

Stephen Coombes · Peter beim Graben
Roland Potthast · James Wright *Editors*

Neural Fields

Theory and Applications



Springer

Neural Fields

Stephen Coombes • Peter beim Graben •
Roland Potthast • James Wright
Editors

Neural Fields

Theory and Applications



Springer

Editors

Stephen Coombes
School of Mathematical Sciences
University of Nottingham
Nottingham, United Kingdom

Peter beim Graben
Department of German Studies
and Linguistics
Humboldt-Universität zu Berlin
Berlin, Germany

Roland Potthast
Department of Mathematics
University of Reading
Reading, United Kingdom

James Wright
School of Medicine
University of Auckland
Auckland, New Zealand

ISBN 978-3-642-54592-4

ISBN 978-3-642-54593-1 (eBook)

DOI 10.1007/978-3-642-54593-1

Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014942371

© Springer-Verlag Berlin Heidelberg 2014

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The Neural Field: A Framework for Brain Data Integration?

This book presents a perspective on the advancing subject of neural fields—that is, theories of brain organization and function in which the interaction of billions of neurons is treated as a continuum. The intention is to reduce the enormous complexity of neuronal interactions to simpler, population properties that are tractable by analytical mathematical tools. By so doing, it is hoped that the theory of brain function can be reduced to its essence, without becoming lost in a wealth of inessential detail. Naturally, this begs the questions of what the “essence” is, and what detail is inessential [1]. The questions themselves are timely for more than neural field theory. Putting aside the most profound of philosophical issues—the existential relation between objective brain function and subjective consciousness—at the cellular level, research has achieved detailed knowledge of individual neuron physiology, and at the gross level, considerable knowledge of sensory processing, the generation of movement and the functional locations in the brain of memory, learning, emotion and decision-making. Yet our knowledge of the functional details of all these processes remains vague, and little surpasses the views held by Sherrington [4]. The ever-accumulating body of experimental data, gathered with ever-improving observational techniques, continues to promise that fundamental understanding of the modes of operation of the brain may be possible—yet the goal seems also to move away, like a mirage, because, despite the mass of data, there is no agreed means to achieve the needed integration. A crisis of confidence looms. It is to be hoped that such a crisis is a healthy state—the darkness before the dawn—analogous to the problems of systematic biology before Darwin, or of astronomy before Kepler, or, more recently, of atomic physics before Bohr, but hope alone will not suffice.

Aware of the risk of becoming trapped in an overwhelming mass of undigested detail, large groups of scientists are joining forces to address the problems of integration. While organizing collaborative efforts of scale unprecedented in neuroscience, all concerned agree on the importance both of technological advances

and of theoretical development, but there are many differences of opinion on the best and shortest route to success. In Europe the *Human Brain Project* [6] is aimed at large scale simulation of the brain, employing very detailed cellular properties. In the United States, the *Brain Activity Map* [5] seeks to establish a functional connectome of the entire brain, and the *MindScope Project* [3] intends to obtain a complete model of the mouse visual cortex. The *BRAIN (Brain Research through Advancing Innovative Neurotechnologies) Initiative* [7] aims to accelerate techniques for study of the brain.

Unresolved questions and fears, around which controversy centres, are:

Do we yet have enough detailed data on structure? How much knowledge of exact connectivity in the brain is enough? Established anatomical techniques are not depleted of possibility to resolve more detail, and very sophisticated new technology is being deployed to add further to this. Yet the capacity of individuals to undergo profound brain damage or deformity of brain development without loss of essential function makes the need for such precise detail seem questionable.

Might some crucial type of data still be missing? Controversy over the role of electrical coupling of neurons, and that of glial cells, over and above signal transmission via axon-synaptic couplings, continues to simmer. Might there be rules of synaptic connection that are not apparent, because the pattern cannot be ascertained within the billions of neurons involved?

To reveal essential patterns of activity, do different types of data have to be obtained using concurrent recording methods? All existing techniques offer a window on brain function limited in scale or in resolution in space or time. That is, only a comparatively few cells can be observed at once, the brain's electric and magnetic fields are relatively blurred in space, and the brain's blood flow, as observed by functional magnetic resonance imaging, is limited to relatively slow variations. None match the scale, speed, and detail relevant to cognition, and the task of making sufficient conjoint observations, in realistic waking contexts, is daunting to say the least.

What then is a reasonably observable explainable unit of the brain? Professor Eric Kandel advocates the complete analysis of a fly or worm brain, as an initial step in the mega-collaborations [2], but in what way, exactly, is a worm's brain more fundamental than, say, a sympathetic ganglion, or a fly's brain than a sensory-motor reflex?

If all the most important observable data is already available, or will become so, will sufficient computer power enable a working brain to be simulated? If this were achieved, would we be any the wiser, or simply unable to understand the functioning of the simulation, just as we cannot understand that upon which the simulation would be based? And would the simulation not, itself, be a person? Thus making our justification for subjecting it to manipulation and interference in the interests of science a little ethically questionable?

Obviously there is no way of knowing the answers to such questions without already having a sufficient unified theoretical understanding of brain function, within which old and new observations can be seen in context. Neural Field Theory hopes to discover such a unification, using as its guiding light explanation of the

large scale observable fields of brain activity, and expecting as this account proceeds an emergent insight into neural information processing. In contrast to its close relative, neural network theory, it seeks explanations beyond the interaction of smaller numbers of neurons, depending instead on the properties of small neural groups to define the properties of the continuum. The layout of this book reflects these intents.

After a brief tutorial, in the first half of the book and beginning from an historical perspective, differing approaches to formulating and analysing equations for neural fields are presented, and in their variety also revealing an underlying unity of conception. Stochastic dynamics are discussed, as well as means of introducing more anatomically and physiologically realistic properties to neural field equations.

The second half of the book begins by addressing the question of embodiment of universal computation within neural fields, and moves on to cognitive processes. Detailed models with cortical connectivity approaching that of the mammalian brain and the relationship to the large-scale electrical fields of the brain follow, and the book concludes with an attempt to show how fundamental field dynamics may play a part in the brain's embryonic development.

Thus a preliminary framework is discernible—methods now exist with the potential to unify material drawn from many branches of neuroscience, guiding their synthesis towards working models that can be tested against observable physical and cognitive properties of the working brain. The framework remains frail, and although the concepts involved seem largely internally consistent, in detail—for instance in the choice of parameters applied in different work—the work reported here is not entirely so. It is not yet possible to say the elusive “essence” referred to in the first paragraph has been captured. But the hopes held at the dawn of this subject appear to have been justified, and future prospects encouraging.

Stephen Coombes
Peter beim Graben
Roland Potthast
James Wright

References

1. Beim Graben, P., Wright, J. J.: From McCulloch-Pitts neurons toward biology. *Bull. Math. Biol.* **73**, 261–265 (2011)
2. Kandel, E.R., Markram, H., Mathews, P.M., Yuste, R., Koch, C.: Neuroscience thinks big (and collaboratively). *Nat. Rev. Neurosci.* **14**, 659–664 (2013)
3. nice.sandia.gov/documents/Christof_Koch_SandiaFEB'13.pdf
4. Sherrington, C.S.: *Man on His Nature*. Gifford Lecture, Edinburgh, 1937–1938. Cambridge University Press, Cambridge (1940)
5. www.columbia.edu/cu/biology/faculty/yuste/bam.html?
6. www.humanbrainproject.eu
7. www.nih.gov/science/brain/

Contents

1 Tutorial on Neural Field Theory	1
Stephen Coombes, Peter beim Graben, and Roland Potthast	
Part I Theory of Neural Fields	
2 A Personal Account of the Development of the Field Theory of Large-Scale Brain Activity from 1945 Onward	47
Jack Cowan	
3 Heaviside World: Excitation and Self-Organization of Neural Fields	97
Shun-ichi Amari	
4 Spatiotemporal Pattern Formation in Neural Fields with Linear Adaptation	119
G. Bard Ermentrout, Stefanos E. Folias, and Zachary P. Kilpatrick	
5 PDE Methods for Two-Dimensional Neural Fields	153
Carlo R. Laing	
6 Numerical Simulation Scheme of One- and Two-Dimensional Neural Fields Involving Space-Dependent Delays	175
Axel Hutt and Nicolas Rougier	
7 Spots: Breathing, Drifting and Scattering in a Neural Field Model	187
Stephen Coombes, Helmut Schmidt, and Daniele Avitabile	
8 Heterogeneous Connectivity in Neural Fields: A Stochastic Approach	213
Chris A. Brackley and Matthew S. Turner	

9	Stochastic Neural Field Theory	235
	Paul C. Bressloff	
10	On the Electrodynamics of Neural Networks.....	269
	Peter beim Graben and Serafim Rodrigues	

Part II Applications of Neural Fields

11	Universal Neural Field Computation.....	299
	Peter beim Graben and Roland Potthast	
12	A Neural Approach to Cognition Based on Dynamic Field Theory ...	319
	Jonas Lins and Gregor Schöner	
13	A Dynamic Neural Field Approach to Natural and Efficient Human-Robot Collaboration	341
	Wolfram Erlhagen and Estela Bicho	
14	Neural Field Modelling of the Electroencephalogram: Physiological Insights and Practical Applications	367
	David T. J. Liley	
15	Equilibrium and Nonequilibrium Phase Transitions in a Continuum Model of an Anesthetized Cortex.....	393
	D. Alistair Steyn-Ross, Moira L. Steyn-Ross, and Jamie W. Sleigh	
16	Large Scale Brain Networks of Neural Fields	417
	Viktor Jirsa	
17	Neural Fields, Masses and Bayesian Modelling	433
	Dimitris A. Pinotsis and Karl J. Friston	
18	Neural Field Dynamics and the Evolution of the Cerebral Cortex....	457
	James J. Wright and Paul D. Bourke	
Index		483

Chapter 1

Tutorial on Neural Field Theory

Stephen Coombes, Peter beim Graben, and Roland Potthast

Abstract The tools of dynamical systems theory are having an increasing impact on our understanding of patterns of neural activity. In this tutorial chapter we describe how to build tractable tissue level models that maintain a strong link with biophysical reality. These models typically take the form of nonlinear integro-differential equations. Their non-local nature has led to the development of a set of analytical and numerical tools for the study of spatiotemporal patterns, based around natural extensions of those used for local differential equation models. We present an overview of these techniques, covering Turing instability analysis, amplitude equations, and travelling waves. Finally we address inverse problems for neural fields to train synaptic weight kernels from prescribed field dynamics.

1.1 Background

Ever since Hans Berger made the first recording of the human electroencephalogram (EEG) in 1924 [8] there has been a tremendous interest in understanding the physiological basis of brain rhythms. This has included the development of mathematical models of cortical tissue – which are often referred to as neural field models. One of the earliest of such models is due to Beurle [9] in the 1950s, who developed a continuum description of the proportion of active neurons in a randomly connected network. This was followed by work of Griffith [40, 41]

S. Coombes (✉)

School of Mathematical Sciences, University of Nottingham, NG7 2RD Nottingham, UK
e-mail: stephen.coombes@nottingham.ac.uk

P.B. Graben

Bernstein Center for Computational Neuroscience, 10115 Berlin, Germany

R. Potthast

Department of Mathematics, University of Reading, RG6 6AX Reading, UK
Deutscher Wetterdienst, Offenbach, Germany

in the 1960s, who also published two books that still make interesting reading for modern practitioners of mathematical neuroscience [42, 43]. However, it were Wilson and Cowan [88, 89], Nunez [67] and Amari [3] in the 1970s who provided the formulations for neural field models that is in common use today (see Chaps. 2 and 3 in this book). Usually, neural field models are conceived as neural mass models describing population activity at spatiotemporally coarse-grained scales [67, 89]. They can be classified as either activity-based [89] or voltage-based [3, 67] models (see [14, 64] for discussion).

For their activity-based model Wilson and Cowan [88, 89] distinguished between excitatory and inhibitory sub-populations, as well as accounted for refractoriness. This seminal model can be written succinctly in terms of the pair of partial integro-differential equations:

$$\begin{aligned}\frac{\partial E}{\partial t} &= -E + (1 - r_E E) S_E [w_{EE} \otimes E - w_{EI} \otimes I], \\ \frac{\partial I}{\partial t} &= -I + (1 - r_I I) S_I [w_{IE} \otimes E - w_{II} \otimes I].\end{aligned}\quad (1.1)$$

Here $E = E(\mathbf{r}, t)$ is a temporal coarse-grained variable describing the proportion of excitatory cells firing per unit time at position \mathbf{r} at the instant t . Similarly the variable I represents the activity of an inhibitory population of cells. The symbol \otimes represents spatial convolution, the functions $w_{ab}(\mathbf{r})$ describe the weight of all synapses to the a th population from cells of the b th population a distance $|\mathbf{r}|$ away, and r_a is proportional to the refractory period of the a th population (in units of the population relaxation rate). The nonlinear function S_a describes the expected proportion of neurons in population a receiving at least threshold excitation per unit time, and is often taken to have a sigmoidal form. In many modern uses of the Wilson-Cowan equations the refractory terms are often dropped. For exponential or Gaussian choices of the connectivity function the Wilson-Cowan model is known to support a wide variety of solutions, including spatially and temporally periodic patterns (beyond a Turing instability), localised regions of activity (bumps and multi-bumps) and travelling waves (fronts, pulses, target waves and spirals), as reviewed in [19, 20, 32] and in Chaps. 4, 5, 7 or 8.

Further work on continuum models of neural activity was pursued by Nunez [67] and Amari [2, 3] under natural assumptions on the connectivity and firing rate function. Amari focused on local excitation and distal inhibition which is an effective model for a mixed population of interacting inhibitory and excitatory neurons with typical cortical connections (commonly referred to as Mexican hat connectivity), and formulated a single population (scalar) voltage-based model (without refractoriness) for activity $u = u(\mathbf{r}, t)$ of the form

$$\frac{\partial u}{\partial t} = -u + w \otimes f(u), \quad (1.2)$$

for some sigmoidal firing rate function f and connectivity function w . For the case that f is a Heaviside step function he showed how exact results for localised states (bumps and travelling pulses) could be obtained.

Since the original contributions of Wilson, Cowan, Nunez and Amari similar models have been used to investigate a variety of neural phenomena, including electroencephalogram (EEG) and magnetoencephalogram (MEG) rhythms [51, 52, 64, 68] (cf. Chaps. 10, 14, and 17), geometric visual hallucinations [16, 33, 84], mechanisms for short term memory [62, 63], feature selectivity in the visual cortex [7], motion perception [39], binocular rivalry [54], and anaesthesia [65] (cf. Chaps. 14 and 15). Neural field models have also found applications in autonomous robotic behaviour [30] (Chap. 13), embodied cognition [81] (Chap. 12), and Dynamic Causal Modelling [28] (Chap. 17), as well as being studied from an inverse problems perspective [5, 75]. As well as an increase in the applications of models like (1.1) and (1.2) in neuroscience, there has been a push to develop a deeper mathematical understanding of their behaviour. This has led to results in one spatial dimension about the existence and uniqueness of bumps [58] and waves [34] with smooth sigmoidal firing rates, as well as some constructive arguments that generalise the original ideas of Amari for a certain class of smoothed Heaviside firing rate functions [23, 69]. Other mathematical work has focused on geometric singular perturbation analysis as well as numerical bifurcation techniques to analyse solutions in one spatial dimension [62, 72, 73]. More explicit progress has been possible for the case of Heaviside firing rate functions, especially as regards the stability of solutions using Evans functions [22]. The extension of results from one to two spatial dimensions has increased greatly in recent years [24, 37, 56, 60, 61, 70, 85] (see Chap. 7). This style of work has also been able to tackle physiological extensions of minimal neural field models to account for axonal delays [21, 48, 50, 67] (included in the original Wilson-Cowan model and then dropped for simplicity), dendritic processing [15], and synaptic depression [55]. In contrast to the analysis of spontaneously generated patterns of activity, relatively little work has been done on neural fields with forcing. The exceptions perhaps being the work in [38] (for localised drive) and global period forcing in [78]. However, much of the above work exploits idealisations of the original models (1.1) and (1.2), especially as regards heterogeneity and noise, to make mathematical progress. More recent work that tackles heterogeneity (primarily using simulations) can be found in [11] (also in Chap. 8), whilst perturbation theory and homogenisation techniques are developed in [13, 24, 80], and functional analytic results in [36]. The treatment of stochastic neural field models is a very new area, and we refer the reader to the recent review by Bressloff [14] and to Chaps. 2 and 9, which also covers methods from non-equilibrium statistical physics that attempt to move beyond the mean-field rate equations of the type exemplified by (1.1) and (1.2). However, it is fair to say that the majority of neural field models in use today can trace their roots back to the seminal work of Wilson and Cowan, Nunez and Amari.

In this chapter we will develop the discussion of a particular neural field model that incorporates much of the spirit of (1.1) and (1.2), though with refinements that make a stronger connection to models of both synaptic and dendritic processing.

We will then show how to analyse these models with techniques from dynamical systems before going on to discuss inverse problems in neural field theory.

1.1.1 Synaptic Processing

At a synapse, presynaptic firing results in the release of neurotransmitters that causes a change in the membrane conductance of the postsynaptic neuron. This postsynaptic current may be written

$$I_s = g(V - V_s), \quad (1.3)$$

where V is the voltage of the postsynaptic membrane, V_s is its reversal potential and g is a conductance. This is proportional to the probability that a synaptic receptor channel is in an open conducting state. This probability depends on the presence and concentration of neurotransmitter released by the presynaptic neuron. The sign of V_s relative to the resting potential (assumed to be zero) determines whether the synapse is excitatory ($V_s > 0$) or inhibitory ($V_s < 0$).

The effect of some synapses can be described with a function that fits the shape of the postsynaptic response due to the arrival of action potential at the presynaptic release site. A postsynaptic conductance change $g(t)$ would then be given by

$$g(t) = \bar{g}\eta(t - T), \quad t \geq T, \quad (1.4)$$

where T is the arrival time of a presynaptic action potential and $\eta(t)$ fits the shape of a realistic postsynaptic conductance. A common (normalised) choice for $\eta(t)$ is a difference of exponentials:

$$\eta(t) = \left(\frac{1}{\alpha} - \frac{1}{\beta} \right)^{-1} [e^{-\alpha t} - e^{-\beta t}] H(t), \quad (1.5)$$

or the α -function:

$$\eta(t) = \alpha^2 t e^{-\alpha t} H(t), \quad (1.6)$$

where H is a Heaviside step function. The conductance change arising from a train of action potentials, with firing times T_m , is given by

$$g(t) = \bar{g} \sum_m \eta(t - T_m). \quad (1.7)$$

We note that both the forms for $\eta(t)$ above can be written as the Green's function of a linear differential operator, so that $Q\eta = \delta$, where

$$Q = \left(1 + \frac{1}{\alpha} \frac{d}{dt}\right) \left(1 + \frac{1}{\beta} \frac{d}{dt}\right), \quad (1.8)$$

for (1.5) and one simply sets $\beta = \alpha$ to obtain the response describing an α -function.

1.1.2 Dendritic Processing

Dendrites form the major components of neurons. They are complex branching structures that receive and process thousands of synaptic inputs from other neurons. It is well known that dendritic morphology plays an important role in the function of dendrites. A nerve fibre consists of a long thin, electrically conducting core surrounded by a thin membrane whose resistance to transmembrane current flow is much greater than that of either the internal core or the surrounding medium. Injected current can travel long distances along the dendritic core before a significant fraction leaks out across the highly resistive cell membrane. Conservation of electric current in an infinitesimal cylindrical element of nerve fibre yields a second-order linear partial differential equation (PDE) known as the *cable equation*. Let $V(x, t)$ denote the membrane potential at position x along a uniform cable at time t measured relative to the resting potential of the membrane. Let τ be the cell membrane time constant, λ the space constant and r the membrane resistance, then the basic uniform (infinite) cable equation is

$$\tau \frac{\partial V(x, t)}{\partial t} = -V(x, t) + \lambda^2 \frac{\partial^2 V(x, t)}{\partial x^2} + rI(x, t), \quad x \in (-\infty, \infty), \quad (1.9)$$

where we include the source term $I(x, t)$ corresponding to external input injected into the cable. Diffusion along the dendritic tree generates an effective spatiotemporal distribution of delays as expressed by the associated Green's function of the cable equation in terms of the diffusion constant $D = \lambda^2/\tau$. In response to a unit impulse at x' at $t = 0$ and taking $V(x, 0) = 0$ the dendritic potential behaves as $V(x, t) = G_\infty(x - x', t)$, where

$$G_\infty(x, t) = \frac{1}{\sqrt{4\pi Dt}} e^{-t/\tau} e^{-x^2/(4Dt)} H(t). \quad (1.10)$$

The Green's function $G_\infty(x, t)$ (derived in Appendix 1) determines the linear response to an instantaneous injection of unit current at a given point on the tree. Using the linearity of the cable equation one may write the general solution as

$$\begin{aligned} V(x, t) &= \int_{-\infty}^t dt' \int_{-\infty}^{\infty} dx' G_\infty(x - x', t - t') I(x', t') \\ &\quad + \int_{-\infty}^{\infty} dx' G_\infty(x - x', t) V(x', 0). \end{aligned} \quad (1.11)$$

Note that for notational simplicity we have absorbed a factor of r/τ within the definition of the source term $I(x, t)$. For example, assuming the soma is at $x = 0$, $V(x, 0) = 0$ and the synaptic input is a train of spikes at $x = x'$, $I(x, t) = \delta(x - x') \sum_m \delta(t - T_m)$ we have that

$$V(0, t) = \sum_m G_\infty(x', t - T_m). \quad (1.12)$$

1.2 Tissue Level Firing Rate Models with Axo-Dendritic Connections

At heart modern biophysical theories assert that EEG signals from a single scalp electrode arise from the coordinated activity of $\sim 10^6$ pyramidal cells in cortex [27]. These are arranged with their dendrites in parallel and perpendicular to the cortical surface. When synchronously activated by synapses at the proximal dendrites extracellular current flows (parallel to the dendrites), with a net membrane current at the synapse. For excitatory (inhibitory) synapses this creates a sink (source) with a negative (positive) extracellular potential. Because there is no accumulation of charge in the tissue the proximal synaptic current is compensated by other currents flowing in the medium causing a distributed source in the case of a sink and vice-versa for a synapse that acts as a source. Hence, at the population level the potential field generated by a synchronously activated population of cortical pyramidal cells behaves like that of a dipole layer. Although the important contribution that single dendritic trees make to generating extracellular electric field potentials has been realised for some time, and can be calculated using Maxwell equations [71], they are typically not accounted for in neural field models. The exception to this being the work of Bressloff, reviewed in [15] and in Chap. 10.

In many neural population models it is assumed that the interactions are mediated by firing rates rather than action potentials (spikes) per se. To see how this might arise we rewrite (1.7) in the equivalent form

$$Qg = \bar{g} \sum_m \delta(t - T_m). \quad (1.13)$$

If we perform a short-time average of (1.13) over some time-scale Δ and assume that η is sufficiently slow so that $\langle Qg \rangle_t$ is approximately constant, where

$$\langle x \rangle_t = \frac{1}{\Delta} \int_{t-\Delta}^t x(s) ds, \quad (1.14)$$

then we have that $Qg = f$, where f is the instantaneous firing rate (number of spikes per time Δ). For a single neuron (real or synthetic) experiencing a constant drive it is natural to assume that this firing rate is a function of the drive alone. If for

the moment we assume that a neuron spends most of its time close to rest such that $V_s - V \approx V_s$, and absorb a factor V_s into g , then for synaptically interacting neurons this drive is directly proportional to the conductance state of the presynaptic neuron. Thus for a single population with self-feedback we are led naturally to equations like:

$$Qg = w_0 f(g), \quad (1.15)$$

for some strength of coupling w_0 . A common choice for the *population* firing rate function is the sigmoid

$$f(g) = \frac{1}{1 + \exp(-\beta(g - h))}, \quad (1.16)$$

which saturates to one for large g . This functional form, with threshold h and steepness parameter β , is approximately obtained for a unimodal distribution of firing thresholds among the population [88]. Note that the notion of a slow response would also be expected in a large globally coupled network which was firing asynchronously (so that mean field signals would be nearly constant).

