

# 空间认知的吸引子动力学

王子群\*, 王 涛\*, 刘 锋†

南京大学物理系和脑科学研究院, 南京 210093

**摘要:** 哺乳动物的导航系统包含多种类型的神经元, 负责位置感知和空间路径规划, 涉及多信息整合。吸引子动力学理论作为一个能统一解释记忆、决策等复杂认知功能的脑理论, 可以解释导航系统中神经元的特异性放电和路径整合机制。本文综述了空间认知领域吸引子动力学的最新研究进展。首先, 概要介绍了计算神经科学和吸引子动力学的一般理论。接着, 以哺乳动物导航系统的连续吸引子动力学为核心, 深入探讨了头朝向细胞和网格细胞的放电动力学特征及其功能意义。在此基础上, 对导航系统吸引子理论进行了拓展与展望。

**关键词:** 计算神经科学; 吸引子动力学; 导航; 头朝向细胞; 网格细胞

中图分类号: Q61

文献标识码: A

DOI: 10.13725/j.cnki.pip.2023.06.003

## 目 录

I. 引言	188
II. 计算神经科学——神经科学中的数学物理	189
III. 吸引子动力学	190
IV. 导航系统中的吸引子动力学	191
A. 头朝向细胞的特征放电与吸引子动力学	191
B. 网格细胞的特征放电与吸引子动力学	192
C. 导航系统中的连续吸引子模型	192
D. 吸引子动力学下的 theta 振荡机制	195
E. 吸引子动力学视角下的人类导航功能损伤	196
V. 总结与展望	197
致 谢	198
参考文献	198

## I. 引言

人脑是个体行为、认知和情感的基础, 是极其复杂的生命系统。计算神经科学运用计算模型、数理方法和数值模拟, 探索神经系统信息处理规律, 对阐明脑的工作原理具有重要意义。30 多年来, 计算神经科学取得了许多重要成果<sup>[1]</sup>, 如建立大量的神经元和神经网络模型,

提出不少新原则 (如“除法归一化”)、概念 (如“均衡态”) 和理论 (如强化学习) 等, 提供大量有关认知功能 (如工作记忆、抉择、注意等) 的回路计算机制, 推动了脑成像、人工智能、脑-机接口和大脑控制等技术的发展, 为精神疾病、神经退行性疾病等的治疗、预防以及药物筛选提供了理论支持。这些成就共同推动了神经科学的飞速发展, 为深入理解大脑功能和行为提供了新途径。

吸引子动力学理论在神经科学中起着重要作用, 帮助我们理解神经网络的动态行为和信息处理机制。神经网络被视为一个动态系统, 其状态随时间演变。把网络的特定状态或模式称为“吸引子”, 可代表特定的认知状态、记忆、决策或其他神经活动模式。吸引子动力学研究系统如何在这些状态间切换, 以及在各种条件或输入下保持这些状态。阐明吸引子动力学对于理解神经系统的运行原理具有重要意义<sup>[2-6]</sup>。本文将着重介绍空间认知行为的吸引子动力学机制。

位置感知和空间导航是人脑最基本的功能之一。人和动物能够确定自身位置并规划路径, 这对个体生存至关重要。这种能力是以脑内特定神经元群的放电活动为基础的。1971 年, O'Keefe 等在大鼠海马区发现位置细胞 (place cell), 其放电野 (firing field) 具有空间位置特异性<sup>[7]</sup>。随后, 对空间朝向敏感的头朝向细胞 (head direction cell)<sup>[8]</sup>、网格状放电野的网格细胞 (grid cell)<sup>[9,10]</sup>、分别编码速度和环境边界的速度细胞 (speed cell)<sup>[11]</sup> 和边界细胞 (border cell)<sup>[12]</sup> 等导航相关神经元陆续被发现, 且哺乳动物特化的海马体 (hippocampus) 及其旁

收稿日期: 2023-10-06

\* 共同第一作者

† E-mail: fliu@nju.edu.cn

区<sup>[13]</sup>等导航相关脑区被确认。对环境中特定位置与头朝向的编码表明哺乳动物存在对环境的整体认知,被称为“认知地图 (cognitive map)”。个体基于自身运动实时更新头朝向和位置认知的过程被称为路径积分 (path integration)<sup>[14]</sup>,它是路径规划的基础;动物经由自身运动与环境的交互认知,连续整合多种感知信息,实现对空间位置表征的持续更新。实验表明,导航神经元的放电行为呈现连续吸引子特征。在吸引子动力学框架下,阐述路径积分和空间导航的神经回路机制和计算原理获得了很大成功,得到了实验的充分验证。

在本篇综述中,我们将详细介绍导航系统的吸引子动力学。先描述神经网络吸引子动力学的一般特性,然后系统解析导航系统中的连续吸引子动力学,并阐述信息如何整合形成特定的神经放电模式。最后,对吸引子动力学的研究进行了总结与展望。

## II. 计算神经科学——神经科学中的数学物理

探究生理解剖结构如何产生复杂多样的生命功能是生命科学的核心研究内容。在各种先进技术的支持下,神经科学能够清晰地呈现大脑在亚细胞、神经元、神经环路、全脑等不同尺度上精细的结构、动力学和功能特征。但是,各个尺度上、跨尺度间,结构、动力学和功能之间的联系往往极为复杂,单纯从实验数据出发,很难定量理解大脑信息处理机制。为了推动对大脑和认知过程的理解,数学、物理学、信息学等多学科的知识和方法被广泛引入到神经科学的研究中,促成了计算神经科学的诞生和发展,成为对实验研究的重要补充。通过定量解析从神经元电生理到神经环路、跨脑区等不同层次的动力学机制,计算神经科学已建立不同尺度神经系统的数理模型,实现对实验数据的定量表征和多领域、多层次之间的交叉,建立起统一的概念和原理,来理解不同类型研究中所揭示的神经信号处理机制及其联系与区别。

计算神经科学的研究涉及了神经科学的各个方面。以研究对象的尺度为纵轴,在亚细胞和神经元层面,计算神经科学通过建模和仿真,揭示了突触和神经元的动态行为和特性<sup>[15]</sup>,为理解神经信号的传递和加工提供了理论支撑;在神经环路层面,模拟神经网络行为,揭示了大量神经元如何通过特定的组织结构和编码方式处理外界信息,实现各种复杂的认知功能<sup>[3]</sup>;在全脑尺度,解析不同脑区间的交互作用和信息传递过程<sup>[16]</sup>,对于揭示脑功能和相关疾病的病理机制具有重要意义。计算神经科学还通过模拟大脑与外部环境的相互作用,为理解和干预大脑与环境的交互提供了可能。

以研究的问题为横轴,计算神经科学可以回答三类问题,即是什么 (what)、怎么办 (how) 和为什么 (why),建立相应的描述型模型、机制型模型和解释型模型。描述型模型能够完整且精确地概括大量实验数据,刻画神经元或神经环路做了什么。这类模型与神经系统的生物物理、解剖生理等特征的联系相对较弱,其主要作用仅是描述实验现象,例如空间导航系统中的“网格细胞”、“头朝向细胞”、“位置细胞”等,以及对于神经振荡的分类描述,如 delta、theta、alpha、beta、gamma 波等,都可以被看作是描述型模型。相反,机制型模型依据已知的解剖、生理等知识,能够回答神经系统如何工作这一问题。这类模型往往能够建立起多个描述型模型之间的联系,如揭示包括工作记忆、空间导航等大脑功能的连续吸引子网络模型<sup>[17,18]</sup>。解释型模型综合利用计算和信息理论中的一些原理,从多个角度来解释神经系统功能特征的行为和认知意义,阐明神经系统为什么以当下方式运行,如从信息论角度阐述不同神经元放电模式的成因等等。此外,随着近些年人工智能领域的巨大发展,大量机器学习的方法也被引入到神经科学,用于处理实验数据、挖掘其内禀特征<sup>[19]</sup>。总的来说,计算神经科学以其独特的理论和方法,为神经科学的深入研究提供了强大的工具和平台,使我们能够更好地理解和解释神经系统的工作原理和行为表现。

与其他领域的计算研究类似,计算神经科学的一个显著优势在于其纯粹性和简洁性。从实验提炼出的核心假设出发,构建计算模型,有效排除与研究目标不相关的干扰因素,对真实系统进行粗粒化模拟,从而更清晰地展示假设与实验现象的内在关联。针对不同的研究尺度和研究内容,模型的粗粒化程度会有所不同,能够复现的实验细节以及能够回答的具体问题也各不相同。这种灵活性和针对性,使得计算神经科学在解决复杂问题时,能够达到更高的效率和精确度。

下面,我们以基础的神经元建模为例,说明计算神经科学研究的特点。常见的神经元模型可大致归结为三类,我们主要介绍各类型中的典型模型。一是霍奇金-赫胥黎模型 (Hodgkin-Huxley model; H-H 模型)<sup>[20]</sup>:

$$C_m \frac{dV(t)}{dt} = -g_{Cl}(V(t) - V_{Cl}) - g_{Na}(V)(V(t) - V_{Na}) - g_K(V)(V(t) - V_K) + I(t), \quad (1)$$

