Spatial coding and attractor dynamics of grid cells in the entorhinal cortex

Yoram Burak

Recent experiments support the theoretical hypothesis that recurrent connectivity plays a central role within the medial entorhinal cortex, by shaping activity of large neural populations, such that their joint activity lies within a continuous attractor. This conjecture involves dynamics within each population (module) of cells that share the same grid spacing. In addition, recent theoretical works raise a hypothesis that, taken together, grid cells from all modules maintain a sophisticated representation of position with uniquely large dynamical range, when compared with other known neural codes in the brain. To maintain such a code, activity in different modules must be coupled, within the entorhinal cortex or through the hippocampus.

Introduction

Grid cells in the dorso-caudal medial entorhinal cortex (dMEC) provide a captivating glimpse into the inner workings of neural computation in the brain. Their geometric firing patterns in the animal’s environment, and their synaptic projections to place cells in the hippocampus, link grid cells to spatial computation and representation. Their discovery in the entorhinal cortices of rats [1], mice [2], bats [3], monkeys [4], and most recently in humans [5], suggests that their computational role was highly conserved during evolution.

Here I review how theoretical models have recently contributed to the understanding of activity in dMEC, focusing on two topics: the properties of grid cell activity as a code for position, and network dynamics within modules — with a focus on attractor models. Studies related to development, and to the precise timing of spikes relative to the theta oscillation in rodents, are mentioned only briefly, within these contexts. Several salient properties of grid cell activity are summarized in Box 1 and in Fig. 1.

The grid cell code for position

The periodic spatial activity pattern of grid cells raises questions about the representation of position in the entorhinal cortex. Recently, it has become clear [6,7] that grid cells are functionally grouped into large sub-populations (modules), in which all cells share identical spacing and orientation, and differ in the phase of their periodic response (Box 1 and Fig. 1). Because of the periodic nature of the grid response, the neural activity of all cells within a module encodes position up to an arbitrary discrete displacement (Figs. 1 and 2a), and may thus seem redundant and ambiguous. However, the activity of several modules, taken together, can encode a large range of positions without ambiguity [7–10] (Fig. 2b). Below, we refer to the range of unambiguously represented locations as the capacity of the code, and to the accuracy in which position can be inferred from the neural activity as its resolution. Increasing the number of neurons by adding modules typically results in exponential increase of the capacity, whereas adding neurons in a place-cell representation leads only to linear increase of the capacity [9]. Thus, the grid cell code has a much larger representable range than a place-cell code with the same resolution and number of neurons. This range can exceed by far the behavioral foraging range in rodents [9].

What are the possible benefits of having such a huge capacity? One possibility is that the huge capacity confers the neural code with error-correcting capabilities [9,11]; for example, a severely corrupted phase in one of the modules will, in general, throw off the represented position to an unrealistic location, allowing for the error to be identified and corrected based on phases of the other modules. Another possible benefit of the large capacity is that grid cells may enable global remapping in the hippocampus [12]: independently shifting the phases in different modules, or rotating the axes of the grids relative to each other, can produce distinct representations for different environments by place cells, with negligible likelihood for overlap between the neural codes — a property that is inherited by place cells if their firing field is determined by inputs from the entorhinal cortex [7–9,13,14].

Coupling between modules

The neural representation in each module is probably prone not only to occasional severe errors, but also to
Box 1 Essential properties of grid cell activity

1. Grid cells fire in multiple locations within the animal’s environment. These locations are approximately arranged on vertices of a triangular lattice that tiles the plane [1]. Grid spacing of a given cell remains similar in different environments [1] (Fig. 1a).

2. Grid cells in dM EC are grouped in discrete modules [6**,29]. Cells in the same module share the same grid spacing, orientation, and precise structure of the lattice [1,6**,28**], but differ in phase (Fig. 1b). All phases are (probably) densely represented.