To obtain a tissue level model in one spatial dimension we simply consider $g = g(x, t)$, with $x \in \mathbb{R}$, and introduce a coupling function and integrate over the domain to obtain

$$Qg = \int_{-\infty}^{\infty} w(x, y) f(g(y, t - D(x, y)/v)) dy, \quad (1.17)$$

or equivalently

$$g(x, t) = \int_{-\infty}^t ds \eta(t - s) \int_{-\infty}^{\infty} w(x, y) f(g(y, s - D(x, y)/v)) dy. \quad (1.18)$$

Here we have allowed for a communication delay, that arises because of the finite speed, v , of the action potential, where $D(x, y)$ measures the length of the axonal fibre between points at x and y . The coupling function $w(x, y)$ represents anatomical connectivity, and is often assumed to be homogeneous so that $w(x, y) = w(|x - y|)$. It is also common to assume that $D(x, y) = |x - y|$.

Following the original work of Bressloff (reviewed in [15]) we now develop the cable modelling approach of Rall [82] to describe a firing rate cortical tissue model with axo-dendritic patterns of synaptic connectivity. For simplicity we shall consider only an effective single population model in one (somatic) spatial dimension to include a further dimension representing position along a (semi-infinite) dendritic cable. The firing rate in the somatic (cell body) layer is taken to be a smooth function of the cable voltage at the soma, which is in turn determined by the spatiotemporal pattern of synaptic currents on the cable. For an illustration see Fig. 1.1.

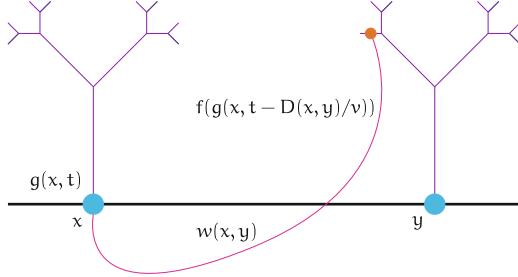


Fig. 1.1 Diagram of a one dimensional neural field model. In this illustration the dendritic tree is drawn with a branched structure. For the sake of simplicity the neural field model is only developed here for unbranched dendrites. However, this can be naturally generalised using the “sum-over-trips” approach of Abbott et al. for passive dendrites [1] and Coombes et al. [25] for resonant dendrites

The voltage $V(\xi, x, t)$ at position $\xi \geq 0$ along a semi-infinite passive cable with somatic coordinate $x \in \mathbb{R}$ can then be written:

$$\frac{\partial V}{\partial t} = -\frac{V}{\tau} + D \frac{\partial^2 V}{\partial \xi^2} + I(\xi, x, t). \quad (1.19)$$

Here, $I(\xi, x, t)$ is the synaptic input (and remember that we absorb within this a factor r/τ), and we shall drop shunting effects and take this to be directly proportional to a conductance change, which evolves according to the usual neural field prescription (cf. Eq. (1.18)) as

$$g(\xi, x, t) = \int_{-\infty}^t ds \eta(t-s) \int_{-\infty}^{\infty} dy W(\xi, x, y) f(h(y, s - D(x, y)/v)). \quad (1.20)$$

The function $W(\xi, x, y)$ describes the axo-dendritic connectivity pattern and the field h is taken as a measure of the drive at the soma. As a simple model of h we shall take it to be the somatic potential and write $h(x, t) = V(0, x, t)$. For no flux boundary conditions $\partial V(\xi, x, t)/\partial \xi|_{\xi=0} = 0$, and assuming vanishing initial data, the solution to (1.19) at $\xi = 0$ becomes

$$V(\xi = 0, x, t) = \kappa(G \otimes g)(\xi = 0, x, t), \quad G = 2G_{\infty} \quad (1.21)$$

for some constant of proportionality $\kappa > 0$, where $G_{\infty}(x, t)$ is given by (1.10) and here the operator \otimes denotes spatiotemporal convolution over the (ξ, t) coordinates. Note that in obtaining (1.21) we have used the result that the Green’s function (between two points ξ and ξ') for the semi-infinite cable with no flux boundary conditions can be written as $G_{\infty}(\xi - \xi', t) + G_{\infty}(\xi + \xi', t)$ [1, 86].

Further assuming that the axo-dendritic weights can be decomposed in the product form $W(\xi, x, y) = P(\xi)w(|x - y|)$ then the equation for h takes the form

$$h(x, t) = \kappa \int_{-\infty}^t ds F(t-s) \int_{-\infty}^s ds' \eta(s-s') \int_{-\infty}^{\infty} dy w(|x-y|) f(h(y, s' - D(x, y)/v)), \quad (1.22)$$

where

$$F(t) = \int_0^{\infty} d\xi P(\xi) G(\xi, t). \quad (1.23)$$

We regard Eq. (1.22) as a natural extension of the Amari model (1.2) to include synaptic and dendritic processing as well as axonal delays. Note that the Amari model is recovered from (1.22) in the limit $v \rightarrow \infty$, $\eta(t) = e^{-t} H(t)$, and $F(t) = \delta(t)/\kappa$.

1.2.1 Turing Instability Analysis

To assess the pattern forming properties of the model given by (1.22) it is useful to perform a Turing instability analysis. This describes how a spatially homogeneous state can become unstable to spatially heterogeneous perturbations, resulting in the formation of periodic patterns. To illustrate the technique consider the one-dimensional model without dendrites or axonal delays, obtained in the limit $v \rightarrow \infty$ and $F(t) \rightarrow \delta(t)$:

$$h(x, t) = \kappa \int_0^{\infty} ds \eta(s) \int_{-\infty}^{\infty} dy w(|y|) f(h(x-y, t-s)). \quad (1.24)$$

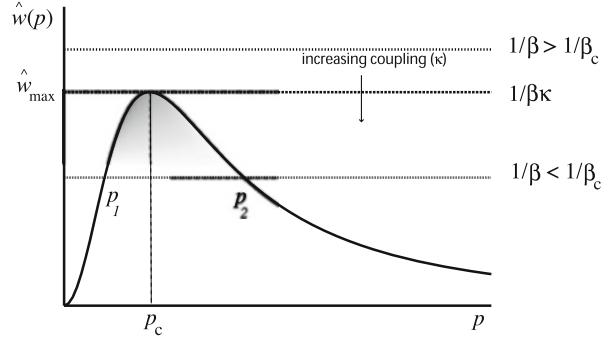
One solution of the neural field equation is the spatially uniform resting state $h(x, t) = h_0$ for all x, t , defined by

$$h_0 = \kappa f(h_0) \int_{-\infty}^{\infty} w(|y|) dy. \quad (1.25)$$

Here we have used the fact that η is normalised, namely that $\int_0^{\infty} ds \eta(s) = 1$. We linearise about this state by letting $h(x, t) \rightarrow h_0 + h(x, t)$ so that $f(h) \rightarrow f(h_0) + f'(h_0)u$ to obtain

$$h(x, t) = \kappa \beta \int_0^{\infty} ds \eta(s) \int_{-\infty}^{\infty} dy w(y) h(x-y, t-s), \quad \beta = f'(h_0). \quad (1.26)$$

Fig. 1.2 A plot of the Fourier transform of the weight kernel $\hat{w}(p)$ illustrating how its shape (with a maximum away from the origin) can determine a Turing instability defined by the condition $\kappa\beta\hat{w}(p_c) = 1$



This has solutions of the form $e^{\lambda t} e^{ipx}$, with a dispersion curve:

$$1 = \kappa\beta\tilde{\eta}(\lambda)\hat{w}(p), \quad \hat{w}(p) = \int_{-\infty}^{\infty} dy w(|y|)e^{-ipy}, \quad \tilde{\eta}(\lambda) = \int_0^{\infty} ds \eta(s)e^{-\lambda s}. \quad (1.27)$$

We recognise \hat{w} as the Fourier transform of w and $\tilde{\eta}$ as the Laplace transform of η . The uniform steady state is linearly stable if $\operatorname{Re}\lambda(p) < 0$ for all $p \in \mathbb{R}$, $p \neq 0$. For the choice $\eta(t) = \alpha e^{-\alpha t} H(t)$ (so that $\mathcal{Q} = (1 + \alpha^{-1}d/dt)$) then $\tilde{\eta}(\lambda) = (1 + \lambda/\alpha)^{-1}$. In this case, since $w(x) = w(-x)$ then $\hat{w}(p)$ is a real even function of p and the stability condition is simply

$$\hat{w}(p) < \frac{1}{\beta\kappa}, \quad \text{for all } p \in \mathbb{R}, \quad p \neq 0. \quad (1.28)$$

Now consider the case that $\hat{w}(p)$ has a positive maximum \hat{w}_{\max} at $p = \pm p_c$, that is $\hat{w}(p_c) = \hat{w}_{\max}$ and $\hat{w}(p) < \hat{w}_{\max}$ for all $p \neq p_c$. For $\beta < \beta_c$, where $\beta_c = 1/(\kappa\hat{w}_{\max})$, we have $\kappa\hat{w}(p) \leq \kappa\hat{w}_{\max} < 1/\beta$ for all p and the resting state is linearly stable. At the critical point $\beta = \beta_c$ (see Fig. 1.2) we have $\beta_c\kappa\hat{w}(p_c) = 1$ and $\beta_c\kappa\hat{w}(p) < 1$ for all $p \neq p_c$. Hence, $\lambda(p) < 0$ for all $p \neq p_c$, but $\lambda(p_c) = 0$. This signals the point of a *static* instability due to excitation of the pattern $e^{\pm ip_c x}$. Beyond the bifurcation point, $\beta > \beta_c$, $\lambda(p_c) > 0$ and this pattern grows with time. In fact there will typically exist a range of values of $p \in (p_1, p_2)$ for which $\lambda(p) > 0$, signalling a set of growing patterns. As the patterns grow, the linear approximation breaks down and nonlinear terms dominate the behaviour. The saturating property of f tends to create patterns with finite amplitude, that scale as $\sqrt{\beta - \beta_c}$ close to bifurcation and have wavelength $2\pi/p_c$. If $p_c = 0$ then we would have a *bulk instability* resulting in the formation of another homogeneous state.

A common choice for $w(x)$ is a Mexican hat function which represents short-range excitation and long-range inhibition. An example of such a function is a difference of two exponentials:

$$w(x) = \Lambda \left[e^{-\gamma_1|x|} - \Gamma e^{-\gamma_2|x|} \right], \quad (1.29)$$

with $\Gamma < 1$, $\gamma_1 > \gamma_2 > 0$ and $\Lambda = +1$. (The case $\Lambda = -1$, which represents short-range inhibition and long-range excitation will be considered below in the full model.) The Fourier transform $\hat{w}(p)$ is calculated as:

$$\hat{w}(p) = 2\Lambda \left[\frac{\gamma_1}{\gamma_1^2 + p^2} - \Gamma \frac{\gamma_2}{\gamma_2^2 + p^2} \right], \quad (1.30)$$

from which we may determine p_c as

$$p_c^2 = \frac{\gamma_1^2 \sqrt{\Gamma \gamma_2 / \gamma_1} - \gamma_2^2}{1 - \sqrt{\Gamma \gamma_2 / \gamma_1}}. \quad (1.31)$$

Hence, $p_c \neq 0$ when $\Gamma > (\gamma_2 / \gamma_1)^3$. Note that for $\Lambda = -1$ then $p_c = 0$ and a static Turing instability does not occur.

For the full model (1.22) with $D(x, y) = |x - y|$ the homogeneous steady state, $h(x, t) = h_0$ for all x, t , satisfies

$$h_0 = \kappa f(h_0) \int_0^\infty F(s) ds \int_{-\infty}^\infty dy w(|y|), \quad (1.32)$$

and the spectral equation takes the form

$$1 = \kappa \beta \hat{w}(p, \lambda) \tilde{\eta}(\lambda) \tilde{F}(\lambda), \quad \hat{w}(p, \lambda) = \int_{-\infty}^\infty dy w(|y|) e^{-ipy} e^{-\lambda|y|/v}, \quad \beta = f'(h_0). \quad (1.33)$$

Compared to (1.27) it is now possible for complex solutions for λ to be supported – allowing for the possibility of *dynamic* (as opposed to static) Turing instabilities to occur. These occur when $\text{Im } \lambda \neq 0$ at the bifurcation point.

For example, in the limit $v \rightarrow \infty$ then $\hat{w}(p, \lambda) \rightarrow \hat{w}(p)$ and for $\eta(t) = \alpha e^{-\alpha t} H(t)$ we have that

$$1 + \lambda/\alpha = \kappa \beta \hat{w}(p) \tilde{F}(\lambda). \quad (1.34)$$

A necessary condition for a dynamic instability ($\text{Re } \lambda = 0$ and $\text{Im } \lambda \neq 0$) is that there exists a pair $\omega, p \neq 0$ such that

$$1 + i\omega/\alpha = \kappa \beta \hat{w}(p) \tilde{F}(i\omega). \quad (1.35)$$

Equating real and imaginary parts (and using the fact that $\hat{w}(p) \in \mathbb{R}$) gives us the pair of simultaneous equations

$$1 = \kappa \beta \hat{w}(p) C(\omega), \quad \omega/\alpha = \kappa \beta \hat{w}(p) S(\omega), \quad (1.36)$$

where $C(\omega) = \operatorname{Re} \tilde{F}(i\omega)$ and $S(\omega) = \operatorname{Im} \tilde{F}(i\omega)$. Note that $C(\omega) = \int_0^\infty ds F(s) \cos(\omega s) \leq |C(0)|$. Hence (dividing the above equations) if there is a non-zero solution to

$$\frac{\omega_c}{\alpha} = \mathcal{H}(\omega_c), \quad \mathcal{H}(\omega_c) \equiv \frac{S(\omega_c)}{C(\omega_c)}, \quad (1.37)$$

then the bifurcation condition, $\beta = \beta_d$, for a dynamic instability is defined by

$$\beta_d \kappa \hat{w}(p_{\min}) = \frac{1}{C(\omega_c)}, \quad (1.38)$$

which should be contrasted with the bifurcation condition, $\beta = \beta_s$, for a static instability, namely

$$\beta_s \kappa \hat{w}(p_{\max}) = \frac{1}{C(0)}, \quad (1.39)$$

where

$$\hat{w}(p_{\min}) = \min_p \hat{w}(p), \quad \hat{w}(p_{\max}) = \max_p \hat{w}(p). \quad (1.40)$$

Assuming that $\hat{w}(p_{\min}) < 0 < \hat{w}(p_{\max})$, a dynamic Turing instability will occur if $\beta < \beta_s$ and $p_{\min} \neq 0$, whereas a static Turing instability will occur if $\beta_s < \beta$ and $p_{\max} \neq 0$.

For the Mexican hat function (1.29) with $\Lambda = +1$ (short-range excitation, long-range inhibition), a dynamic Turing instability is not possible since $p_{\min} = 0$. However, it is possible for bulk oscillations to occur instead of static patterns when

$$\hat{w}(p_c) < -\frac{C(\omega_c)}{C(0)} |\hat{w}(0)|, \quad (1.41)$$

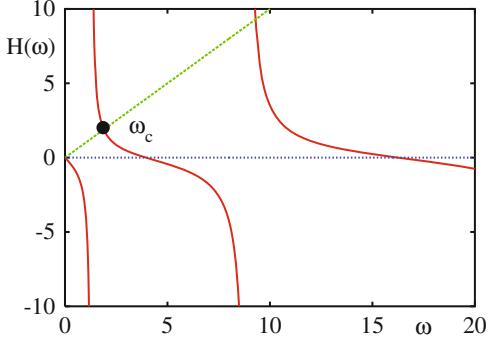
with p_c given by (1.31). On the other hand, when $\Lambda = -1$ (short-range inhibition, long-range excitation) a dynamic instability can occur since $p_{\min} = p_c$ and $p_{\max} = 0$, provided that

$$\hat{w}(0) < -\frac{C(\omega_c)}{C(0)} |\hat{w}(p_c)|. \quad (1.42)$$

As an explicit example consider the choice $P(\xi) = \delta(\xi - \xi_0)$ (so that the synaptic contact occurs at a fixed distance $\xi_0 > 0$ from the soma). In this case $F(t) = G(\xi_0, t)$ with Laplace transform (calculated in Appendix 2):

$$\tilde{F}(\lambda) = \frac{e^{-\gamma(\lambda)\xi_0}}{D\gamma(\lambda)}, \quad \gamma^2(\lambda) = (1/\tau + \lambda)/D. \quad (1.43)$$

Fig. 1.3 A plot of the function $H(\omega)$ for $D = \tau = 1$ with $\xi_0 = 2$, showing a non-zero solution of (1.37) for $\alpha = 1$. This highlights the possibility of a dynamic Turing instability ($\omega_c \neq 0$) in a dendritic neural field model with short-range inhibition and long-range excitation



In this case we may calculate the real and imaginary parts of $\tilde{F}(i\omega)$ as

$$C(\omega) = \frac{1}{\sqrt{1/\tau^2 + \omega^2}} e^{-A_+(\omega)\xi_0} [A_+(\omega) \cos(A_-(\omega)\xi_0) - A_-(\omega) \sin(A_-(\omega)\xi_0)] \quad (1.44)$$

$$S(\omega) = -\frac{1}{\sqrt{1/\tau^2 + \omega^2}} e^{-A_+(\omega)\xi_0} [A_+(\omega) \sin(A_-(\omega)\xi_0) + A_+(\omega) \cos(A_-(\omega)\xi_0)], \quad (1.45)$$

where $\sqrt{(1/\tau + i\omega)/D} = A_+(\omega) + iA_-(\omega)$ and

$$A_{\pm}(\omega) = \sqrt{[\sqrt{1/(\tau D)^2 + \omega^2/D^2} \pm 1/(\tau D)]/2}. \quad (1.46)$$

A plot of $H(\omega)$ is shown in Fig. 1.3, highlighting the possibility of a non-zero solution of (1.37) for a certain parameter set (and hence the possibility of a dynamic instability).

For a discussion of dynamic Turing instabilities with finite v we refer the reader to [87]. For the treatment of more general forms of axo-dendritic connectivity (that do not assume a product form) we refer the reader to [12, 17].

The extension of the above argument to two dimensions shows that the linearised equations of motion have solutions of the form $e^{\lambda t} e^{i\mathbf{p}\cdot\mathbf{r}}$, $\mathbf{r}, \mathbf{p} \in \mathbb{R}^2$, with $\lambda = \lambda(p)$, $p = |\mathbf{p}|$ as determined by (1.33) with

$$\hat{w}(p, \lambda) = \int_{\mathbb{R}^2} d\mathbf{r} w(|\mathbf{r}|) e^{-i\mathbf{p}\cdot\mathbf{r}} e^{-\lambda|\mathbf{r}|/v}. \quad (1.47)$$

Near bifurcation we expect spatially periodic solutions of the form $\exp i[p_1 x + p_2 y]$, $p_c^2 = p_1^2 + p_2^2$. For a given p_c there are an infinite number of choices for p_1 and p_2 . It is therefore convenient to restrict attention to doubly periodic solutions that tessellate the plane. These can be expressed in terms of the basic symmetry groups of

hexagon, square and rhombus. Solutions can then be constructed from combinations of the basic functions $e^{ip_c \mathbf{R} \cdot \mathbf{r}}$, for appropriate choices of the basis vectors \mathbf{R} . If ϕ is the angle between two basis vectors \mathbf{R}_1 and \mathbf{R}_2 , we can distinguish three types of lattice according to the value of ϕ : square lattice ($\phi = \pi/2$), rhombic lattice ($0 < \phi < \pi/2$) and hexagonal ($\phi = \pi/3$). Hence, all doubly periodic functions may be written as a linear combination of plane waves

$$h(\mathbf{r}) = \sum_j A_j e^{ip_c \mathbf{R}_j \cdot \mathbf{r}} + \text{cc}, \quad |\mathbf{R}_j| = 1, \quad (1.48)$$

where cc stands for complex-conjugate. For hexagonal lattices we use $\mathbf{R}_1 = (1, 0)$, $\mathbf{R}_2 = (-1, \sqrt{3})/2$, and $\mathbf{R}_3 = (1, \sqrt{3})/2$. For square lattices we use $\mathbf{R}_1 = (1, 0)$, $\mathbf{R}_2 = (0, 1)$, while the rhombus tessellation uses $\mathbf{R}_1 = (1, 0)$, $\mathbf{R}_2 = (\cos \zeta, \sin \zeta)$.

1.2.2 Weakly Nonlinear Analysis: Amplitude Equations

A characteristic feature of the dynamics of systems beyond an instability is the slow growth of the dominant eigenmode, giving rise to the notion of a *separation of scales*. This observation is key in deriving the so-called *amplitude equations*. In this approach information about the short-term behaviour of the system is discarded in favour of a description on some appropriately identified slow time-scale. By Taylor-expansion of the dispersion curve near its maximum one expects the scalings $\text{Re } \lambda \sim \beta - \beta_c$, $p - p_c \sim \sqrt{\beta - \beta_c}$, close to bifurcation, where β is the bifurcation parameter. Since the eigenvectors at the point of instability are of the type $A_1 e^{i(\omega_c t + p_c x)} + A_2 e^{i(\omega_c t - p_c x)} + \text{cc}$, for $\beta > \beta_c$ emergent patterns are described by an infinite sum of unstable modes (in a continuous band) of the form $e^{v_0(\beta-\beta_c)t} e^{i(\omega_c t + p_c x)} e^{ip_0 \sqrt{\beta-\beta_c}x}$. Let us denote $\beta = \beta_c + \epsilon^2 \delta$ where ϵ is arbitrary and δ is a measure of the distance from the bifurcation point. Then, for small ϵ we can separate the dynamics into fast eigen-oscillations $e^{i(\omega_c t + p_c x)}$, and slow modulations of the form $e^{v_0 \epsilon^2 t} e^{ip_0 \epsilon x}$. If we set as further independent variables $\tau = \epsilon^2 t$ for the modulation time-scale and $\chi = \epsilon x$ for the long-wavelength spatial scale (at which the interactions between excited nearby modes become important) we may write the weakly nonlinear solution as $A_1(\chi, \tau) e^{i(\omega_c t + p_c x)} + A_2(\chi, \tau) e^{i(\omega_c t - p_c x)} + \text{cc}$. It is known from the standard theory [47] that weakly nonlinear solutions will exist in the form of either travelling waves ($A_1 = 0$ or $A_2 = 0$) or standing waves ($A_1 = A_2$).

We are now in a position to derive the amplitude equations for patterns emerging beyond the point of an instability for a neural field model. These are also useful for determining the sub- or super-critical nature of the bifurcation. For clarity we shall first focus on the case of a static instability, and consider the example system given by (1.18) with $\eta(t) = e^{-t} H(t)$ and $v \rightarrow \infty$, equivalent to the Amari model (1.2). In this case the model is conveniently written as an integro-differential equation:

$$\frac{\partial g}{\partial t} = -g + w \otimes f(g), \quad (1.49)$$

where the symbol \otimes denotes spatial convolution (assuming $w(x, y) = w(|x - y|)$).