描述了神经元膜电位  $V$  的时间变化率。其中,  $C_m$  为膜电容,  $I$  为输入电流,  $g_{Na}$ 、 $g_K$ 、 $g_{Cl}$  分别为  $Na^+$ 、 $K^+$ 、 $Cl^-$  通道的电导,且  $g_{Na}$  与  $g_K$  为  $V$  的函数,  $V_{Na}$ 、 $V_K$ 、 $V_{Cl}$  则为相应的平衡电位。H-H 模型准确刻画了神经元的许多电学特性,包括动作电位的产生、传导过程等,并且容易扩展,可通过引入特异性离子通道来描述不同类型的神

经元。H-H 模型的劣势则在于其计算复杂度较高，需要较强的算力，有时很难通过实验确定全部的参数值。这些都增加了模型的使用难度，限制了它在复杂神经网络模型中的应用。

二是整合发放模型 (leaky integrate and fire model; LIF 模型)<sup>[21]</sup>：

$$C_m \frac{dV(t)}{dt} = -g_L(V(t) - V_L) - g_{syn}(V(t) - V_{syn}) + I(t), \quad (2)$$

其中， $g_L$  与  $V_L$  分别是漏电导与静息电位， $g_{syn}$  为突触电导， $V_{syn}$  为反转电压， $I$  为输入电流。当膜电位达到发放阈值时，产生一次放电，并将膜电位保持为重置电压直至不应期结束。LIF 模型忽略了复杂的动作电位产生过程，保留了对动作电位的记录，能刻画神经元在多种刺激方式和条件下的放电行为，但不及 H-H 模型精细。LIF 模型对算力的要求较低，适合研究神经网络的放电动力学。

三是频率模型或门变量模型 (firing-rate model 或 gating-variable model)<sup>[22]</sup>：

$$\tau_i \frac{dS_i}{dt} + S_i = F_i \left( \sum_j W_{ij} S_j \right), \quad (3)$$

以神经元活化程度或某个代表性突触通道开放程度为变量  $S$ ， $W_{ij}$  为神经元突触连接强度， $F_i$  是刻画输入—输出关系的非线性函数，如线性整流函数 (linear rectification function)。由于极为粗粒化，频率模型可以快速计算大量神经元的行为，计算效率高。频率模型的非线性仅来自非线性函数  $F(x)$ ，相对简单，因而可以建立起系统的理论框架，为模型的分析 and 求解提供支持。频率模型忽略了神经元的大量细节，无法刻画神经元行为的复杂特性，在不少情况下无法适用。

从三类神经元模型的特点可以看出，应根据研究目的和内容的不同，选择合适的建模层次，揭示不同层面的放电特性和信息处理机制。确实，人们通过建立不同复杂度和形式的计算模型，模拟神经元和神经网络的动态行为，揭示神经系统的结构、动力学和功能及其联系，阐述神经系统的工作原理。

计算神经科学的重要性不言自明。一个鲜明的例子就是，在实验数据暗示果蝇的头朝向神经环路具有环状连续吸引子网络特征后，实验人员基于连续吸引子理论，很快就确定了果蝇头朝向神经元的连接结构、动力学特性，实实在在地验证了头朝向表征的连续吸引子机制<sup>[23]</sup>。在人工智能领域，计算神经科学的理论和方法有广泛应用，如用于构建高效的机器学习模型，改善语

音识别、图像识别和自然语言处理等应用效果<sup>[24]</sup>。此外，计算神经科学促进了神经系统疾病的研究和新治疗方法的开发。例如，在具有阿尔茨海默病风险基因的人群中，网格细胞的放电模式不稳定<sup>[25]</sup>，这为疾病的早期诊断提供了线索。总之，计算神经科学具有独特的优势，能够解答一些实验难以回答的问题，对于理解神经系统和开发新应用具有重要作用。

### III. 吸引子动力学

吸引子网络是计算神经科学中的重要概念。这类网络中，神经元连接具有特定的对称性和反馈机制，使得网络能够自组织和维持特定的放电模式，即所谓的“吸引子”。形成吸引子网络需要满足两个基本条件。一是神经元之间有兴奋性的回返性连接 (recurrent connection)，即使没有外界输入，仅靠神经元之间的正反馈，网络也能维持放电活动。二是网络中要有抑制性连接，避免因持续的正反馈而导致系统活动失控。在适当的兴奋与抑制作用下，网络就会形成吸引子态。如果神经元的连接没有特殊对称性，网络通常只支持离散的吸引子态；若连接具有空间平移不变性，网络可能展现一簇连续而非孤立的吸引子态。这些吸引子在状态的相空间中紧密排列，构成一个连续的状态子空间，被称为“连续吸引子”。

从相空间的角度来看，吸引子是其中的一个特殊子集。当系统状态接近此吸引子时，系统状态随时间演化逐渐趋向吸引子态，呈现被“吸引”现象，如图 1 (a) 所示。该吸引子态在相空间中的吸引区域被称为其吸引域 (basin of attraction)<sup>[2]</sup>。离散吸引子在相空间中表现为离散的点，而连续吸引子则表现为光滑的线或更复杂的几何形状，图 1 (b)，被称为吸引子流形 (attractor manifold)。从能量角度来看，离散吸引子可视为能量面上的若干深坑，而连续吸引子则呈现为一条连续的沟渠。

许多中枢神经系统的功能可以通过吸引子网络得到解释。最初，离散吸引子网络被用于解释联想记忆功能。联想记忆是指个体唤起物品或概念间联系的能力，而这可通过 Hopfield 网络来实现。Hopfield 网络是离散吸引子神经网络的典型代表<sup>[22]</sup>，通过突触学习法则下神经元连接的改变，网络将收敛至与输入模式相关的离散稳态，即吸引子态。每个吸引子态对应了某个输入模式，并且具有残缺输入可唤起完整激活模式的能力，实现联想记忆功能。此外，通过神经元集群侧抑制构建的赢家全胜 (winner-take-all) 网络<sup>[26]</sup>，作为一类特殊的 Hopfield 网络，其吸引子态为某一神经元集群的持续兴奋与其他神经元集群的抑制。由于其清晰的神经集群

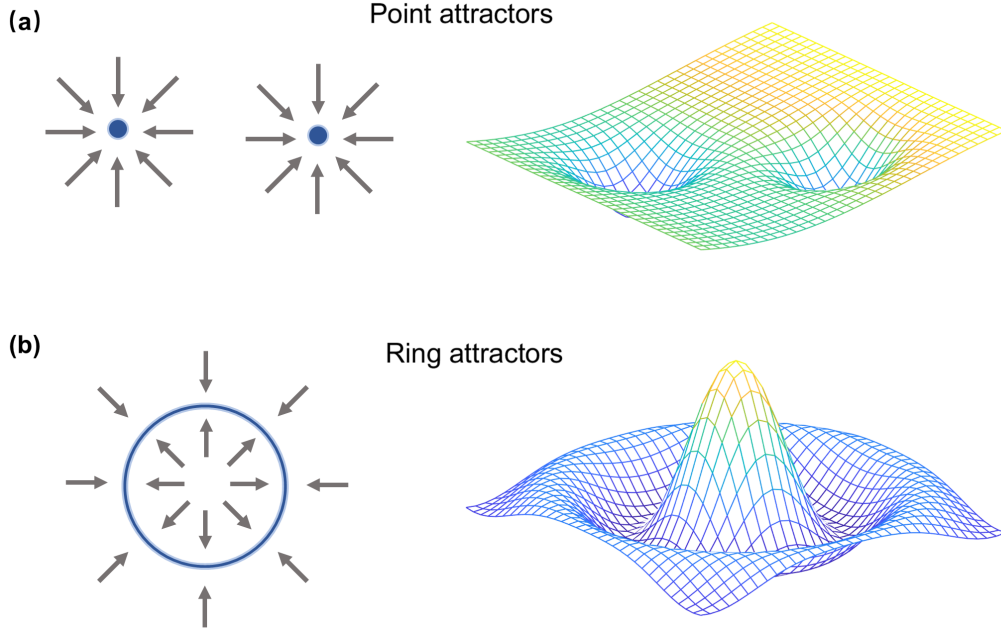


图 1. 吸引子动力学示意图。(a) 点吸引子 (离散吸引子) 在相空间中的表现 (左) 与能量景观 (energy landscape) 示意图 (右)。(b) 环吸引子 (一维连续吸引子) 在相空间中的特征 (左) 与能量景观示意图 (右)。

分类和吸引子态之间的竞争特性, 这类网络能够完成决策 (decision-making)<sup>[6]</sup>、竞争学习 (competition learning)<sup>[27]</sup>、模式识别 (pattern recognition)<sup>[28]</sup>、目标分类 (target classification) 等多种功能。

在连续吸引子网络中, 系统状态可以在连续的吸引子流形上迁移, 并始终保持稳态, 是一种随遇平衡状态<sup>[29]</sup>。通过将连续吸引子态与真实世界任务变量联系起来, 神经系统能实现对连续刺激和连续特征的编码<sup>[30,31]</sup>, 如前额叶皮层 (prefrontal cortex; PFC) 和后顶叶皮层 (posterior parietal cortex; PPC) 中与一维视觉空间刺激位置相关的工作记忆<sup>[32]</sup>。值得注意的是, 这种记忆与 Hopfield 网络下的联想记忆不同, 其记忆过程通过神经元群体持续兴奋的吸引子态实现, 并不涉及吸引子流形的变化, 更多用于阐释短期记忆相关机制。此外, 连续吸引子动力学还能够完成诸如信息整合<sup>[4,5]</sup>、导航<sup>[19,33]</sup>、决策<sup>[34]</sup>等多种认知功能。除离散与连续吸引子外, 神经系统还存在极限环 (limit cycle) 等更复杂的吸引子动力学, 承担模式生成<sup>[35]</sup>、昼夜节律<sup>[36]</sup>等重要功能。

