), 211–221.

Goodridge, J. P., & Touretzky, D. S. (2000). Modeling attractor deformation in the rodent head direction system. Journal of Neurophysiology, 83(6), 3402–3410.

Hartley, T., Burgess, N., Lever, C., Cacucci, F., & O'Keefe, J. (2000). Modeling place fields in terms of the cortical inputs to the hippocampus. Hippocampus, 10, 369–379.

Káli, S., & Dayan, P. (2000). The involvement of recurrent connections in area CA3 in establishing the properties of place fields: a model. Journal of Neuroscience, 20(19), 7463–7477.

Knierim, J. J. (2002). Dynamic interactions between local surface cues, distal landmarks, and intrinsic circuitry in hippocampal place cells. Journal of Neuroscience, 22(14), 6254–6264.

Lukashin, A. V., Amirikian, B. R., Mozhaev, V. L., Wilcox, G. L., & Georgopoulos, A. P. (1996). Modeling motor cortical operations by an attractor network of stochastic neurons. Biological Cybernetics, 74(3), 255-261.

Muller, R. U., Kubie, J. L., & Ranck, J. B., Jr. (1987). Spatial firing patterns of hippocampal complex-spike cells in a fixed environment. Journal of Neuroscience, 7, 1935-1950.

Pinto, D. J., Brumberg, J. C., Simons, D. J., & Ermentrout, G. B. (1996). A quantitative population model of whisker barrels: Re-examining the Wilson-Cowan equations. Journal of Computational Neuroscience, 3(3), 247-264.

Pouget, A., Deneve, S., & Duhamel, J.-R. (2002). A computational perspective on multi-sensory spatial representations. Nature Reviews Neuroscience, 3, 741–747.

Redish, A. D. (1999). Beyond the cognitive map: From place cells to episodic memory. Cambridge, MA: MIT Press.

Redish, A. D., Elga, A. N., & Touretzky, D. S. (1996). A coupled attractor model of the rodent head direction system. Network: Computation in Neural Systems, 7(4), 671-685.

Redish, A. D., & Touretzky, D. S. (1998). The role of the hippocampus in solving the Morris water maze. Neural Computation, 10(1), 73-111.

Samsonovich, A., & McNaughton, B. L. (1997). Path integration and cognitive mapping in a continuous attractor neural network model. Journal of Neuroscience, 17(15), 5900-5920.

Seung, H. S. (1996). How the brain keeps the eyes still. Proceedings of the National Academy of Science, USA, 93, 13339–13344.

Tanila, H., Shapiro, M. L., & Eichenbaum, H. (1997). Discordance of spatial representation in ensmbles of hippocampal place cells. Hippocampus, 7(6), 613-623.

Taube, J. S., Muller, R. I., & Ranck, J. B., Jr. (1990a). Head direction cells recorded from the postsubiculum in freely moving rats. II. Effects of environmental manipulations. Journal of Neuroscience, 10, 436-447.

Taube, J. S., Muller, R. I., & Ranck, J. B., Jr. (1990b). Head direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis. Journal of Neuroscience, 10, 420-435.

Zhang, K. (1996). Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: A theory. Journal of Neuroscience, 16(6), 2112-2126.

Statistic Measured	Standard Configuration	Cards Apart by 25°	Cards Together by 25°
mean peak activation level	1.5374	1.3852	1.3517
standard deviation	0.0889	0.0739	0.0860
maximum peak activation	1.9733	1.7993	1.7554
minimum peak activation	1.3252	0.9707	0.9430

Table 1

Distribution of peak activation levels (in dimensionless units) for all 1129 place cells within the arena in the attractor network model, with the cue cards in standard position or rotated apart or together by 25°.

.

Figure Captions

Figure 1. Calculated displacement vectors when cards are rotated closer together by 25°. (a) Rotational displacement D_{rot} from Equation 3. (b) Translational displacement D_{trans} from Equation 4. (c) Total displacement D_{tot} from Equation 5.