We first Taylor expand the nonlinear firing rate around the steady state g_0 :

$$f(g) = f(g_0) + \beta_1(g - g_0) + \beta_2(g - g_0)^2 + \beta_3(g - g_0)^3 + \dots, \quad (1.50)$$

where $\beta_1 = f'(g_0)$, $\beta_2 = f''(g_0)/2$ and $\beta_3 = f'''(g_0)/6$. We also adopt the perturbation expansion

$$g = g_0 + \epsilon g_1 + \epsilon^2 g_2 + \epsilon^3 g_3 + \dots. \quad (1.51)$$

After rescaling time according to $\tau = \epsilon^2 t$ and setting $\beta_1 = \beta_c + \epsilon^2 \delta$, where β_c is defined by the bifurcation condition $\beta_c = 1/\hat{w}(p_c)$, we then substitute into (1.49). Equating powers of ϵ leads to a hierarchy of equations:

$$g_0 = f(g_0) \int_{-\infty}^{\infty} w(|y|) dy, \quad (1.52)$$

$$0 = \mathcal{L}g_1, \quad (1.53)$$

$$0 = \mathcal{L}g_2 + \beta_2 w \otimes g_1^2, \quad (1.54)$$

$$\frac{dg_1}{d\tau} = \mathcal{L}g_3 + \delta w \otimes g_1 + 2\beta_2 w \otimes g_1 g_2 + \beta_3 w \otimes g_1^3, \quad (1.55)$$

where

$$\mathcal{L}g = -g + \beta_c w \otimes g. \quad (1.56)$$

The first equation fixes the steady state g_0 . The second equation is linear with solutions $g_1 = A(\tau)e^{ip_c x} + cc$ (where p_c is the critical wavenumber at the static bifurcation). Hence the null space of \mathcal{L} is spanned by $e^{\pm ip_c x}$. A dynamical equation for the complex amplitude $A(\tau)$ (and we do not treat here any slow spatial variation) can be obtained by deriving solvability conditions for the higher-order equations, a method known as the Fredholm alternative. These equations have the general form $\mathcal{L}g_n = v_n(g_0, g_1, \dots, g_{n-1})$ (with $\mathcal{L}g_1 = 0$). We define the inner product of two periodic functions (with periodicity $2\pi/p_c$) as

$$\langle U, V \rangle = \frac{p_c}{2\pi} \int_0^{2\pi/p_c} U^*(x) V(x) dx, \quad (1.57)$$

where $*$ denotes complex conjugation. It is simple to show that \mathcal{L} is self-adjoint with respect to this inner product (see Appendix 3), so that

$$\langle g_1, \mathcal{L}g_n \rangle = \langle \mathcal{L}g_1, g_n \rangle = 0. \quad (1.58)$$

Hence we obtain the set of solvability conditions

$$\langle e^{\pm ip_c x}, v_n \rangle = 0, \quad n \geq 2. \quad (1.59)$$

The solvability condition with $n = 2$ is automatically satisfied, since $w \otimes g_1^2 = \hat{w}(2p_c)[A^2 e^{2ip_c x} + cc] + 2|A|^2 \hat{w}(0)$, and we make use of the result $\langle e^{imp_c x}, e^{inp_c x} \rangle = \delta_{m,n}$. For $n = 3$ the solvability condition (projecting onto $e^{+ip_c x}$) is

$$\langle e^{ip_c x}, \frac{dg_1}{dt} - \delta w \otimes g_1 \rangle = \beta_3 \langle e^{ip_c x}, w \otimes g_1^3 \rangle + 2\beta_2 \langle e^{ip_c x}, w \otimes g_1 g_2 \rangle. \quad (1.60)$$

The left-hand side is easily calculated, using $w \otimes g_1 = \hat{w}(p_c)[A e^{ip_c x} + cc]$, as

$$\frac{dA}{d\tau} - \delta \hat{w}(p_c) A = \frac{dA}{d\tau} - \beta_c^{-1} \delta A, \quad (1.61)$$

where we have made use of the bifurcation condition $\beta_c = 1/\hat{w}(p_c)$. To evaluate the right-hand side we use the result that $w \otimes g_1^3 = \hat{w}(p_c)[A^3 e^{3ip_c x} + cc] + 3|A|^2 \hat{w}(p_c)[A e^{ip_c x} + cc]$, to obtain

$$\langle e^{ip_c x}, w \otimes g_1^3 \rangle = 3\beta_c^{-1} A |A|^2. \quad (1.62)$$

The next step is to determine g_2 . From (1.54) we have that

$$-g_2 + \beta_c w \otimes g_2 = -\beta_2 \{ \hat{w}(2p_c)[A^2 e^{2ip_c x} + cc] + 2|A|^2 \hat{w}(0) \}. \quad (1.63)$$

We now set

$$g_2 = A_+ e^{2ip_c x} + A_- e^{-2ip_c x} + A_0 + \phi g_1. \quad (1.64)$$

The constant ϕ remains undetermined at this order of perturbation but does not appear in the amplitude equation for $A(\tau)$. Substitution into (1.63) and equating powers of $e^{ip_c x}$ gives

$$A_0 = \frac{2\beta_2 |A|^2 \hat{w}(0)}{1 - \beta_c \hat{w}(0)}, \quad A_+ = \frac{\beta_2 A^2 \hat{w}(2p_c)}{1 - \beta_c \hat{w}(2p_c)}, \quad A_- = A_+^*, \quad (1.65)$$

where we have used the result that $w \otimes g_2 = \hat{w}(2p_c)[A_+ e^{2ip_c x} + A_- e^{-2ip_c x}] + \hat{w}(0)A_0 + \phi[A \hat{w}(p_c)e^{ip_c x} + cc]$. We then find that

$$\langle e^{ip_c x}, w \otimes g_1 g_2 \rangle = \hat{w}(p_c)[A_+ A^* + A_0 A]. \quad (1.66)$$

Combining (1.61), (1.62) and (1.66) we obtain the Stuart-Landau equation

$$\beta_c \frac{dA}{d\tau} = A(\delta - \Phi |A|^2), \quad (1.67)$$

where

$$\Phi = -3\beta_3 - 2\beta_2^2 \left[\frac{\hat{w}(2p_c)}{1 - \beta_c \hat{w}(2p_c)} + \frac{2\hat{w}(0)}{1 - \beta_c \hat{w}(0)} \right]. \quad (1.68)$$

Introducing $A = R e^{i\theta}$ we may rewrite Eq. (1.67) as

$$\beta_c \frac{dR}{d\tau} = \delta R - \Phi R^3, \quad \frac{d\theta}{d\tau} = 0. \quad (1.69)$$

Hence, the phase of A is arbitrary ($\theta = \text{const}$) and the amplitude has a pitchfork bifurcation which is super-critical for $\Phi > 0$ and sub-critical for $\Phi < 0$.

Amplitude equations arising for systems with a dynamic instability are treated in [87]. The appropriate amplitude equations are found to be the coupled mean-field Ginzburg–Landau equations describing a Turing–Hopf bifurcation with modulation group velocity of $O(1)$.

1.2.2.1 Amplitude Equations for Planar Neural Fields

In two spatial dimensions the same ideas go across and can be used to determine the selection of patterns, say stripes vs. spots [31]. In the hierarchy of Eqs. (1.52)–(1.55) the symbol \otimes now represents a convolution in two spatial dimensions. The two dimensional Fourier transform \hat{w} takes the explicit form

$$\hat{w}(p_1, p_2) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy w(x, y) e^{i(p_1 x + p_2 y)}, \quad (1.70)$$

and the inner product for periodic scalar functions defined on the plane is taken as

$$\langle U, V \rangle = \frac{1}{|\Omega|} \int_{\Omega} U^*(\mathbf{r}) V(\mathbf{r}) d\mathbf{r}, \quad (1.71)$$

with $\Omega = (0, 2\pi/p_c) \times (0, 2\pi/p_c)$. We shall assume a radially symmetric kernel so that $\hat{w}(p_1, p_2) = \hat{w}\left(\sqrt{p_1^2 + p_2^2}\right)$. One composite pattern that solves the linearised equations is

$$g_1(x, y, \tau) = A_1(\tau) e^{ip_c x} + A_2(\tau) e^{ip_c y} + \text{cc}. \quad (1.72)$$

For $A_1 = 0$ and $A_2 \neq 0$ we have a *stripe*, while if both A_1 and A_2 are non-zero, and in particular equal, we have a *spot*. Here p_c is defined by the condition $\beta_c = 1/\hat{w}(p_c)$. The null space of \mathcal{L} is spanned by $\{e^{\pm ip_c x}, e^{\pm ip_c y}\}$, and we may proceed as for the one dimensional case to generate a set of coupled equations for the amplitudes A_1 and A_2 . It is simple to show that

$$\langle e^{ip_c x}, w \otimes g_1^3 \rangle = 3\beta_c^{-1} A_1 (|A_1|^2 + 2|A_2|^2). \quad (1.73)$$

Assuming a representation for g_2 as

$$g_2 = \alpha_0 + \alpha_1 e^{2ip_c x} + \alpha_2 e^{-2ip_c x} + \alpha_3 e^{2ip_c y} + \alpha_4 e^{-2ip_c y} + \alpha_5 e^{ip_c(x+y)} \\ + \alpha_6 e^{-ip_c(x+y)} + \alpha_7 e^{ip_c(x-y)} + \alpha_8 e^{-ip_c(x-y)} + \phi g_1, \quad (1.74)$$

allows us to calculate

$$\langle e^{ip_c x}, w \otimes g_1 g_2 \rangle = \beta_c^{-1} [\alpha_0 A_1 + \alpha_1 A_1^* + \alpha_5 A_2^* + \alpha_7 A_2]. \quad (1.75)$$

Balancing terms in (1.54) gives

$$\alpha_0 = \frac{2\beta_2(|A_1|^2 + |A_2|^2)\hat{w}(0)}{1 - \beta_c \hat{w}(0)}, \quad \alpha_1 = \frac{\beta_2 A_1^2 \hat{w}(2p_c)}{1 - \beta_c \hat{w}(2p_c)}, \quad (1.76)$$

$$\alpha_5 = \frac{2\beta_2 A_1 A_2 \hat{w}(\sqrt{2}p_c)}{1 - \beta_c \hat{w}(\sqrt{2}p_c)}, \quad \alpha_7 = \frac{2\beta_2 A_1 A_2^* \hat{w}(\sqrt{2}p_c)}{1 - \beta_c \hat{w}(\sqrt{2}p_c)}. \quad (1.77)$$

Combining the above yields the coupled amplitude equations:

$$\beta_c \frac{dA_1}{dt} = A_1 (\delta - \Phi |A_1|^2 - \Psi |A_2|^2), \quad (1.78)$$

$$\beta_c \frac{dA_2}{dt} = A_2 (\delta - \Phi |A_2|^2 - \Psi |A_1|^2), \quad (1.79)$$

where

$$\Phi = -3\beta_3 - 2\beta_2^2 \left[\frac{2\hat{w}(0)}{1 - \beta_c \hat{w}(0)} + \frac{\hat{w}(2p_c)}{1 - \beta_c \hat{w}(2p_c)} \right], \quad (1.80)$$

$$\Psi = -6\beta_3 - 4\beta_2^2 \left[\frac{\hat{w}(0)}{1 - \beta_c \hat{w}(0)} + \frac{2\hat{w}(\sqrt{2}p_c)}{1 - \beta_c \hat{w}(\sqrt{2}p_c)} \right]. \quad (1.81)$$

The stripe solution $A_2 = 0$ and $|A_1| = \sqrt{\delta/\Phi}$ (or vice versa) is stable if and only if $\Psi > \Phi > 0$. The spot solution $|A_1| = |A_2| = \sqrt{\delta/(\Phi + \Psi)}$ is stable if and only if $\Phi > \Psi > 0$. Hence, stripes and spots are mutually exclusive as stable patterns. In the absence of quadratic terms in f , namely $\beta_2 = 0$, then $\Psi = -6\beta_3$ and $\Phi = -3\beta_3$ so that for an odd firing rate function like $f(x) = \tanh x \simeq x - x^3/3$ then $\beta_3 < 0$ and so $\Psi > \Phi$ and stripes are selected over spots. The key to the appearance of spots is non-zero quadratic terms, $\beta_2 \neq 0$, in the firing rate function; without these terms spots can never stably exist. For a Mexican hat connectivity then $\hat{w}(\sqrt{2}p_c) > \hat{w}(2p_c)$ and the quadratic term of Ψ is larger than that of Φ so that as $|\beta_2|$ increases then spots will arise instead of stripes.

The technique above can also be used to determine amplitude equations for more general patterns of the form

$$g_1(\mathbf{r}, \tau) = \sum_{j=1}^N A_j(\tau) e^{ip_c \mathbf{R}_j \cdot \mathbf{r}}. \quad (1.82)$$

For further discussion we refer the reader to [32, 84].

1.2.3 Brain Wave Equations

Given the relatively few analytical techniques for investigating neural field models one natural step is to make use of numerical simulations to explore system dynamics. For homogeneous models we may exploit the convolution structure of interactions to develop fast Fourier methods to achieve this. Indeed we may also exploit this structure further to obtain equivalent PDE models [61] (see also Chap. 5), and recover the brain wave equation often used in EEG modelling [49, 67].

For example consider a one-dimensional neural field model with axonal delays:

$$\mathcal{Q}g = \psi, \quad \psi(x, t) = \int_{-\infty}^{\infty} dy w(|x - y|) f(g(y, t - |x - y|)/v). \quad (1.83)$$

The function $\psi(x, t)$ may be expressed in the form

$$\psi(x, t) = \int_{-\infty}^{\infty} ds \int_{-\infty}^{\infty} dy G(x - y, t - s) \rho(y, s), \quad (1.84)$$

where

$$G(x, t) = \delta(t - |x|/v) w(x), \quad (1.85)$$

can be interpreted as another type of Green's function, and we use the notation

$$\rho(x, t) = f(g(x, t)). \quad (1.86)$$

Introducing Fourier transforms of the following form

$$\psi(x, t) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i(kx + \omega t)} \psi(k, \omega) dk d\omega, \quad (1.87)$$

allows us to write

$$\psi(k, \omega) = G(k, \omega) \rho(k, \omega), \quad (1.88)$$

assuming the Fourier transform of $f(u)$ exists. It is straightforward to show that the Fourier transform of (1.85) is

$$G(k, \omega) = v(\omega/v + k) + v(\omega/v - k), \quad (1.89)$$

where

$$v(E) = \int_0^\infty w(x)e^{-iEx}dx. \quad (1.90)$$

We shall focus on a common form of (normalised) exponential synaptic footprint:

$$w(x) = \frac{1}{2\sigma} \exp(-|x|/\sigma), \quad v(E) = \frac{1}{2\sigma} \frac{1}{\sigma^{-1} + iE}, \quad \sigma > 0. \quad (1.91)$$

We now exploit the product structure of (1.88) and properties of (1.89) to re-formulate the original integral model in terms of a PDE. Using (1.89) and (1.91) we see that

$$G(k, \omega) = \frac{1}{\sigma} \frac{A(\omega)}{A(\omega)^2 + k^2}, \quad A(\omega) = \frac{1}{\sigma} + i \frac{\omega}{v}. \quad (1.92)$$

We may now write (1.88) as $(A(\omega)^2 + k^2)\psi(k, \omega) = A(\omega)\rho(k, \omega)/\sigma$, which upon inverse Fourier transforming gives the PDE:

$$[\mathcal{A}^2 - \partial_{xx}] \psi = \frac{1}{\sigma} \mathcal{A} \rho, \quad \mathcal{A} = \left(\frac{1}{\sigma} + \frac{1}{v} \partial_t \right). \quad (1.93)$$

This is a type of damped wave equation with an inhomogeneity dependent on (1.86). This equation has previously been derived by Jirsa and Haken [49] and studied intensively in respect to the brain-behaviour experiments of Kelso et al. [53]. For the numerical analysis of travelling wave solutions to (1.93) we refer the reader to [21].

The same approach can also be used in two spatial dimensions, and here $G(k, \omega) = G(\mathbf{k}, \omega)$ would be interpreted as the three dimensional integral transform:

$$G(k, \omega) = \int_{-\infty}^{\infty} ds \int_{\mathbb{R}^2} d\mathbf{r} G(\mathbf{r}, s) e^{-i(\mathbf{k} \cdot \mathbf{r} + \omega s)}. \quad (1.94)$$

For the choice $w(r) = e^{-r/\sigma}/(2\pi)$, we find that (see Appendix 4)

$$G(k, \omega) = \frac{A(\omega)}{(A^2(\omega) + k^2)^{3/2}}, \quad (1.95)$$

which, unlike in one spatial dimension, is not a ratio of polynomials in k and ω . The problem arises as of how to interpret $[\mathcal{A}^2 - \nabla^2]^{3/2}$. In the long-wavelength approximation one merely expands $G(k, \omega)$ around $k = 0$ for small k , yielding a

“nice” rational polynomial structure which is then manipulated as described above to give the PDE:

$$\left(\mathcal{A}^2 - \frac{3}{2}\nabla^2\right)\psi = \rho. \quad (1.96)$$

This model has been intensively studied by a number of authors in the context of EEG modelling, see for example [10, 77, 83]. For an alternative approximation to the long-wavelength one we refer the reader to [26].

1.3 Travelling Waves and Localised States

As well as global periodic patterns neural field models are able to support localised solutions in the form of standing bumps and travelling pulses of activity. For clarity of exposition we shall focus on a one-dimensional neural field model with axonal delays:

$$g(x, t) = \int_{-\infty}^{\infty} dy w(y) \int_0^{\infty} ds \eta(s) f \circ g(x - y, t - s - |y|/v). \quad (1.97)$$

Following the standard approach for constructing travelling wave solutions to PDEs, such as reviewed by Sandstede [79], we introduce the coordinate $\xi = x - ct$ and seek functions $U(\xi, t) = g(x - ct, t)$ that satisfy (1.97). In the (ξ, t) coordinates, the integral equation (1.97) reads

$$U(\xi, t) = \int_{-\infty}^{\infty} dy w(y) \int_0^{\infty} ds \eta(s) f \circ U(\xi - y + cs + c|y|/v, t - s - |y|/v). \quad (1.98)$$

The travelling wave is a stationary solution $U(\xi, t) = q(\xi)$ (independent of t), that satisfies

$$q(\xi) = \int_0^{\infty} \eta(s) \psi(\xi + cs) ds, \quad \psi(\xi) = \int_{-\infty}^{\infty} w(y) f \circ q(\xi - y + c|y|/v) dy. \quad (1.99)$$

To determine stability we linearise (1.98) about the steady state $q(\xi)$ by writing $U(\xi, t) = q(\xi) + u(\xi, t)$, and Taylor expand, to give

$$u(\xi, t) = \int_{-\infty}^{\infty} dy w(y) \int_0^{\infty} ds \eta(s) f'(q(\xi - y + cs + c|y|/v)) u(\xi - y + cs + c|y|/v, t - s - |y|/v). \quad (1.100)$$

Of particular importance are bounded smooth solutions defined on \mathbb{R} , for each fixed t . Thus one looks for solutions of the form $u(\xi, t) = u(\xi)e^{\lambda t}$. This leads to the eigenvalue equation $u = \mathcal{L}u$:

$$u(\xi) = \int_{-\infty}^{\infty} dy w(y) \int_{\xi-y+c|y|/v}^{\infty} \frac{ds}{c} \eta(-\xi/c + y/c - |y|/v + s/c) e^{-\lambda(-\xi/c + y/c + s/c)} f'(q(s)) u(s). \quad (1.101)$$

Let $\sigma(\mathcal{L})$ be the spectrum of \mathcal{L} . We shall say that a travelling wave is linearly stable if

$$\max\{\operatorname{Re}(\lambda) : \lambda \in \sigma(\mathcal{L}), \lambda \neq 0\} \leq -K, \quad (1.102)$$

for some $K > 0$, and $\lambda = 0$ is a simple eigenvalue of \mathcal{L} . In general the normal spectrum of the operator obtained by linearising a system about its travelling wave solution may be associated with the zeros of a complex analytic function, the so-called Evans function. This was originally formulated by Evans [35] in the context of a stability theorem about excitable nerve axon equations of Hodgkin-Huxley type. Next we show how to calculate the properties of waves and bumps for the special case of a Heaviside firing rate function, $f(g) = H(g - h)$, for some constant threshold h .

1.3.1 Travelling Front

As an example consider travelling front solutions such that $q(\xi) > h$ for $\xi < 0$ and $q(\xi) < h$ for $\xi > 0$. It is then a simple matter to show that

$$\psi(\xi) = \begin{cases} \int_{\xi/(1-c/v)}^{\infty} w(y) dy & \xi \geq 0 \\ \int_{\xi/(1+c/v)}^{\infty} w(y) dy & \xi < 0 \end{cases}. \quad (1.103)$$

The choice of origin, $q(0) = h$, gives an implicit equation for the speed of the wave as a function of system parameters. The construction of the Evans function begins with an evaluation of (1.101). Under the change of variables $z = q(s)$ this equation may be written

$$\begin{aligned} u(\xi) &= \int_{-\infty}^{\infty} dy w(y) \int_{q(\xi-y+c|y|/v)}^{q(\infty)} \frac{dz}{c} \eta(q^{-1}(z)/c - \xi/c + y/c - |y|/v) \\ &\times e^{-\lambda(q^{-1}(z)/c - \xi/c + y/c)} \frac{\delta(z-h)}{|q'(q^{-1}(z))|} u(q^{-1}(z)). \end{aligned} \quad (1.104)$$

For the travelling front of choice we note that when $z = h$, $q^{-1}(h) = 0$ and (1.104) reduces to

$$u(\xi) = \frac{u(0)}{c|q'(0)|} \int_{-\infty}^{\infty} dy w(y) \eta(-\xi/c + y/c - |y|/v) e^{-\lambda(y-\xi)/c}. \quad (1.105)$$

From this equation we may generate a self-consistent equation for the value of the perturbation at $\xi = 0$, simply by setting $\xi = 0$ on the left hand side of (1.105). This self-consistent condition reads

$$u(0) = \frac{u(0)}{c|q'(0)|} \int_{-\infty}^{\infty} dy w(y) \eta(y/c - |y|/v) e^{-\lambda y/c}. \quad (1.106)$$

Importantly there are only nontrivial solutions if $\mathcal{E}(\lambda) = 0$, where

$$\mathcal{E}(\lambda) = 1 - \frac{1}{c|q'(0)|} \int_{-\infty}^{\infty} dy w(y) \eta(y/c - |y|/v) e^{-\lambda y/c}. \quad (1.107)$$

From causality $\eta(t) = 0$ for $t \leq 0$ and physically $c < v$ so

$$\mathcal{E}(\lambda) = 1 - \frac{1}{c|q'(0)|} \int_0^{\infty} dy w(y) \eta(y/c - y/v) e^{-\lambda y/c}. \quad (1.108)$$

We identify (1.108) with the Evans function for the travelling front solution of (1.97). The Evans function is real-valued if λ is real. Furthermore, (i) the complex number λ is an eigenvalue of the operator \mathcal{L} if and only if $\mathcal{E}(\lambda) = 0$, and (ii) the algebraic multiplicity of an eigenvalue is equal to the order of the zero of the Evans function.

