

## Grid cells on the ball

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

J. Stat. Mech. (2013) P03013

(<http://iopscience.iop.org/1742-5468/2013/03/P03013>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

### Download details:

IP Address: 132.239.1.231

The article was downloaded on 05/04/2013 at 16:44

Please note that [terms and conditions apply](#).

## Grid cells on the ball

Federico Stella<sup>1</sup>, Bailu Si<sup>2</sup>, Emilio Kropff<sup>3</sup> and  
Alessandro Treves<sup>1</sup>

<sup>1</sup> SISSA, Cognitive Neuroscience, via Bonomea 265, I-34136 Trieste, Italy

<sup>2</sup> Weizmann Institute of Science, Department of Neurobiology, Rehovot, Israel

<sup>3</sup> Laboratory of Neuronal Plasticity, Leloir Institute (IIBBA-CONICET),  
Buenos Aires, Argentina

E-mail: [fstella@sissa.it](mailto:fstella@sissa.it), [bailu.si@weizmann.ac.il](mailto:bailu.si@weizmann.ac.il), [kropff@gmail.com](mailto:kropff@gmail.com)  
and [ale@sissa.it](mailto:ale@sissa.it)

Received 9 July 2012

Accepted 15 November 2012

Published 12 March 2013

Online at [stacks.iop.org/JSTAT/2013/P03013](http://stacks.iop.org/JSTAT/2013/P03013)

[doi:10.1088/1742-5468/2013/03/P03013](https://doi.org/10.1088/1742-5468/2013/03/P03013)

**Abstract.** What sort of grid cells do we expect to see in rodents who have spent their developmental period inside a large spherical cage? Or, in a different experimental paradigm, toddling on a revolving ball, with virtual reality simulating a coherently revolving surround? We consider a simple model of grid firing map formation, based on firing rate adaptation, that we have earlier analyzed when playing out on a flat environment. The model predicts that whether experienced on the outside or inside, a spherical environment induces one of a succession of grid maps realized as combinations of spherical harmonics, depending on the relation of the radius to the preferred grid spacing, itself related to the parameters of firing rate adaptation. Numerical simulations concur with analytical predictions.

**Keywords:** neuronal networks (theory), pattern formation (theory), neural code, computational neuroscience

---

**Contents**

<b>1. Introduction</b>	<b>2</b>
<b>2. A model based on adaptation</b>	<b>3</b>
<b>3. Life on a sphere: a mathematical analysis</b>	<b>4</b>
<b>4. Numerical simulations of grid development</b>	<b>9</b>
4.1. Network model . . . . .	10
4.2. Results . . . . .	11
<b>5. Conclusions</b>	<b>13</b>
<b>References</b>	<b>15</b>

---

**1. Introduction**

The study of navigation in rodents has revealed the presence of specific groups of neurons devoted to the representation of space in the brain, with the most conspicuous among them the place cells [1] in the hippocampus, and the grid cells [2, 3] in the neighboring medial entorhinal cortex (mEC). These neurons are selective for certain portions of the environment: they get activated, and emit spikes, only when the animal occupies one of these so called *fields*. Place cells and grid cells thus share the same behavioral correlate, i.e. the position of the animal in space, but in other respects they are rather different. A place cell is generally selective for only one location, its (often idealized) place field, or for a few locations randomly distributed in the environment. In a different environment, the same place cell will often be inactive. In contrast, a grid cell is active in any environment, and in each environment not in one but in many preferred positions—or grid fields. The spatial distribution of its grid fields is strikingly regular, as they form a periodic triangular grid that covers all the 2D space accessible to the animal.

Another characteristic of grid cell activity is that the representation of space a grid cell provides is almost invariant across different environments. The typical hexagonal tiling appears to be internally generated and to be applied to any environment the rodent visits, with only distinct reference frames to differentiate the activity of an ensemble of cells between different environments, i.e. a common translation and rotation of the grids [4]. Here an ensemble is a group of cells recorded simultaneously and located next to each other in the tissue, although their grid fields are typically not overlapping beyond chance. Whatever the relation between the grids expressed by different cells, it is maintained across environments. These properties are in marked contrast with those of place cell maps in the hippocampus. Place cell representations are multiple, in that such relations are not maintained at all. Considering two cells that are active at the same location in one environment, for example, moving to a new environment, one may find that they are both inactive, or only one of the two is active, but at an unrelated location.

Place cells, particularly those in the CA3 region of the hippocampus, are thought to participate in cell assemblies that can be reactivated by partial cues through attractor dynamics [5, 6]. Such assemblies would be organized in multiple quasi-continuous attractors [7], which might be rapidly acquired during exploration. Grid cells, instead, might participate in a single continuous attractor, which is then imposed to represent whatever environment, and which is established much more slowly, through a gradual process, likely during animal development. According to the most recent experimental data [8, 9], the critical period for such a slow process should be somewhere between the 15th and 35th postnatal days.

## 2. A model based on adaptation

This scenario is described in a model that shows how the grids may emerge spontaneously from disorder under the influence of firing rate adaptation, a very general property of cortical neurons. Such temporal modulation induces a self-organization process that, over time, sculpts the activity of single units to produce a triangular spatial pattern [10].

The gradual process can be observed in simulations of a model network representing the mEC, which receives spatially tuned inputs from the hippocampus or from parahippocampal cortices [10]. The combination of adaptation in mEC units and plasticity on the feed-forward connectivity is enough to generate increasingly regular grid patterns over time scales corresponding to a couple of weeks of rodent development.

The same process, at the single-unit level, can be described in analytical terms as an unsupervised optimization process, if one neglects the collateral interactions that are presumed to align the grids expressed by interacting units [11]. The simplified version of the model that can be analyzed mathematically is very abstract, and does not specify most of the parameters necessary to the simulations. Nevertheless, it indicates which states are the asymptotic ones that should be reached by the system after having evolved for a certain time.

The asymptotic states are defined in terms of a variational principle amounting to the minimization of a cost function of the form

$$H = H_K + H_A = \int d\chi [\nabla\Psi(\chi)]^2 + \gamma \int d\chi \int dt' \Psi(\chi(t))K(t-t')\Psi(\chi(t')), \quad (1)$$

where  $\chi$  is the spatial coordinate and  $\Psi$  represents the firing rate of the neuron across the environment.

The functional is defined based on the hypothesis that the activity of the units reflects only two simple constraints.