Figure 2. Calculated displacement vectors when cards are rotated apart by 25°. (a) Rotational displacement D_{rot} from Equation 3. (b) Translational displacement D_{trans} from Equation 4. (c) Total displacement D_{tot} from Equation 5.

Figure 3. Maximum likelihood estimation of firing field movement when the cards are rotated closer together by 25°, using (a) distance to landmarks, (b) angles between pairs of landmarks, (c) allocentric bearings to landmarks, or (d) a combination of distance and angle information.

Figure 4. Maximum likelihood estimation of firing field movement when the cards are rotated apart by 25°, using (a) distance to landmarks, (b) angles between pairs of landmarks, (c) allocentric bearings to landmarks, or (d) a combination of distance and angle information.

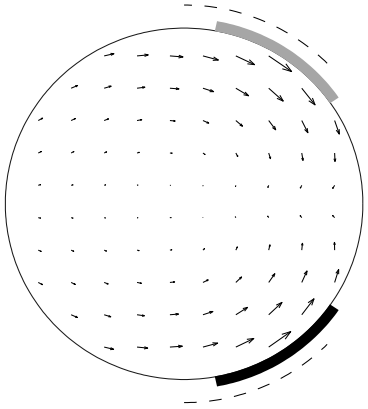
Figure 5. Feature detector activity projected onto the place cell grid with the cue cards in standard position. Left: distance-based detectors produce gaussian arcs centered on the four card edges. Right: allocentric bearing-based detectors produce gaussian cones emanating from the card edges.

Figure 6. Map distortion in the attractor model with the cards rotated closer together by 25°.

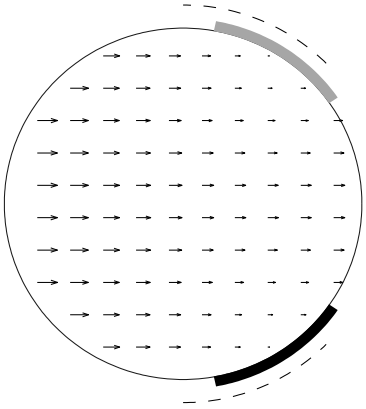
Figure 7. Map distortion in the attractor model with the cards rotated farther apart by 25°.

Figure 8. Firing fields of three place cells with cue cards (left) in standard position, (center) rotated apart by 25°, and (right) rotated together by 25°. The top row shows a cell whose field, when the cards are in standard position, is located at the center of the arena. The cell in the middle row has a field near the black cue card, and the cell in the bottom row has a field at the west edge of the arena.