Consider the choice $\eta(t) = e^{-t} H(t)$ and $w(x) = e^{-|x|}/2$. Assuming $c > 0$ the travelling front (1.99) is given in terms of (1.103) which takes the explicit form

$$\psi(\xi) = \begin{cases} \frac{1}{2} e^{m-\xi} & \xi \geq 0 \\ 1 - \frac{1}{2} e^{m+\xi} & \xi < 0 \end{cases}, \quad m_{\pm} = \frac{v}{c \pm v}. \quad (1.109)$$

The speed of the front is determined from the condition $q(0) = h$ as

$$c = \frac{v(2h-1)}{2h-1-2hv}. \quad (1.110)$$

The Evans function is easily calculated as

$$\mathcal{E}(\lambda) = \frac{\lambda}{c + (1 - c/v) + \lambda}, \quad (1.111)$$

where we use the result from (1.99) and (1.103) that $cq'(0) = \int_0^{\infty} \eta(y/c - y/v) w(y) dy$. The equation $\mathcal{E}(\lambda) = 0$ only has the solution $\lambda = 0$. We also have that $\mathcal{E}'(0) > 0$ showing that $\lambda = 0$ is a simple eigenvalue. Hence, the travelling wave front for this example is linearly stable.

We refer the reader to [22] for further examples of wave calculations in other neural field models.

1.3.2 Stationary Bump

Here we construct static (time-independent) patterns of the form $g(x, t) = q(x)$ for all t . Using (1.97) gives

$$q(x) = \int_{-\infty}^{\infty} w(x - y) H(q(y) - h) dy. \quad (1.112)$$

We shall assume the synaptic kernel has a Mexican hat shape given by

$$w(x) = (1 - |x|) e^{-|x|}, \quad (1.113)$$

and look for solutions of the form $\lim_{x \rightarrow \pm\infty} q(x) = 0$, with $q(x) \geq h$ for $x_1 < x < x_2$ (see inset of Fig. 1.4). In this case the exact solution is given simply by

$$q(x) = \int_{x_1}^{x_2} w(x - y) dy. \quad (1.114)$$

For the Mexican hat function (1.113), a simple calculation gives

$$q(x) = \begin{cases} g(x - x_1) - g(x - x_2) & x > x_2 \\ g(x_2 - x) + g(x - x_1) & x_1 \leq x \leq x_2, \\ g(x_2 - x) - g(x_1 - x) & x < x_1 \end{cases} \quad (1.115)$$

where $g(x) = x e^{-x}$. The conditions $q(x_1) = h$ and $q(x_2) = h$ both lead to the equation

$$\Delta e^{-\Delta} = h, \quad (1.116)$$

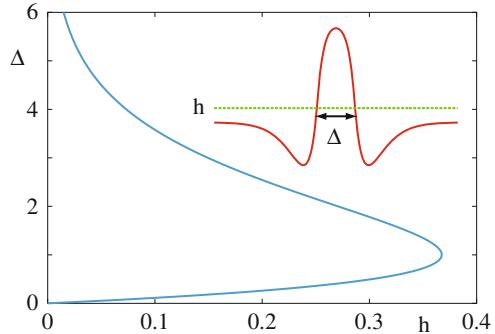
describing a family of solutions with $\Delta = (x_2 - x_1)$. Hence, homoclinic bumps are only possible if $h < 1/e$. The full branch of solutions for $\Delta = \Delta(h)$ is shown in Fig. 1.4. Here we see a branch of wide solutions ($\Delta > 1$) and a branch of thinner solutions ($\Delta < 1$) that connect in a saddle-node bifurcation at $\Delta = 1$.

One of the translationally invariant solutions may be picked by imposing a *phase* condition $q(0) = q_0$, where $q_0 \leq \max q(x) = \Delta e^{-\Delta/2}$ (which occurs at $(x_1 + x_2)/2$ using (1.115)).

To determine stability we use the result that $q'(x) = w(x - x_1) - w(x - x_2)$ so that the eigenvalue equation (1.101), with $c = 0$, reduces to

$$u(x) = \frac{\tilde{\eta}(\lambda)}{|w(0) - w(\Delta)|} [w(x - x_1)u(x_1)e^{-\lambda|x-x_1|/v} + w(x - x_2)u(x_2)e^{-\lambda|x-x_2|/v}], \quad (1.117)$$

Fig. 1.4 Bump width as a function of h , as determined by Eq. (1.116) for a step function firing rate function and Mexican hat kernel (1.113). The inset shows the shape of a bump. A linear stability analysis shows that the upper branch of solutions is stable, and the lower branch unstable



where $\tilde{\eta}(\lambda) = \int_0^\infty \eta(s)e^{-\lambda s} ds$. The equality in (1.117) implies that if $u(x_{1,2}) = 0$ then $u(x) = 0$ for all x . We now examine the matrix equation obtained from (1.117) at the points $x = x_{1,2}$,

$$\begin{bmatrix} u(x_1) \\ u(x_2) \end{bmatrix} = \frac{\tilde{\eta}(\lambda)}{|w(0) - w(\Delta)|} \begin{bmatrix} w(0) & w(\Delta)e^{-\lambda|\Delta|/v} \\ w(\Delta)e^{-\lambda|\Delta|/v} & w(0) \end{bmatrix} \begin{bmatrix} u(x_1) \\ u(x_2) \end{bmatrix}. \quad (1.118)$$

Non-trivial solutions are only possible if

$$\frac{1}{\tilde{\eta}(\lambda)} = \frac{w(0) \pm w(\Delta)e^{-\lambda|\Delta|/v}}{|w(0) - w(\Delta)|}. \quad (1.119)$$

Solutions will be stable if $\text{Re}(\lambda) < 0$. It is a simple matter to check that there is always one solution with $\lambda = 0$. For an exponential synapse $\tilde{\eta}(\lambda) = (1 + \lambda/\alpha)^{-1}$ and we see that as $v \rightarrow \infty$ solutions are only stable if $(w(0) + w(\Delta))/|w(0) - w(\Delta)| < 1$. This is true if $w(\Delta) = \exp(-\Delta)(1 - \Delta) < 0$. Hence, of the two possible branches of $\Delta = \Delta(h)$, it is the one with largest Δ that is stable. In the other extreme where $v \rightarrow 0$ it is simple to show that real and positive values for λ can only occur when $w(\Delta) > 0$. Hence, as v is varied there are no new instabilities due to real eigenvalues passing through zero. However, for finite conduction velocities it is possible that solutions may destabilise via a Hopf bifurcation. By writing $\lambda = i\omega$ the conditions for a Hopf bifurcation, where $\text{Re}(\lambda) = 0$ and $\text{Im}(\lambda) \neq 0$, are obtained from the simultaneous solution of

$$1 = \frac{w(0) + w(\Delta) \cos(\omega|\Delta|/v)}{|w(0) - w(\Delta)|}, \quad \frac{\omega}{\alpha} = -\frac{w(\Delta) \sin(\omega|\Delta|/v)}{|w(0) - w(\Delta)|}. \quad (1.120)$$

Eliminating $\sin(\omega|\Delta|/v)$ between these two equations gives

$$\omega^2 = \frac{\alpha^2}{|w(0) - w(\Delta)|^2} [w^2(\Delta) - \{ |w(0) - w(\Delta)| - w(0) \}^2]. \quad (1.121)$$

The condition that $\omega \neq 0$ requires the choice $w(\Delta) > w(0)$, which does not hold for (1.113). Hence, the presence of axonal delays neither affects the existence or stability of bump solutions.

1.3.3 Interface Dynamics

For zero axonal delays ($v \rightarrow \infty$) and $\eta(t) = \alpha e^{-\alpha t} H(t)$ the existence and stability of bumps may be investigated using the alternative techniques of Amari [3]. We use the threshold condition $g(x_i, t) = h$ to define the locations $x_{1,2}$ from which we obtain the equations of motion of the boundary as

$$\frac{dx_i}{dt} = -\left. \frac{g_t}{g_x} \right|_{x=x_i(t)}. \quad (1.122)$$

It is then possible to calculate the rate of change of the interval $\Delta(t) = x_2(t) - x_1(t)$ using

$$\frac{1}{\alpha} g_t(x, t) = -g(x, t) + \int_{x_1(t)}^{x_2(t)} w(x-y) dy. \quad (1.123)$$

Then (for a single bump)

$$\frac{d\Delta}{dt} = \alpha \left(\frac{1}{c_1} + \frac{1}{c_2} \right) \left[\int_0^\Delta w(|y|) dy - h \right], \quad (1.124)$$

where

$$c_1 = \frac{\partial g(x_1, t)}{\partial x}, \quad -c_2 = \frac{\partial g(x_2, t)}{\partial x}. \quad (1.125)$$

Hence, making the convenient assumption that $\partial g(x_i, t)/\partial x$ is roughly constant, the equilibrium solution defined by $\int_0^\Delta w(|y|) dy = h$ is stable if

$$\frac{d}{d\Delta} \int_0^\Delta w(|y|) dy = w(\Delta) < 0. \quad (1.126)$$

We thus recover the result obtained with the Evans function method (widest bump is stable).

It is also possible to recover results about travelling fronts in the interface framework (at least for $v \rightarrow \infty$). In this case it is natural to define a pattern boundary as the interface between a high and low activity state. Let us assume that the front is such that $g(x, t) > h$ for $x < x_0(t)$ and $g(x, t) \leq h$ for $x \geq x_0(t)$ then (1.97), for the choice $\eta(t) = e^{-t} H(t)$, reduces to

$$g_t(x, t) = -g(x, t) + \int_{x-x_0(t)}^{\infty} w(y) dy. \quad (1.127)$$

Introducing $z = u_x$ and differentiating (1.127) with respect to x gives

$$z_t(x, t) = -z(x, t) - w(x - x_0(t)). \quad (1.128)$$

Integrating (1.128) from $-\infty$ to t (and dropping transients) gives

$$z(x, t) = -e^{-t} \int_{-\infty}^t e^s w(x - x_0(s)) ds. \quad (1.129)$$

We may now use the interface dynamics, $g(x_0(t), t)$, defined by:

$$\dot{x}_0 = -\frac{g_t}{g_x} \Big|_{x=x_0(t)}, \quad (1.130)$$

to study the speed $c > 0$ of a front, defined by $\dot{x}_0 = c$. In this case $x_0(t) = ct$ (where without loss of generality we set $x_0(0) = 0$) and from (1.127) and (1.128) we have that

$$g_t|_{x=x_0(t)} = -h + \tilde{w}(0), \quad g_x|_{x=x_0(t)} = -\tilde{w}(1/c)/c, \quad (1.131)$$

where $\tilde{w}(\lambda) = \int_0^\infty e^{-\lambda s} w(s) ds$. Hence from (1.130) the speed of the front is given implicitly by the equation

$$h = \tilde{w}(0) - \tilde{w}(1/c). \quad (1.132)$$

To determine stability of the travelling wave we consider a perturbation of the interface and an associated perturbation of g . Introducing the notation $\hat{\cdot}$ to denote perturbed quantities to a first approximation we will set $\hat{g}_x|_{x=\hat{x}_0(t)} = g_x|_{x=ct}$, and write $\hat{x}_0(t) = ct + \delta x_0(t)$. The perturbation in g can be related to the perturbation in the interface by noting that the perturbed and unperturbed boundaries are defined by a level set condition so that $g(x_0, t) = h = \hat{g}(\hat{x}_0, t)$. Introducing $\delta g(t) = g|_{x=ct} - \hat{g}|_{x=\hat{x}_0(t)}$ we thus have the condition that $\delta g(t) = 0$ for all t . Integrating (1.127) (and dropping transients) gives

$$g(x, t) = e^{-t} \int_{-\infty}^t ds e^s \int_{x-x_0(s)}^{\infty} dy w(y), \quad (1.133)$$

and \hat{g} is obtained from (1.133) by simply replacing x_0 by \hat{x}_0 . Using the above we find that δg is given (to first order in δx_0) by

$$\delta g(t) = \frac{1}{c} \int_0^\infty ds e^{-s/c} w(s) [\delta x_0(t) - \delta x_0(t - s/c)] = 0. \quad (1.134)$$

This has solutions of the form $\delta x_0(t) = e^{\lambda t}$, where λ is defined by $\mathcal{E}(\lambda) = 0$, with

$$\mathcal{E}(\lambda) = 1 - \frac{\tilde{w}((1 + \lambda)/c)}{\tilde{w}(1/c)}. \quad (1.135)$$

A front is stable if $\text{Re}\lambda < 0$.

As an example consider the choice $w(x) = e^{-|x|}/2$, for which $\tilde{w}(\lambda) = (\lambda + 1)^{-1}/2$. In this case the speed of the wave is given from (1.132) as

$$c = \frac{1 - 2h}{2h}, \quad (1.136)$$

and

$$\mathcal{E}(\lambda) = \frac{\lambda}{1 + c + \lambda}. \quad (1.137)$$

Note that these results recover Eqs. (1.110) and (1.111) in the limit $v \rightarrow \infty$, as expected. Hence, the travelling wave front for this example is neutrally stable.

For a recent extension of the Amari interface dynamics to planar neural field models we refer the reader to [24].

1.4 Inverse Neural Modelling

While neural modelling is concerned with the study of the dynamics which arise in the framework of neural activity on the basis of some given connectivity function w , *inverse neural modelling* studies the construction of such connectivity kernels $w(x, y)$ given a prescribed dynamics of the activity fields $u(x, t)$.

1.4.1 Inverse Problems

Here, we focus on the investigation of a kernel construction for the Amari equation (1.2) with a homogeneous kernel function w or its inhomogeneous version

$$\tau \frac{\partial u}{\partial t}(x, t) = -u(x, t) + \int_D w(x, y) f(u(y, t)) dy, \quad x \in D, t \geq 0, \quad (1.138)$$

with some constant $\tau > 0$, where $D \subset \mathbb{R}^m$ is some bounded domain in space dimension $m \in \mathbb{N}$. Let us assume that on D the field $u(x, t)$ is given for $x \in D, t \in [0, T]$ with some final time $T > 0$. This is usually called the *full field inverse neural problem*. The task is to find a kernel $w(x, y)$ such that Eq. (1.138) with kernel w

and initial condition $u(x, 0)$ has $u(x, t)$ as its unique solution for all $x \in D$ and $t \in [0, T]$.

Since $u(x, t)$ is given, we can define

$$\psi(x, t) := \tau \frac{\partial u}{\partial t}(x, t) + u(x, t), \quad x \in D, t \in [0, T] \quad (1.139)$$

and

$$\varphi(x, t) := f(u(x, t)), \quad x \in D, t \in [0, T]. \quad (1.140)$$

With the functions (1.139) and (1.140) we transform the dynamics (1.138) into

$$\psi(x, t) = \int_D w(x, y)\varphi(y, t)dy, \quad x \in D, t \in [0, T]. \quad (1.141)$$

Equation (1.141) is an integral equation for the unknown connectivity function w .

We first remark that since the integration in (1.141) is carried out with respect to the variable y , and since the kernel $\varphi(y, t)$ of the integral operator

$$(Kg)(t) := \int_D \varphi(y, t)g(y)dy, \quad t \in [0, T], \quad (1.142)$$

is dependent on y and t only, the spatial coordinate $x \in D$ can be considered as a parameter, i.e. the integral equation is indeed a *family* of integral equations

$$Kw_x = \psi(x, \cdot), \quad x \in D, \quad (1.143)$$

with different left-hand sides $\psi(x, \cdot)$ for $x \in D$, where we use the notation $w_x := w(x, \cdot)$.

We need to answer questions of *uniqueness*, *existence* and *stability* for the inverse neural task. It will turn out, that the inverse neural field problem shares basic features with many other inverse problems as described for example in [18, 29, 57, 74]. Most inverse problems are *ill-posed* in the sense of Hadamard [44]. A problem is called *well-posed*, if

1. For given input data it has at most one solution (*uniqueness*),
2. For any input data there exists a solution (*existence*),
3. The solution depends continuously on the input data (*stability*).

A problem which is not well-posed, is called *ill-posed*. Clearly, ill-posedness depends on the spaces under consideration, an appropriate condition can make non-unique problems uniquely solvable, the choice of appropriate norms can make an unstable problem stable. But often, the spaces are dependent on the particular applied setup, and it is not possible to control input functions for example in spaces which need to take care of an infinite number of derivatives. This means that ill-posedness

naturally appears in many inverse problems and needs to be taken care of. This is true also for the inverse neural field problem above.

It is well-known [59] that a compact operator $K : X \rightarrow Y$ from an infinite dimensional normed space X into an infinite dimensional normed space Y cannot have a bounded inverse, since otherwise the identity operator $I = K^{-1} \circ K$ would be compact, which cannot be the case. Integral operators with continuous or weakly-singular kernels on bounded sets are compact in the spaces of continuous or L^2 -integrable functions, since we find a sequence of finite-dimensional approximations to the kernels. As a consequence we obtain compactness of the operator K defined in (1.142) in the spaces $BC(D)$ and $L^2(D)$. Thus, the *inverse neural field* problem under consideration is *ill-posed* in the sense described by the conditions (1)–(3). The kernel w does not depend stably on the right-hand side $\psi \in BC(D)$ or $\psi \in L^2(D)$. We need to stabilize the solution to the inverse problem, which is usually carried out by some *regularization method*. We will introduce regularization further down.

Here, we first show that the inverse neural field problem is neither uniquely solvable nor is it exactly solvable at all in the general case where a function $u(x, t)$ is prescribed (cf. [5]). To see its *non-uniqueness* we consider the case where $u(x, t) = 0$ for $x \in M$ and $t \in [0, T]$ on some set $M \subset D$. Let $\tilde{w} \neq 0$ be a kernel which is zero for $x, y \notin M$ and arbitrary for $x, y \in M$. Then, for a solution $w_1(x, y)$ of the inverse problem the kernel $w_2 := w_1 + \tilde{w} \neq w_1$ is another different solution, since it does not change the dynamical behaviour of u . Thus, the inverse problem is non-unique.

In general, the non-uniqueness can be characterized as follows. We define

$$V := \text{span}\{\varphi(\cdot, t) : t \in [0, T]\} \quad (1.144)$$

as a subset of $L^2(D)$. Then, $Kg = 0$ can be written as

$$\langle v, g \rangle = 0 \quad \text{forall } v \in V \Leftrightarrow g \in V^\perp. \quad (1.145)$$

Thus, the inverse neural field problem is unique in $L^2(D)$ if and only if $\overline{V} = L^2(D)$, which is equivalent to $V^\perp = \{0\}$. It means that the time dynamics of the field $\varphi(\cdot, t) = f(u(\cdot, t))$ covers all dimensions of the space $L^2(D)$. Any element \tilde{w} of V^\perp can be added to a solution kernel w and will not change the dynamics of the neural field u .

Existence of solutions in a space X with for example $X = L^2(D)$ or $X = BC(D)$ is equivalent to the condition

$$\psi(x, \cdot) \in K(X) \quad \text{for all } x \in D. \quad (1.146)$$

Assume that $u(x, \cdot) \in C^n([0, T])$, but $u(x, \cdot) \notin C^{n+1}([0, T])$ and let f be analytic. Then $Kg \in C^n([0, T])$ for any $g \in L^2(D)$, but $\psi = \tau \partial u / \partial t + u \in C^{n-1}([0, T])$, $\psi \notin C^n([0, T])$. As a consequence there cannot be a solution to the inverse neural field equation in $L^2(D)$ in this case. In general the inverse neural field problem does

not have a solution. But we will be able to construct approximate solutions. To this end we need to introduce regularization techniques.

Consider a compact linear operator $K : X \rightarrow Y$ from a Hilbert space X into a Hilbert space Y . We denote its adjoint by K^* . Then, we denote the nonnegative square roots μ_n of the eigenvalues of the self-adjoint compact operator K^*K as *singular values* of the operator K , see for example [59]. We assume that μ_n are ordered according to its size, such that $\mu_1 \geq \mu_2 \geq \dots$. The singular vector for μ_n is denoted by $v_n, n \in \mathbb{N}$. Then, the set of vectors

$$g_n := \frac{1}{\mu_n} K v_n, \quad n \in \mathbb{N}, \quad (1.147)$$

is an orthonormal system in Y , such that we have

$$Kv = \sum_{n=1}^{\infty} \mu_n \langle v, v_n \rangle g_n \quad (1.148)$$

for elements $v \in X$. Equation (1.148) is known as the *spectral representation* of K with respect to its singular system. We also note that

$$Kv_n = \mu_n g_n, \quad K^* g_n = \mu_n v_n \quad (1.149)$$

for all $n \in \mathbb{N}$. If K is injective, then the inverse of the operator K is given by

$$K^{-1}g = \sum_{n=1}^{\infty} \frac{1}{\mu_n} \langle g, g_n \rangle v_n \quad (1.150)$$

for an element $g \in Y$.

If the operator K is compact, then the singular values μ_n need to tend to zero for $n \rightarrow \infty$. This also means that $1/\mu_n$ tends to infinity. The spectral coefficients $\langle g, g_n \rangle$ of the function $g \in Y$ are multiplied by a large number $1/\mu_n$, which reflects the unboundedness of the operator K^{-1} . The image space of K in Y is given by the elements $g \in Y$ for which we have

$$\sum_{n=1}^{\infty} \frac{1}{\mu_n^2} \langle g, g_n \rangle^2 < \infty. \quad (1.151)$$

Picard's theorem (cf. [18]) states that $Kv = g$ has a solution in X if and only if (1.151) is satisfied for $g \in Y$.

The basic idea of regularization is to replace the unbounded term $1/\mu_n$ in (1.150) by a bounded term. The method is called *spectral damping*, since it damps the unbounded modes of $K^{-1}g$. A *regularization operator* $R_\alpha : Y \rightarrow X$ is defined by

$$R_\alpha g := \sum_{n=1}^{\infty} q_n^{(\alpha)} \langle g, g_n \rangle v_n. \quad (1.152)$$

Here, $\alpha > 0$ is denoted as *regularization parameter* and the damping factors $q_n^{(\alpha)}$ need to be bounded for $n \in \mathbb{N}$. The choice

$$q_n^{(\alpha)} := \begin{cases} \frac{1}{\mu_n}, & n \leq 1/\alpha \\ 0, & \text{otherwise,} \end{cases} \quad (1.153)$$

is known as *spectral cut-off*. The damping

$$q_n^{(\alpha)} := \frac{\mu_n}{\alpha + \mu_n^2}, \quad n \in \mathbb{N}, \quad (1.154)$$

is known as *Tikhonov regularization*. In operator terms, Tikhonov regularization can be written as

$$R_\alpha = (\alpha I + K^* K)^{-1} K^*, \quad \alpha > 0. \quad (1.155)$$

Tikhonov regularization can be equivalently obtained by minimizing the *Tikhonov functional*

$$\mu_\alpha(v) := \|v\|^2 + \|Kv - g\|^2 \quad (1.156)$$

over $v \in X$. The minimizer v_α of (1.156) is calculated by $v_\alpha = R_\alpha g$. The operator R_α is a bounded operator from Y to X . It converges pointwise to the inverse K^{-1} , i.e. we have

$$R_\alpha g \rightarrow K^{-1}g, \quad \alpha \rightarrow \infty \quad (1.157)$$

for each $g \in K(X)$, see for example [59], where one can also find the norm estimate

$$\|R_\alpha\| \leq \frac{1}{2\sqrt{\alpha}}, \quad \alpha > 0. \quad (1.158)$$

From (1.154) we also observe that $\|R_\alpha\| \rightarrow \infty$ for $\alpha \rightarrow 0$.