- (i) Minimization of the variability of the maps across space, i.e. a preference for smooth maps. Such smoothness is expected to stem from the smoothness of the spatial inputs and the neuronal transfer function. This constraint is expressed in the first term of the functional, the *kinetic* one.
- (ii) Penalization of maps that require a neuron to fire for prolonged periods of time, despite the presence of neuronal fatigue. The second term of the functional, the adaptation term, represents this constraint.

The parameter  $\gamma$  parametrizes the relative importance of the two constraints. In the second term, taking into account the averaging effect of a long run over many trajectories, we

can substitute the time-dependent kernel  $K(\Delta t)$  with an effective space-dependent one,  $K(\chi' - \chi)$ ,

$$H = H_K + H_A = \frac{1}{A} \int_S d\chi [\nabla \Psi(\chi)]^2 + \gamma \frac{1}{A} \int_S d\chi \Psi(\chi) \int_S d\chi' \Psi(\chi') K(|\chi' - \chi|), \quad (2)$$

where we have also made explicit the normalization by the area  $A$  of the environment  $S$ .

In [10] it is shown that the grid pattern exhibited by the mEC neurons minimizes this cost function, as it is the optimal compromise solution (among periodic and fully symmetric solutions) between the two opposite constraints, for a number of reasonable choices of the kernel  $K$ . Since the phenomena described by this model are presumed to extend over a long developmental time, it is expected that the topological properties of the environment in which the animal develops affect the shape and the appearance of the grid maps [12]–[15]. The possibility is currently being explored of running experiments in which rodents develop, during the critical period for grid cell formation, on a revolving sphere [16], with virtual reality provisions that should give them the visual impression of living and running on the surface of a tiny planet—literally a small world. In an independent approach, rats are being raised inside a large comfortable spherical cage, which can be partially revolved from time to time [24]. What grid cell firing maps would we expect to observe after such training protocols, according to the model previously described?

### 3. Life on a sphere: a mathematical analysis

Here we discuss a formulation of the adaptation model on the sphere. This formulation is both experimentally relevant and mathematically treatable. While on a 2D plane the solutions to the minimization problem were formulated in terms of Fourier components, a similar approach can be applied on a sphere in terms of spherical harmonics [17].

Using the following set of coordinates to parametrize the unitary sphere ( $R = 1$ ):

$$x = \sin(\theta) \cos(\varphi), \quad y = \sin(\theta) \sin(\varphi), \quad z = \cos(\theta),$$

we define a function  $\Psi(\theta, \varphi)$  whose values on the sphere represent the firing rate of a unit in the different locations.

We normalize this function such that

$$\Psi(\theta, \varphi) \geq 0, \quad (3)$$

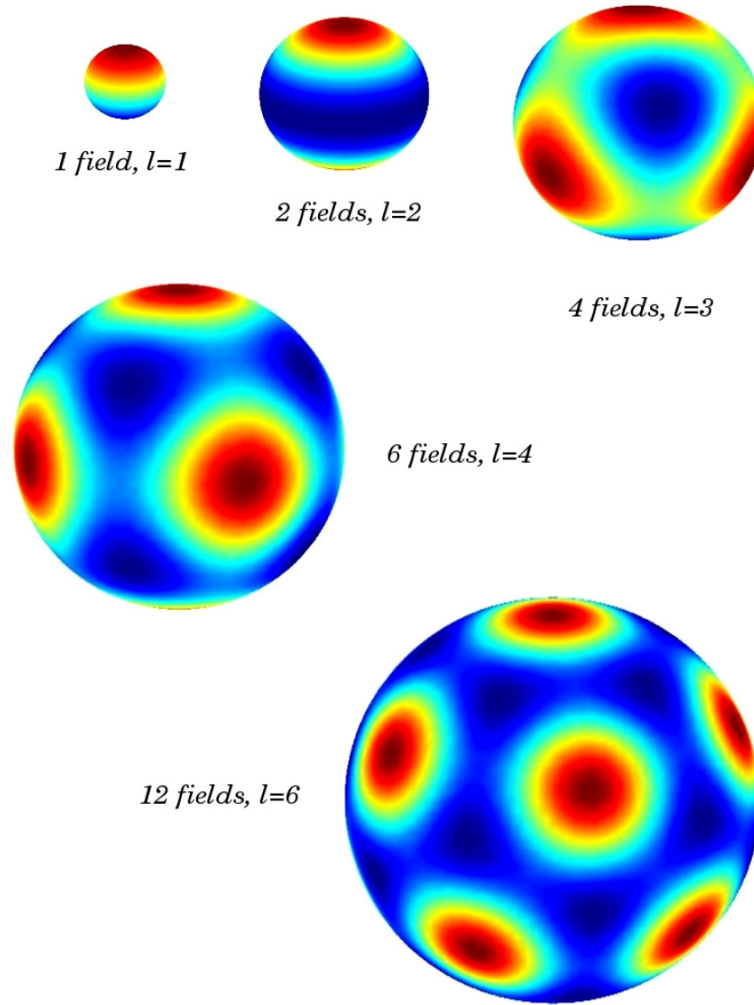
$$\int \Psi(\theta, \varphi) = 1, \quad (4)$$

corresponding to allowing only positive firing rates and to fixing the mean rate.

We start by considering the linear combinations of spherical harmonics that generate an ordered and symmetrical grid-like arrangement of firing fields. Among these, we then look for potential solutions to the minimization problem.

The number of fields obtained in the solution depends on the angular momentum ( $l$ ) of the harmonics used in the linear combination. Solutions for  $l$  up to 6 are as follows (see figure 1). The solution with a single field is obtained with  $l = 1$  spherical harmonics,

$$\Psi_1(\theta, \varphi) = d[cY_0^0 + Y_1^0] = \frac{1}{4\pi} * (1 + \cos(\theta)). \quad (5)$$



**Figure 1.** The analytical solutions considered. Larger spheres lead to more fields, given an approximately constant spacing.