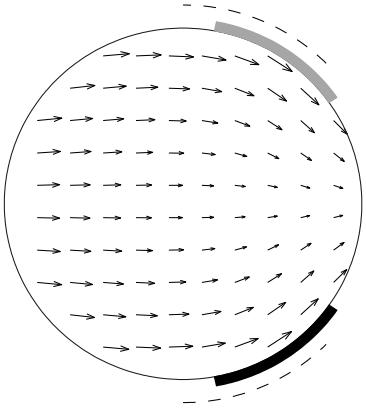
Rotational Displacement



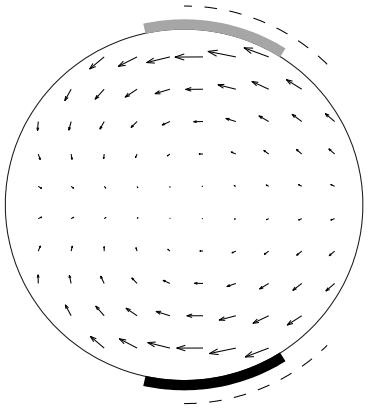
Translational Displacement



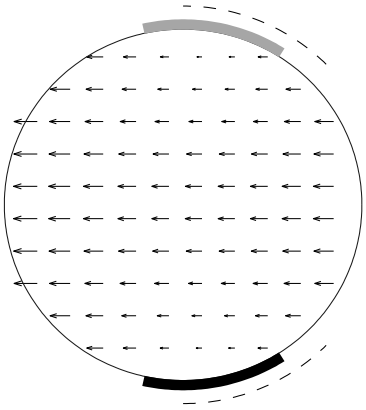
Total Displacement



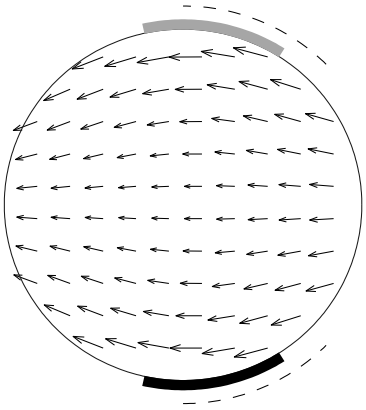