3. Cells from modules with increasing spacing are found in increasingly ventral locations along the dorsoventral axis of dM EC [1,6**] (Fig. 1c). In rats, spacings range from approximately 30 cm up to (at least) several meters [21]. The discrete spacings form, approximately, a geometric series with a ratio $\sim$ 1.4 between successive modules, averaged across module pairs and animals [6**].

4. Conjunctive cells: A sub-population of grid cells is selective also to head direction [31]. The activity of many cells in the entorhinal cortex is also modulated by the animal’s velocity.

5. Phase precession: In rodents, spike timing is modulated temporally, in synchrony with the LFP theta oscillation. The timing of spikes gradually recedes relative to the LFP theta oscillation as the animal crosses a firing field [22,24], similarly to place cells in the hippocampus.

6. Grid phases and orientations are closely matched when repeatedly placing the same animal within a given environment. Different animals often express similar grid orientations within the same environment (T. Stensola et al., abstract 769.15/JJ34, Neuroscience 2013, San Diego 2013).

The neural circuitry can potentially solve this problem if the modules are coupled to each other, in a way that prevents them from drifting independently, thus suppressing the independent components of the noise (Fig. 2c). Such an interlocking mechanism might be implemented through feedback from place cells in the hippocampus, as recently proposed in [16**], or perhaps by connectivity within dM EC. Hippocampal inputs may be instrumental also in resetting the grid cell representation based on sensory inputs [17], and in correcting large errors in individual phases. Overall, coupling the modules through the hippocampus may result in a neural encoding of position that achieves exquisite stability, alongside an exponentially large capacity [16**].

Optimization of the grid spacings

Several recent works [18**,19,20**] considered how the resolution of the grid cell code depends on the number of neurons and the allocation of grid spacings to modules, while keeping the maximal grid spacing fixed. In Refs. [18**,19] the resolution was estimated by evaluating the Fisher information (a lower bound on precision in which position can be decoded from the spiking activity), as well as the performance of maximum-likelihood decoders, while assuming that grid cells fire independently as Poisson spike generators. The Fisher information was maximized by choosing grid spacings that form a geometric series. Under these conditions, the resolution of the code is related to the smallest grid spacing, and is thus exponentially smaller than the maximal spacing. Tiling an area comparable to the largest grid spacing with a place cell code of the same resolution as the grid cell code, would require an exponentially large number of neurons.

Recent characterization of grid spacings in rats [6**] points to an approximately fixed ratio between successive grid spacings, in accord with the above optimization principle. Moreover, simple scaling arguments [20**], somewhat different from those proposed in [18**,19], further support the optimality of a geometric ratio of grid spacings.
Figure 2

(a) All neurons in a single module respond to position with the same grid periodicity. Hence, their joint activity represents position only up to an arbitrary discrete displacement. (b) In a one dimensional (1D) analog of the grid cell code, all neurons in a single module represent position as a 1D phase, relative to the grid spacing. The thin black bars represent the encoded phase in each module (1st, 2nd, and 3rd), and the red, blue, and green bars represent the possible locations compatible with the activity in each module. The width of the colored bars represents schematically the accuracy of readout. Combining the activity from all neurons in the three modules can greatly reduce the ambiguity. Here, the joint activity from the three modules is compatible with a single location (yellow bar) within the range shown. The full range over which positions can be unambiguously inferred is the capacity of the code, whereas the width of the yellow bar roughly represents the resolution of the code. (c) In the absence of sensory cues, the phase represented in each module will probably drift gradually, relative to the value that matches the true position of the animal. These drifts may occur independently in different modules (top). Alternatively, an interlocking mechanism might constrain these drifts to occur in unison, in compatibility with motion within the local environment of the animal (bottom).

spacings, and predict that to minimize the number of encoding neurons, while keeping the resolution and maximal grid spacing fixed, this ratio should be $\sim 1.4$ or $\sqrt{2} \approx 1.65$ in 2D (depending on details of the decoding scheme used to extract phase from the neural activity) — in close agreement with the observed value [6**].