$l = 2$  harmonics lead to a 2-field map,

$$\Psi_2(\theta, \varphi) = d[cY_0^0 + Y_2^0] = \frac{3}{4\pi} \cos^2(\theta). \quad (6)$$

For  $l = 3$ –4 fields

$$\begin{aligned} \Psi_3(\theta, \varphi) &= d [cY_0^0 + [aY_3^0 + b(Y_3^3 - Y_3^{-3})]] = \frac{2\sqrt{2}}{\pi(16\sqrt{2} + 3\pi)} \\ &\times [2 + \cos(\theta)(\frac{5}{2} \cos(2\theta) - \frac{1}{2}) + \sqrt{2} \sin^3(\theta)(2 \cos(2\varphi) + 1)]. \end{aligned} \quad (7)$$

For  $l = 4$ –6 fields

$$\begin{aligned} \Psi_4(\theta, \varphi) &= d [cY_0^0 + [aY_4^0 + b(Y_4^4 + Y_4^{-4})]] \\ &= \frac{1}{8\pi} [2 + 3(\cos^4(\theta) + \sin^4(\theta) \cos^4(\varphi) + \sin^4(\theta) \sin^4(\varphi)) \\ &\quad - 9(\cos^2(\theta) \sin^2(\theta) + \sin^4(\theta) \cos^2(\varphi) \sin^2(\varphi))]. \end{aligned} \quad (8)$$

For  $l = 6-12$  fields, i.e., a soccer ball,

$$\begin{aligned}\Psi_6(\theta, \varphi) &= d[cY_0^0 + [aY_6^0 + \frac{b}{2}(Y_6^{-5} - Y_6^5)]] \\ &= d \left[ c \frac{1}{2} \sqrt{\frac{1}{\pi}} + \left[ \sqrt{\frac{143}{137\pi}} \frac{1}{32} * (231 \cos^6(\theta) - 315 \cos^4(\theta) \right. \right. \\ &\quad \left. \left. + 105 \cos^2(\theta) - 5) + \frac{143}{137\pi} \frac{21}{6} * \cos(5\varphi) \sin^5(\theta) \cos(\theta) \right] \right].\end{aligned}\quad (9)$$

Here,  $d$  and  $c$  are used to satisfy conditions (3) and (4) while the  $a$  and  $b$  values are chosen in each case so that the solutions are fully symmetric. We now want to calculate the value of the ‘energy’, i.e., the cost function associated with these solutions. Consider our cost function  $H$ ,

$$H = H_K + H_A = -\frac{1}{A} \int_S d\chi [\nabla \Psi(\chi)]^2 + \gamma \frac{1}{A} \int_S d\chi \Psi(\chi) \int_S d\chi' \Psi(\chi') K(|\chi' - \chi|). \quad (10)$$

To evaluate this function for every  $l$ , we expand our solutions on the spherical harmonics,

$$\Psi_{l^*}(r) = \sum_{l=0, l^*} \sum_{m=0, m^*} a_{lm} Y_l^m(r). \quad (11)$$

Then, using the spherical harmonics defining property

$$R^2 \nabla^2 Y_l^m(r) = -l(l+1) Y_l^m, \quad (12)$$

we can write the kinetic term  $L_1$  as

$$H_K = \frac{1}{4\pi} \frac{1}{R^2} l^*(l^*+1) (a_{l^*,0}^2 + 2a_{l^*,m^*}^2). \quad (13)$$

To calculate the adaptation term  $H_A$  we project the adaptation kernel on a basis of Lagrange polynomials. Suppose we consider an exponential kernel

$$K(\vec{x} \cdot \vec{y}) = \frac{1}{v\tau_L} e^{-q/v\tau_L} - \rho * \frac{1}{v\tau_S} e^{-q/v\tau_S}, \quad (14)$$

which expresses the fact that adaptation effects become significant over time differences  $\tau_S$  and decay away after time differences  $\tau_L$ . It means that, after a neuron has emitted a spike in a certain position, the spatial region of strong adaptation is comprised in a ring between radii  $v\tau_S$  and  $v\tau_L$  around that position.

In our case  $q$  is the distance between two vectors  $\vec{x}, \vec{y}$  having coordinates  $(\theta, \varphi)$  and  $(\theta', \varphi')$  on the sphere. We write the kernel in terms of the angle between these two vectors,

$$K(\vec{x} \cdot \vec{y}) = \frac{1}{v\tau_L} e^{-(1/v\tau_L)*R*\arccos(\vec{x}\cdot\vec{y})} - \rho * \frac{1}{v\tau_S} e^{-(1/v\tau_S)*R*\arccos(\vec{x}\cdot\vec{y})}. \quad (15)$$

The kernel is thus a function in the  $[-1, 1]$  domain and can therefore be projected on a set of Lagrange polynomials which are complete in that range,

$$K(q) = \sum_l k_l P_l(q). \quad (16)$$

$H_A$  is thus

$$\begin{aligned} H_A &= \frac{1}{4\pi} \sum_{l=0, l^*} \sum_{m=0, m^*} a_{lm} \sum_{l'=0, l^*} \sum_{m'=0, m^*} a_{l'm'} \int d\Omega Y_l^m(r) \int d\Omega' Y_{l'}^{m'}(r') K(\vec{x} \cdot \vec{y}) \\ &= \frac{1}{4\pi} \sum_{l=0, l^*} \sum_{m=0, m^*} \sum_{l'=0, l^*} \sum_{m'=0, m^*} a_{lm} a_{l'm'} \int d\Omega Y_l^m \int d\Omega' Y_{l'}^{m'} \sum_{l''} k_{l''} P_{l''}(\vec{x} \cdot \vec{y}). \end{aligned}$$

Using the addition theorem [18]

$$P_l(\vec{x} \cdot \vec{y}) = \frac{4\pi}{2l+1} \sum_{m=-l}^l (-1)^m Y_l^m(\vec{x}) Y_l^{-m}(\vec{y}) \quad (17)$$

the previous expression becomes

$$\begin{aligned} H_A &= \frac{1}{4\pi} \sum_{l, l'=0, l^*} \sum_{m, m'=0, m^*} a_{lm} a_{l'm'} \sum_{l''} k_{l''} \frac{4\pi}{2l''+1} \\ &\quad \times \sum_{\mu=-l''}^{l''} (-1)^\mu \int d\Omega Y_l^m(\vec{x}) Y_{l''}^\mu(\vec{x}) \int d\Omega' Y_{l'}^{m'}(\vec{y}) Y_{l''}^{-\mu}(\vec{y}). \end{aligned} \quad (18)$$

Given the orthonormality rules for spherical harmonics,

$$\int d\Omega Y_l^m(\vec{x}) Y_{l'}^{m'}(\vec{x})^* = \delta_{l, l'} \delta_{m, m'}, \quad (19)$$

and that  $l, l' = 0, l^*$  and  $m, m' = 0, m^*$ , the surviving combinations in the sum are

$$H_A = \frac{1}{2l^*+1} k_{l^*} (a_{l^*, 0}^2 + 2a_{l^*, m^*}^2) + k_0 (a_0^2). \quad (20)$$

The last term is identical for all solutions, as  $k_0$  and  $a_0$  are always the same.