Rotational Displacement



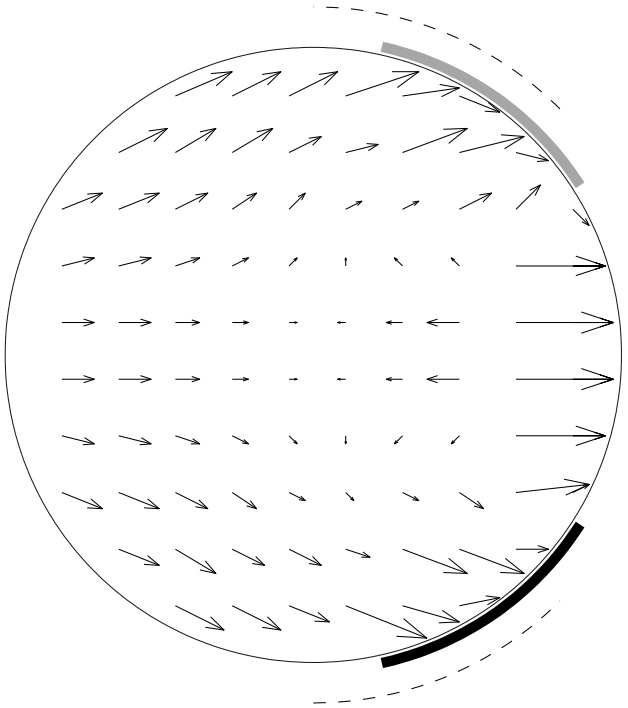
Translational Displacement



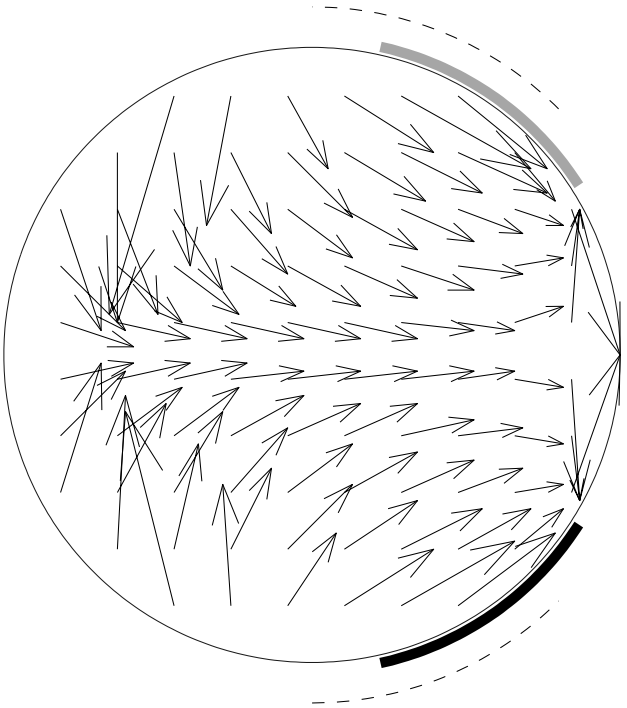
Total Displacement



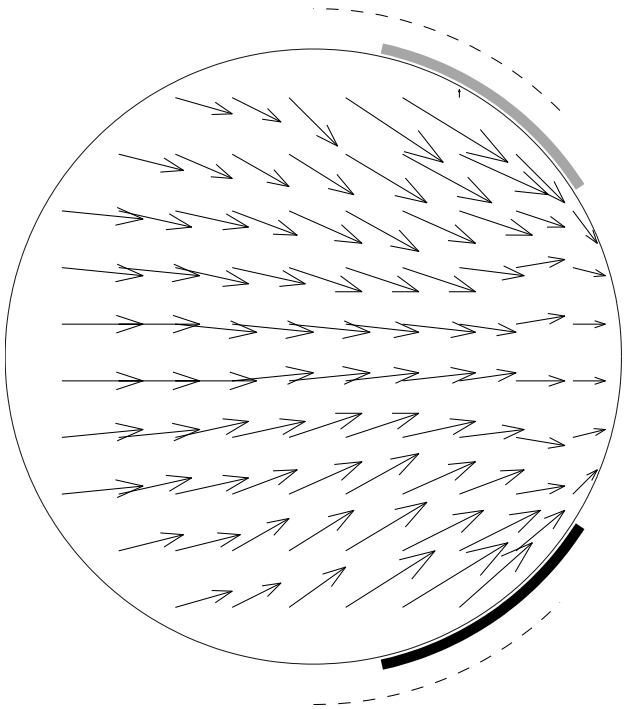
Distance



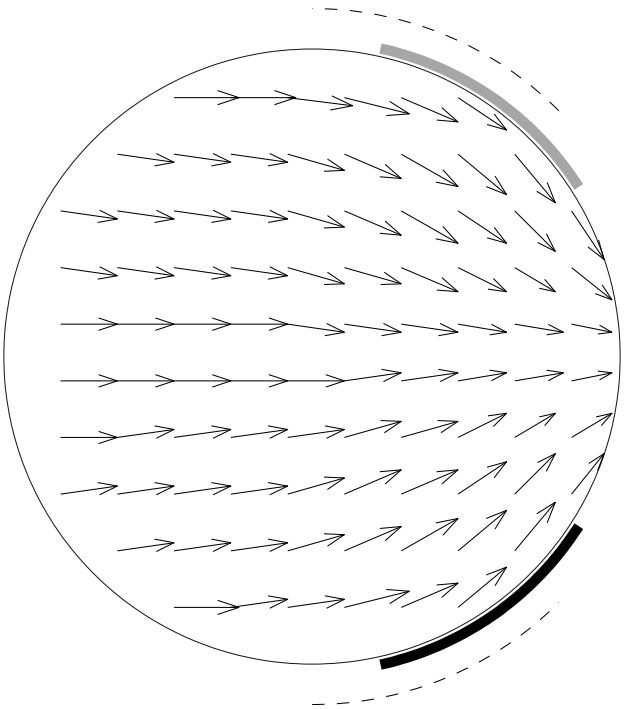