吸引子网络在中枢神经系统的广泛存在凸显了其在认知过程中的重要性, 为研究大脑的计算功能提供了有效的理论框架。吸引子动力学的优势主要源于其相空间低维特性对信息编码和计算的便利性。由于吸引子流形的维度  $K$  远远小于相空间维度 (如神经元个数  $N$ ), 其动力学体现了对信息的降维表征。同时, 正交于吸引

子流形的噪声被消除, 噪声大小变为原来的  $\sqrt{K/N}$ , 增强了系统的信息处理能力和编码的准确性<sup>[2,29]</sup>。吸引子动力学的优势还在于吸引子流形的多样性。通过与真实世界任务变量的锚定, 吸引子动力学能够灵活完成记忆、抉择、分类、积分等复杂多样的认知任务。然而, 吸引子动力学仍面临难以解释稀疏放电下的神经动力学编码与快速动力学机制<sup>[17,37]</sup>等方面的挑战。

## IV. 导航系统中的吸引子动力学

### A. 头朝向细胞的特征放电与吸引子动力学

头朝向细胞是哺乳动物编码头朝向的神经元, 当动物朝向某一特定方向时剧烈发放 (图 2 (a)), 广泛分布在后旁下托 (post-subiculum)<sup>[8]</sup>、丘脑前核 (anterodorsal thalamic nucleus; ADN)<sup>[38]</sup>、乳头体外侧核 (lateral mammillary nuclei; LMN)<sup>[39]</sup>等脑区。头朝向细胞群体的放电模式表征了头朝向, 其在角速度与感知觉信号输入下被持续更新。一方面, 来自前庭区域 (vestibular region) 的运动信息对头朝向细胞的持续放电与角度路径积分 (angular path integration) 的实现十分重要<sup>[8,39,40]</sup>; 另一方面, 头朝向表征也依赖于视觉和嗅觉等感官信息的输入<sup>[41]</sup>, 这种导航方式被称为地标导航 (landmark navigation)。头朝向细胞整合多种信息、连续更新头朝向表征的能力暗示了连续吸引子动力学的

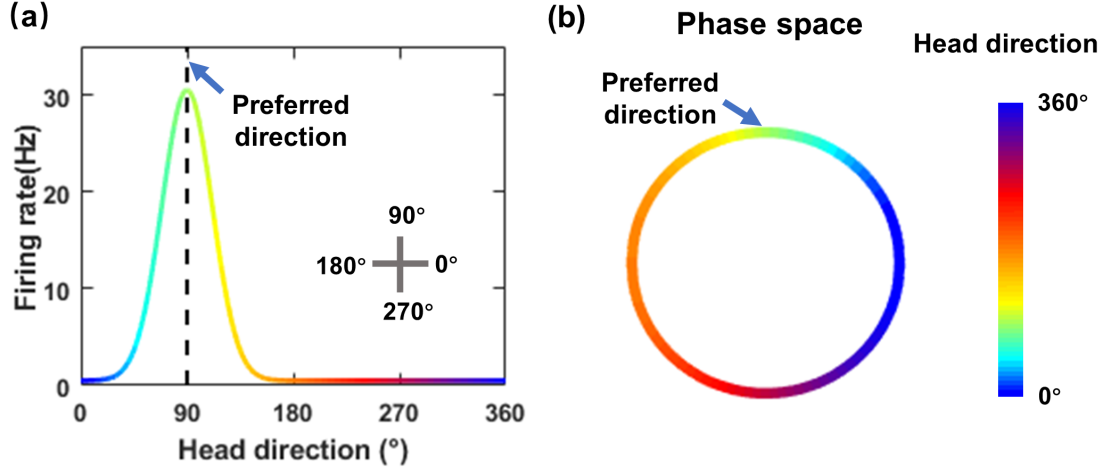


图 2. 头朝向细胞放电曲线与群体放电特征。(a) 单个细胞的放电调谐曲线, 显示当头朝向  $90^\circ$  时该细胞剧烈放电。(b) 细胞群体放电模式在相空间中呈一维流形。图中颜色标识不同的头朝向。

可能性。

除了功能表现, 头朝向细胞的动力学也符合吸引子动力学。(1) 单细胞的偏好方向唯一且在神经元群体中连续均匀分布<sup>[8]</sup>; 动物朝某方向时, 部分头朝向细胞的持续放电可看作系统的吸引子态, 如图 2 (b) 所示。(2) 虽然头朝向细胞存在环境特异性的偏好方向漂移, 但不同细胞的偏好方向差保持不变<sup>[8]</sup>, 表明头朝向的编码方式稳定不变。这既暗示头朝向的整合计算在局域网络内完成, 也与吸引子态的不变性相符。头朝向细胞的动力学表现与其头朝向编码形式的保守性暗示了可能依赖吸引子动力学的信息整合机制。

## B. 网络细胞的特征放电与吸引子动力学

网络细胞是哺乳动物内侧内嗅皮层 (medial entorhinal cortex; MEC) 中一类特殊神经元, 其放电野呈正六边形点阵铺满整个二维平面<sup>[9,10]</sup>, 如图 3 (a) 所示。网络细胞放电野的形成也涉及多信息整合。头朝向细胞所在的后下托区域对 MEC 有直接的投射<sup>[42]</sup>; 在 MEC 区 III–VI 层、前下托及旁下托区域, 存在同时具有网络放电特性与头朝向方向敏感性的联合细胞<sup>[43]</sup>, 表明 MEC 区网络细胞接收来自下托区域的头朝向信息。网络细胞的放电野会随环境的变化而改变, 如在特定环境下会发生畸变<sup>[44]</sup>, 或者在地标指示下发生重映射 (remapping)<sup>[45]</sup>, 表明网络细胞同样接收来自感官信息的输入。总之, 网络细胞放电野是多种信息整合计算的结果, 其对复杂信息的连续整合可能由连续吸引子动力学来实现。

网络细胞的放电动力学也与连续吸引子动力学假

设相符。(1) 细胞的放电特征只与空间位置有关, 当动物位于二维空间特定位置时, 对应网络细胞会持续兴奋<sup>[9,10]</sup>, 可视为系统的吸引子态, 如图 3 (b)。(2) 网络细胞具有模块化 (modularization) 特征, 可沿腹背轴分为 4 个离散模块<sup>[46]</sup>, 同一模块内的细胞具有相似的网格间距与网格朝向, 这提示其特异性放电可能依赖多个离散的局域网络。(3) 网络细胞的放电模式高度保守: 同一模块内的细胞始终保持稳定的网格相位差<sup>[45]</sup>, 且放电关联性在不同环境、不同状态下保持不变<sup>[47–49]</sup>。神经活动形式的高度保守性与吸引子态的不变性相符, 再次说明网络细胞的特殊放电形式由特定连接下的局域网络产生。

诸多实验从多方面证实了网络细胞与头朝向细胞的放电活动符合连续吸引子动力学。同时, 连续吸引子理论确保了网络细胞与头朝向细胞的信息整合功能的稳定性。可以说, 吸引子动力学是揭示导航计算原理的关键性机制。

## C. 导航系统中的连续吸引子模型

基于头朝向细胞与网络细胞清晰的连续吸引子动力学特征, 人们建立连续吸引子网络 (Continuous attractor network; CAN) 模型来阐述其信息整合与表征机制。在头朝向细胞的 CAN 模型<sup>[30]</sup>中, 假设  $N$  个头朝向细胞沿圆环均匀排列, 用 LIF 模型来刻画:

$$C_m \frac{dV_i(t)}{dt} = -g_L(V_i(t) - V_L) - \sum_{j=1}^N W_{ij} \tilde{g}_{\text{syn},j}(t - t_j) \cdot (V_i(t) - V_{\text{syn}}), \quad (4)$$



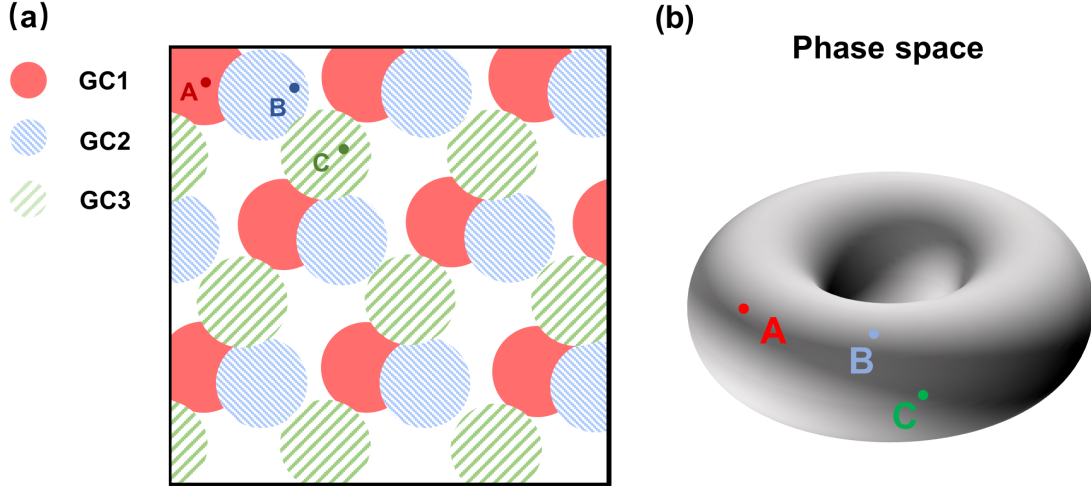


图 3. 网格细胞的放电模式。(a) 3 个相位不同的网格细胞的放电野示意图。放电野呈正六边形点阵，铺满全场。(b) 网格细胞群体放电模式在相空间中呈二维流形。A-C 三点分别代表动物处于 (a) 中 A-C 位置时的神经元放电活动。