The total energy is

$$H = (a_{l^*, 0}^2 + 2a_{l^*, m^*}^2) * \left( \frac{1}{4\pi} \frac{1}{R^2} l^*(l^*+1) + \gamma \frac{1}{2l^*+1} k_{l^*} \right) + \gamma k_0 (a_0^2). \quad (21)$$

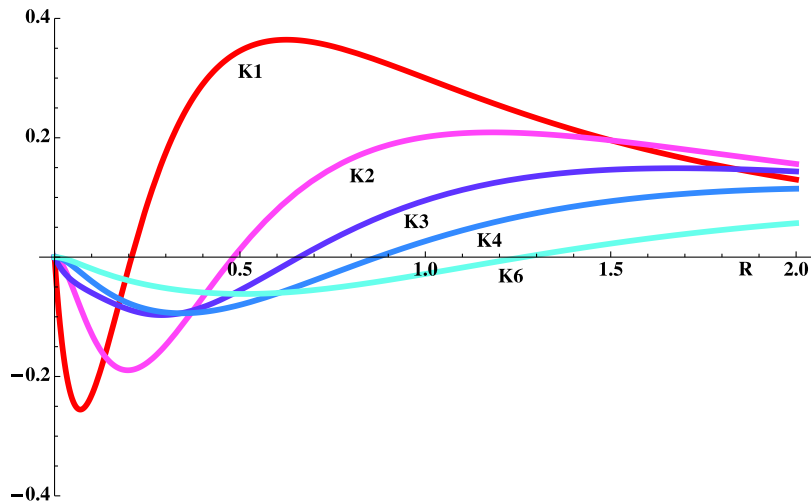
In the region of parameters where  $((1/4\pi)(1/R^2)l^*(l^*+1) + \gamma(1/(2l^*+1))k_{l^*}) > 0$  any solution is worse than the constant solution with all the  $a_{lm} = 0$ .

Since the first term in  $((1/4\pi)(1/R^2)l^*(l^*+1) + \gamma(1/(2l^*+1))k_{l^*})$  is always positive, for the sum to be negative it has to be at the very least that  $k_{l^*} < 0$ . As shown in figure 2, the different components of the kernel projection have a qualitatively similar dependence on the radius, but the range in which they take negative values is progressively shifted towards larger radii. This result does not depend on the particular choice of the kernel. Choosing a kernel consisting of radially symmetric Gaussians,

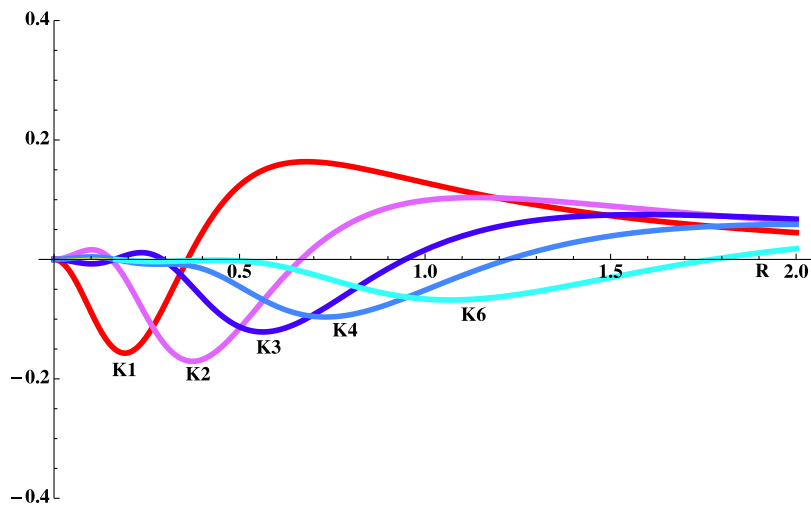
$$K(\vec{x} \cdot \vec{y}) = \frac{1}{v\tau_L \sqrt{2\pi}} e^{-1/2(v\tau_L)^2 * (R * \arccos(\vec{x} \cdot \vec{y}))^2} - \rho * \frac{1}{v\tau_S \sqrt{2\pi}} e^{-1/2(v\tau_S)^2 * (R * \arccos(\vec{x} \cdot \vec{y}))^2}, \quad (22)$$

produces the same behavior, figure 3. The choice of the adaptation time scale ( $\tau_L$ ,  $\tau_S$ ) determines the spatial scale of these transitions. Slower adaptation dynamics means that all curves are shifted towards larger values of  $R$ . This allows us to conclude that the solutions associated with increasing values of the angular momentum become energetically





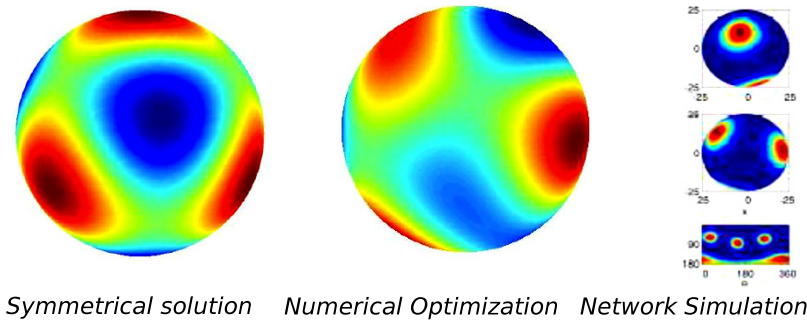
**Figure 2.** Lagrange projections of the exponential kernel.



**Figure 3.** Lagrange projections of the Gaussian kernel.

favored one after the other as the size of the sphere increases. The precise radius at which the switch occurs between two solutions depends also on the values of the components  $a_{lm}$ , on the particular value of  $\gamma$  and on the choice of the kernel, among other factors, but the point of transition from negative to positive values of the corresponding component of the kernel provides an upper bound for the validity of the solution. Therefore, the different solutions are not clearly separated and it is very possible that the energy minima are quite shallow. Transitions from one solution to the next may thus be gradual, and allow for intermediate spurious solutions consisting of a mix of different locally minimal configurations. These questions cannot be analytically addressed and are thus left for numerical simulations.

These calculations are not conclusive with regard to the effective optimality of the combinations of spherical harmonics we have defined. Once we set a certain radius we know that solutions including spherical harmonics with a certain angular momentum are



**Figure 4.** Comparison of the symmetrical solution for  $l = 3$  (left) with the result of the numerical minimization of the energy functional (center) and with the asymptotical state of network simulations (right, see section 4).

avored. The particular combination that minimizes the cost function is determined by the values of the  $a_{lm}$  associated to that solution by the constraints in equation (4). What about other periodical but non-symmetrical solutions?