Angle



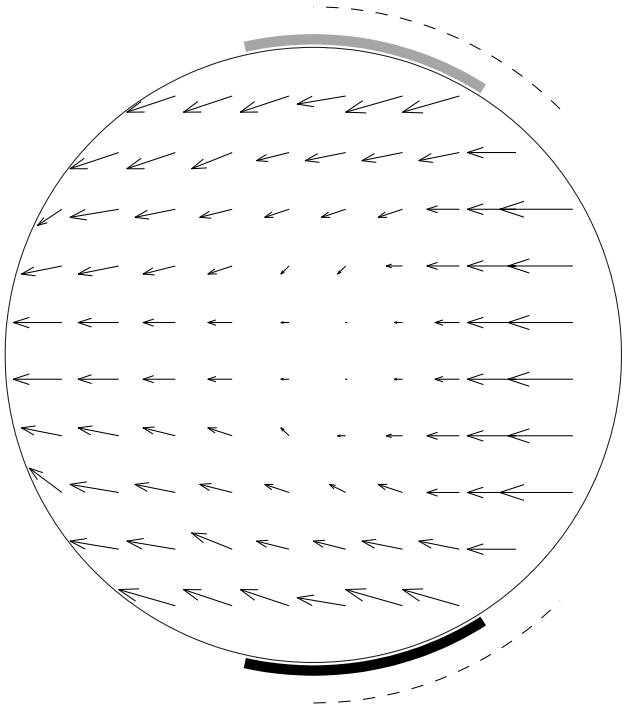
Bearing



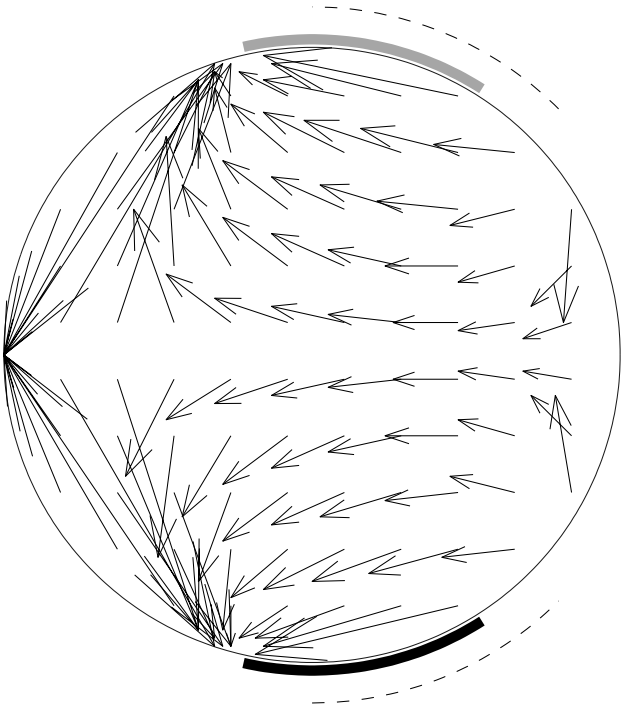
Distance + Angle



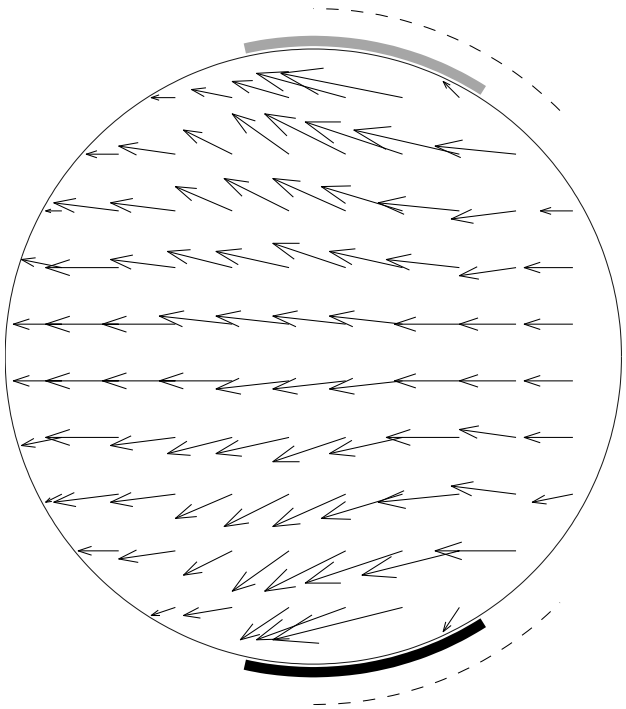
Distance



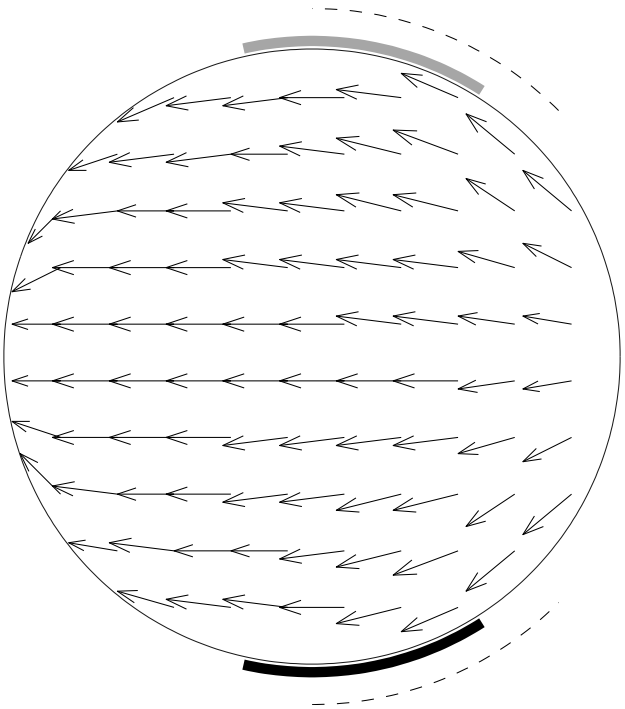
Angle