神经元间的连接强度  $W_{ij}$  与突触前电导函数  $\tilde{g}_{\text{syn},j}$  的乘积表示突触电导  $g_{\text{syn}}$ ， $\tilde{g}_{\text{syn},j}$  与突触前神经元  $j$  的发放相关。突触连接强度与突触前后神经元的角度位置差  $\Delta\theta$  相关，见图 4 (a)：

$$W_{ij} = W(\Delta\theta) = J_- + (J_+ - J_-)\exp(-\frac{\Delta\theta^2}{2\sigma^2}) \quad (5)$$

$J_- < 0$ ,  $J_+ > 0$ ，即细胞间存在广泛的抑制连接，但角度位置相近的细胞间存在较强的兴奋连接，连接形式具有平移不变性。网络的吸引子态表现为部分细胞的持续放电 (图 4 (b))。随后，引入对称性破缺下的角度积分机制，将连续吸引子态与头朝向锚定；具体来说，在模型中引入两组与角速度相关的神经元群体，它们分别接收顺时针和逆时针方向的角速度输入，以及来自头朝向神经元的输入，并且与头朝向神经元有非对称性连接 (图 4 (c))。动物转向时，转向相关的神经元群体被激活，非对称连接将引导 CAN 中的吸引子沿流形运动，实现角速度输入下的吸引子态转变，完成真实头朝向与吸引子态的锚定 (图 4 (d))。

头朝向细胞的 CAN 模型提出了依赖吸引子动力学的信息整合与头朝向表征更新机制。其中，角速度输入下头朝向表征的更新机制已在果蝇中被证实<sup>[50]</sup>；哺乳动物 ADN 和 LMN 区域角速度与头朝向共调谐神经元<sup>[38,39]</sup>的存在，也与模型中的角速度输入神经元群体相互印证。同时记录的大规模神经元放电数据与拓扑分析表明，头朝向细胞群体放电模式在相空间中确实呈现一维连续吸引子流形。这一流形在快速眼动睡眠 (rapid eye movement; REM) 中保持不变<sup>[19]</sup>，符合吸引子理论中态的不变性预测。

网格细胞的二维 CAN 模型可以通过对头朝向细胞 CAN 模型做拓展而得到<sup>[18]</sup>。网格细胞具有不同相位，其放电活动仅与动物在二维空间的位置相关。因此，将动物位于某一位置时网格细胞群体的放电模式作为系统的吸引子态，动物位置的改变导致系统状态沿吸引子流形演化，这些构成了模型的基本假设。我们以 Burak 与 Fiete 的 CAN 模型<sup>[18]</sup>为例，介绍模型的基本架构与网格状放电野的产生机制。

网格细胞排列在神经元片层中，其位置由网格相位决定，边缘网格细胞的连接满足周期性边界条件 (图 5 (a))。每个网格细胞的活动用门变量模型描述：

$$\tau \frac{dS_i}{dt} + S_i = f \left[ \sum_j W_{ij} S_j + B_i \right], \quad (6)$$

$S_i$  为门变量， $W_{ij}$  为网格细胞间的连接强度， $B_i$  为其他输入信息，如速度、头朝向信息输入等。传递函数  $f[x]$  采用线性整流函数定义：当  $x > 0$  时， $f[x] = x$ ；当  $x \leq 0$  时， $f[x] = 0$ 。模型包含  $N = 128^2$  个神经元。若不考虑速度与头朝向信息输入， $B_i = 1$  为固定输入。为了使网络出现自发且连续的稳态，需要构建具有平移不变性的竞争网络。假设网格细胞对距离较近的神经元产生兴奋或去抑制作用，对中等距离的产生抑制作用，对较远距离的影响较弱，则连接矩阵可表示为

$$W_{ij}(\mathbf{x}) = \alpha e^{-\gamma|\mathbf{x}|^2} - e^{-\beta|\mathbf{x}|^2}, \quad (7)$$

其中， $\mathbf{x} = \mathbf{x}_i - \mathbf{x}_j$  为细胞间的位移向量， $W$  通常为墨西哥草帽形式。特定的竞争性连接条件下，网络产生正六边形的活跃点阵，其连续分布、平移不变，可视为网络的连续吸引子态 (图 5 (b))。

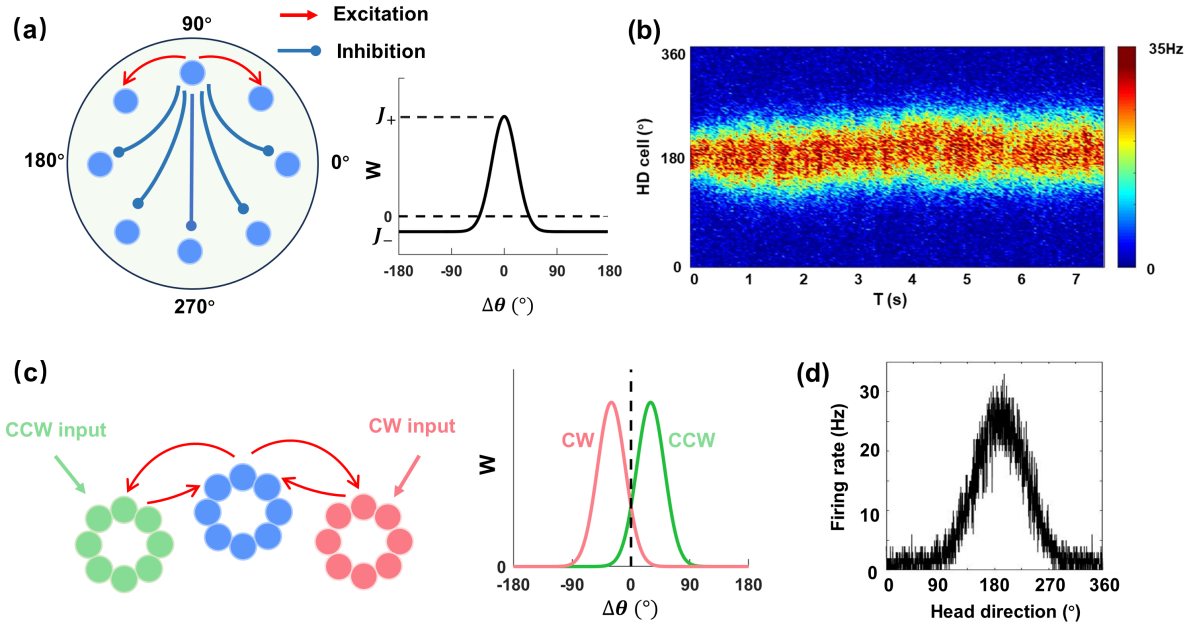


图 4. 头朝向细胞的一维 CAN 模型。(a) 头朝向细胞按角度位置排列，位置相近的神经元间相互兴奋，其他神经元间则为广泛的抑制性连接。连接强度曲线见右图。(b) 头朝向细胞的群体放电时空图。特殊连接形式下，一群持续放电的神经元编码头朝向。(c) 角速度信号的非对称性输入。两个角速度神经元群分别接受逆时针 (绿色) 和顺时针 (粉色) 角速度输入，与头朝向细胞有非对称性连接 (右图)，以此牵引头朝向吸引子网络的兴奋位置移动。(d) 引入角速度输入下的吸引子态迁移后，CAN 内神经元编码真实的头朝向。