We have used a numerical approach to explore the existence of other solutions representing global minima for our functional. We have looked for any combination of the  $l + 1$  real spherical harmonics associated with each eigenvalue  $l$ . The optimal solution for a given  $l$  was defined as maximizing the energy function

$$H_{\text{Num}} = a_{l,0}^2 + 2 * \sum_{m=1}^l a_{l,m}^2 - \lambda \int d\chi \Theta(-\Psi_{a_{l,m}}(\chi)), \quad (23)$$

where  $\Psi_{a_{l,m}}$  is the solution corresponding to a particular set of coefficients,  $\lambda \gg 1$ , and the last term is used to enforce constraint (3).

With this approach, we find that for every  $l$  the solutions reached by the numerical optimization procedure are equivalent to, or strongly resemble, our symmetric ansatz, as shown in the example of figure 4.

We limit this approach to the single  $l$  case for computational reasons. Not only would the parameter space significantly grow when any combination of different  $l$  was allowed, but also the energy minima would become dependent on the particular choice of the adaptation kernel, making a comprehensive approach impossible. This situation is similar to the one described for the planar case in [10], where a single wavevector length  $k$  is selected to produce the solutions.

#### 4. Numerical simulations of grid development

The solutions we have studied in section 3 should correspond to the asymptotic behavior of a more complex model, which explicitly expresses the effects of adaptation on neural dynamics. We now proceed to study a more detailed model with numerical simulations. If our analytical approach is robust, we expect the firing maps generated by our simulated model units to correspond to the combinations of spherical harmonics we have previously described.

#### 4.1. Network model

In our simulations, time is discretized in steps of length  $\Delta t = 0.01$  s. The virtual rat moves on surfaces of spheres of different radii with a constant speed of  $v = 0.4$  m s<sup>-1</sup>. For simplicity, the change in running direction between two consecutive steps of the virtual rat is sampled from a Gaussian distribution with zero mean and standard deviation  $\sigma_h = 0.15$  radians. The virtual rat always runs along the great circle determined by its running direction. Our model comprises two layers: the input network represents, e.g., the CA1 region of the hippocampus and contains  $N_{\text{inp}} = 4\pi R^2 \rho$  units, where  $R$  is the radius of the sphere and  $\rho = 8000/m^2$  is the density of input units; the output network represents a population of  $N_{\text{mEC}} = 100$  would-be grid units in the mEC, all with the same adaptation parameters.

Each mEC unit receives afferent spatial inputs which, as already discussed in [10], we take for simplicity to arise from regularly arranged place cells, although they could also arise from spatially modulated units in the adjacent cortices. The input to unit  $i$  at time  $t$  is then given by  $h_i^t$ ,

$$h_i^t = \sum_j W_{ij}^{t-1} r_j^t. \quad (24)$$

The weight  $W_{ij}^t$  connects input unit  $j$  to mEC unit  $i$ . We will assume that at the time the mEC units develop their firing maps, spatially modulated or place cell-like activity is already present, either in the parahippocampal cortex or in the hippocampus. The network model works in the same way with any kind of spatially modulated input, but the place-cell assumption reduces the time necessary for learning. Moreover, as recent studies have shown [8, 9], it is entirely plausible that place cells develop adult-like spatial code earlier than grid cells do. We thus model the place field as an exponential bump centered in the place cell preferred position  $\vec{x}_{j0}$ ,

$$r_j^t = \exp\left(-\frac{\|\vec{x}^t - \vec{x}_{j0}\|^2}{2\sigma_p^2}\right), \quad (25)$$

where  $\vec{x}^t$  is the position at time  $t$  of the simulated rat,  $\sigma_p = 0.05m$  is the width of the firing field and  $\|a - b\|$  is the great-circle distance between two points  $a$  and  $b$  on a sphere. The firing rate  $\Psi_i^t$  of mEC unit  $i$  is determined by a non-linear transfer function

$$\Psi_i^t = \frac{\pi}{2} \arctan [g^t(\alpha_i^t - \mu^t)] \Theta(\alpha_i^t - \mu^t), \quad (26)$$

which is normalized to have a maximal firing rate equal to 1 (in arbitrary units), and  $\Theta(\cdot)$  is the Heaviside function.

The variable  $\alpha_i^t$  represents the adaptation-mediated input to the unit  $i$ . It is related to the  $h_i^t$  as follows:

$$\alpha_i^t = \alpha_i^{t-1} + b_1(h_i^{t-1} - \beta_i^{t-1} - \alpha_i^{t-1}), \quad \beta_i^t = \beta_i^{t-1} + b_2(h_i^{t-1} - \beta_i^{t-1}), \quad (27)$$

where  $\beta_i$  has slower dynamics than  $\alpha_i$ , with  $b_2 = b_1/3$ ,  $b_1 = 0.1$ . This adaptive threshold makes it more difficult for a neuron to fire for prolonged periods of time, and corresponds to the kernel  $K$  considered in the analytical treatment. The gain  $g^t$  and threshold  $\mu^t$  are iteratively adjusted by equations (28) at every time step to fix the mean activity  $a = \sum_i \Psi_i^t / N_{\text{mEC}}$  and the sparsity  $s = (\sum_i \Psi_i^t)^2 / (N_{\text{mEC}} \sum_i \Psi_i^{t2})$  within a 10% relative error

bound from pre-specified values,  $a_0 = 0.1$  and  $s_0 = 0.3$  respectively. If  $k$  is indexing the iteration process,

$$\mu^{t,k+1} = \mu^{t,k} + b_3(a^k - a_0), \quad g^{t,k+1} = g^{t,k} + b_4g^{t,k}(s^k - s_0). \quad (28)$$

$b_3 = 0.01$  and  $b_4 = 0.1$  are the sizes of the iteration steps.  $a^k$  and  $s^k$  are the values of the mean activity and sparsity determined by  $\mu^{t,k}$  and  $g^{t,k}$  in the intermediate iteration steps. The iteration stops once the gain and threshold have been brought within the 10% error range, and the activity of mEC units is determined by the final values of the gain and threshold in equation (26).