From a theoretical perspective it is interesting to consider, as an optimization goal for a neural code of a continuous variable, its dynamic range: the ratio between the capacity and the resolution. The works discussed above [9,16**,**18*,19,20*] demonstrate that the dynamic range of the grid-cell code can be exponential in the number of neurons. However, optimizing the dynamic range is a different goal from the one considered in [18*,20*], where resolution was minimized while keeping the largest grid spacing fixed — and these two objectives may perhaps lead to different predictions for the optimal sequence of grid spacings.

Within all theoretical treatments discussed above, the resolution of the grid-cell code is essentially set by the lowest grid spacing, which is known to be approximately 30 cm in rats (further reduction of the resolution is only linear in the number of neurons). It is less clear how the capacity depends on the precise grid spacings and orientations. If the ratio between successive spacings is precisely $\sqrt{2}$ in rats [6**] then many of the spacings are integer multiples of each other, which in 1D would severely limit the capacity of the code [9]. However, small deviations from this ratio might be sufficient to restore the large capacity of the code. Moreover, in 2D the capacity depends also on the relative grid orientations in different modules. Thus, obtaining a well grounded estimate of capacity in rats requires further experimental and theoretical analysis. Even if the grid cell code in rats can represent an exponentially large range of positions relative to the maximal grid spacing ($\sim 10$ m [6**,21]), it is unknown whether the brain makes full use of this large capacity — an important question from the biological and behavioral perspectives.

Two recent works addressed spike timing of grid cells, relative to the LFP theta cycle in rodents. Spike timing of individual cells, recorded in rats running through linear tracks [22], were found to contribute more to Fisher information on position than spike counts [23*]. Phase precession was recently demonstrated also in open environments [24]. These observations should be taken into account when assessing the resolution of the code. Qualitatively, it is unlikely that they affect conclusions about the exponential capacity of the representation [9,16**,**18*], because the noise observed in the temporal coding is largely independent in subsequently visited firing fields [23*].

Attractor dynamics of grid cell activity

The main goal of this section is to review the interplay between theoretical ideas and experiment, since much of the progress in the past two years has been at this interface. Theoretical models for emergence of grid-like firing were recently reviewed in [25*,26] (see also [11]). In this section we focus on populations of cells that share the same grid spacing and orientation.

Does grid-cell activity lie within a two-dimensional attractor manifold?

Attractor models of grid cell activity [7,8,15,27] propose that the activity of multiple grid cells is tightly
coordinated by recurrent connectivity, such that it lies within a two-dimensional attractor (a continuous manifold of stable steady states). An example is sketched in Fig. 3: cells are arranged on a two-dimensional layer, and ‘mexican-hat’ connectivity between them leads to activity patterns with grid-like structure within the neural sheet (Fig. 3a). All translations of the pattern correspond to possible steady states of the dynamics. Sensory inputs, in conjunction with inputs representing self-motion, select one of these phases in correspondence with motion in the environment. As a result, single-neuron responses form a grid pattern as a function of the animal’s position (Fig. 3b). Different locations in space, whose separation matches the discrete periodicity of the grid are represented by the same population activity pattern.

Attractor models were partly motivated by measurements from nearby grid cells, recorded from the same tetrode, which display identical grid spacing and orientation, but different spatial phases [1,3*] — a feature that arises naturally in these models [7,8,15]. Recent analysis and experiments confirmed a number of other salient predictions of these models [15]: grid spacings and orientations are discretized in modules [6**]; deviations from a precise triangular-lattice are common to all cells within a module [6**,15,28**]; transient modifications in the grid structure (for example, during rescaling of the environment [29]) are tightly correlated in cells from the same module [6**,28**]; and the relative spatial phases in cell pairs are remarkably stable [28**]. Inputs from the hippocampus are not a likely source for these precise constraints, since they persist even in novel environments, and under hippocampal remapping [28**]. Each of the above predictions, if refuted, would have posed a critical challenge for the attractor hypothesis, whereas models based on feed-forward inputs to dMEC alone do not explain or predict these features.