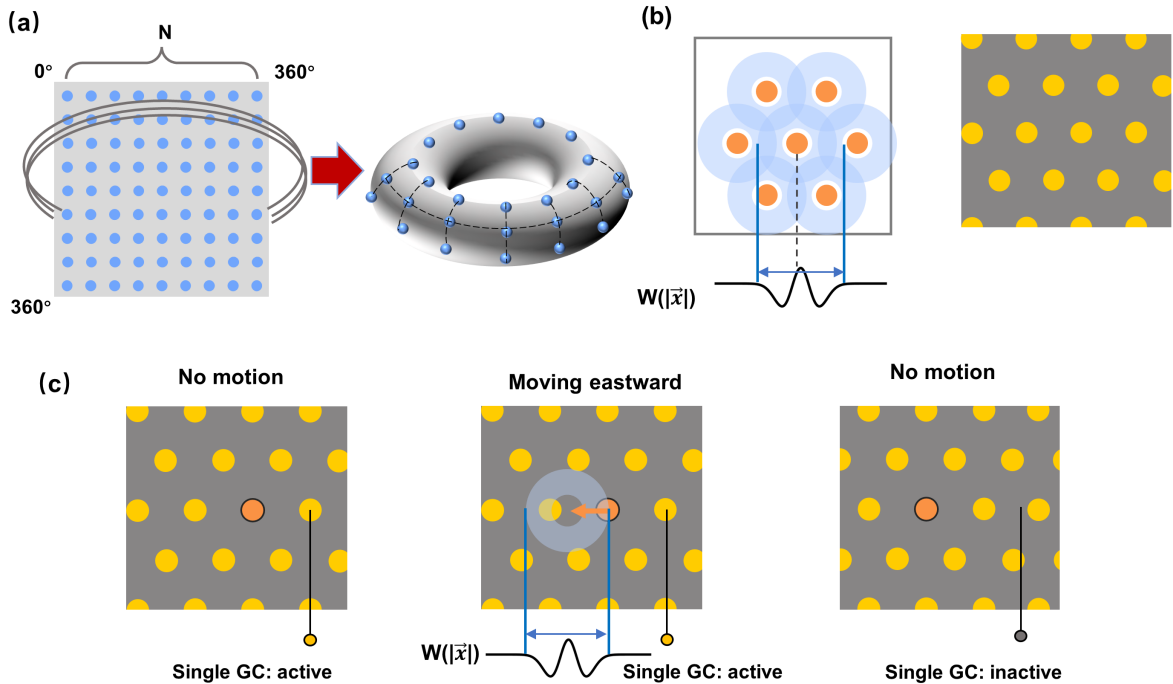


图 5. 网格细胞的二维 CAN 模型。(a) 二维网格细胞片层满足周期性边界条件，即边缘神经元与对边相连 (左图)；从拓扑的角度看，神经元位于圆环面上，神经元间的连接强度依赖于它们在圆环面上的距离 (右图)。改编自文献<sup>[14]</sup> 的图 3。(b) 特定的连接形式下，网格细胞片层产生正六边形的激活点阵。左：连接结构示意图，橙色内圆标识兴奋或去抑制区域，蓝色环形为抑制区域。右：六边形激活点阵，黄色圆表示激活的细胞群体。(c) 网格细胞方向敏感性下的对称性破缺引入。以向东运动为例，动物运动激活东向运动敏感细胞，通过向左偏移的连接形式牵引激活区域向左移动，完成吸引子态的转变。标出的橙色圆为被牵引的示例活跃区域。黑线指被选中的单个网格细胞。

通过引入对称性破缺,吸引子态与真实空间的位置锚定,实现网格细胞的位置编码。假设网格细胞具有方向敏感性,影响其信息输入,并使其对其他神经元的连接产生敏感方向的偏移。此时,输入信息与连接矩阵的变化可描述为:

$$B_j = 1 + \alpha \hat{\mathbf{e}}_{\theta_j} \cdot \mathbf{v}, \quad (8)$$

$$W_{ij}(\mathbf{x}) = \alpha e^{-\gamma|\mathbf{x}-l\hat{\mathbf{e}}_{\theta_j}|^2} - e^{-\beta|\mathbf{x}-l\hat{\mathbf{e}}_{\theta_j}|^2}, \quad (9)$$

其中,  $\hat{\mathbf{e}}_{\theta_j}$  为网格细胞  $j$  敏感方向的单位向量,  $l$  为固定的连接偏移参数,  $\mathbf{v}$  为动物运动速度。动物运动时,运动方向敏感的细胞群体高度兴奋,使网络兴奋区域沿其敏感方向迁移,完成吸引子态的转变(图 5 (c))。还可以引入速度矢量相关神经元来实现对称性破缺:假设存在同时对位置与速度矢量敏感的神经元群,它们以非对称形式连接网格细胞片层。当特定运动出现时,对应神经元片层兴奋,牵引网格细胞激活位置迁移,完成吸引子态与位置空间的映射。这两种引入对称性破缺的方式原理相同,区别仅在于对网络吸引子态的牵引是通过其他神经元群,还是由网格细胞自身完成。此外,在将位置空间映射到连续吸引子态时,由于吸引子态的激活点阵呈正六边形,单个网格细胞的放电野也按正六边形周期排列。

二维 CAN 模型成功复现了网格细胞的特异性放电模式,实现了对头朝向信息与运动信息的整合,其合理性被后续实验证实。首先,解剖学和电生理实验结果显示,MEC 第 II 层的星形细胞(即网格细胞)之间缺少直接的兴奋性连接,而是通过中间神经元产生抑制性连接<sup>[51]</sup>,与 CAN 广域抑制的假设相符。其次,2015 年的实验发现了编码线速度的速度细胞<sup>[11]</sup>,其放电频率与动物速度呈线性相关,符合 CAN 模型的预测。最后,网格细胞放电模式在相空间中呈现圆环面流形<sup>[19,27]</sup>,且流形在不同环境下与行为状态中保持稳定<sup>[27]</sup>,与吸引子态的不变性相符。

CAN 模型成功再现了头朝向细胞与网格细胞的特异性放电行为,实现了对多重输入信息的整合与导航信息的表征,其理论预测与诸多实验结果高度一致。因此,CAN 模型被广泛用于探究导航计算机制。

#### D. 吸引子动力学下的 theta 振荡机制

除吸引子动力学外,哺乳动物空间导航神经环路还呈现显著的节律特征,如网格细胞与位置细胞均存在清晰的 theta 节律性(4–12 Hz)。网格细胞的阈下膜电位(subthreshold membrane potential)<sup>[52]</sup>、神经元发放<sup>[53]</sup>与局域场电位(local field potential; LFP)<sup>[54]</sup>均存在

theta 振荡,其对放电野的形成十分重要<sup>[55]</sup>。网格细胞的放电活动与 LFP 存在相位关联:MEC 第 II 层细胞的放电与 LFP 振荡呈现相位进动(phase precession)现象,即当动物沿某一方向穿过放电野时,细胞放电时刻在 LFP 的相位逐渐提前<sup>[56]</sup>;在 MEC 深层中,二者主要呈现锁相关系,网格细胞的发放锁定于 LFP 某一固定相位<sup>[57]</sup>。相位进动现象同样出现在位置细胞中,并显著影响位置编码的准确性<sup>[58]</sup>。位置细胞的发放还存在相位退动(phase procession)现象(图 6 (a)),其相位进动和退动交替出现,可能实现了对位置编码的“前瞻”(prospective mode)与“回顾”(retrospective mode)<sup>[59]</sup>。

上述相位关系表明导航系统存在多类型、多来源的节律性相互作用,也暗示了可能的时空信息转化过程,如在相位进动或退动中,动物空间位置可以由发放相位的时间信息读出,完成空间信息到时间信息的转化(图 6 (b))。因此,了解导航细胞振荡的产生机制是理解导航时空信息交互的基础。有观点认为,导航系统 theta 振荡可能源自其他脑区的振荡输入,如内侧隔核-布罗卡斜带(medial septum and diagonal band of Broca; MS-DBB)对 MEC 和海马区的振荡输入。但有新的建模工作指出,CAN 中能形成内源性的 theta 振荡<sup>[62]</sup>。接下来,我们介绍两项 CAN 模型对导航系统振荡机制的阐释工作,作为对连续吸引子理论的拓展。

2023 年,我们发表了关于网格细胞 CAN 模型中内禀 theta 振荡的工作,指出 theta 振荡可能来自局域 CAN 特异性突触连接结构<sup>[62]</sup>。建立了包含主神经元(E)与中间神经元(I)、类网格细胞的一维 CAN 模型,实现两个或多个吸引子与 theta 振荡共存。由于 E-I 和 I-E 突触连接强度的最大值分别出现在相位差  $0^\circ$  和  $90^\circ$  处,导致中间神经元的分组分工,使吸引子与 theta 振荡共存。一方面,吸引子内主神经元通过吸引子内的中间神经元产生竞争性正反馈作用,维持了吸引子态的稳定;另一方面,吸引子内的主神经元通过吸引子外的中间神经元产生负反馈作用,并在 NMDA 受体介导的慢突触电流下产生了 theta 波段的振荡(图 7)。该工作提出了网格细胞 theta 节律产生的新机制,为吸引子框架下探讨振荡在路径积分与空间导航中的作用提供了基础。

2023 年吴思教授团队在 CAN 模型中,引入神经元的适应性反应,阐释了位置细胞的特殊相位表现与位置编码<sup>[63]</sup>。通过在一维 CAN 中引入外部输入与反馈抑制的竞争性作用,模型成功复现了位置细胞对于输入位置的振荡追踪,产生了交替出现、向前向后的 theta 振荡序列,并再现了单神经元的相位进动和退动现象。该研究对吸引子框架下振荡相关的精准位置编码具有深远的影响。通过对吸引子与振荡特征的统一描述,位



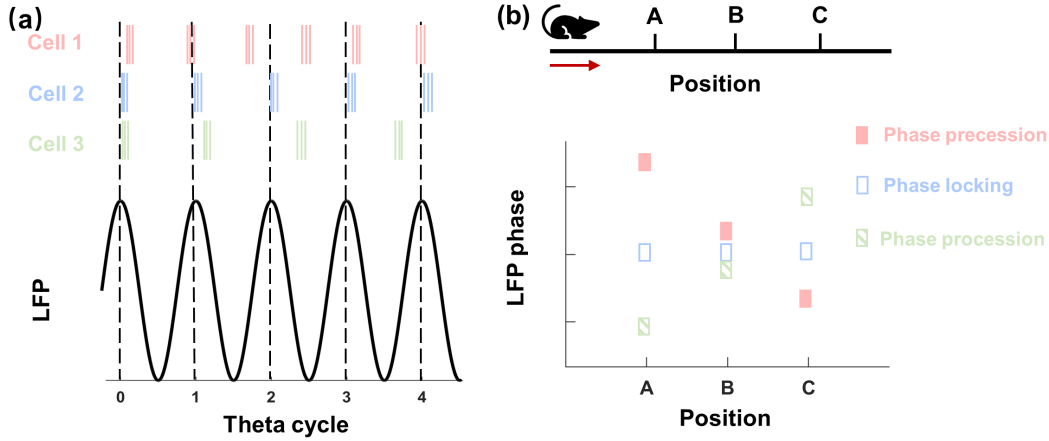


图 6. 导航神经元的相位表现。(a) 导航神经元相位进动 (Cell 1), 锁相 (Cell 2) 与相位退后 (Cell 3)。(b) 相位相关的时空信息转换。动物穿越放电野时, 导航神经元的发放在 LFP 相位上逐渐提前 (相位进动) 或逐渐退后 (相位退后)。LFP: Local field potential。