We can now apply regularization to the solution of the integral equation family (1.143). With R_α defined via (1.155) we calculate a regularized reconstruction kernel $w_\alpha(x, y)$ on $x, y \in D$ by

$$w_\alpha(x, \cdot) := R_\alpha \psi(x, \cdot), \quad x \in D. \quad (1.159)$$

Numerically, the reconstruction can be carried out using a grid $x_k, k = 1, \dots, N_D$ in the domain D and some time discretization $t_j, j = 1, \dots, N_T$. Then, K is approximated by a matrix \mathbf{K} in $\mathbb{R}^{N_T \times N_D}$ and $\psi(x, t)$ by a matrix Ψ in $\mathbb{R}^{N_D \times N_T}$, where each column corresponds to a time $t_j, j = 1, \dots, N_T$. The kernel w is approximated by a matrix \mathbf{W} in $\mathbb{R}^{N_D \times N_D}$, with entries $w_{k,\ell}$ which reflect the connectivity from y_ℓ to

x_k for $k, \ell \in \{1, \dots, N_D\}$. Now, the integral equation (1.143) corresponds to the matrix equation

$$\mathbf{K}\mathbf{W}^T = \boldsymbol{\Psi}^T. \quad (1.160)$$

The regularized solution of (1.160) is calculated by

$$\mathbf{W}_\alpha^T = (\alpha I + \mathbf{K}^T \mathbf{K})^{-1} \mathbf{K}^T \boldsymbol{\Psi}^T. \quad (1.161)$$

Further regularization methods such as a gradient scheme, also known as *backpropagation algorithm* for neural learning, can be found in [5].

We close this methodological part by looking into a slightly more general form than (1.143). The task is to construct an integral operator

$$(Wg)(x) := \int_D w(x, y)g(y)dy, \quad x \in D \quad (1.162)$$

which maps a given family $v_j \in L^2(D)$, $j \in J$, of functions with some index set J into its images $g_j \in L^2(D)$, $j \in J$, i.e. we search W such that

$$Wv_j = g_j, \quad j \in J. \quad (1.163)$$

Now consider some orthonormal system $(\varphi_n)_{n \in \mathbb{N}}$. If W is an integral operator, we can develop its kernel $w(x, \cdot)$ into its Fourier series with respect to $(\varphi_n)_{n \in \mathbb{N}}$, i.e.

$$\begin{aligned} w(x, y) &= \sum_{n=1}^{\infty} \left(\int_D w(x, y)\varphi_n(y)dy \right) \varphi_n(y) \\ &= \sum_{n=1}^{\infty} (W\varphi_n)(x)\varphi_n(y) \\ &= \sum_{n=1}^{\infty} \psi_n(x)\varphi_n(y) \quad x, y \in D \end{aligned} \quad (1.164)$$

with $\psi_n(x) := (W\varphi_n)(x)$. This means that for every orthonormal system $(\varphi_n)_{n \in \mathbb{N}}$ the kernel w is given by the sum over the products of $\varphi_n(y)$ with its image element $\psi_n(x)$, $x, y \in D$. The sum (1.164) is known as *Hebbian learning rule* in the framework of neural networks. It has originally been suggested by Hebb [46] as a mechanism based on physiological arguments. Here, it is a consequence of the Fourier theorem.

Equation (1.164) shows that the Hebbian learning rule is exact for learning the dynamics on orthonormal patterns. Of course, in general training patterns are no longer orthonormal and then the Hebb rule inhibits strong errors due to cross-talk between the patterns under consideration. In this case, the above approach based

on regularization is the correct modification and extension of this approach which yields better and stable results.

One important restriction for the task (1.163) needs to be noted at this point. Clearly, we have transformed the inverse problem into a linear equation. Assume that we have a state v which is linearly dependent on the states v_1, \dots, v_N . Then, we are not completely free to choose the image of v , but it is determined by the images of v_j , for $j = 1, \dots, N$ via the *consistency condition A*

$$Wv = W\left(\sum_{j=1}^N \alpha_j v_j\right) = \sum_{j=1}^N \alpha_j Wv_j. \quad (1.165)$$

Only if the consistency condition is satisfied, we can expect to obtain solvability of Eq. (1.163).

A second important *consistency condition B* is coming from the time-dynamical aspect of the inverse problem. Of course, to excite some pulse at time $t \in [0, T]$, we need other pulses to be active in a neighborhood of t . The activity needs to be large enough such that the threshold given by the function f is reached. Thus, in the form (1.138) we cannot expect to obtain a solvable inverse problem, if the consistency conditions of type A and B are violated.

We also refer to [75] for an interpretation of the above approach in terms of bi-orthogonal basis functions, which are numerically realized by the Moore-Penrose pseudoinverse given by $K^\dagger := (K^T K)^{-1} K^T$ and which is stabilized by Tikhonov regularization (1.155) when K is an ill-posed operator as for the inverse neural field problem.

1.4.2 Cognitive Modelling

Finally, we apply the inverse methods and its theory to problems of Dynamic Cognitive Modeling as suggested in [4, 5, 75]. The basic task here is to gain a better understanding of cognitive processes by mapping them into a dynamic field environment. The dynamical fields are then understood as neural activity patterns, which leads to the above inverse neural field or neural kernel construction problem.

The first step of dynamic cognitive modeling is to formulate some cognitive process in an abstract environment. In the easiest possible case we can consider a sequence of mental representations s_j , $j = 1, 2, \dots, N_j$. These states live in some space Z , the mental state space. We now construct a structure preserving mapping $\rho : Z \rightarrow X$ into the Hilbert space $X = L^2(D)$ of functions on a domain D . Here, we obtain a discrete sequence of states v_j , $j = 1, \dots, N_j$ in X . A continuous time dynamics can then be obtained by regarding the states v_j as saddle nodes that are connected along a stable heteroclinic sequence (SHS) [76], using Haken's order parameter approach [45]. Here, we employ either compactly supported functions $\chi : \mathbb{R} \rightarrow \mathbb{R}$ of class C^2 with $\chi(0) = 1$ and $\chi(s) = 0$ for $|s| \geq \Delta t$ or Gaussians

$$\chi(s) = e^{-\sigma s^2}, \quad s \in \mathbb{R}, \quad (1.166)$$

and define a temporal dynamics by

$$u(x, t) := \sum_{j=1}^{N_j} v_j(x) \chi(t - t_j), \quad x \in D, t \in [0, T], \quad (1.167)$$

where t_j are the points in time where the states v_j have maximal excitation. We then solve the inverse neural field problem (1.141) with (1.167) as input. The result is a neural kernel w which generates the dynamical representation of the above cognitive process in activation space. An application of this approach to syntactic language processing and its tentative relation to event-related brain potentials can be found in [6].

More complicated processes can be carried out in a similar way. For example, we can study elementary logical activity such as OR, AND or XOR. The logical representation consists of variables a_1 and a_2 as input variables and an output variable b , which takes states in $\{0, 1\}$ depending on the particular logical gate which we consider. The logical tables constitute our abstract logical space Z . As second step we map this space into a dynamical environment and construct a time-continuous dynamics. This is carried out by identifying the variables with coefficients of particular states v_1, v_2 and v_{out} in $X = L^2(D)$. The dynamics can now be constructed as in (1.167), i.e. we choose times $t = 0$ on which the variables a_1 and a_2 take their input states, i.e. they are either equal to 1 or equal to 0 for v_ξ , $\xi = 1, 2$. The output time is set to $t = T$.

Consistency. We now have to decide whether a direct transition from states at $t = 0$ to $t = T$ leads to a consistent inverse problem. For the AND and for the OR operator this is feasible, but XOR cannot be realized directly due to the inconsistency of the corresponding image vectors as described in (1.165). The XOR logic in its simplest form tries to map

$$(0, 0) \mapsto 0, \quad (1, 0) \mapsto 1, \quad (0, 1) \mapsto 1, \quad (1, 1) \mapsto 0$$

But the coefficients of the state $(1, 1)$ of v_1 and v_2 are the linear combination of $(1, 0)$ and $(0, 1)$, i.e. we cannot reach 0 as its image coefficient for v_{out} when we map $(1, 0)$ onto 1 and $(0, 1)$ onto 1 as well. We cannot expect to model the nonlinear XOR dynamics based on a linear equation. This observation, called the XOR problem, has played a large role historically, when Minsky and Papert investigated the capabilities of the perceptron [66]. As shown in [75] in a neural field environment the problem has a simple solution by extending the dynamics. In a discrete environment this would correspond to the introduction of multi-layer networks, as it has been carried out for neural networks. Here, we employ two further states v_3 and v_4 , such that v_1, \dots, v_4, v_{out} are linearly independent. Then, we choose $t_1 := T/2$ and decompose the task into the OR map from v_1, v_2 onto v_3 at time t_1 , the AND map from v_1, v_2 onto v_4 at time t_1 , a plain transition from v_3

to v_{out} and an inhibition from v_4 to v_{out} . Our dynamics within a time interval $T/2$ based on five linearly independent states can be written as

$$\begin{array}{ll} (1) (0, 0, 0, 0, 0) \mapsto (0, 0, 0, 0, 0), & (5) (0, 0, 1, 0, 0) \mapsto (0, 0, 0, 0, 1), \\ (2) (1, 0, 0, 0, 0) \mapsto (0, 0, 1, 0, 0), & (6) (0, 0, 0, 1, 0) \mapsto (0, 0, 0, 0, 0), \\ (3) (0, 1, 0, 0, 0) \mapsto (0, 0, 1, 0, 0), & (7) (0, 0, 1, 1, 0) \mapsto (0, 0, 0, 0, 0), \\ (4) (1, 1, 0, 0, 0) \mapsto (0, 0, 1, 1, 0), & \end{array}$$

The time dynamics is now realized by an application of the rules in the interval $[0, t_1]$ first and then again in the interval $[t_1, T]$ based on the order parameter approach (1.167). We obtain a set of consistent pairs of states and their neuro-dynamical images. The result is a set of four dynamical activation patterns or neural fields $u_1(x, t), \dots, u_4(x, t)$ representing XOR, which are consistent with a neuro dynamical environment. We solve the inverse problem (1.141) or (1.160), respectively, to construct a kernel $w(x, y)$ which now realizes the neural dynamics on the given states v_1, \dots, v_4 and v_{out} .

We show one possible realization with states which are Gaussian functions on the domain $D = [0, 10] \times [-5, 5]$

$$v_k(x) = e^{-\tau|x-z^{(k)}|^2}, \quad \text{for } k = 1, 2, \quad v_{out}(x) = e^{-\tau|x-z^{(out)}|^2}, \quad (1.168)$$

for $x \in D$ with points $z^{(1)} = (0.2, 3)$, $z^{(2)} = (0.2, -3)$ and $z^{(out)} = (9, 0)$ and $\tau = 2$ as well as

$$\begin{aligned} v_3(x) &= 0.5 \cdot (1 + \sin(2x_2) \cdot \sin(2x_1)) \cdot e^{-\tilde{\tau}|x_1-z_1^{(3)}|^2}, \\ v_4(x) &= 0.5 \cdot (1 + \cos(2x_2) \cdot \cos(2x_1)) \cdot e^{-\tilde{\tau}|x_1-z_2^{(3)}|^2}, \end{aligned} \quad (1.169)$$

for $x \in D$ with $\tilde{\tau} = 0.3$ and $z_1^{(3)} = 5$. The time interval has been selected as $[0, T]$ with $T = 10$ and a choice of the order parameter functions (1.166) with $\sigma = 2$.

The states v_1 , v_2 and v_{out} are *localized* states in neural space. The states v_3 and v_4 are distributed space over some region of the neural space, which we have kept separate from the support of v_1 , v_2 and v_{out} here for better visibility. This demonstrates that we can solve the inverse neural field problem with macrostates which are distributed over a large region of our neural space. We are not bound to localized neural pulses, as it is often used for neural modelling.

For the numerical tests we have chosen $n_t = 40$ time steps for the training and a discretization of $n_1 \times n_2 = 55 \times 56$ points in our neural space. Then, the calculations can be carried out in some seconds in MATLAB®. We show three time slices of the both the training pulses as well as their reconstruction via the neural field equation (1.138) in Fig. 1.5. The plots are taken at the times $t = 6 \cdot \frac{T}{n_t} = 1.53$, $t = 18 \cdot \frac{T}{n_t} = 4.62$ and $t = 36 \cdot \frac{T}{n_t} = 9.23$.

To avoid an *inverse crime* (cf. [18]), i.e. the testing of an inversion method on the same grid which was used to generate the data, here we have chosen a finer time-grid with $n_t = 80$ for the neural field equation for carrying out the simulation after construction of the kernel w .

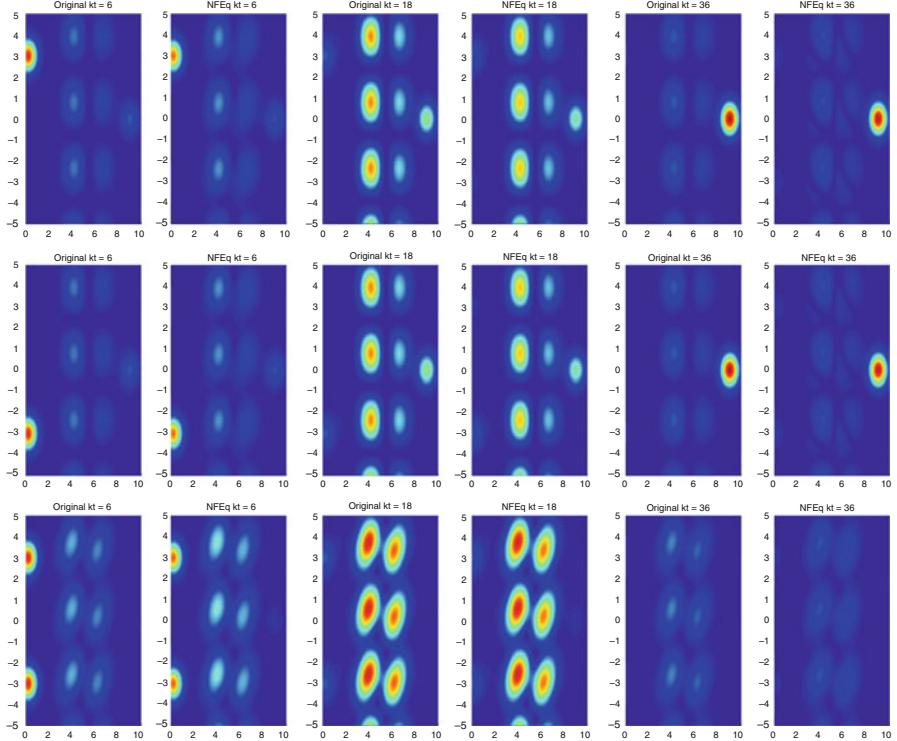


Fig. 1.5 We show three time slices of two pulses which enter the activation space at time $t = 0$ at states v_1 and v_2 . They lead to the excitation of the states v_3 and v_4 and then v_{out} is excited or inhibited. The images show the original pattern (*left*) and the pattern which is generated by the neural field equation (*right*), where we used a time discretization which was finer than the training pattern to avoid an inverse crime

Acknowledgements SC would like to thank Ruth Smith for a careful reading of the material in this chapter. PbG gratefully acknowledges support from a DFG Heisenberg fellowship (GR 3711/1-2). The authors also thank EPSRC and the Centre for Cognitive Neuroscience and Neurodynamics (CINN) of the University of Reading, UK, for supporting the discussion on this tutorial by funding the *2nd International Conference on Neural Field Theory* in April 2012 in Reading.

Appendix 1

The Green's function for the infinite cable equation satisfies the PDE

$$G_t = -\frac{G}{\tau} + DG_{xx}, \quad G(x, 0) = \delta(x). \quad (1.170)$$

If we introduce the Fourier transform

$$G(k, t) = \int_{-\infty}^{\infty} e^{-ikx} G(x, t) dx, \quad G(x, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} G(k, t) dk \quad (1.171)$$

then we may obtain the ordinary differential equation

$$G_t(k, t) = -\epsilon(k)G(k, t), \quad G(k, 0) = 1, \quad \epsilon(k) = \frac{1}{\tau} + Dk^2, \quad (1.172)$$

with solution $G(k, t) = G(k, 0) \exp(-\epsilon(k)t)$. Performing an inverse Fourier transform gives

$$G(x, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ikx} e^{-\epsilon(k)t} = e^{-t/\tau} e^{-x^2/4Dt} \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-Dt[k+ix/2Dt]^2} dk \quad (1.173)$$

$$= \frac{1}{\sqrt{4\pi Dt}} e^{-t/\tau} e^{-x^2/(4Dt)}, \quad (1.174)$$

where we complete the square in (1.173) and use the fact that $\int_{-\infty}^{\infty} \exp(-x^2) dx = \sqrt{\pi}$.

Appendix 2

Introducing the Laplace transform of $G(x, t)$ as $G(x, \lambda)$, where

$$G(x, \lambda) = \int_0^{\infty} ds e^{-\lambda s} G(x, s), \quad (1.175)$$

means that we may transform (1.85) to obtain

$$[\gamma^2(\lambda) - d_{xx}] G(x, \lambda) = \delta(x)/D, \quad \gamma^2(\lambda) = (1/\tau + \lambda)/D. \quad (1.176)$$

This second order ordinary differential equation has the solution

$$G(x, \lambda) = \frac{e^{-\gamma(\lambda)|x|}}{2D\gamma(\lambda)}. \quad (1.177)$$

Appendix 3

Introduce periodic functions $U(x)$ and $V(x)$ and write using Fourier series as $U(x) = \sum_{n=-\infty}^{\infty} U_n e^{2\pi i n x / \Lambda}$ and $V(x) = \sum_{n=-\infty}^{\infty} V_n e^{2\pi i n x / \Lambda}$, where $\Lambda = 2\pi/p_c$.

In this case,

$$\begin{aligned}
 [w \otimes V](x) &= \int_{-\infty}^{\infty} dy w(|y|) V(x-y) \\
 &= \sum_n \int_{-\infty}^{\infty} dy w(|y|) V_n e^{2\pi i n(x-y)/\Lambda} \\
 &= \sum_n V_n e^{2\pi i n x / \Lambda} w_n, \quad w_n = \int_{-\infty}^{\infty} dy w(|y|) e^{-2\pi i n y / \Lambda}. \quad (1.178)
 \end{aligned}$$

Hence, we may now calculate

$$\begin{aligned}
 \langle U, w \otimes V \rangle &= \frac{1}{\Lambda} \int_0^\Lambda \sum_n U_n^* e^{-2\pi i n x / \Lambda} \sum_m V_m w_m e^{2\pi i m x / \Lambda} \\
 &= \sum_m U_m^* V_m w_m, \quad (1.179)
 \end{aligned}$$

where we make use of the result that $\int_0^\Lambda e^{2\pi i x(m-n)/\Lambda} / \Lambda = \delta_{n,m}$. Similarly we find

$$\langle w \otimes U, V \rangle = \sum_m U_m^* V_m w_m^* = \langle U, w \otimes V \rangle, \quad (1.180)$$

where we make use of the fact that $w_m = w_m^*$. From (1.56) we have that

$$\langle U, \mathcal{L}V \rangle = -\langle U, V \rangle + \beta_c \langle U, w \otimes V \rangle = -\langle U, V \rangle + \beta_c \langle w \otimes U, V \rangle = \langle \mathcal{L}U, V \rangle. \quad (1.181)$$

Hence \mathcal{L} is self-adjoint.

Appendix 4

For $w(r) = \exp(-r/\sigma)/(2\pi)$ we may calculate Eq. (1.94) as

$$\begin{aligned}
 G(k, \omega) &= (2\pi)^{-1} \int_{-\infty}^{\infty} ds \int_{\mathbb{R}^2} d\mathbf{r} e^{-|\mathbf{r}|/\sigma} \delta(s - |\mathbf{r}|/v) e^{-i(\mathbf{k}\cdot\mathbf{r} + \omega s)}, \\
 &= (2\pi)^{-1} \int_0^{2\pi} \int_0^{\infty} e^{-ikr \cos \theta} e^{-Ar} r dr d\theta \\
 &= -(2\pi)^{-1} \frac{\partial}{\partial A} \int_0^{2\pi} \frac{1}{A + ik \cos \theta} d\theta, \quad (1.182)
 \end{aligned}$$

where $A(\omega) = 1/\sigma + i\omega/v$. This may be evaluated using a (circular) contour integral in the complex plane to give

$$G(k, \omega) = \frac{A(\omega)}{(A(\omega)^2 + k^2)^{3/2}}. \quad (1.183)$$

References

1. Abbott, L.F., Fahri, E., Gutmann, S.: The path integral for dendritic trees. *Biol. Cybern.* **66**, 49–60 (1991)
2. Amari, S.: Homogeneous nets of neuron-like elements. *Biol. Cybern.* **17**, 211–220 (1975)
3. Amari, S.: Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.* **27**, 77–87 (1977)
4. Beim Graben, P., Pinotsis, D., Saddy, D., Potthast, R.: Language processing with dynamic fields. *Cogn. Neurodyn.* **2**(2), 79–88 (2008)
5. Beim Graben, P., Potthast, R.: Inverse problems in dynamic cognitive modeling. *Chaos* **19**(1), 015103 (2009)
6. Beim Graben, P., Potthast, R.: A dynamic field account to language-related brain potentials. In: Rabinovich, M., et al. (eds.) *Principles of Brain Dynamics: Global State Interactions*. MIT, Cambridge (2012)
7. Ben-Yishai, R., Bar-Or, L., Sompolinsky, H.: Theory of orientation tuning in visual cortex. *Proc. Natl. Acad. Sci. USA* **92**, 3844–3848 (1995)
8. Berger, H.: Über das Elektroenzephalogramm des Menschen. *Archiv für Psychiatrie* **87**, 527–570 (1929)
9. Beurle, R.L.: Properties of a mass of cells capable of regenerating pulses. *Philos. Trans. R. Soc. Lond. B* **240**, 55–94 (1956)
10. Bojak, I., Liley, D.T.J.: Modeling the effects of anesthesia on the electroencephalogram. *Phys. Rev. E* **71**, 041,902 (2005)
11. Brackley, C.A., Turner, M.S.: Random fluctuations of the firing rate function in a continuum neural field model. *Phys. Rev. E* **75**, 041,913 (2007)
12. Bressloff, P.C.: New mechanism for neural pattern formation. *Phys. Rev. Lett.* **76**, 4644–4647 (1996)
13. Bressloff, P.C.: Traveling fronts and wave propagation failure in an inhomogeneous neural network. *Phys. D* **155**, 83–100 (2001)
14. Bressloff, P.C.: Spatiotemporal dynamics of continuum neural fields. *J. Phys. A* **45**, 033,001 (2012)
15. Bressloff, P.C., Coombes, S.: Physics of the extended neuron. *Int. J. Mod. Phys. B* **11**, 2343–2392 (1997)
16. Bressloff, P.C., Cowan, J.D., Golubitsky, M., Thomas, P.J., Wiener, M.: Geometric visual hallucinations, Euclidean symmetry and the functional architecture of striate cortex. *Philos. Trans. R. Soc. Lond. B* **40**, 299–330 (2001)
17. Bressloff, P.C., Souza, B.D.: Neural pattern formation in networks with dendritic structure. *Phys. D* **115**, 124–144 (1998)
18. Colton, D., Kress, R.: *Inverse Acoustic and Electromagnetic Scattering Theory*. Applied Mathematical Sciences, vol. 93. Springer, Berlin (1998)
19. Coombes, S.: Waves, bumps, and patterns in neural field theories. *Biol. Cybern.* **93**, 91–108 (2005)
20. Coombes, S.: Large-scale neural dynamics: simple and complex. *NeuroImage* **52**, 731–739 (2010)