Recurrent synaptic connectivity in dMEC

The evidence in support of a low-dimensional attractor suggests a central role for recurrent connectivity in dMEC. An early study in slice preparations [30] found extensive excitatory transmission within layers III and V of the entorhinal cortex. However, no monosynaptic transmission was found between excitatory neurons in layer II, where pure grid cells are the most prevalent [26,31]. More recent work [32*,33*] confirmed the latter observation, but found extensive disynaptic connectivity between stellate cells in layer II, mediated by inhibitory interneurons.

Attractor models [7,8,15,27] can be formulated with purely inhibitory synaptic transmission, as shown explicitly in Ref. [15]. The developmental course of the transmission between stellate cells in layer II and inhibitory interneurons roughly matches the maturation of adult-like grid cell activity [32*,34,35], in agreement with the hypothesis that these recurrent connections contribute to the grid-like response. However, the strengths of effective inhibitory connections between stellate cells are probably more narrowly distributed than assumed previously in theoretical models (Fig. 3c in [32*]). Attractor models [15,27] can be adjusted to explicitly account for the dynamics of inhibitory and excitatory cells [32*,33*], while complying with the latter observation [32*].

Topographical and functional organization of the synaptic connectivity

Attractor models have mostly assumed synaptic connectivity that depends on the distance between neurons within a 2D surface, leading to smooth variation of the grid phase within this surface. This prediction is challenged by observations of widely variable grid phases in different cells recorded from the same tetrode [1].

A recent model of grid cell development (Widloski et al., unpublished) proposes that the two-dimensional metric is inherited during development of the grid-cell network from the hippocampal representation, which at this phase is driven by sensory inputs. Within this model, spike time dependent plasticity leads to radially symmetric synaptic connectivity within dMEC, in terms of a 2D metric which is unrelated to the location of grid cells on the cortical sheet (see also Refs. [7,36,37]). Another possibility is that an underlying topographical organization does exist, but is not clearly reflected in the location of individual cell bodies.

Low dimensional attractor dynamics might arise from connectivity that differs significantly from the radially
symmetric structure (within a topographical or functional metric) proposed in Refs. [8,15,27] (see, for example Refs. [37,38]). However, to support attractor dynamics, the connectivity must be related to the functional properties of the cells: their grid phases, and possibly their inputs from other brain areas.

In general, it is difficult to map the detailed connectivity in recurrent neural networks [39]. Nonetheless, characterization of functional and anatomical cell types in dMEC and their connectivity [40] will be crucial for constraining the space of possible theoretical models. For example, if inhibitory interneurons mediate the recurrent connections responsible for the attractor dynamics [32*], their synaptic connections with stellate cells are expected to depend on the grid phases of the synaptic partners. In addition, the synaptic inputs from inhibitory cells to excitatory grid cells should contribute significantly to the grid response of excitatory cells, and the inhibitory neurons are expected to display a clear spatial response of their own [33*]. These aspects of the organization in layer II (and similar aspects of the excitatory connectivity in deeper layers [30]) are not yet characterized.

The functional coupling between cells within a module is likely reflected in differences between inter-module and intra-module connectivity, another aspect of anatomy in dMEC that has not yet been characterized experimentally (an issue of importance also for the global properties discussed in the previous section, “The grid cell code for position”).

Is grid-cell activity driven by slowly varying synaptic inputs, or by membrane potential oscillations?

Most attractor models rely on slow, asynchronous modulation of synaptic inputs for maintenance of the attractor state, whereas other models hypothesize that grid cell activity is modulated by interference of temporally oscillating currents ([41] and references therein; [42,43]). To distinguish between these scenarios, whole cell recordings were recently collected from grid cells in freely behaving mice, navigating through a virtual linear track [44**,45**]. Both studies concluded that grid cell activity is primarily related to the slow modulation of the membrane potential, as expected within attractor models, while observing only weak correlation with the amplitude of sub-threshold membrane oscillations. It is conceivable, however, that interference takes part in the distal dendritic tree, and is not apparent in the membrane potential at the soma. A separate challenge to oscillatory interference models has arisen from the discovery of grid cells in bats [37], without continuous theta oscillations [37,46] (but see [25,47,48,49] and the author’s response).