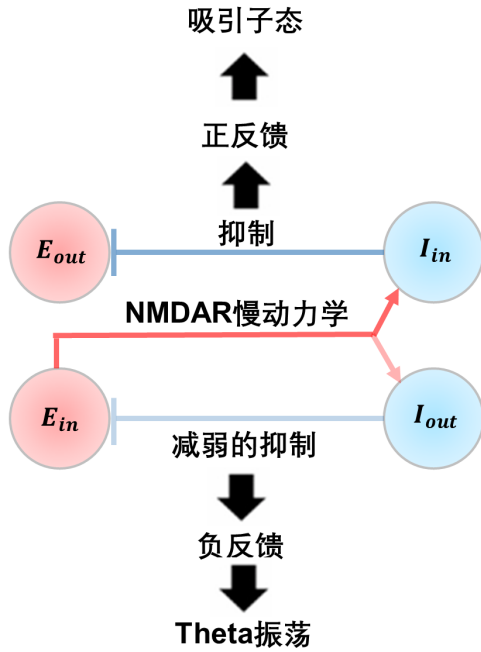


图 7. 吸引子网络中的内源性 theta 振荡<sup>[62]</sup>。在特定网络连接结构与突触动力学下, 由中间神经元分组而导致的吸引子与 theta 振荡共存。

置细胞的时空信息被巧妙地结合起来, 共同完成位置编码。

许多研究指出, 导航环路中的吸引子和振荡两种动力学在整合多种感知觉信息的过程中具有独特作用, 二者结合能够更好地编码与记忆空间位置信息及其先后顺序。吸引子理论框架下对振荡机制的研究, 整合了导航神经元的空间编码与时序编码, 为导航时空信息交互、脑区间信息传递的研究奠定了基础, 对导航系统吸引子理论的发展及导航计算机制的解析有重要意义。

### E. 吸引子动力学视角下的人类导航功能损伤

与动物类似, 人类的导航功能同样依赖于内嗅皮层 (entorhinal cortex; EC) 与海马等脑区。大量网格细胞具有相同的放电野点阵朝向 (图 3 (a)), 细胞群体的空间激活模式具有六重旋转对称性。利用功能磁共振成像 (functional magnetic resonance imaging; fMRI), 人们在体素 (Voxel) 水平上研究网格细胞的激活特征, 即当人类被试在虚拟现实环境中运动时, EC 内体素的血氧水平依赖 (blood oxygenation level dependent; BOLD) 信号具有依赖运动方向的六重旋转对称性<sup>[64]</sup>。基于沿网格朝向快速运动时神经元群体更活跃, 可计算每个体素内网格细胞放电野的朝向。对于正常被试, 实验发现 EC 内网格朝向的分布非常集中, 并且在时间上很稳定, 不同时间段的数据显示基本一致的网格朝向<sup>[25]</sup>。在局域吸引子网络中, 神经元的网格朝向都是相同且稳定的, 而神经元分布在较大的 EC 范围内, 跨多个 fMRI 体素, 这与实验结果是一致的。此外, 基于癫痫病人的颅内脑电记录, 研究人员发现其 EC 与腹内侧前额叶皮层 (ventromedial prefrontal cortex; vmPFC) 的 theta 振荡功率也出现了依赖被试运动方向的六重旋转对称性<sup>[65,66]</sup>。

研究发现, 携带阿尔兹海默症 (Alzheimer's disease; AD) 风险基因 APOE4 的人群, 其在 fMRI 下的网格表征较正常人群明显减弱, 并且导航能力显著下降<sup>[25]</sup>。风险基因人群 EC 体素计算得到的网格朝向的时间稳定性很差, 不同时段数据给出的网格朝向几乎毫无关联, 但在固定时段内, 不同体素的网格朝向分布集中度较正常人群并未明显降低<sup>[25]</sup>。在导航表现上, 风险基因人群更倾向于依靠环境边界或者地标附近运动<sup>[25]</sup>。进一步研

究表明, 风险基因人群依靠自运动信息的纯路径积分能力较正常人群明显减弱, 但依靠边界或地标的导航能力则无明显差别<sup>[67]</sup>。从吸引子动力学的角度来看, 在每个时段内, 风险基因人群不同 EC 体素的网格朝向比较集中, 暗示了神经元的群体放电模式没有变化, 依然服从吸引子动力学, 与运动方向的联系并没有被破坏, 这间接说明风险人群依然能通过自运动信息进行路径积分。而 EC 体素网格朝向的时间稳定性差, 则只能理解为吸引子动力学的稳定性降低了。这可能有两种解释: 吸引子内部的连接减弱, 特别是兴奋性连接的减弱, 使吸引子的随机漂移增大, 导致网格朝向不稳定; 或者, 尽管吸引子的随机漂移没有明显增强, 但是利用视觉等外界输入信息校正吸引子态的能力降低了, 导致随机漂移的误差不断累积, 出现网格朝向不断改变。病理学上, AD 的出现与 tau 蛋白的神经纤维缠结 (neurofibrillary tangles; NFTs) 和 A $\beta$  蛋白的沉积相关; 其中, tau 蛋白的病理学改变在早期病程中就已经出现在 EC, 可能与 AD 患者的导航能力下降相关<sup>[68]</sup>。此外, 表达突变 tau 蛋白的转基因小鼠呈现 EC 兴奋性神经元损伤, 导致网格表征能力的下降<sup>[69]</sup>。这支持了前述第一种可能性, 后续的动物模型应进一步探究该可能性与 AD 的因果联系。当然, 也应该设计实验来验证第二种可能性。总之, 基于吸引子动力学的分析, 人们能够更准确地确定导航功能损伤的神经计算环节。

通过对导航过程中吸引子动力学的研究, 人们能够更好地理解 AD 风险人群导航功能的损伤, 揭示相关病理的特异性动力学变化机制, 将此前的相关性研究向因果性研究推进, 对 AD 的早期诊断与靶向药物设计具有重大意义。

## V. 总结与展望

本文以吸引子理论为基础, 着重描述了头朝向细胞与网格细胞放电行为的吸引子机制及相关的 CAN 模型, 介绍了 CAN 模型中的振荡行为及其功能意义, 并分析了人类导航认知损伤的动力学机制。对导航系统吸引子动力学的深入研究有助于理解空间认知行为及其计算原理。吸引子理论在导航系统的成功应用不仅丰富了这一理论框架, 也拓展了其在中枢神经系统的应用范围, 为探索普适的神经计算原理提供了支撑。

在成功复现导航细胞的放电特征后, 有必要进一步探究导航信息的处理机制。这是一个更具挑战性的课题。首先, 连续吸引子对沿着吸引子流形的噪声抵抗较弱, 容易产生空间编码的随机漂移, 但这种漂移可通过边界信

息或感知觉信息的输入来有效规避<sup>[41,45]</sup>; 多模态信息如何在神经环路层面整合, 产生准确而稳健的空间表征将是导航研究的重点。其次, 导航系统存在多个吸引子网络, 包括网格细胞与头朝向细胞网络, 但它们的相互作用和信息传递方式仍不明确。有研究表明, 多模块吸引子网络的存在可能保证了位置编码的准确性<sup>[70]</sup>。最后, 目前对导航系统的研究仍集中在空间位置信息的编码与解码, 对导航路径规划的研究相对较少。选择合适的研究对象, 先行开展研究, 是有效策略。研究发现, 果蝇通过对比目标朝向与当前朝向, 决定其角速度与线速度<sup>[71]</sup>。研究果蝇等低维导航系统的路径规划机制, 有助于理解哺乳动物的导航行为。

吸引子动力学在中枢神经系统的广泛存在, 暗示这可能是大脑信息处理的一种普适机制。然而, 吸引子动力学对神经环路的结构有较高要求, 如维持稳态的兴奋/抑制关系、连续吸引子动力学要求的平移不变性等。生物神经网络是否具有这些特征, 又是如何形成的? 有研究发现, 在突触可塑性的作用下, 导航网络可以在感知觉输入的影响下, 形成准吸引子的连接形式、特定对称性破缺的信息输入模式<sup>[72,73]</sup>。除此之外, 凭借视觉线索, 动物还能够实现路径积分的重新调整<sup>[74]</sup>。在突触可塑性机制下, 神经网络能灵活调整其相空间吸引子流形, 增强吸引子动力学编码的灵活性。研究可塑性机制下吸引子动力学的产生与优化, 将帮助我们深入理解外界环境与脑功能的交互影响。