21. Coombes, S., Lord, G.J., Owen, M.R.: Waves and bumps in neuronal networks with axo-dendritic synaptic interactions. *Phys. D* **178**, 219–241 (2003)
22. Coombes, S., Owen, M.R.: Evans functions for integral neural field equations with Heaviside firing rate function. *SIAM J. Appl. Dyn. Syst.* **34**, 574–600 (2004)
23. Coombes, S., Schmidt, H.: Neural fields with sigmoidal firing rates: approximate solutions. *Discret. Contin. Dyn. Syst. Ser. A* **28**, 1369–1379 (2010)
24. Coombes, S., Schmidt, H., Bojak, I.: Interface dynamics in planar neural field models. *J. Math. Neurosci.* **2**(9), 1–27 (2012)
25. Coombes, S., Timofeeva, Y., Svensson, C.M., Lord, G.J., Josic, K., Cox, S.J., Colbert, C.M.: Branching dendrites with resonant membrane: a “sum-over-trips” approach. *Biol. Cybern.* **97**, 137–149 (2007)
26. Coombes, S., Venkov, N.A., Shiu, L., Bojak, I., Liley, D.T.J., Laing, C.R.: Modeling electrocortical activity through improved local approximations of integral neural field equations. *Phys. Rev. E* **76**, 051901 (2007)
27. da Silva, F.L., Rotterdam, A.V.: Biophysical aspects of EEG and Magnetoencephalogram generation. In: *Electroencephalography: Basic Principles, Clinical Applications and Related Fields*, pp. 107–126. Lippincott Williams & Wilkins, Philadelphia (2005)
28. Daunizeau, J., Kiebel, S.J., Friston, K.J.: Dynamic causal modelling of distributed electromagnetic responses. *NeuroImage* **47**, 590–601 (2009)
29. Engl, H.W., Hanke, M., Neubauer, A.: *Regularization of Inverse Problems. Mathematics and its Applications*, vol. 375. Kluwer, Dordrecht (1996)
30. Erlhagen, W., Bicho, E.: The dynamic neural field approach to cognitive robotics. *J. Neural Eng.* **3**, R36–R54 (2006)
31. Ermentrout, B.: Stripes or spots? Nonlinear effects in bifurcation of reaction-diffusion equations on the square. *Proc. R. Soc. Lond. Ser. A* **434**, 413–417 (1991)
32. Ermentrout, G.B.: Neural nets as spatio-temporal pattern forming systems. *Rep. Prog. Phys.* **61**, 353–430 (1998)
33. Ermentrout, G.B., Cowan, J.D.: A mathematical theory of visual hallucination patterns. *Biol. Cybern.* **34**, 137–150 (1979)
34. Ermentrout, G.B., McLeod, J.B.: Existence and uniqueness of travelling waves for a neural network. *Proc. R. Soc. Edinb.* **123A**, 461–478 (1993)
35. Evans, J.: Nerve axon equations: IV the stable and unstable impulse. *Indiana Univ. Math. J.* **24**, 1169–1190 (1975)
36. Faugeras, O., Grimbert, F., Slotine, J.J.: Absolute stability and complete synchronization in a class of neural fields models. *SIAM J. Appl. Math.* **69**, 205–250 (2008)
37. Folias, S.E., Bressloff, P.C.: Breathing pulses in an excitatory neural network. *SIAM J. Appl. Dyn. Syst.* **3**, 378–407 (2004)
38. Folias, S.E., Bressloff, P.C.: Breathers in two-dimensional neural media. *Phys. Rev. Lett.* **95**(1–4), 208,107 (2005)
39. Geise, M.A.: *Neural Field Theory for Motion Perception*. Kluwer, Norwell (1999)
40. Griffith, J.S.: A field theory of neural nets: I: derivation of field equations. *Bull. Math. Biophys.* **25**, 111–120 (1963)
41. Griffith, J.S.: A field theory of neural nets: II: properties of field equations. *Bull. Math. Biophys.* **27**, 187–195 (1965)
42. Griffith, J.S.: *A View of the Brain*. Oxford University Press, Oxford (1967)
43. Griffith, J.S.: *Mathematical Neurobiology*. Academic, London/New York (1971)
44. Hadamard, J.: *Lectures on Cauchy’s Problem in Linear Partial Differential Equations*. Yale University Press, New Haven (1923)
45. Haken, H.: *Principles of Brain Functioning*. Springer Series in Synergetics, vol. 67. Springer, Berlin (1996)
46. Hebb, D.O.: *The Organization of Behavior*. Wiley, New York (1949). Partly reprinted in Anderson, J.A., Rosenfeld, E., pp. 45ff. (1988)
47. Hoyle, R.: *Pattern Formation: An Introduction to Methods*. Cambridge University Press, Cambridge (2006)

48. Hutt, A.: Effects of nonlocal feedback on traveling fronts in neural fields subject to transmission delay. *Phys. Rev. E* **60**(1–4), 052,902 (2004)
49. Jirsa, V.K., Haken, H.: Field theory of electromagnetic brain activity. *Phys. Rev. Lett.* **77**, 960–963 (1996)
50. Jirsa, V.K., Haken, H.: A derivation of a macroscopic field theory of the brain from the quasi-microscopic neural dynamics. *Phys. D* **99**, 503–526 (1997)
51. Jirsa, V.K.V., Jantzen, K.J., Fuchs, A., Kelso, J.A.S.: Neural field dynamics on the folded three-dimensional cortical sheet and its forward EEG and MEG. In: *Information Processing in Medical Imaging*, pp. 286–299. Springer, Berlin (2001)
52. Jirsa, V.K., Jantzen, K.J., Fuchs, A., Kelso, J.A.S.: Spatiotemporal forward solution of the EEG and MEG using network modeling. *IEEE Trans. Med. Imaging* **21**(5), 493–504 (2002)
53. Kelso, J.A.S., Bressler, S.L., Buchanan, S., Deguzman, G.C., Ding, M., Fuchs, A., Holroyd, T.: A phase-transition in human brain and behaviour. *Phys. Lett. A* **169**, 134–144 (1992)
54. Kilpatrick, Z.P., Bressloff, P.C.: Binocular rivalry in a competitive neural network with synaptic depression. *SIAM J. Appl. Dyn. Syst.* **9**, 1303–1347 (2010)
55. Kilpatrick, Z.P., Bressloff, P.C.: Effects of synaptic depression and adaptation on spatiotemporal dynamics of an excitatory neuronal network. *Phys. D* **239**, 547–560 (2010)
56. Kilpatrick, Z.P., Bressloff, P.C.: Spatially structured oscillations in a two-dimensional excitatory neuronal network with synaptic depression. *J. Comput. Neurosci.* **28**, 193–209 (2010)
57. Kirsch, A.: *An Introduction to the Mathematical Theory of Inverse Problems*. Applied Mathematical Sciences, vol. 120. Springer, New York (1996)
58. Kishimoto, K., Amari, S.: Existence and stability of local excitations in homogeneous neural fields. *J. Math. Biol.* **7**, 303–318 (1979)
59. Kress, R.: *Linear Integral Equations*. Springer, Berlin (1989)
60. Laing, C.R.: Spiral waves in nonlocal equations. *SIAM J. Appl. Dyn. Syst.* **4**, 588–606 (2005)
61. Laing, C.R., Troy, W.C.: PDE methods for nonlocal models. *SIAM J. Appl. Dyn. Syst.* **2**, 487–516 (2003)
62. Laing, C.R., Troy, W.C.: Two bump solutions of Amari-type models of working memory. *Phys. D* **178**, 190–218 (2003)
63. Laing, C.R., Troy, W.C., Gutkin, B., Ermentrout, G.B.: Multiple bumps in a neuronal model of working memory. *SIAM J. Appl. Math.* **63**, 62–97 (2002)
64. Liley, D.T.J., Cadusch, P.J., Dafilis, M.P.: A spatially continuous mean field theory of electrocortical activity. *Netw. Comput. Neural Syst.* **13**, 67–113 (2002)
65. Liley, D.T.J., Foster, B.L., Bojak, I.: A mesoscopic modelling approach to anaesthetic action on brain electrical activity. In: *Sleep and Anesthesia*, pp. 139–166. Springer, New York (2011)
66. Minsky, M., Papert, S.: *Perceptrons*. MIT, Cambridge (1969). Partly reprinted in Anderson, J.A., Rosenfeld, E., pp. 161ff (1988)
67. Nunez, P.L.: The brain wave equation: a model for the EEG. *Math. Biosci.* **21**, 279–297 (1974)
68. Nunez, P.L.: *Neocortical Dynamics and Human EEG Rhythms*. Oxford University Press, New York (1995)
69. Oleynik, A., Posnov, A., Wyller, J.: Iterative schemes for bump solutions in a neural field model. *J. Differ. Equ. Dyn. Syst.* doi:10.1007/s12591-013-0191-5 (2013)
70. Owen, M.R., Laing, C.R., Coombes, S.: Bumps and rings in a two-dimensional neural field: splitting and rotational instabilities. *New J. Phys.* **9**, 378 (2007)
71. Pettersen, K.H., Einevoll, G.T.: Amplitude variability and extracellular low-pass filtering of neuronal spikes. *Biophys. J.* **94**, 784–802 (2008)
72. Pinto, D.J., Ermentrout, G.B.: Spatially structured activity in synaptically coupled neuronal networks: I. Travelling fronts and pulses. *SIAM J. Appl. Math.* **62**, 206–225 (2001)
73. Pinto, D.J., Ermentrout, G.B.: Spatially structured activity in synaptically coupled neuronal networks: II. Lateral inhibition and standing pulses. *SIAM J. Appl. Math.* **62**, 226–243 (2001)
74. Potthast, R.: *Point Sources and Multipoles in Inverse Scattering Theory*. Chapman & Hall/CRC Research Notes in Mathematics, vol. 427. Chapman & Hall/CRC, Boca Raton (2001)
75. Potthast, R., Beim Graben, P.: Inverse problems in neural field theory. *SIAM J. Appl. Dyn. Syst.* **8**(4), 1405–1433 (2009)

76. Rabinovich, M.I., Huerta, R., Varona, P., Afraimovich, V.S.: Transient cognitive dynamics, metastability, and decision making. *PLoS Comput. Biol.* **4**(5), e1000,072 (2008)
77. Robinson, P.A., Rennie, C.J., Wright, J.J., Bahramali, H., Gordon, E., Rowe, D.L.: Prediction of electroencephalographic spectra from neurophysiology. *Phys. Rev. E* **63**, 021,903 (2001)
78. Rule, M., Stoffregen, M., Ermentrout, B.: A model for the origin and properties of flicker-induced geometric phosphenes. *PLoS Comput. Biol.* **7**(9), e1002,158 (2011)
79. Sandstede, B.: Stability of travelling waves. In: *Handbook of Dynamical Systems II*, pp. 983–1055. Elsevier, Amsterdam (2002)
80. Schmidt, H., Hutt, A., Schimansky-Geier, L.: Wave fronts in inhomogeneous neural field models. *Phys. D* **238**, 1101–1112 (2009)
81. Schönér, G., Dineva, E.: Dynamic instabilities as mechanisms for emergence. *Dev. Sci.* **10**, 69–74 (2007)
82. Segev, I., Rinzel, J., Shepherd, G.M. (eds.): *The theoretical foundations of dendritic function: selected papers of Wilfrid Rall with commentaries*. MIT, Cambridge (1995)
83. Steyn-Ross, M.L., Steyn-Ross, D.A., Sleigh, J.W., Liley, D.T.J.: Theoretical electroencephalogram stationary spectrum for a white-noise-driven cortex: evidence for a general anesthetic-induced phase transition. *Phys. Rev. E* **60**, 7299–7311 (1999)
84. Tass, P.: Cortical pattern formation during visual hallucinations. *J. Biol. Phys.* **21**, 177–210 (1995)
85. Taylor, J.G.: Neural ‘bubble’ dynamics in two dimensions: foundations. *Biol. Cybern.* **80**, 393–409 (1999)
86. Tuckwell, H.C.: *Introduction to Theoretical Neurobiology: Linear Cable Theory and Dendritic Structure*, vol. 1. Cambridge University Press, Cambridge (1988)
87. Venkov, N.A., Coombes, S., Matthews, P.C.: Dynamic instabilities in scalar neural field equations with space-dependent delays. *Phys. D* **232**, 1–15 (2007)
88. Wilson, H.R., Cowan, J.D.: Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.* **12**, 1–24 (1972)
89. Wilson, H.R., Cowan, J.D.: A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* **13**, 55–80 (1973)

Part I

Theory of Neural Fields

Chapter 2

A Personal Account of the Development of the Field Theory of Large-Scale Brain Activity from 1945 Onward

Jack Cowan

Ah, but a man's reach should exceed his grasp, Or what's a heaven for?

Robert Browning (1855)

Abstract In this paper I give my personal perspective on the development of a field theory of large-scale brain activity. I review early work by Pitts, Wiener, Beurle and others, and give an account of the development of the mean-field Wilson-Cowan equations. I then outline my reasons for trying to develop a stochastic version of these equations, and recall the steps leading to the discovery that one can use field theory and the van Kampen system-size expansion of a neural master equation to obtain stochastic Wilson-Cowan equations. I then describe how stochastic neural field theory led to the discovery that there is a directed percolation phase transition in large-scale brain activity, and how the stochastic Wilson-Cowan equations can provide insight into many aspects of large-scale brain activity, such as the generation of fluctuation-driven avalanches and oscillations.

2.1 Introduction

Sometime between 1945 and 1948 Walter Pitts (1923–1969) wrote a rough draft of a paper entitled “Statistical mechanics of the nervous system” (Fig. 2.1).

He never published anything on this topic, but in 1961 he loaned me a draft to look at and to hand-copy. No one else except Norbert Wiener (1894–1964), and perhaps Warren McCulloch (1898–1969) and Jerry Lettvin (1920–), has ever seen this draft. It is written using the notation in the original 1943 McCulloch-Pitts

J. Cowan (✉)

Department of Mathematics, University of Chicago, 5734 S. University Ave., Chicago, IL 60637 USA

Fig. 2.1 Walter Pitts c. 1946



paper [50], and used combinatorics and probabilistic reasoning to compute statistics associated with the moments of large-scale activity. Figure 2.2 shows the opening page of the draft:

The only result I could find in the draft is that there are no equilibrium points in a large random net. (This was probably because Pitts looked only at the linearized net.)

Perhaps it was another draft that Wiener saw, because in 1946 he wrote the following comment on Pitts' work:

Besides the long-distance connections of the brain, made by tracts of white matter, there are also the random connections of short fibers between nearby cells. The anatomical picture of the cortex suggests that we may usefully employ statistical methods in the study of its function. This work has been taken up brilliantly by Walter Pitts. He finds that, under many conditions, the approximation to the activity of the cortex should not be made on a basis of rest and the zero basal activity. Proceeding from this, he has developed the wave equations of the cortex. These promise to give an interpretation of the electroencephalogram. He has also suggested the exploration of the connections between the different parts of the cortex by suitable stimulation of regions and the observance of the frequency-response at several points.

In 1964 when I was a visitor at the Electrical Engineering Department at Imperial College, London (supported by a contract with the US Department of the Navy, Office of Naval Research) Wiener and I decided to work together on these problems, but unfortunately he died before we got going, so I carried on alone (Fig. 2.3).

I had started earlier, as part of my tenure as a research assistant in MIT's Research Laboratory of Electronics in the Neurophysiology Group run by McCulloch, Pitts, and Lettvin. During conversations with McCulloch and Pitts, and Wiener in 1962, I asked McCulloch and Pitts what kind of mathematics they thought was needed to represent large-scale brain activity. McCulloch was happy with the McCulloch-Pitts model, but to my surprise Pitts told me to look for differential equations that could be treated analytically. Wiener on the other hand suggested I look at a stochastic treatment involving path integrals, about which I had listened to in a series of lectures Wiener gave several years earlier [77]. Unfortunately, at the time, I didn't really understand Wiener's lectures or remarks. It took me the best part of 40 years to

Statistical Mechanics of the Nervous System

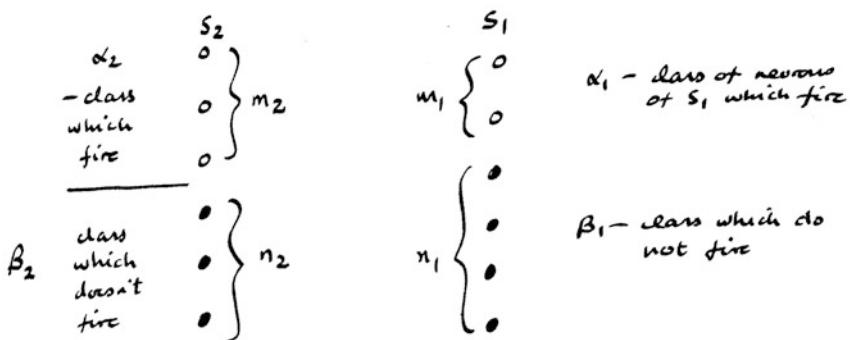
Walter H. Pitts
 Massachusetts Institute of Technology
 c. 1945-1948
 (rewritten by J. D. Cowan Sept 1961)

I. Assumptions of Analysis

1. That all thresholds are unity
2. absolute inhibition

In general there will be 3^N sets of connections of N neurons into themselves, i.e., 3^N realizable transformations of the manifold of states into itself.

What proportion will take a given state s into a state s' ?

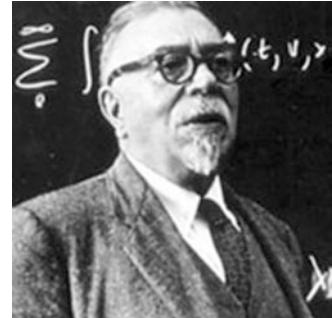


Suppose connections independent of one another: s_i may be connected with s_j either

- (a) by nothing (?) with a probability $P_0 = P$
- (b) by an excitatory axon and no inhibiting
— with a prob. P_+

Fig. 2.2 Facsimile of the first page of Pitts' unpublished manuscript

Fig. 2.3 Norbert Wiener
c. 1958



fully understand the deep insights that Wiener tried to convey to me! So I followed up Pitts' suggestion which I did understand! In line with this the first problem I confronted was that, as Pitts had understood, the McCulloch-Pitts neuron model is too simple a representation of realistic neural properties. The leaky “integrate-and-fire” model that succeeded it [72], is a much more realistic model, however neither model generates analytically tractable equations.

2.2 Lotka-Volterra Dynamics

Instead I used an analogy with Lotka-Volterra ecological dynamics. This dynamics was explored by [42]. Kerner started from the Lotka-Volterra equations [49, 75], which can be written in the form:

$$\frac{dN_r}{dt} = \epsilon_r N_r + \frac{1}{\beta_r} \sum_s \alpha_{sr} N_s N_r \quad (2.1)$$

where $N_r(t)$ represents the number of individuals of the r th species at time t , and where the sum runs over $2n$ interacting species, of whom n are predators, and the remaining n are prey. This implies that for predators $\epsilon_r < 0$, whereas for prey $\epsilon_r > 0$, and also that

$$\alpha_{sr} + \alpha_{rs} = 0, \quad \alpha_{rr} = 0 \quad (2.2)$$

The stationary states q_r of this system of equations satisfy the condition

$$\epsilon_r + \frac{1}{\beta_r} \sum_s \alpha_{sr} q_s = 0 \quad (2.3)$$

Kerner showed that if one introduces the new variable v_r such that

$$N_r = q_r e^{v_r} \quad (2.4)$$

then Eq. (2.1) can be rewritten in the form

$$\beta_r \frac{dv_r}{dt} = \sum_s \alpha_{sr} q_s (e^{v_s} - 1) \quad (2.5)$$

from which Kerner then showed, using Eq. (2.2) that

$$\sum_r \tau_r (e^{v_r} - v_r) = G = \text{constant} \quad (2.6)$$

where $\tau_r = \beta_r q_r$, and therefore that Eq. (2.5) can be rewritten in the form

$$\frac{dv_r}{dt} = \left(\sum_s \gamma_{sr} \frac{\partial}{\partial v_s} \right) G, \quad \gamma_{sr} = \frac{\alpha_{sr}}{\beta_s \beta_r} = -\gamma_{rs} \quad (2.7)$$

Thus the Lotka-Volterra equations are a form of Hamilton's equations of classical mechanics in disguise, and G is the Hamiltonian.

It follows that one can treat the network activity with the methods of equilibrium statistical mechanics. Let P_r be the probability that one species will have its v_r in the interval $(v_r, v_r + dv_r)$. Kerner showed that P_r is given by

$$P_r dv_r = \frac{e^{-G_r/\theta}}{Z_r} dv_r \quad (2.8)$$

where θ is the analog of temperature, equal to the mean square fluctuations of v_r about q_r . This probability can be rewritten in terms of the variable $n_r = N_r/q_r$ as

$$P(n_r) dn_r = \frac{n_r^{\alpha \tau_r - 1} e^{-\alpha \tau_r n_r}}{\alpha \tau_r^{-\alpha \tau_r} \Gamma(\alpha \tau_r)} dn_r \quad (2.9)$$

where $\alpha = 1/\theta$. As Kerner notes in his paper, this is the analog of the Maxwell-Boltzmann distribution. In this case $P(n_r)$ is the well-known *gamma*-density defined on the interval $(0, \infty)$.

2.2.1 A Neural Analogy

Of course the Lotka-Volterra equations are an imperfect representation of population dynamics. In particular they lack the self-interaction term $\alpha_{rr} N_r^2$, the Verhulst-Pearl term which essentially dissipates the “energy” G . The dynamics induced by such a term differs greatly from that of the Hamiltonian form described above. However Kerner's analysis of the Lotka-Volterra dynamics suggested to me sometime during the summer of 1963, a way to write down an analogous equation for interacting

populations of excitatory and inhibitory neurons. Thus inhibitory neurons would be analogs of predators and excitatory ones of prey. However without interactions the growth equation is simple exponential growth, which is not a good representation of the neural firing rate/current equation. But the Verhulst-Pearl term can be used to solve this problem. Consider the equation

$$\frac{dN_r}{dt} = \epsilon_r N_r - \alpha_{rr} N_r^2 = \epsilon_r N_r (1 - \delta_r N_r) \quad (2.10)$$

where $\delta_r = \alpha_{rr}/\epsilon_r$. This is a form of the Verhulst-Pearl or *logistic* equation [60, 74]. It can be reduced to the very simple form

$$\frac{dv_r}{dt} = \epsilon_r, \quad v_r = \ln \frac{N_r}{1 - \delta_r N_r} \quad (2.11)$$

with solution

$$N_r(t) = \frac{1}{\delta_r + e^{-v_r(t)}} \quad (2.12)$$

This is one form of the logistic curve. I therefore introduced a logistic variant of the Lotka-Volterra equations as a neural network model., i.e.

$$\frac{dN_r}{dt} = \left(\epsilon_r + \frac{1}{\beta_r} \sum_s \alpha_{sr} N_s \right) N_r (1 - N_r) \quad (2.13)$$

where $N_r(t)$ after appropriate scaling now represents the *rate* at which the r th neuron is activated, and where the constants ϵ_r and β_r are suitably redefined.