The learning process modifies the strength of the feed-forward connections according to a Hebbian rule

$$\tilde{W}_{ij}^t = W_{ij}^{t-1} + \epsilon(\Psi_i^t r_j^t - \bar{\Psi}_i^{t-1} \bar{r}_j^{t-1}), \quad (29)$$

with a rate  $\epsilon = 0.002$ .  $\bar{\Psi}_i^t$  and  $\bar{r}_j^t$  are estimated mean firing rates of mEC unit  $i$  and place unit  $j$  that are adjusted at each time step of the simulation,

$$\bar{\Psi}_i^t = \bar{\Psi}_i^{t-1} + \eta(\Psi_i^t - \bar{\Psi}_i^{t-1}), \quad \bar{r}_j^t = \bar{r}_j^{t-1} + \eta(r_j^t - \bar{r}_j^{t-1}), \quad (30)$$

with  $\eta = 0.05$  being a time averaging factor.

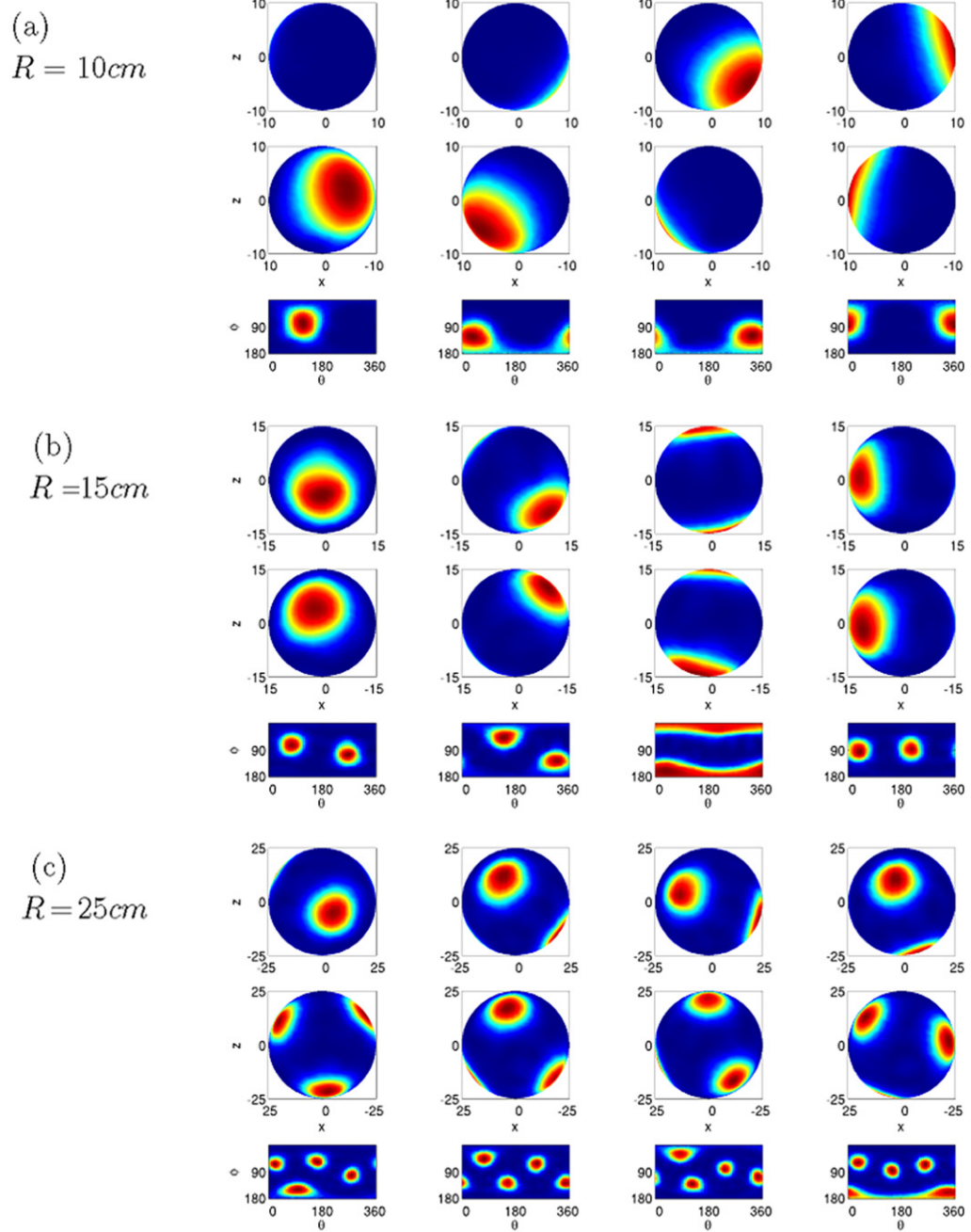
After each learning step, the provisional  $\tilde{W}_{ij}^t$  weights are normalized into the unitary norm,

$$\sum_j W_{ij}^t = 1. \quad (31)$$

Units that win during competitive learning manage to establish strong connections with units that provide strong inputs. As the learning proceeds the units establish fields where they receive strong inputs and at the same time are recovering from adaptation. The emergence of the grid map is the product of averaging over a period of time corresponding to some hours of running for the animal. It remains to be assessed whether the mechanism we propose is suitable to explain the formation of grid representations in each new environment the animal encounters, or if, instead, it can only be applied to the developmental period.