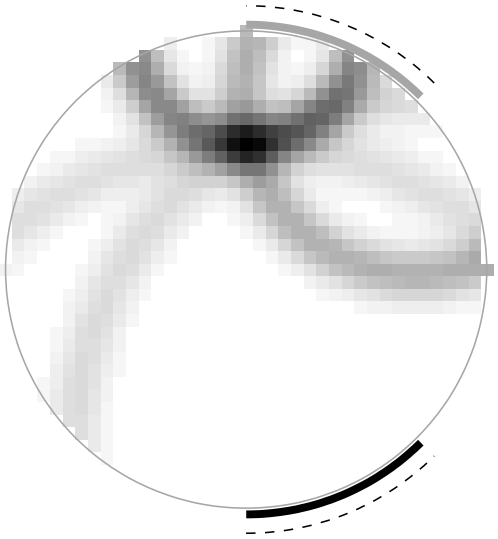
Bearing



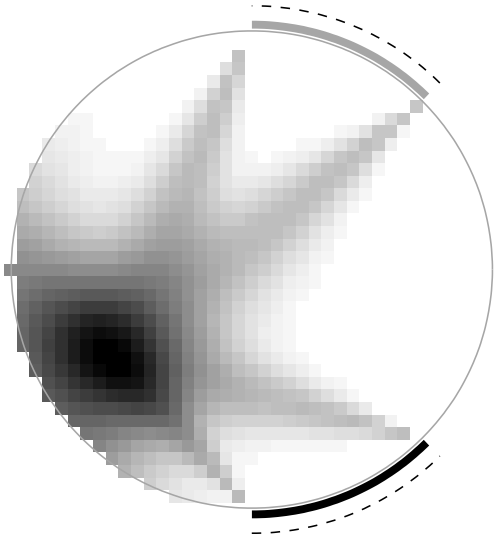
Distance + Angle



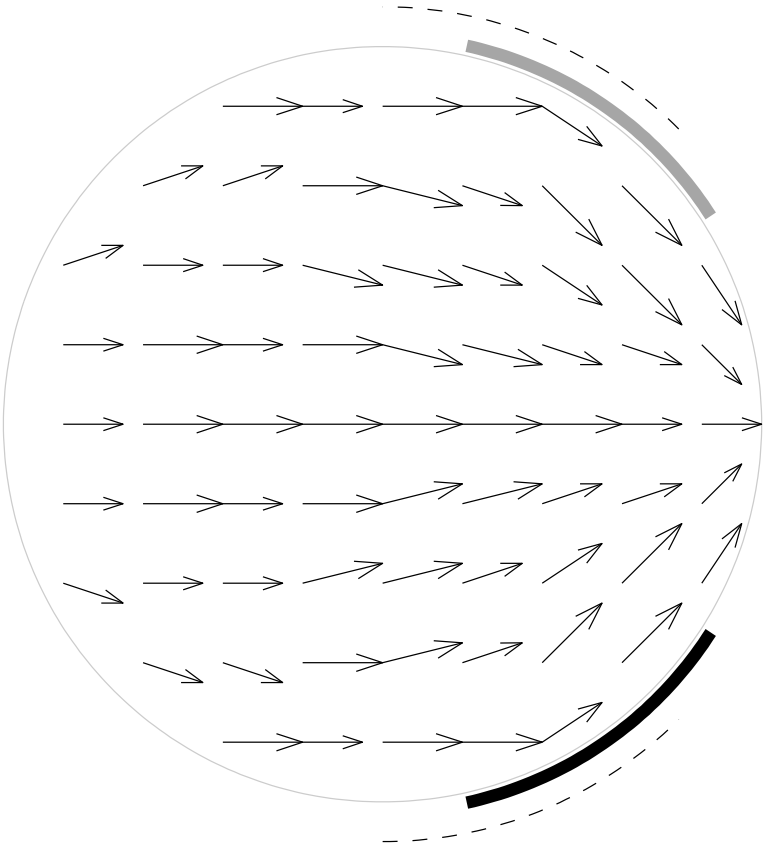
Distances



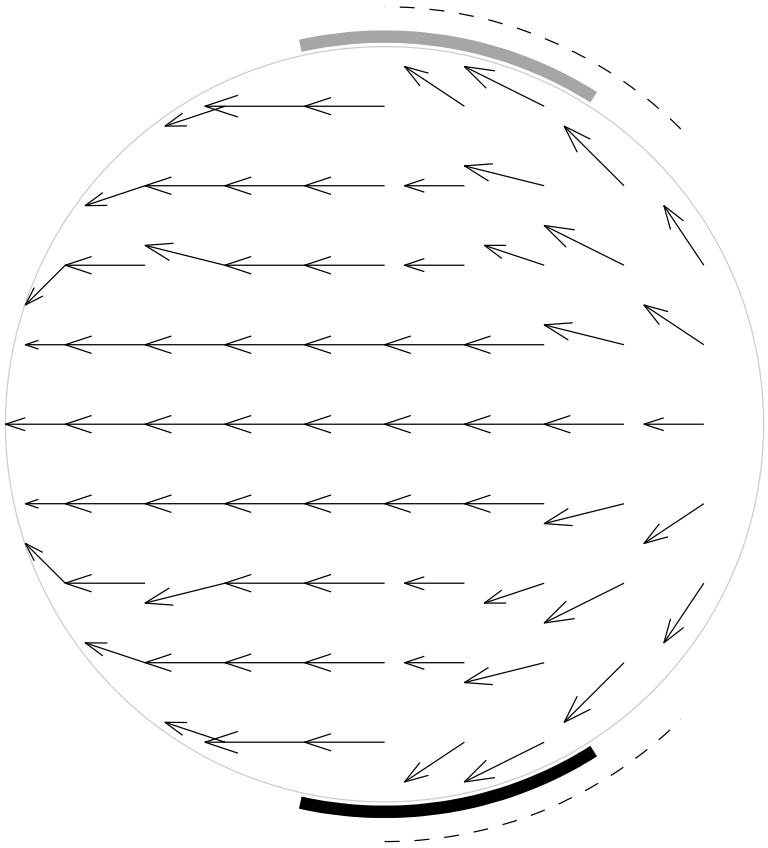
Bearings