The stationary states of this system still follow Eq. (2.3), and Eq. (2.13) reduces to

$$\frac{dv_r}{dt} = \left(\sum_s \gamma_{sr} \frac{\partial}{\partial v_s} \right) G \quad (2.14)$$

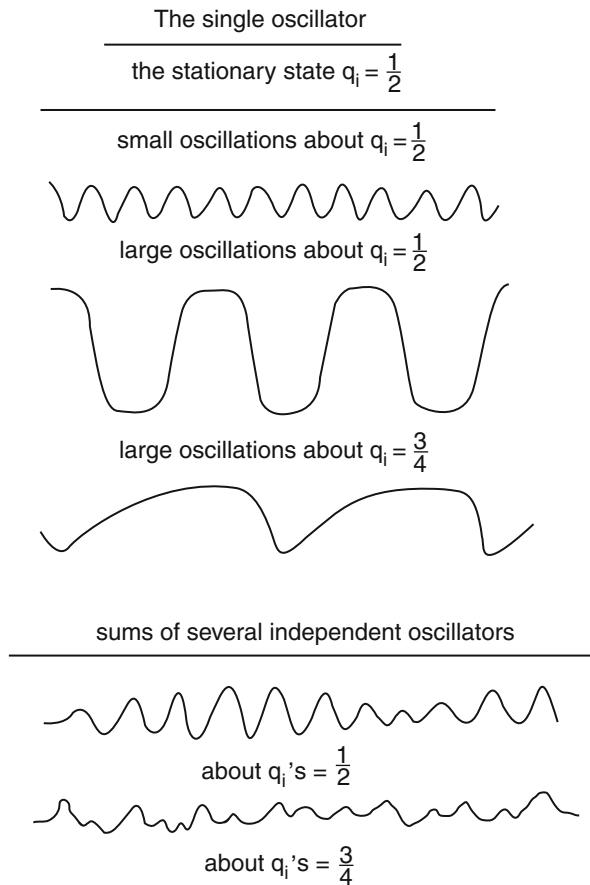
as before, where

$$\frac{N_r}{1 - N_r} = q_r e^{v_r} \quad (2.15)$$

and

$$\sum_r \beta_r [\ln(1 + q_r e^{v_r}) - q_r v_r] = G = \text{constant} \quad (2.16)$$

Fig. 2.4 Marginally stable oscillations in the logistic network (Reprinted from [21])



Thus there is also a Hamiltonian mechanics form of the logistic variant of the Lotka-Volterra equations, and an analog for a network of couple excitatory and inhibitory neurons.

However, note that the coupling coefficients α_{sr} are still skew-symmetric, i.e., they satisfy the conditions of Eq. (2.2). This implies that the eigenvalues of the weight matrix defining the dynamics of perturbations of the stationary states q_r are pure imaginary. Thus the q_r are only marginally stable. The network equations lack a true Verhulst-Pearl “friction” or damping term that would make the stationary states asymptotically stable. The effects of this skew-symmetry are seen in Fig. 2.4.

It will be seen that small oscillations are approximately sinusoidal, but large oscillations become more like periodic square waves. This is the saturating effect of the Verhulst-Pearl term. Although it is not shown in the figure, the effects of noise on the simulation is to convert sinusoidal waves eventually into square waves,

since the dynamics is marginally stable. Of more interest are the traces of the sum of several independent oscillators showing bursts of oscillatory activity that look remarkably like the *alpha* bursts found in EEG recordings.

It now follows from the analogy that one can also write down the analog of the Maxwell-Boltzmann distribution for such a network of coupled neurons. It takes the form given in Eq. (2.8) or in terms of the variable N_r

$$P(N_r)dN_r = \frac{N_r^{\alpha\tau_r-1}(1-N_r)^{\alpha(\beta_r-\tau_r)-1}}{B[\alpha\tau_r, \alpha(\beta_r-\tau_r)]}dN_r \quad (2.17)$$

where $B[x, y]$ is the *beta* function, and $P(N_r)$ is the *beta*-density defined on the interval $(0,1)$. Various “observables” can be deduced from this expression, including the “temperature”, and the effective coupling strengths of the network. An account of this work can be found in [23].

The logistic network I had introduced seemed promising. But of course the fatal flaw in the formulation, was that the dynamics is non-generic in the mathematical sense. Any perturbations of the stationary states will result eventually in large-amplitude fluctuations and oscillations. Sometime in 1967 I therefore recast the logistic model in a slightly different format [22]. I formulated it as the difference equation

$$N_r(t + \tau) = f\left[\sum_{s=1}^N w_{rs}N_s + h_r\right] \quad (2.18)$$

for a network of N neurons coupled with weights w_{rs} external stimuli h_r , and the firing rate function

$$f[x] = (1 + e^{-x})^{-1} \quad (2.19)$$

suitably normalized and scaled.

A differential equation version of this model is

$$\tau \frac{dN_r}{dt} = -N_r + f\left[\sum_{s=1}^N w_{rs}N_s + h_r\right] \quad (2.20)$$

This is the *sigmoid* or “canonical” [41] or firing-rate model. I did not realize at the time that I had invented a neural model that could be used to solve the perceptron training problem [64, 66], or to formalize the analogy between neural networks and magnetic systems suggested by [25, 40, 48] in part because I was completely focused on trying to develop a statistical mechanics of neural networks.

The important point is that the network represented in Eq. (2.20) is much more generic than the skew-symmetric weight network (or for that matter any symmetric

weight network). It has arbitrary weight values, and a decay constant, and therefore the possibility that its stationary states are asymptotically stable to perturbations. However it seemed difficult to analyze (although [5, 6] did obtain some results on the generic statistical mechanics problem.) I therefore started to look at a more statistical approach to the problem.

2.3 Population Dynamics

I had noticed in the course of this work that the structure of Eq. (2.20) is closely related to an equation for neural population activity introduced by [10]. The method introduced by Beurle was very interesting. He focused, not on the activation of individual neurons, but on the fraction of neurons becoming active per unit time in a given volume element of a slab of model cortex consisting of randomly connected neurons. Thus Beurle's formulation used a *continuum* approximation of the network I was considering. This means that the quantity $N_r(t)$ maps into $n(\mathbf{x}, t)$ a firing-rate density. Beurle then introduced the ansatz that

$$n(\mathbf{x}, t + \tau) = q(\mathbf{x}, t)g[n(\mathbf{x}, t)] \quad (2.21)$$

where $q(\mathbf{x}, t)$ is the density of quiescent neurons in the given volume element, and $g[x]$ is the fraction of neurons receiving *exactly* threshold excitation. (There is an implicit assumption that individual neurons are of the integrate-and-fire variety.)

There are three points to note here.

1. By assuming that $n(t + \tau) = q(t)g[n(t)]$ Beurle ignored the effects of fluctuations and correlations on the dynamics. It is not true that q and $g[n]$ are statistically independent quantities, as was first pointed out by [69].
2. The ansatz is incorrect. $g[n]$ should be the fraction of neurons receiving *at least* threshold excitation, as was first noted by [71]. I deduced this myself in 1966 and subsequently confirmed this point in discussions with Violet Cane (1916–2008) a distinguished mathematical statistician then at Newnham College and the Cambridge Statistical Laboratory. But the correct $g[n]$ is essentially a continuum version of $f[N_r]$, and so the corrected version of Beurle's ansatz takes the form

$$\begin{aligned} n(\mathbf{x}, t + \tau) &= q(\mathbf{x}, t) f[n(\mathbf{x}, t)] \\ &= \left(1 - \int_{t-r}^t n(\mathbf{x}, t')\right) \\ &\quad \times f\left[\int_{-\infty}^t dt' \int_{-\infty}^{\infty} d\mathbf{x}' \alpha(t - t')[\beta(\mathbf{x} - \mathbf{x}')n(\mathbf{x}', t') + h(\mathbf{x}, t')]\right] \end{aligned} \quad (2.22)$$

where

$$q(\mathbf{x}, t) = 1 - \int_{t-r}^t n(\mathbf{x}, t') \quad (2.23)$$

where $r = 1\text{ ms}$ is the (absolute) refractory period or width of the action potential, and

$$\alpha(t - t') = \alpha_0 e^{-(t-t')/\tau}, \quad \beta(\mathbf{x} - \mathbf{x}') = b e^{-|\mathbf{x}-\mathbf{x}'|/\sigma} \quad (2.24)$$

are, respectively, the impulse response function and spatially homogeneous weighting function of the continuum model.

3. Beurle's formulation does not explicitly incorporate a role for inhibitory neurons.

Despite these flaws in the formulation, Beurle's ansatz, when corrected seemed to be closely related to the sigmoid model, especially if we rewrite Eq. (2.22) in network form, in which case it becomes

$$n_r(t + \tau) = \left(1 - \int_{t-r}^t n_r(t')\right) f \left[\int_{-\infty}^t dt' \alpha(t - t') \left[\sum_{s=1}^N w_{rs} n_s + h_r(t') \right] \right] \quad (2.25)$$

This was the state of my understanding of the problem when I worked at Imperial College, London from 1962–1967, where my (nominal) supervisor was the Hungarian physicist and inventor Dennis Gabor (1900–1979), who received the 1971 Nobel prize in Physics for his invention of holography. In the fall of 1967 I became Professor and Chairman of the Committee on Mathematical Biology at the University of Chicago, succeeding Nicholas Rashevsky (1899–1984). I was very fortunate to have been mentored at Imperial College initially in 1956–1957 by Arthur Porter (1910–2010), then by Gabor, who had deep insights into many problems in physics, communication and information theory, and cybernetics; and of course I was also very fortunate in being mentored by McCulloch and Pitts at MIT from 1960–1962, and having teachers there like Wiener and Claude Shannon (1916–2001), the founder of Information Theory.

2.3.1 The Wilson-Cowan Equations

In the summer of 1969 Hugh Wilson joined me as a Postdoctoral Fellow. Wilson was a recent graduate of the Chemistry department at the University of Chicago, was very well informed about statistical approaches to chemical and biological problems, was a pleasure to work with and we set to work on this problem. The results of our efforts were published in [79, 80]. Essentially what we did, was to

time coarse-grain Eq. (2.25) [43], and add an explicit inhibitory term, to obtain (in network form)

$$\begin{aligned} \tau \frac{dE_r}{dt} &= -E_r(t) + (1 - rE_r(t)) f_E \left[\sum_{s=1}^M w_{rs}^{EE} E_s - \sum_{s=1}^N w_{rs}^{EI} I_s + h_r^E(t) \right] \\ \tau \frac{dI_r}{dt} &= -I_r(t) + (1 - rI_r(t)) f_I \left[\sum_{s=1}^M w_{rs}^{IE} E_s - \sum_{s=1}^N w_{rs}^{II} I_s + h_r^I(t) \right] \end{aligned} \quad (2.26)$$

or in continuum form

$$\begin{aligned} \tau \frac{\partial E(\mathbf{x}, t)}{\partial t} &= -E(\mathbf{x}, t) + (1 - rE(\mathbf{x}, t)) \\ &\quad \times f_E \left[\int_{-\infty}^{\infty} \rho_E d\mathbf{x}' \beta_{EE}(\mathbf{x} - \mathbf{x}') E(\mathbf{x}', t) \right. \\ &\quad \left. - \int_{-\infty}^{\infty} \rho_I d\mathbf{x}' \beta_{EI}(\mathbf{x} - \mathbf{x}') I(\mathbf{x}', t) + h_E(\mathbf{x}, t) \right] \\ \tau \frac{\partial I(\mathbf{x}, t)}{\partial t} &= -I(\mathbf{x}, t) + (1 - rI(\mathbf{x}, t)) \\ &\quad \times f_I \left[\int_{-\infty}^{\infty} \rho_E d\mathbf{x}' \beta_{IE}(\mathbf{x} - \mathbf{x}') E(\mathbf{x}', t) \right. \\ &\quad \left. - \int_{-\infty}^{\infty} \rho_I d\mathbf{x}' \beta_{II}(\mathbf{x} - \mathbf{x}') I(\mathbf{x}', t) + h_I(\mathbf{x}, t) \right] \end{aligned} \quad (2.27)$$

where M and ρ_E , and N and ρ_I are, respectively, the numbers and packing densities of excitatory and inhibitory cells in the cortical slab. These are the Wilson-Cowan equations.

However, there is an important difference between the continuum version and the network version. In the former version we define $f[x]$ as the fraction of cells receiving at least threshold excitation, per unit time in the appropriate volume element. This can be represented as

$$f[x] = \int_{-\infty}^x P(x_{\text{TH}}) dx_{\text{TH}} = \int_{-\infty}^{\infty} \vartheta[x - x_{\text{TH}}] P(x_{\text{TH}}) dx_{\text{TH}} = \langle \vartheta[x] \rangle \quad (2.28)$$

where $\vartheta[x]$ is the Heaviside step function and $\langle \vartheta[x] \rangle$ is the average of $\vartheta[x]$ over the probability distribution of thresholds $P(x_{\text{TH}})$. This indicates that the Wilson-Cowan equations represent the dynamics of a population of integrate-and-fire neurons

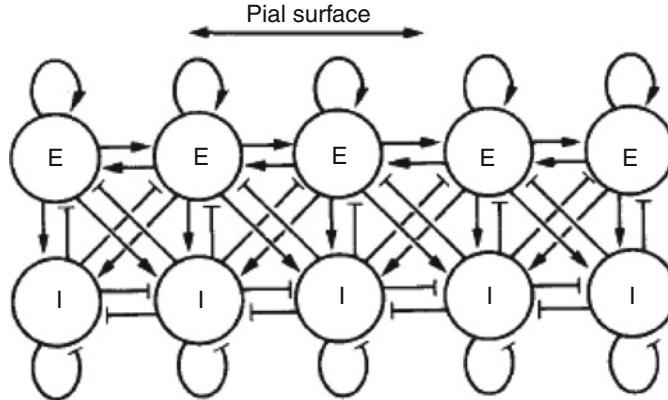


Fig. 2.5 Architecture of an E/I network (Reprinted from [80])

in which there is a random distribution of thresholds, or equivalently, a random distribution of synaptic weights.

The most important new property of the Wilson-Cowan equations was that their dynamics had *attractors*, i.e. equilibria which are asymptotically stable. Figure 2.5 shows the architecture of a typical E/I network, and Fig. 2.6 shows some of these attractors, and their properties.

The first panel (upper-left) of Fig. 2.6 shows the response of a linear array of coupled E/I pairs to an excitatory stimulus of a certain width. The initial response is at the top. It will be seen that after some delay the response shows some edge-enhancement. It turns out that if the parameters of the network are chosen so that the recurrent excitation provided by β_{EE} is not too strong, then the response is transient, whereas if β_{EE} is dominant, then a maintained and localized response occurs, and the activity of the array is *bistable*, as shown in the upper right panel. On the other hand in case the feedback loop characterized by the product $\beta_{EI}\beta_{IE}$ is dominant then the activity exhibits an asymptotically stable *limit-cycle* oscillation that is also localized. The localization property occurs if the range of connectivity specified by the space constants σ_{rs} is such that $\sigma_{IE} > \sigma_{EE}$, i.e., the lateral excitatory to inhibitory range must be greater than the recurrent excitatory to excitatory range. Finally the lower-right panel shows the result of *disinhibition* of the array, e.g., by raising the thresholds for inhibition. The localization property disappears and a local stimulus gives rise to propagating solitary waves of excitation followed by inhibition.

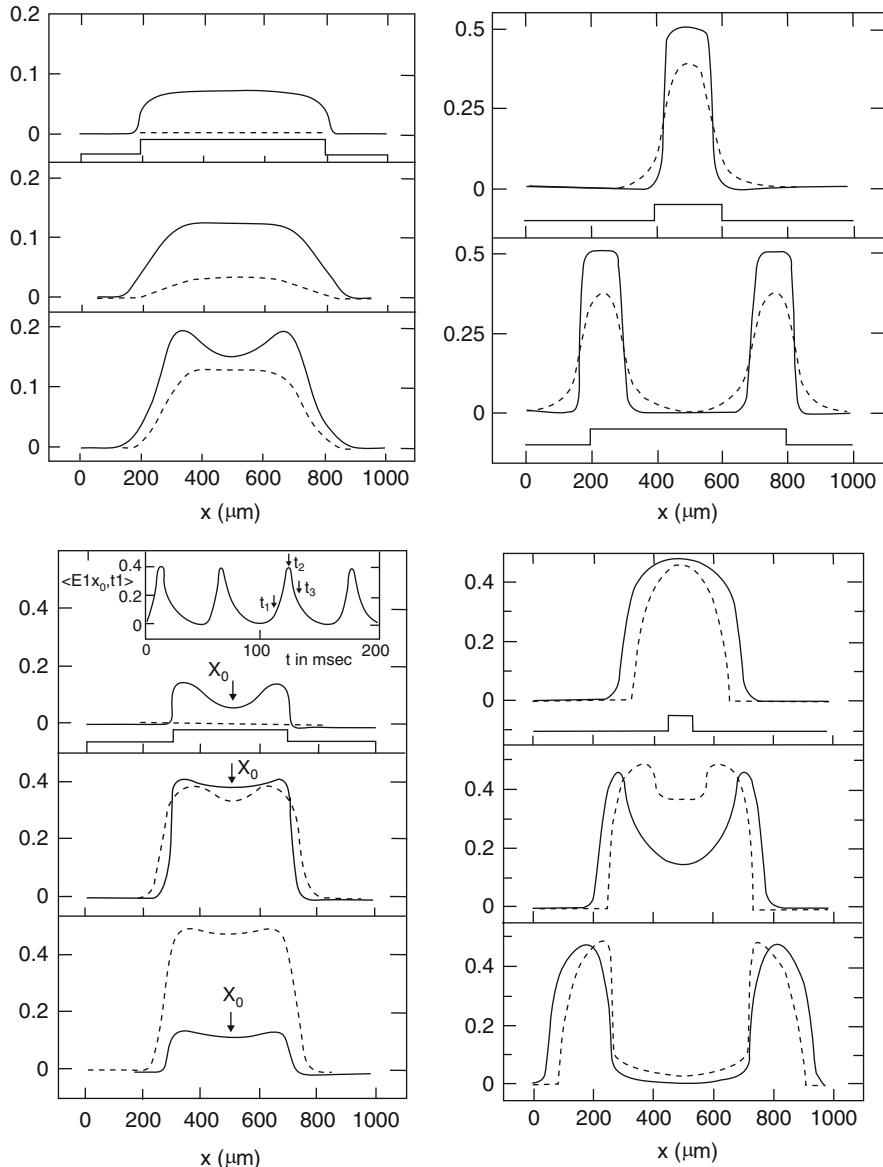


Fig. 2.6 Responses of the E/I network to various excitatory stimuli. Solid curves represent the response of the excitatory population, dashed lines that of the inhibitory population. See text for details (Reprinted from [80])

2.4 A Master Equation Approach

The Wilson-Cowan equations proved to be very influential (see a recent review by [27]). However, there still remained the problem that the Beurle ansatz ignores the effects of correlations and fluctuations, so the Wilson-Cowan equations are *mean field* equations, i.e., the quantities $E(\mathbf{x}, t)$ and $I(\mathbf{x}, t)$, in addition to being time-coarse-grained, are also ensemble averages associated with an underlying probability density of the fractions of activated excitatory and inhibitory neurons in a cortical slab. In 1978 I started to work on this problem. By 1979 I had formulated a “master equation” for the evolution of such a probability density, but refrained from submitting anything for publication then, since I wasn’t completely satisfied with the results.

I considered first the problem of a single spatially homogeneous network of N excitatory cells. Let $P_n(t)$ be the probability that a fraction n/N is active at time t . I then formulated the master equation

$$\begin{aligned} \frac{dP_n(t)}{dt} = & \alpha[(n+1)P_{n+1} - nP_n] + (N-n+1)f[s(n-1)]P_{n-1} \\ & -(N-n)f[s(n)]P_n \end{aligned} \quad (2.29)$$

where α is the rate at which activated neurons become quiescent, and $s(n)$ is the total current or excitation driving each neuron in the population. I assumed in this simplified model that each neuron receives a signal weighted by w_0/N from each of N other neurons in the population, so that

$$s(n) = w_0n + h \quad (2.30)$$

where h is an external current.

I then used a truncation of the Kramers-Moyal expansion of the master equation [45, 55] to derive a Fokker-Planck equation for the evolution of the fraction $e(t) = n(t)/N$ of active neurons, i.e.,

$$\frac{\partial}{\partial t} P(e, t) = -\frac{\partial}{\partial e} [K_1(e)P(e, t)] + \frac{1}{2N} \frac{\partial^2}{\partial e^2} [K_2(e)P(e, t)] \quad (2.31)$$

where $P(e, t) = NP_n(t)$ and

$$K_1(e) = -\alpha e + (1-e)f[s(e)], \quad K_2(e) = \alpha e + (1-e)f[s(e)] \quad (2.32)$$

It was then possible to write down a spatial generalization of Eq.(2.31) as a functional Fokker-Planck equation taking the form

$$\frac{\partial}{\partial t} P(e(\mathbf{x}), t) = \int d\mathbf{x} \left[-\frac{\delta}{\delta e(\mathbf{x})} [K_1(e(\mathbf{x})) P(e(\mathbf{x}), t)] + \frac{1}{2N} \frac{\delta^2}{\delta e(\mathbf{x})^2} [K_2(e(\mathbf{x})) P(e(\mathbf{x}), t)] \right] \quad (2.33)$$

where

$$\begin{aligned} K_1(e(\mathbf{x})) &= -\alpha e(\mathbf{x}) + (1 - e(\mathbf{x})) f[s(e(\mathbf{x}))], \\ K_2(e(\mathbf{x})) &= \alpha e(\mathbf{x}) + (1 - e(\mathbf{x})) f[s(e(\mathbf{x}))] \end{aligned} \quad (2.34)$$

I was then able to show that the first moment equations for $e(\mathbf{x})$ satisfied the Wilson-Cowan equations. However the Wilson-Cowan equations do not contain any second moment terms—they are *closed*. I therefore decided, at the time (1979), not to pursue the Fokker-Planck approach. In retrospect this was a mistake. In recent work with Benayoun, van Drongelen and Wallace (to be discussed later) it is shown that a different Fokker-Planck equation, obtained from the system-size expansion of the master equation [59, 73] does give results that go well beyond those obtained with the Wilson-Cowan equations.

2.5 Field Theory

I therefore started to look for another way to analyze a generalization of the master equation to space-dependent neural interactions. At this time I started spending my summers at the Los Alamos National Laboratories, and ended up in the summer of 1983, introducing two physicists there, David Sharp and Alan Lapedes, to the problem of analyzing neural networks. Later that year Lapedes found two articles by [29, 30] which showed how to analyze classical many-particle systems and diffusion controlled chemical reactions, using the operator techniques of quantum mechanics and quantum field theory. Later we also found another relevant article by [37]. I immediately recognized that there should be a way to analyze neural networks with such techniques. However, it was not until late summer 1984 that I found a way to correctly apply these operator techniques to neural networks.

Consider once again the master equation given by Eq.(2.29), extended to the spatially inhomogeneous case. Let n_r/N_r be the fraction of active cells at time t in the r th population of N_r cells, and let $P[\mathbf{n}, t]$ be the probability of the configuration $\mathbf{n} = \{n_1, n_2, \dots, n_r, \dots, n_\Omega\}$ existing at time t . The extended master equation then takes the form

$$\begin{aligned} \frac{dP[\mathbf{n}, t]}{dt} = \alpha \sum_r & [(n_r + 1)P[\mathbf{n}_{r+}, t] - n_r P[\mathbf{n}, t]] \\ & + \sum_r [(N - n_r + 1)f[s(n_r - 1)]P[\mathbf{n}_{r-}, t] - (N - n_r)f[s(n_r)]P[\mathbf{n}, t]] \end{aligned} \quad (2.35)$$

where $\mathbf{n}_{r\pm} = \{n_1, n_2, \dots, n_r \pm 1, \dots, n_\Omega\}$. Thus there are a total of Ω locally homogeneous populations.

Following [73] I rewrote Eq. (2.35) using the one-step operators

$$\mathcal{E}_r^\pm f(\mathbf{n}) = f(\mathbf{n}_{r\pm}) \quad (2.36)$$

i.e.,

$$\frac{dP[\mathbf{n}, t]}{dt} = \sum_r [\alpha(\mathcal{E}_r^+ - 1)n_r + (\mathcal{E}_r^- - 1)(N_r - n_r)f[s(n_r)]] P[\mathbf{n}, t] \quad (2.37)$$

Note that the total number of cells N_r in the r th population comprises n_r active cells and $N_r - n_r = q_r$ quiescent cells, so that

$$n_r + q_r = N_r \quad (2.38)$$

Thus Eq. (2.37) can be rewritten slightly in the form

$$\frac{dP[\mathbf{n}, t]}{dt} = \sum_r [\alpha(\mathcal{E}_r^+ - 1)n_r + (\mathcal{E}_r^- - 1)q_r f[s(n_r)]] P[\mathbf{n}, t] \quad (2.39)$$

This master equation is in a form suitable for mapping into the operator formalism used by Doi and Grassberger-Scheunert. I was not aware that there was also similar work by [63] who was also working near Lapedes and Sharp in T-division at the Los Alamos National Laboratory!