大脑的信息编码模式是脑科学的重要研究内容。吸引子动力学最直接的编码方式就是群体矢量编码 (population vector coding), 而振荡节律的存在则暗示了基于放电时刻相对于场电位的位相或者不同节律互作用的相位编码 (phase coding)。两种模式在记忆与导航等认知过程中普遍存在, 但始终存在争议。研究发现, 两种编码模式可能同时存在并交互关联, 如位置细胞在进入与离开放电野时的慢速与快速 gamma 振荡分别参与空间记忆的编码与检索<sup>[75]</sup>, 其 theta 相位进动则编码了动物在放电野的具体位置, 相位编码与群体矢量编码的结合将产生更准确的位置编码与记忆。近期研究还发现, 基于电突触的短时程易化 (short-term facilitation), 神经环路可以由早期的高频放电编码演变为后期的低频神经元群体编码, 体现了大脑动态的编码策略<sup>[76]</sup>。对编码模式的研究将帮助我们揭示多模态编码模式的转化与信息整合, 探索大脑编码的不同策略。

近年来, 人工神经网络的发展与导航系统的 CAN 动力学研究相互促进。在完成导航任务的人工神经网络中, 出现了类似吸引子网络的架构和动力学行为<sup>[77,78]</sup>, 且其导航表现显著优于其他人工神经网络, 这有助于揭

示导航行为的核心机制。另一方面, 预设了吸引子动力学的人工神经网络能够更快速、准确地完成导航相关任务, 如借鉴生命系统导航信息整合过程的人工神经网络能够精准地将自我中心运动信息转换为异我中心的位置信息, 并在多种输入下准确解码智能体的位置<sup>[79]</sup>。特别地, 预设了连续吸引子架构的天玑芯片成功实现了自行车的自动行驶, 实现对目标的灵活追踪与自动避障<sup>[80]</sup>, 表明连续吸引子动力学在智能导航领域拥有广泛的应用情景。

总之, 探究空间导航网络中的吸引子动力学, 能够帮助我们理解人和动物在空间环境中的导航和决策行为, 理解神经元如何通过复杂的放电模式编码和处理空间信息, 辅助开发新的导航和决策算法 (如基于神经网络的机器人导航系统), 理解与空间认知有关疾病的发病机制, 例如方向感缺失、空间认知障碍等, 为这些疾病的诊断和治疗提供新的可能途径。

## 致 谢

本文部分内容来自王子群的博士毕业论文。本文受科技部科技创新 2030-2021ZD0201300 项目的支持。

## 参考文献

- [1] WANG X J, HU H, HUANG C, et al. Computational neuroscience: a frontier of the 21st century[J]. Natl. Sci. Rev., 2020, 7(9): 1418.
- [2] KHONA M, FIETE I R. Attractor and integrator networks in the brain[J]. Nat. Rev. Neurosci., 2022, 23(12): 744.
- [3] LI N, DAIE K, SVOBODA K, et al. Robust neuronal dynamics in premotor cortex during motor planning[J]. Nature, 2016, 532(7600): 459.
- [4] SEUNG H S. How the brain keeps the eyes still[J]. Proc. Natl. Acad. Sci. U S A, 1996, 93(23): 13339.
- [5] AKSAY E, GAMKRELIDZE G, SEUNG HS, et al. In vivo intracellular recording and perturbation of persistent activity in a neural integrator[J]. Nat. Neurosci., 2001, 4(2): 184.
- [6] KRIENER B, CHAUDHURI R, FIETE I R. Robust parallel decision-making in neural circuits with nonlinear inhibition[J]. Proc. Natl. Acad. Sci. U S A, 2020, 117(41): 25505.
- [7] O'KEEFE J, DOSTROVSKY J. The hippocampus as a spatial map. preliminary evidence from unit activity in the freely-moving rat[J]. Brain Res., 1971, 34(1): 171.
- [8] TAUBE J S, MULLER R U, RANCK J B JR. Head-direction cells recorded from the postsubiculum in freely moving rats. I. description and quantitative analysis[J]. J. Neurosci., 1990, 10(2): 420.
- [9] FYHN M, MOLDEN S, WITTER M P, et al. Spatial representation in the entorhinal cortex[J]. Science, 2004, 305(5688): 1258.
- [10] HAFTING T, FYHN M, MOLDEN S, et al. Microstructure of a spatial map in the entorhinal cortex[J]. Nature, 2005, 436(7052): 801.
- [11] KROPFF E, CARMICHAEL J E, MOSER M B, et al. Speed cells in the medial entorhinal cortex[J]. Nature, 2015, 523(7561): 419.
- [12] BARRY C, LEVER C, HAYMAN R, et al. The boundary vector cell model of place cell firing and spatial memory[J]. Rev. Neurosci., 2006, 17(1-2): 71.
- [13] MOSER E I, MOSER M B, MCNAUGHTON B L. Spatial representation in the hippocampal formation: a history[J]. Nat. Neurosci., 2017, 20(11): 1448.
- [14] MCNAUGHTON B L, BATTAGLIA F P, JENSEN O, et al. Path integration and the neural basis of the 'cognitive map' [J]. Nat. Rev. Neurosci., 2006, 7(8): 663.
- [15] VAN ROSSUM M C W, BI G Q, TURRIGIANO G G. Stable hebbian learning from spike timing-dependent plasticity[J]. J. Neurosci., 2000, 20(23): 8812.
- [16] MEJIAS J F, WANG X J. Mechanisms of distributed working memory in a large-scale network of macaque neocortex[J]. eLife, 2022, 11: e72136.
- [17] WANG X J. Synaptic basis of cortical persistent activity: the importance of NMDA receptors to working memory[J]. J. Neurosci., 1999, 19(21): 9587.
- [18] BURAK Y, FIETE I R. Accurate path integration in continuous attractor network models of grid cells[J]. PLoS Comput. Biol., 2009, 5(2): e1000291.
- [19] CHAUDHURI R, GERÇEK B, PANDEY B, et al. The intrinsic attractor manifold and population dynamics of a canonical cognitive circuit across waking and sleep[J]. Nat. Neurosci., 2019, 22(9): 1512.
- [20] HODGKIN A L, HUXLEY A F. A quantitative description of membrane current and its application to conduction and excitation in nerve[J]. J. Physiol., 1952, 117(4): 500.
- [21] BURKITT A N. A review of the integrate-and-fire neuron model: I. homogeneous synaptic input[J]. Biol. Cybern., 2006, 95(1): 1.
- [22] HOPFIELD J J. Neurons with graded response have collective computational properties like those of two-state neurons[J]. Proc. Natl. Acad. Sci. U S A, 1984, 81(10): 3088.
- [23] KIM S S, ROUAULT H, DRUCKMANN S, et al. Ring attractor dynamics in the Drosophila central brain[J].

- Science, 2017, 356(6340): 849.
- [24] DIEHL P U, COOK M. Unsupervised learning of digit recognition using spike-timing-dependent plasticity[J]. *Front. Comput. Neurosci.*, 2015, 9: 99.
- [25] KUNZ L, SCHRÖDER T N, LEE H, et al. Reduced grid-cell-like representations in adults at genetic risk for Alzheimer's disease[J]. *Science*, 2015, 350(6259): 430.
- [26] FANG Y, COHEN M A, KINCAID T G. Dynamics of a winner-take-all neural network[J]. *Neural Netw.*, 1996, 9(7): 1141.
- [27] SAJIKUMAR S, MORRIS R G M, KORTE M. Competition between recently potentiated synaptic inputs reveals a winner-take-all phase of synaptic tagging and capture[J]. *Proc. Natl. Acad. Sci. U S A*, 2014, 111(33): 12217.
- [28] SHEN J Y, DING T Y, ZHANG Y X, et al. Real-time pattern recognition system based on a bipolar winner-take-all model with a threshold[J]. *Opt. Lett.*, 1994, 19(24): 2158.
- [29] WU S, HAMAGUCHI K, AMARI S. Dynamics and computation of continuous attractors [J]. *Neural Comput.*, 2008, 20(4): 994.
- [30] ZHANG K. Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory[J]. *J. Neurosci.*, 1996, 16(6): 2112.
- [31] WU S, AMARI S. Computing with continuous attractors: stability and online aspects[J]. *Neural Comput.*, 2005, 17(10): 2215.
- [32] COMPTE A, BRUNEL N, GOLDMAN-RAKIC P S, et al. Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model[J]. *Cereb. Cortex*, 2000, 10(9): 910.
- [33] GARDNER R J, HERMANSEN E, PACHITARIU M, et al. Toroidal topology of population activity in grid cells[J]. *Nature*, 2022, 602(7895): 123.
- [34] XUE C, LIU F. Structured synaptic inhibition has a critical role in multiple-choice motion-discrimination tasks[J]. *J. Neurosci.*, 2014, 34(40): 13444.
- [35] MARDER E, BUCHER D. Central pattern generators and the control of rhythmic movements[J]. *Curr. Biol.*, 2001, 11(23): R986.
- [36] ADAMANTIDIS A R, GUTIERREZ HERRERA C, GENT T C. Oscillating circuitries in the sleeping brain[J]. *Nat. Rev. Neurosci.*, 2019, 20(12): 746.
- [37] LUNDQVIST M, HERMAN P, MILLER E K. Working memory: delay activity, yes! persistent activity? maybe not[J]. *J. Neurosci.*, 2018, 38(32): 7013.
- [38] TAUBE J S. Head direction cells recorded in the anterior thalamic nuclei of freely moving rats[J]. *J. Neurosci.*, 1995, 15(1): 70.
- [39] STACKMAN R W, TAUBE J S. Firing properties of rat lateral mammillary single units: head direction, head pitch, and angular head velocity[J]. *J. Neurosci.*, 1998, 18(21): 9020.
- [40] STACKMAN R W, CLARK A S, TAUBE J S. Hippocampal spatial representations require vestibular input[J]. *Hippocampus*, 2002, 12(3): 291.
- [41] GOODRIDGE J P, DUDCHENKO P A, WORBOYS K A, et al. Cue control and head direction cells[J]. *Behav. Neurosci.*, 1998, 112(4): 749.
- [42] CABALLERO-BLEDA M, WITTER M P. Regional and laminar organization of projections from the pre-subiculum and parasubiculum to the entorhinal cortex: an anterograde tracing study in the rat[J]. *J. Comp. Neurol.*, 1993, 328(1): 115.
- [43] SARGOLINI F, FYHN M, HAFTING T, et al. Conjunctive representation of position, direction, and velocity in entorhinal cortex[J]. *Science*, 2006, 312(5774): 758.
- [44] KRUPIC J, BAUZA M, BURTON S, et al. Grid cell symmetry is shaped by environmental geometry[J]. *Nature*, 2015, 518(7538): 232.
- [45] FYHN M, HAFTING T, TREVES A, et al. Hippocampal remapping and grid realignment in entorhinal cortex[J]. *Nature*, 2007, 446(7132): 190.
- [46] STENSOLA H, STENSOLA T, SOLSTAD T, et al. The entorhinal grid map is discretized[J]. *Nature*, 2012, 492(7427): 72.
- [47] DUNN B, MØRREAUNET M, ROUDI Y. Correlations and functional connections in a population of grid cells[J]. *PLoS Comput. Biol.*, 2015, 11(2): e1004052.
- [48] TRETTEL S G, TRIMPER J B, HWAUN E, et al. Grid cell co-activity patterns during sleep reflect spatial overlap of grid fields during active behaviors[J]. *Nat. Neurosci.*, 2019, 22(4): 609.
- [49] GARDNER R J, LU L, WERNLE T, et al. Correlation structure of grid cells is preserved during sleep[J]. *Nat. Neurosci.*, 2019, 22(4): 598.
- [50] TURNER-EVANS D, WEGENER S, ROUAULT H, et al. Angular velocity integration in a fly heading circuit[J]. *eLife*, 2017, 6: e23496.
- [51] COUEY J J, WITOELAR A, ZHANG S J, et al. Recurrent inhibitory circuitry as a mechanism for grid formation[J]. *Nat. Neurosci.*, 2013, 16(3): 318.
- [52] GIOCOMO L M, ZILLI E A, FRANSEN E, et al. Temporal frequency of subthreshold oscillations scales with entorhinal grid cell field spacing[J]. *Science*, 2007, 315(5819): 1719.
- [53] JEEWAJEE A, BARRY C, O'KEEFE J, et al. Grid cells and theta as oscillatory interference: electrophysiological data from freely moving rats[J]. *Hippocampus*, 2008, 18(12): 1175.