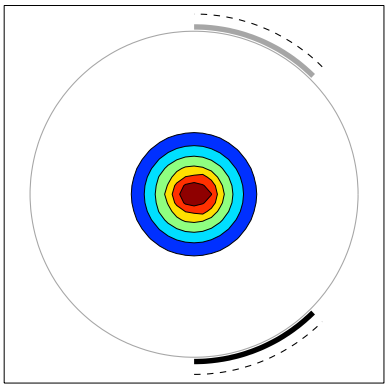
Together by 25^{o}



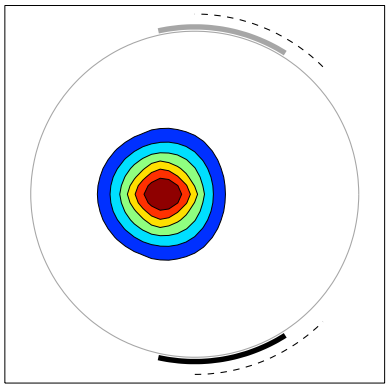
Apart by 25^{o}



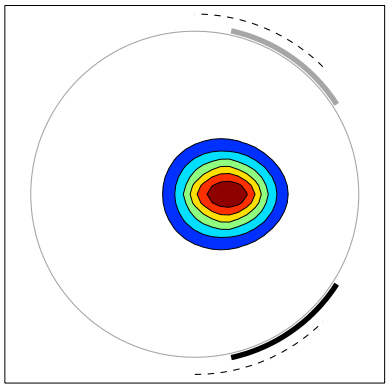
Peak 1.4217



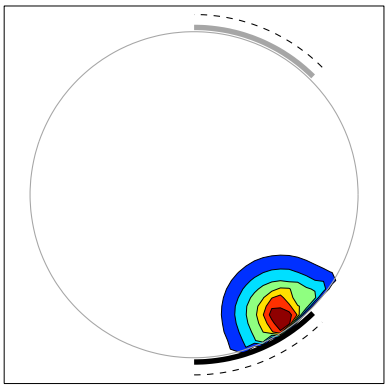
Peak 1.3518



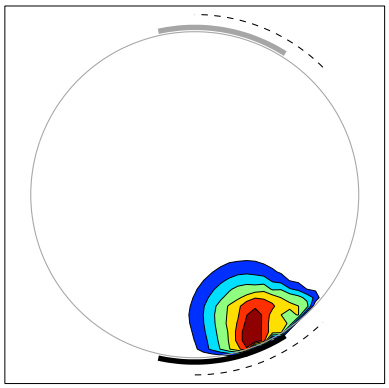