The studies discussed above [44**,45**] do not unequivocally establish that the confinement of grid cell firing activity to a 2D attractor arises within dMEC, because the source of the slowly varying currents is not yet identified. In principle, these currents could arise from feed-forward inputs to dMEC (e.g., from the hippocampus [36] or from cells with stripe-like responses [50,51]). Within the attractor picture, they are expected to arise primarily from recurrent connectivity.

Theta oscillations and phase precession are salient features in the rodent dMEC [22,24,44**,45**]. Several recent works [38,45**,52,53] considered how phase precession can be explained within the attractor picture, in the presence of intrinsic or extrinsic theta drive in dMEC (see also Ref. [33*]).

Conclusions

Experiments in the past two years tested key theoretical hypotheses on the mechanisms responsible for grid cell dynamics. Many of these studies resulted in support for the attractor hypothesis, based on the dynamics of neural activity in pairs of cells [6**,28**], as well as the dynamics of the membrane potential in slice preparations [33*] and in freely behaving mice [44**,45**]. The latter experiments suggest that the attractor state is maintained (at least, primarily) through slowly varying synaptic inputs (see also Ref. [37]). Another line of evidence in support of the attractor models comes from the existence of extensive recurrent connectivity between stellate cells in layer II [32*,33*]. However, the role of recurrent connections in shaping the grid cell response is not yet established.

While several experimental studies reported on agreement with specific attractor models [6**,33*,44**,45**], the main support so far is for the broader notion, that network dynamics within modules is restricted to lie (anywhere) within a two-dimensional attractor [6**,28**], likely through recurrent connectivity within dMEC. On the other hand, many details in existing models may or may not be valid: For example, very little is known from experiments about mechanisms responsible for updating the attractor state based on sensory inputs and self motion. In the foreseeable future, new techniques will perhaps enable direct tests for some of the theoretical ideas, for example by selectively manipulating the inputs to dMEC that encode head direction and velocity, or the activity of conjunctive cells within dMEC.

The interplay between theoretical and experimental work demonstrates that the entorhinal cortex, together with the hippocampus, provides a rich arena for mechanistic studies of neural computation in the association cortex, at a detailed level that was previously seen mostly in sensory and motor areas.

The grid cell code

Recent theoretical studies [9,16**,18**,20] have proposed that the entorhinal cortex, perhaps jointly with the hippocampus, maintains a sophisticated, highly precise system
for encoding position, with uniquely large dynamical range in comparison with other known coding schemes in the brain. Important theoretical questions remain unanswered, for example: if the entorhinal cortex is responsible for idiothetic path integration (update of the animal’s estimate of position based on self motion), it should be possible to extract from the grid cell code a distance and angle to a known landmark such as the animal’s initial position (a ‘homing vector’). A biologically plausible scheme for implementing this calculation in neural circuitry is not known.

Regardless of the role of grid cells in path integration, the theoretical models surveyed here rely on key assumptions, and raise important questions that can be clarified only by experiments: Do the grid responses, observed in relatively small enclosures, persist also in much larger open environments? (advances in wireless recording [46,54] and virtual environments [5,44**,**45**] may aid in addressing this question). To what extent do insights obtained from open environments, carry through to environments with internal borders [55], or to environments with varying topology or metric [56]; to motion in three dimensions [46,57]; and to the encoding of a position other than the animal’s location [4]? Are the phases in different modules restricted, within a particular environment, to co-vary in a single 2D sub-space, even in the absence of sensory cues? If so, is this functional restriction dependent on the hippocampal representation, as proposed in [16**]? Existence of an such an interlocking mechanism would pose strong constraints on the dynamics of simultaneously recorded cells from different modules in the presence and absence of sensory cues, as well as mechanistic constraints on the neural circuitry.