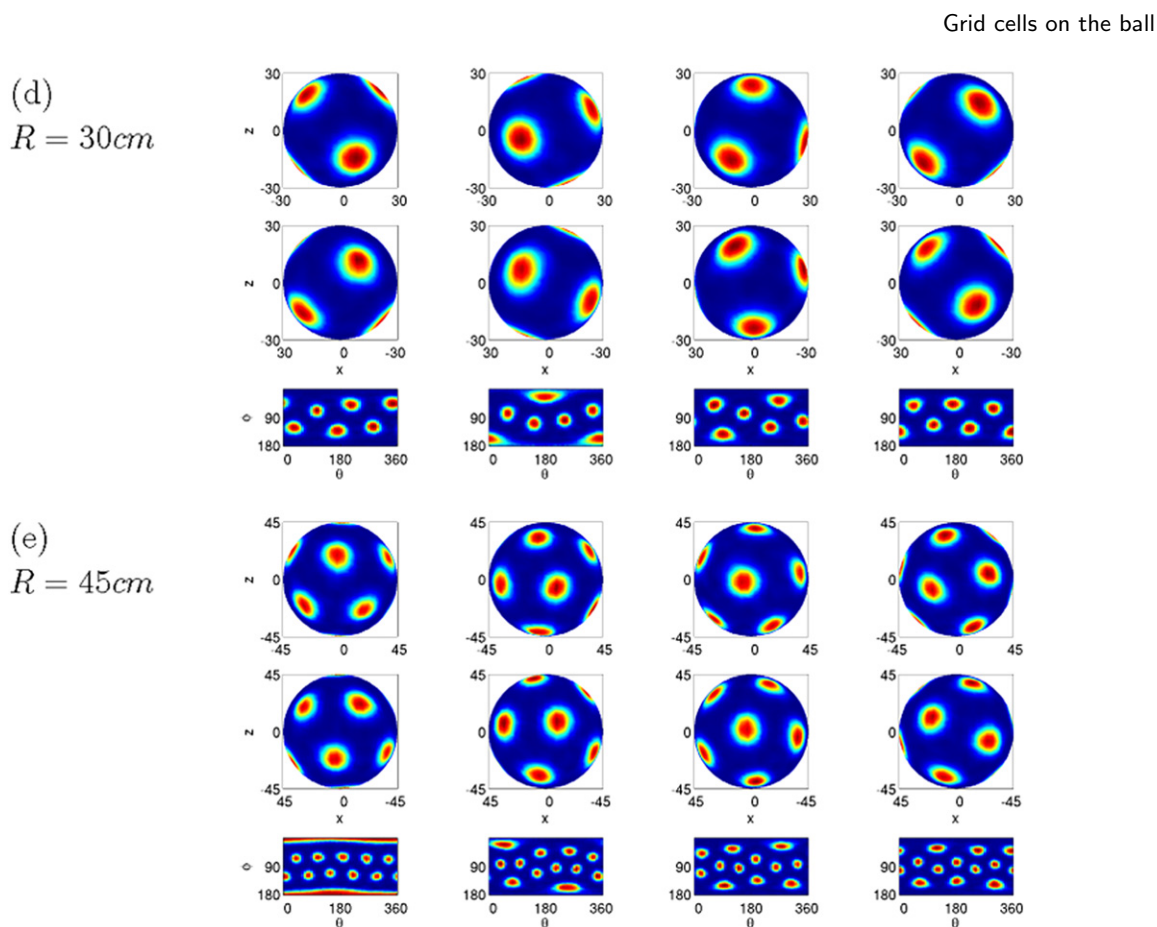
## 4.2. Results

We have run several simulations of  $3 \times 10^7$  learning time steps, i.e. 83 h of animal time, considerably longer than needed to produce clear grids in a flat 2D environment. The fields developed by model mEC units are then expected to be in the asymptotically stable regime, and they are depicted, for different values of the radius of the sphere, in figure 5. The figure presents the firing activity of a random sample of mEC units. The units develop a grid-like pattern of activation and the emerging configurations are in reasonable agreement with those predicted by the analytical model. For each value of the radius, the units typically express the same number of fields. As the radius increases, the units express an increasing number of grid fields, regularly tiling the surface. As already discussed for the analytical approach, the particular association between the radius and the number of fields is the result of the choice of the adaptation time scale. A different value of  $b_1$  (and  $b_2$ ) will result in the appearance of the same configuration of fields



**Figure 5.** Units in an adapting network form grid-like firing maps on spheres. The number of fields increases from 1, 2, 4, 6 to 12 when the radius of the sphere  $R$  increases. In each panel, the first two rows show four examples of firing maps from two opposite views perpendicular to the  $x$ - $z$  plane. The bottom rows of each panel are the Miller cylindrical projections of the corresponding maps. (a)  $R = 10\text{ cm}$ , (b)  $R = 15\text{ cm}$ , (c)  $R = 25\text{ cm}$ , (d)  $R = 30\text{ cm}$  and (e)  $R = 45\text{ cm}$ .

with a smaller or larger radius of the sphere. In the case of a population of cells with heterogeneous adaptation properties, we expect the coexistence of different solutions on the same sphere.



**Figure 5.** (Continued.)

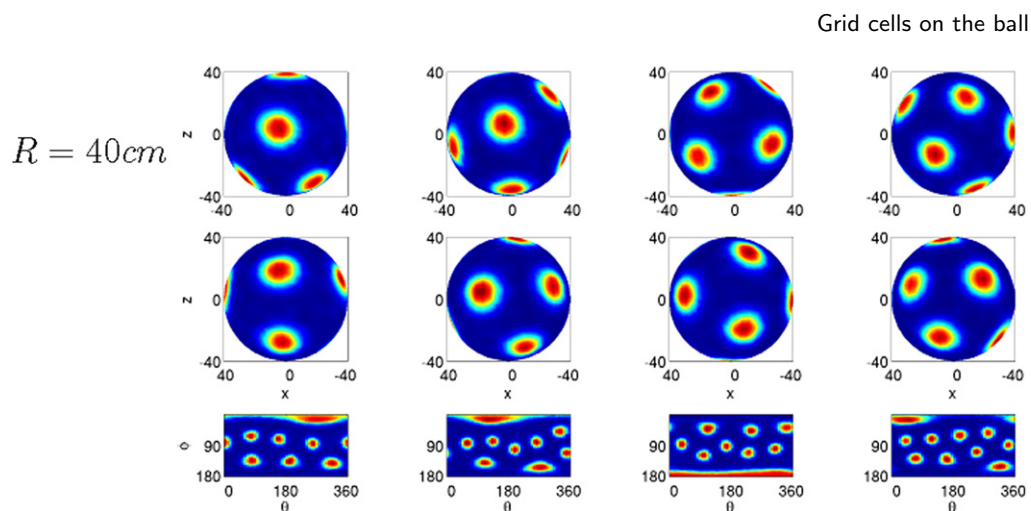
The particular set of adaptation parameters we used would produce a spacing of about 56 cm in the planar case. We notice that this spacing is approximately maintained on the sphere, when its radius is varied. As a consequence of this constraint, the number of fields expressed by the cells grows roughly quadratically with the radius of the sphere.

Generally the configurations correspond to those indicated as asymptotic solutions (1, 2, 4, 6, 12 fields), but in the case of the 40 cm radius, some units exhibit a non-symmetrical configuration of 8–10 fields, figure 6, that emerges at the transition, as it were, from the 6-field to the 12-field solution. The existence and the stability of such a configuration is most probably the expression of the very shallow nature of the energy profile. As previously noted, we cannot exclude the presence of other asymptotically stable solutions. The results of the simulations seem to point in this direction.