Peak 1.3517



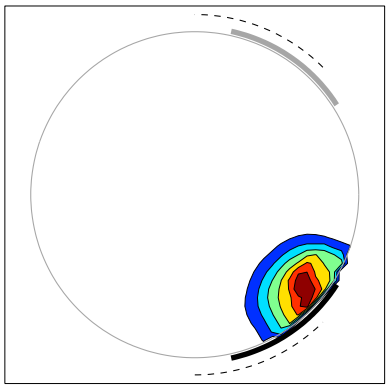
Peak 1.6629



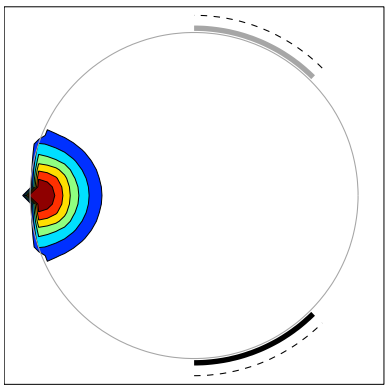
Peak 1.5143



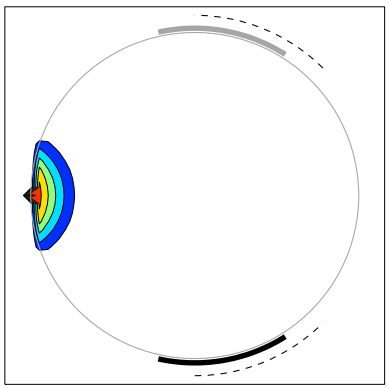
Peak 1.4049



Peak 1.3653



Peak 1.3551



Peak 1.0738