Acknowledgements

I thank Dori Dzedikman, Ila Fiete, Neta Ravid, Haim Sompolinsky, and Nachum Ulanovsky for many helpful comments on the manuscript, and Hannah Monyer for a helpful discussion. Support from the Gatsby Charitable Foundation is gratefully acknowledged.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest


First report of grid cells in the entorhinal cortex of bats, crawling on a 2D surface. Grid cell activity has all the salient properties observed in rodents except for continuous modulation in the theta band.


Evidence that grid cells are organized in tightly coupled sub-populations. Grid spacings are shown for the first time to be discretized in modules (see also earlier evidence in Ref. [28**]). Within modules, grid parameters except phase are matched precisely, and cells respond coherently during grid recoding in a compressed environment, as predicted theoretically [10]. Cells in different modules can respond differently under the latter condition.


Theoretical proposal for suppressing noise through reciprocal connectivity between grid cells and place cells in CA1. The outcome is a robust representation of position with exponentially large capacity, coupling all modules in dMEC and the hippocampus.


Resolution of the grid cell code is analyzed via the Fisher information, under an assumption that grid cells fire as Poisson spike generators. An optimization principle predicts that grid spacings should form a geometric series, leading to exponential dependence of the resolution on the number of modules, when maintaining the largest grid spacing fixed.


A simple and elegant optimization principle predicts that the ratio between subsequent grid spacings should lie between 1.4 and 1.7 in 2D, in striking similarity to the biological observation [28**].


In rats running through a linear track, phase precession is more pronounced at the single neuron level than at the population level, and is informative about position in single trials — even more than the spike count.


An extensive survey of theoretical models for grid cell dynamics (see also Ref. [26]). Computer code, implementing many of the models, is provided as supplementary material.


Evidence for low-dimensional attractor dynamics: transient and sustained deviations from a perfect triangular-lattice grid response, in pairs of cells from the same module are precisely matched, as entailed by attractor models (see also earlier evidence in [13]). Velocity driven perturbations in the neural response (other than the update of position) decay rapidly, as expected.


Recurrent inhibitory circuitry as a mechanism for grid formation. *Nat Neurosci* 2013, 16:318-324.

Simultaneous whole cell recordings from rat entorhinal slices in groups of two or more neurons reveal that stellate cells in layer II are highly interconnected, mainly via inhibitory interneurons. This connectivity develops after eye opening, at least up to postnatal day P28. The coupling between pairs of cells has a bimodal, ‘all or none’ distribution.


Experimental study of nested gamma oscillations within the theta cycle, and their dependence on recurrent inhibitory connectivity in dMEC, using optogenetic stimulation and whole cell recordings in entorhinal slices from mice. Nested gamma oscillations arise also in a continuous attractor model of dMEC, which explicitly includes excitatory and inhibitory neurons and a uniform, theta modulated drive.


The emergence of grid cells: intelligent design or just adaptation? *Hippocampus* 2008, 18:1256-1260.


Computer code, implementing many of the models, is provided as supplementary material.


These two concurrent studies measured the membrane potential of grid cells in freely behaving mice, in a virtual environment. In the second work, a multi-compartment model of a stellate cell was used to evaluate the expected dynamics of the membrane potential, within several published models of grid cells dynamics (see also main text).


Spiking neurons in a hierarchical self-organizing map model can learn to develop spatial and temporal properties of entorhinal grid cells and hippocampal place cells. *PLoS ONE* 2013, 8:e60599.


A wireless multi-channel neural amplifier for freely moving animals. *Nat Neurosci* 2011, 14:263.


Anisotropic encoding of three-dimensional space by place cells and grid cells. *Nat Neurosci* 2011, 14:1182-1188.