## 5. Conclusions

Grid cells, found in rodents and now also in bats [19], are a striking example of the way the brain represents the space we move in. The precise triangular tiling of the environment that is expressed by mEC neurons closely resembles a coordinate system, but how such a regularity can be generated in a neural circuit is still controversial.

We propose a model in which the grid pattern is a product of a self-organization process driven by neural adaptation, with additional recurrent interactions to align the



**Figure 6.** Units develop intermediate numbers of fields if the size of the sphere does not match the adaptation parameters of the network. The radius here is 40 cm, and the four units in the sample shown develop from 8 to 10 fields. Representation as in figure 5.

grids, which can, however, be regarded as second order corrections [11]. This mechanism does not rely on any internal hard-wired structure to work. The final arrangement of the grid fields is a spontaneous effect of the protracted exploration of the environment by the animal. The model thus differs from those based on a pre-wired attractor structure [20, 21], which rely on a specific connectivity established *a priori* to produce the grid pattern of activity. It is not clear to us what such models would predict in the case of rodents raised in or on a sphere, or in fact in any other environment not matching the flat 2D geometry of the pre-wired connectivity (it is likewise not clear what such pre-wired attractor models would predict in the case of bats, flying in 3D). Unless attractor models can be shown to also lead to symmetric combinations of spherical harmonics, or something similar, experiments with spherical environments might help to contrast different models of grid field formation. In parallel, the observation of spherical grid fields with  $l \leq 6$  would immediately lead one to refute models of grid cell formation based on the combination of striped fields at 60 or 120 degrees to each other [22, 23].

In this paper we analyze the specific case of a spherical environment. Using an analytical approach based on the minimization of an energy function, together with more detailed numerical simulations of the system (though still a network of noninteracting units), we study the configurations of activity that develop after a long rearing period.

We find that these configurations are described by linear combinations of spherical harmonics that express a grid-like arrangement of the fields on the surface of the sphere. By combining harmonics of increasing angular momentum one obtains solutions with an increasing number of fields. The radius of the sphere determines which of these is energetically favored, and thus prevails in the asymptotic configuration of the activity, as shown also by simulations.

One problem that remains open is the existence of intermediate, stable states present for certain radii, at the transition between one fully symmetrical solution and the next. Their appearance may reflect the shallow nature of the energy profile described by our

model, or the inability of our approach to successfully identify all the minima of the energy function defining the asymptotic states of the system.

New experiments that make use of virtual reality on a revolving sphere, or in which rats are reared inside a spherical cage, will likely provide a direct test for our model in the near future.

## References

- [1] O'Keefe J and Dostrovsky J, *The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat*, 1971 *Brain Res.* **34** 171
- [2] Fyhn M, Molden S, Witter M, Moser E and Moser M-B, *Spatial representation in the entorhinal cortex*, 2004 *Science* **305** 1258
- [3] Hafting T, Fyhn M, Molden S, Moser M-B and Moser E, *Microstructure of a spatial map in the entorhinal cortex*, 2005 *Nature* **436** 801
- [4] Fyhn M, Hafting T, Treves A, Moser M-B and Moser Edvard, *Hippocampal remapping and grid realignment in entorhinal cortex*, 2007 *Nature* **446** 190
- [5] Wills T, Lever C, Cacucci F, Burgess N and O'Keefe J, *Attractor dynamics in the hippocampal representation of the local environment*, 2005 *Science* **308** 873
- [6] Leutgeb J, Leutgeb S, Treves A, Meyer R, Barnes C, McNaughton B, Moser M-B and Moser E, *Progressive transformation of hippocampal neuronal representations in 'morphed' environments*, 2005 *Neuron* **48** 345
- [7] Cerasti E and Treves A, *How informative are spatial *ca3* representations established by the dentate gyrus?*, 2010 *PLoS Comput. Biol.* **6** e1000759
- [8] Langston R, Ainge J, Couey J, Canto C, Bjerknes T, Witter M, Moser E and Moser M-B, *Development of the spatial representation system in the rat*, 2010 *Science* **328** 1576
- [9] Wills T, Cacucci F, Burgess N and O'Keefe J, *Development of the hippocampal cognitive map in preweanling rats*, 2010 *Science* **328** 1573
- [10] Kropff E and Treves A, *The emergence of grid cells: intelligent design or just adaptation?*, 2008 *Hippocampus* **18** 1256
- [11] Si B, Kropff E and Treves A, *Grid alignment in entorhinal cortex*, 2012 *Biol. Cybernet.* **106** 483
- [12] Barry C, Hayman R, Burgess N and Jeffery K, *Experience-dependent rescaling of entorhinal grids*, 2007 *Nature Neurosci.* **10** 682
- [13] Savelli F, Yoganarasimha D and Knierim J, *Influence of boundary removal on the spatial representations of the medial entorhinal cortex*, 2008 *Hippocampus* **18** 1270
- [14] Solstad T, Boccara C, Kropff E, Moser M-B and Moser E, *Representation of geometric borders in the entorhinal cortex*, 2008 *Science* **322** 1865
- [15] Derdikman D and Moser E, *A manifold of spatial maps in the brain*, 2010 *Trends Cogn. Sci.* **14** 561
- [16] Harvey C, Collman F, Dombeck D and Tank D, *Intracellular dynamics of hippocampal place cells during virtual navigation*, 2009 *Nature* **461** 941
- [17] Courant R and Hilbert D, 1962 *Methods of Mathematical Physics* (New York: Wiley-Interscience)
- [18] Unsöld A, *Beitrage zur Quantenmechanik der Atome*, 1927 *Ann. Phys.* **387** 355
- [19] Yartsev M, Witter M and Ulanovsky N, *Grid cells without theta oscillations in the entorhinal cortex of bats*, 2011 *Nature* **479** 103
- [20] Fuhs M and Touretzky D, *A spin glass model of path integration in rat medial entorhinal cortex*, 2006 *J. Neurosci.* **26** 4266
- [21] Burak Y and Fiete I R, *Accurate path integration in continuous attractor network models of grid cells*, 2009 *PLoS Comput. Biol.* **5** e1000291
- [22] Hasselmo M, Giocomo L and Zilli E, *Grid cell firing may arise from interference of theta frequency membrane potential oscillations in single neurons*, 2007 *Hippocampus* **17** 1252
- [23] Burgess N, Barry C and O'Keefe J, *An oscillatory interference model of grid cell firing*, 2007 *Hippocampus* **17** 801
- [24] Kruge I, Wernle T and Moser M-B, 2012 personal communication