- [54] MIZUSEKI K, SIROTA A, PASTALKOVA E, et al. Theta oscillations provide temporal windows for local circuit computation in the entorhinal-hippocampal loop[J]. *Neuron*, 2009, 64(2): 267.
- [55] KOENIG J, LINDER A N, LEUTGEB J K, et al. The spatial periodicity of grid cells is not sustained during reduced theta oscillations[J]. *Science*, 2011, 332(6029): 592.
- [56] HAFTING T, FYHN M, BONNEVIE T, et al. Hippocampus-independent phase precession in entorhinal grid cells[J]. *Nature*, 2008, 453(7199): 1248.
- [57] NEWMAN E L, HASSELMO M E. Grid cell firing properties vary as a function of theta phase locking preferences in the rat medial entorhinal cortex[J]. *Front. Syst. Neurosci.*, 2014, 8: 193.
- [58] SKAGGS W E, MCNAUGHTON B L, WILSON M A, et al. Theta phase precession in hippocampal neuronal populations and the compression of temporal sequences[J]. *Hippocampus*, 1996, 6(2): 149.
- [59] WANG M, FOSTER D J, PFEIFFER B E. Alternating sequences of future and past behavior encoded within hippocampal theta oscillations[J]. *Science*, 2020, 370(6513): 247.
- [60] ROBINSON J, MANSEAU F, DUCHARME G, et al. Optogenetic activation of septal glutamatergic neurons drive hippocampal theta rhythms[J]. *J. Neurosci.*, 2016, 36(10): 3016.
- [61] VAN CAUTER T, POU CET B, SAVE E. Unstable CA1 place cell representation in rats with entorhinal cortex lesions[J]. *Eur. J. Neurosci.*, 2008, 27(8): 1933.
- [62] WANG Z, WANG T, YANG F, et al. Intrinsic theta oscillation in the attractor network of grid cells[J]. *iScience*, 2023, 26(4): 106351.
- [63] CHU T, JI Z, ZUO J, et al. Firing rate adaptation affords place cell theta sweeps, phase precession and procession[J]. *eLife*, 2023, 12: RP87055
- [64] DOELLER C F, BARRY C, BURGESS N. Evidence for grid cells in a human memory network[J]. *Nature*, 2010, 463(7281): 657.
- [65] CHEN D, KUNZ L, WANG W, et al. Hexadirectional modulation of theta power in human entorhinal cortex during spatial navigation[J]. *Curr. Biol.*, 2018, 28(20): 3310.
- [66] CHEN D, KUNZ L, LV P, et al. Theta oscillations coordinate grid-like representations between ventromedial prefrontal and entorhinal cortex[J]. *Sci. Adv.*, 2021, 7(44): eabj0200.
- [67] BIERBRAUER A, KUNZ L, GOMES C A, et al. Unmasking selective path integration deficits in Alzheimer's disease risk carriers[J]. *Sci. Adv.*, 2020, 6(35): eaba1394.
- [68] COUGHLAN G, LACZÓ J, HORT J, et al. Spatial navigation deficits-overlooked cognitive marker for pre-clinical Alzheimer disease[J]. *Nat. Rev. Neurol.*, 2018, 14(8): 496.
- [69] FU H, RODRIGUEZ G A, HERMAN M, et al. Tau pathology induces excitatory neuron loss, grid cell dysfunction, and spatial memory deficits reminiscent of early Alzheimer's disease[J]. *Neuron*, 2017, 93(3): 533.
- [70] WANG T, YANG F, WANG Z Q, et al. Modularization of grid cells constrained by the pyramidal patch lattice[J]. *iScience*, 2021, 24(4): 102301.
- [71] GREEN J, VIJAYAN V, MUSSELLS PIRES P, et al. A neural heading estimate is compared with an internal goal to guide oriented navigation[J]. *Nat. Neurosci.*, 2019, 22(9): 1460.
- [72] WEBER S N, SPREKELER H. Learning place cells, grid cells and invariances with excitatory and inhibitory plasticity[J]. *eLife*, 2018, 7: e34560.
- [73] VAFIDIS P, OWALD D, D'ALBIS T, et al. Learning accurate path integration in ring attractor models of the head direction system[J]. *eLife*, 2022, 11: e69841.
- [74] JAYAKUMAR R P, MADHAV M S, SAVELLI F, et al. Recalibration of path integration in hippocampal place cells[J]. *Nature*, 2019, 566(7745): 533.
- [75] BIERI K W, BOBBITT K N, COLGIN L L. Slow and fast  $\gamma$  rhythms coordinate different spatial coding modes in hippocampal place cells[J]. *Neuron*, 2014, 82(3): 670.
- [76] LI L, MI Y, ZHANG W, et al. Dynamic information encoding with dynamic synapses in neural adaptation[J]. *Front. Comput. Neurosci.*, 2018, 12: 16.
- [77] BANINO A, BARRY C, URIA B, et al. Vector-based navigation using grid-like representations in artificial agents[J]. *Nature*, 2018, 557(7705): 429.
- [78] SORSCHER B, MEL G C, OCKO S A, et al. A unified theory for the computational and mechanistic origins of grid cells[J]. *Neuron*, 2023, 111(1): 121.
- [79] ZHAO D, ZHANG Z, LU H, et al. Learning cognitive map representations for navigation by sensory-motor integration[J]. *IEEE Trans. Cybern.*, 2022, 52(1): 508.
- [80] PEI J, DENG L, SONG S, et al. Towards artificial general intelligence with hybrid tianjic chip architecture[J]. *Nature*, 2019, 572(7767): 106.



## Attractor Dynamics in Spatial Cognition

WANG Zi-qun <sup>\*</sup>, WANG Tao <sup>\*</sup>, LIU Feng <sup>†</sup>

Department of Physics and Institute for Brain Sciences, Nanjing 210093, China

**Abstract:** The mammalian navigation system comprises various kinds of neurons responsible for position perception and spatial path planning, involving the integration of multiple information sources. As a unified brain theory capable of providing explanations for complex cognitive functions like memory and decision-making, the theory of attractor dynamics can elucidate the firing dynamics of neurons and path integration in the navigation system. This review describes recent advances in attractor dynamics in spatial cognition. First, it provides a brief overview of computational neuroscience and the general theory of attractor dynamics. Subsequently, focusing on the continuous attractor dynamics, it delves into the dynamical characteristics and functional significance of head direction cells and grid cells. Finally, an extension and prospects of the attractor theory for spatial cognition are presented.

**Key words:** computational neuroscience; attractor dynamics; navigation; head direction cells; grid cells

---

<sup>\*</sup> Co-First author

<sup>†</sup> E-mail: fliu@nju.edu.cn