2.5.1 The Operator Map

At around this time I obtained another grant from the ONR, specifically to develop this work, and I now followed Doi in introducing the formalism of quantum mechanics and quantum field theory into the neural network problem. In doing so I recognized that there are essentially two ways to do this, and several differing routes to follow in deriving equations. One way is to assume that individual cortical neurons occupy the vertices of a lattice with spacing L , the other is to assume, as above, that at each vertex there is a homogeneous population of neurons. I first

considered the individual neuron case, and introduced the *occupation number* formalism of quantum mechanics [31].

Let $Q(t)$ and $A(t)$ denote, respectively, the number of quiescent and activated neurons at time t , in a d -dimensional hypercubic lattice containing a neuron at each vertex, where

$$Q(t) + A(t) = N \quad (2.40)$$

Let a neural state vector be denoted by

$$|\Omega\rangle = |\nu_1, \nu_2, \dots, \nu_r, \dots, \nu_N\rangle \quad (2.41)$$

where ν_r indicates that the neuron at the r th vertex is in the *quiescent* “ q ” or else *activated* “ a ” state. This is the occupation number representation, in the Dirac “bra” and “ket” vector notation. Evidently $|\Omega\rangle$ ranges from the configuration $|q_1, q_2, \dots, q_r, \dots, q_N\rangle$ in which every neuron in the lattice is quiescent, to the configuration $|a_1, a_2, \dots, a_r, \dots, a_N\rangle$ in which every neuron in the lattice is activated. Evidently there are 2^N such configurations. Given that there are about 3.10^{11} neurons in human cortex, the number of possible configurations is googleplexian.

Let $P[\Omega(t)]$ be the probability of finding the N -neuron lattice in the state $|\Omega\rangle$ at time t , and let

$$|P[t]\rangle = \sum_{\Omega} P[\Omega(t)]|\Omega\rangle, \quad (2.42)$$

be a neural probability state vector, where

$$\sum_{\Omega} P[\Omega(t)] = 1 \quad (2.43)$$

At this point there were two choices on how to proceed. (1) Follow the Doi method and introduce *creation* and *annihilation* operators, to represent the state transitions in Eq. (2.39), or (2) Use matrices to represent the state transitions. In my notes I used method 1, but in the first published version of this work I used method 2 [24]. In the end the two methods give the same results. In this article we look at both methods. Consider first the Doi method. Let the empty or *vacuum* state be denoted by the vector $|0\rangle$. Let ϕ_r^\dagger and θ_r^\dagger denote, respectively, operators that create active and quiescent neurons at the r th vertex, and let ϕ_r and θ_r denote the corresponding operators that annihilate such states and return the r th vertex to the vacuum state. Thus these operators satisfy the rules:

$$\begin{aligned} \phi_r^\dagger |0\rangle &= a_r, \quad \phi_r |a_r\rangle = 0; \quad \langle 0 | \phi_r^\dagger = 0, \quad \langle 0 | \phi_r = a_r \\ \theta_r^\dagger |0\rangle &= q_r, \quad \theta_r |a_r\rangle = 0; \quad \langle 0 | \theta_r^\dagger = 0, \quad \langle 0 | \theta_r = q_r \end{aligned} \quad (2.44)$$

where $\langle 0 |$ is the dual of $|0\rangle$, and also the *commutation* rules

$$\begin{aligned}\phi_r \phi_{r'}^\dagger - \phi_{r'}^\dagger \phi_r &= [\phi_r, \phi_{r'}^\dagger] = \delta(r - r'), & [\phi_r, \phi_{r'}] &= [\phi_r^\dagger, \phi_{r'}^\dagger] = 0 \\ \theta_r \theta_{r'}^\dagger - \theta_{r'}^\dagger \theta_r &= [\theta_r, \theta_{r'}^\dagger] = \delta(r - r'), & [\theta_r, \theta_{r'}] &= [\theta_r^\dagger, \theta_{r'}^\dagger] = 0\end{aligned}\quad (2.45)$$

Thus these operators are *bosonic* [31].

The next step is to rewrite the one-step operators of Eq. (2.39) in terms of these operators. The translation is easy. It takes the form

$$(\mathcal{E}_r^+ - 1) \rightarrow \theta_r^\dagger \phi_r - \phi_r^\dagger \phi_r, \quad (\mathcal{E}_r^- - 1) \rightarrow \phi_r^\dagger \theta_r - \theta_r^\dagger \theta_r \quad (2.46)$$

whence Eq. (2.39) transforms to

$$\frac{d}{dt} |P[t]\rangle = \sum_r [\alpha(\theta_r^\dagger \phi_r - \phi_r^\dagger \phi_r) + (\phi_r^\dagger \theta_r - \theta_r^\dagger \theta_r) f[s(\phi_r^\dagger \phi_r)]] |P[t]\rangle \quad (2.47)$$

Thus the master equation is now expressed in terms of bosonic operators. The first term expresses the decay of activation at the rate α , the second the increase of activation at the rate $f[s]$.

In reformulating the master equation in such a fashion, at first I found the use of bosonic operators a bit confusing, since in physics bosonic commutation rules allow no exclusion principle. It seemed that many neural states could exist at the same vertex. However the master equation conserves probability equation (2.43), and also the total number of neurons and neural states is fixed at N . These properties imply that if initially $|\Omega(0)\rangle$ is a configuration of initial states, one state to each vertex, then all subsequent configurations will also be “physical”. Thus for master equations bosonic operators do not cause any problems. (In spite of this, Lapedes, Sharp and I had made an earlier attempt in the spring of 1984, to incorporate an exclusion property into the formulation, but we could not get it to work.)

2.5.2 Coherent States

The next step in the Doi method is to introduce the coherent state representation. Let

$$\begin{aligned}|\alpha_a\rangle &= \exp(\alpha_a \sum_r \phi_r^\dagger) |0\rangle, & \langle \alpha_a | &= \langle 0 | \exp(\alpha_a \sum_r \phi_r) \\ |\alpha_q\rangle &= \exp(\alpha_q \sum_r \theta_r^\dagger) |0\rangle, & \langle \alpha_q | &= \langle 0 | \exp(\alpha_q \sum_r \theta_r)\end{aligned}\quad (2.48)$$

be *coherent states* [35]. Such states satisfy the equations

$$\phi_r |\alpha_a\rangle = \alpha_a |\alpha_a\rangle, \quad \langle \alpha_a | \phi_r^\dagger = \alpha_a \langle \alpha_a |, \quad \theta_r |\alpha_q\rangle = \alpha_q |\alpha_q\rangle, \quad \langle \alpha_q | \theta_r^\dagger = \alpha_q \langle \alpha_q | \quad (2.49)$$

i.e., they are *eigenstates* of the creation and annihilation operators ϕ_r and θ_r etc.

I was able to use these coherent states to re-express the master equation in the coherent state representation and to find a coherent state representation of the functional differential equation satisfied by the moment generating functional. I also began the derivation of a coherent state path integral, and so finally began to explore the path that, in retrospect, Wiener had outlined for me, 23 years earlier! However I got a bit bogged down in the details, and in 1987 I started to explore the other approach I had seen was possible, that of using vectors and matrices to express the neural state transitions.

I therefore considered the following encoding of the problem. Let the vectors $|q\rangle$ and $|a\rangle$ denoting quiescent and activated neural states be represented, respectively, by the column vectors

$$|q\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad |a\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (2.50)$$

Now consider the Pauli spin matrices representing the Lie Group $SU(2)$ [33]

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (2.51)$$

and the associated raising and lowering operators

$$\sigma_\pm = \frac{1}{2}(\sigma_1 \pm i\sigma_2) \quad (2.52)$$

It is easy to show that

$$\sigma_+ |q\rangle = |a\rangle, \quad \sigma_- |a\rangle = |q\rangle \quad (2.53)$$

The neural master equation can now be expressed in the form

$$\frac{d}{dt} |P[t]\rangle = \sum_r \left[\alpha(1 - \sigma_{+r})\sigma_{-r} + (1 - \sigma_{-r})\sigma_{+r} f \left[\sum_s (w_{rs}\sigma_{+s}\sigma_{-s}) \right] \right] |P[t]\rangle \quad (2.54)$$

I recognized that the spin operator expression on the rhs of Eq.(2.54) was a generalization of the Regge spin-Hamiltonian of quantum field theory, i.e. of the

spin-Hamiltonian at the core of what is now called Reggeon Field Theory [3, 38]

$$H = \sum_r \left[\alpha(\sigma_{+r} - 1)\sigma_{-r} + (\sigma_{-r} - 1)\sigma_{+r} \kappa \sum_s \sigma_{+s}\sigma_{-s} \right] \quad (2.55)$$

This seemed promising. I had successfully mapped the mathematics of neural networks into the language of spins in quantum field theory, in particular into a generalization of the Regge spin model.

The formal solution of Eqs. (2.47) and (2.54) takes the form

$$|P[t]\rangle = \exp \left[- \int_0^t H(\tau) d\tau \right] |P[0]\rangle \quad (2.56)$$

where $H(t)$ is given, respectively, by the expressions

$$H_1(t) = \sum_r \left[\alpha(\phi_r^\dagger \phi_r - \theta_r^\dagger \theta_r) + (\theta_r^\dagger \theta_r - \phi_r^\dagger \phi_r) f \left[\sum_s w_{rs} \phi_s^\dagger \phi_s \right] \right] \quad (2.57)$$

and

$$H_2(t) = \sum_r \left[\alpha(\sigma_{+r} - 1)\sigma_{-r} + (\sigma_{-r} - 1)\sigma_{+r} f \left[\sum_s w_{rs} \sigma_{+s}\sigma_{-s} \right] \right] \quad (2.58)$$

In retrospect it seems evident that there is a map between these two representations of the master equation. Indeed there is such a map, provided by the Schwinger decomposition [68] of σ_+ and σ_- into, respectively,

$$(\phi^\dagger \theta^\dagger) \sigma_+ \begin{pmatrix} \phi \\ \theta \end{pmatrix} = \phi^\dagger \theta; \quad (\phi^\dagger \theta^\dagger) \sigma_- \begin{pmatrix} \phi \\ \theta \end{pmatrix} = \theta^\dagger \phi \quad (2.59)$$

However in 1987 I was not aware of this, and treated the two approaches separately. I therefore decided to work with the spin representation, and proceeded to investigate the details of the solution given in Eq. (2.56).

2.5.3 Moment Generating Equations and Spin-Coherent States

Extracting information from this integral solution seemed difficult. I decided to follow Doi again, and to introduce a suitable family of coherent states, the *spin-coherent states* of [62] and [39] which take the form

$$|\alpha\rangle = \exp \left(\sum_r \alpha_r^* \sigma_{+r} \right) |0\rangle, \quad \langle \alpha| = \langle 0| \exp \left(\sum_r \alpha_r \sigma_{-r} \right) \quad (2.60)$$

where α is a complex number.

It can be easily shown that

$$\langle \alpha | \Omega \rangle = \alpha_1^{v_1} \alpha_2^{v_2} \cdots \alpha_r^{v_r} \cdots \alpha_N^{v_N} \quad (2.61)$$

and that

$$\langle \alpha | P[t] \rangle = \langle \alpha \sum_{\Omega} P[\Omega(t)] | \Omega \rangle = \sum_{\Omega} P[\Omega(t)] \langle \alpha | \Omega \rangle = G(\alpha_1 \alpha_2 \cdots \alpha_N, t) \quad (2.62)$$

$G(\alpha_1 \alpha_2 \cdots \alpha_N, t)$ is the *moment generating function* for the probability distribution $P[t]$.

I was then able to show that $G(t)$ satisfies the PDE

$$\frac{\partial G}{\partial t} = \left[\sum_r \alpha(D\alpha_r - 1) \frac{\partial}{\partial \alpha_r} + \left(\frac{\partial}{\partial \alpha_r} - 1 \right) D\alpha_r f \left[\sum_s w_{rs} D\alpha_s \frac{\partial}{\partial \alpha_s} \right] \right] G \quad (2.63)$$

where

$$D\alpha_r = \alpha_r \left(1 - \frac{\partial}{\partial \alpha_r} \right) \quad (2.64)$$

i.e., the *moment generating PDE* expressed in the *oscillator algebra* representation [62].

2.5.4 A Neural Network Path Integral

The content of Eqs. (2.63) and (2.64) can be summarized in a Wiener-Feynman path integral [67]. Using standard techniques I was able to write an expression for the transition probability $P[\Omega(t)|\Omega(t')] = G[\Omega, t|\Omega, t']$ as the coherent state path integral

$$\int \prod_r \mathcal{D}\alpha_r(\tau) \exp \left[- \int_{t'}^t d\tau \sum_r \left\{ \frac{1}{2} \left(\frac{\partial}{\partial t} D\alpha_r D\alpha_r^* - D\alpha_r \frac{\partial}{\partial t} D\alpha_r^* \right) - H_2(D\alpha_r, D\alpha_r^*) \right\} \right] \quad (2.65)$$

where

$$\mathcal{D}\alpha_r(\tau) = \left(\frac{2}{\pi} \right)^n \lim_{n \rightarrow \infty} \prod_{j=1}^n \frac{d(\text{Re } \alpha_r(j)) d(\text{Im } \alpha_r(j))}{(1 + \alpha_r(j) \alpha_r^*(j))^3} \quad (2.66)$$

This propagator is sometimes written as an *expectation* wrt the Wiener measure

$$\int \prod_r \mathcal{D}\alpha_r(\tau)$$

as

$$G[\Omega, t | \Omega, t'] = \langle \exp \left[- \int_{t'}^t d\tau L_2(D\alpha_r, D\alpha_r^\star) \right] \rangle \quad (2.67)$$

where

$$L_2(D\alpha_r, D\alpha_r^\star) = \sum_r \left\{ \frac{1}{2} \left(\frac{\partial}{\partial t} D\alpha_r D\alpha_r^\star - D\alpha_r \frac{\partial}{\partial t} D\alpha_r^\star \right) - H_2(D\alpha_r, D\alpha_r^\star) \right\} \quad (2.68)$$

is the neural network Lagrangian. Such a path integral can be used to construct the moment generating functional that generates the statistics of the network activity.

2.6 Back to the Master Equation

However, there remained a major problem: the Wiener measure given in Eq. (2.66) is complicated, and in fact the path integral is defined on a non-Euclidean manifold. In 1990 I did not know how to calculate anything with this integral! At this point I took on a new Physics graduate student, Toru Ohira, who expressed an interest in working on the statistics of neural networks. Although we weren't able to make much immediate progress on the path integral problem, we did look at various other aspects of the general problem, by using the master equation in the spin representation, and the associated moment generating function. We were able to show that the equal-time moments can be obtained as

$$\begin{aligned} \langle v_1(t) v_2(t) \cdots v_r(t) \cdots \rangle &= \sum_{\Omega} (v_1(t) v_2(t) \cdots v_r(t) \cdots) P[\Omega, t] \\ &= \left(\frac{\partial}{\partial \alpha_1} \frac{\partial}{\partial \alpha_2} \cdots \frac{\partial}{\partial \alpha_r} \cdots \right) G(\alpha, t) |_{\alpha=1} \end{aligned} \quad (2.69)$$

and that an analog of the Bogoliubov-Born-Green-Kirkwood-Yvon or BBGKY hierarchy of coupled moment equations exists, i.e.

$$\begin{aligned} \frac{\partial}{\partial t} \langle v_r(t) \rangle &= -\alpha \langle v_r(t) \rangle + \langle (1 - v_r) f \left[\sum_s w_{rs} v_s \right] \rangle, \\ \frac{\partial}{\partial t} \langle v_r(t) v_s(t) \rangle &= -2\alpha \langle v_r(t) v_s(t) \rangle + \langle (1 - v_r) v_s f \left[\sum_u w_{ru} v_u \right] \rangle \\ &\quad + \langle (1 - v_s) v_r f \left[\sum_u w_{su} v_u \right] \rangle, \quad \cdots \end{aligned} \quad (2.70)$$

To gain some insight into this model we assumed that

$$f[x] = x, \quad w_{rs} = w_0(\delta_{r,s-1} + \delta_{r,s+1}) \quad (2.71)$$

i.e., we linearized $f[x]$ and limited the connectivity to nearest neighbors in a one-dimensional lattice.

With these conditions the model became equivalent to the *contact* process [47], whose spin-Hamiltonian is that of Eq.(2.55), modulo the restriction to a one-dimensional lattice, and is therefore isomorphic to Reggeon Field Theory in one dimension. This process has been studied using both analytical methods and computer simulations [28,36]. It was found to have an active steady state for $\alpha < \alpha_c$ and a totally quiescent, vacuum state for $\alpha > \alpha_c$. Thus the 1+1-dimensional contact process has a phase transition at a critical value of the decay constant $\alpha = \alpha_c$.

We investigated this process in our neural network framework by looking at two quantities, $\chi = N^{-1} \sum_r \langle v_r \rangle$ and $\eta = N^{-1} \sum_r \langle v_r v_{r+1} \rangle$ under the assumption that all three-point moments factor into products of one- and two-point moments. This was our moment-closure hypothesis. This allowed us to write coupled moment equations for χ and η

$$\begin{aligned} \frac{\partial}{\partial t} \chi &= (1 - \alpha) \chi - \eta \\ \frac{\partial}{\partial t} \eta &= -(1 + 2\alpha) \eta - \eta \chi + \chi(1 + \chi) \end{aligned} \quad (2.72)$$

For this approximation the steady states satisfy

$$\chi \rightarrow 1 - 2\alpha \ (\alpha < 0.5), \quad \chi \rightarrow 0 \ (\alpha > 0.5) \quad (2.73)$$

However, these steady state values did not match Monte Carlo simulations very well over a broad range of values of the decay constant α .

2.6.1 A System-Size Expansion of the Master Equation

We therefore turned to a more systematic way to approximate the master equation, the system-size expansion introduced by [73]. This approximation provides a systematic small parameter expansion of the master equation given in Eq.(2.29), using the basic assumption that the fluctuations about the value $n(t)$ of the number of active neurons in a population of N neurons are of order \sqrt{N} . Thus let

$$n(t) = N\phi(t) + \sqrt{N}\xi(t) \quad (2.74)$$

where $\phi(t)$ satisfies the *macroscopic* equation

$$\frac{\partial}{\partial t} \phi = -\alpha\phi + (1 - \phi)f[s(\phi)] \quad (2.75)$$

which is a form of the Wilson-Cowan equation for a single neural population. Note that Eq. (2.74) implies that the fluctuations are Gaussian, and that this is valid only for fluctuations about a *stable* fixed point ϕ_0 of the macroscopic equation.

The variable $\xi(t)$ represents Gaussian fluctuations of $n(t)$ with distribution $\Pi(\xi, t)d\xi = P(n, t)dn$. We now expand the van Kampen one-step operators, Eq. (2.36) in terms of ξ using Eq. (2.74) as

$$\mathcal{E}^+ = 1 + \frac{1}{\sqrt{N}} \frac{\partial}{\partial \xi} + \frac{1}{N} \frac{\partial^2}{\partial \xi^2} + \dots, \quad \mathcal{E}^- = 1 - \frac{1}{\sqrt{N}} \frac{\partial}{\partial \xi} + \frac{1}{N} \frac{\partial^2}{\partial \xi^2} - \dots \quad (2.76)$$

On substituting these expressions for \mathcal{E}^\pm into the master equation given in Eq. (2.37) and collecting terms of $O(\sqrt{N})$ we obtain

$$\left\{ \frac{\partial}{\partial t} \phi + \alpha\phi - (1-\phi)f[s(\phi)] \right\} \frac{\partial}{\partial \xi} \Pi = 0 \quad (2.77)$$

assuming that ϕ satisfies the macroscopic equation (2.75).

Similarly collecting terms of $O(1)$ gives the Fokker-Planck equation

$$\frac{\partial}{\partial t} \Pi = -\frac{\partial}{\partial \xi} \{K_1(\phi, \xi)\} \Pi + \frac{1}{2} \frac{\partial^2}{\partial \xi^2} \{K_2(\phi, \xi)\} \Pi \quad (2.78)$$

where

$$K_1(\phi, \xi) = -(\alpha + f[s(\phi)])\xi + (1-\phi)\xi f'[s(\phi)], \quad K_2(\phi, \xi) = \alpha\phi + (1-\phi)f[s(\phi)] \quad (2.79)$$

Note that the coefficients K_1 and K_2 are quite different from those obtained from the Kramers-Moyal expansion, given in Eq. (2.31). In particular K_2 is independent of N .

Let

$$\chi = \frac{n}{N} = \phi + \frac{\xi}{\sqrt{N}} = \phi + \eta \quad (2.80)$$

It now follows from Eqs. (2.75) and (2.78), that

$$\begin{aligned} \frac{d}{dt} \chi &= -\alpha\chi + (1-\chi)f[s(\chi-\eta)] + \eta(1-\chi+\eta)f'[s(\chi-\eta)] + \dots \\ \frac{d}{dt} \eta &= -\alpha\eta - \eta f[s(\chi-\eta)] + \eta(1-\chi+\eta)f'[s(\chi-\eta)] + \dots \end{aligned} \quad (2.81)$$

These equations can be truncated and numerically integrated. Figures 2.7 and 2.8 show comparison of the results for (A) the macroscopic equation, and (B) Eqs. (2.81) under various conditions:

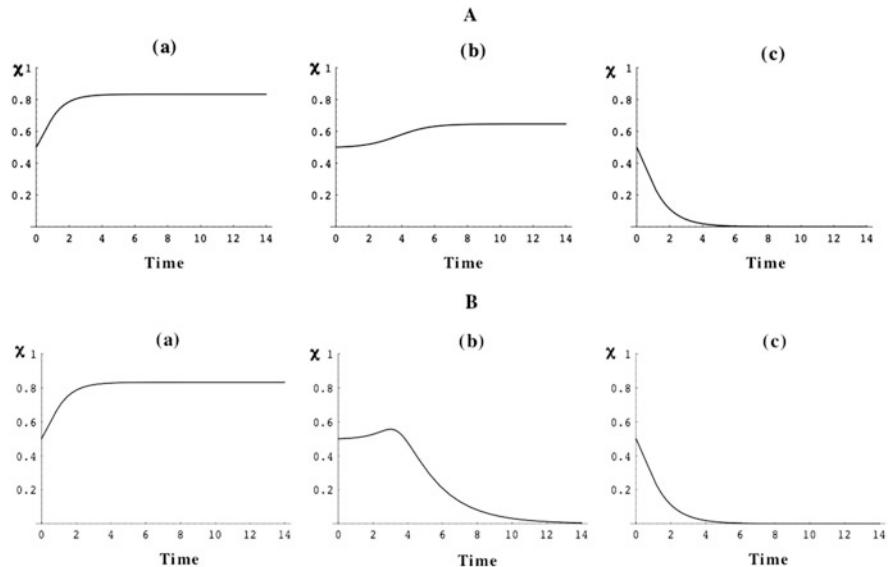


Fig. 2.7 Comparison of numerical solutions of (A) the macroscopic equation, and (B) Eqs. (2.81) for α = (a) 0.2, (b) 0.493, and (c) 0.9, and initial conditions $\chi = 0.5$ and η = (A) 0.01, and (B) 0.0 (Redrawn from [59])

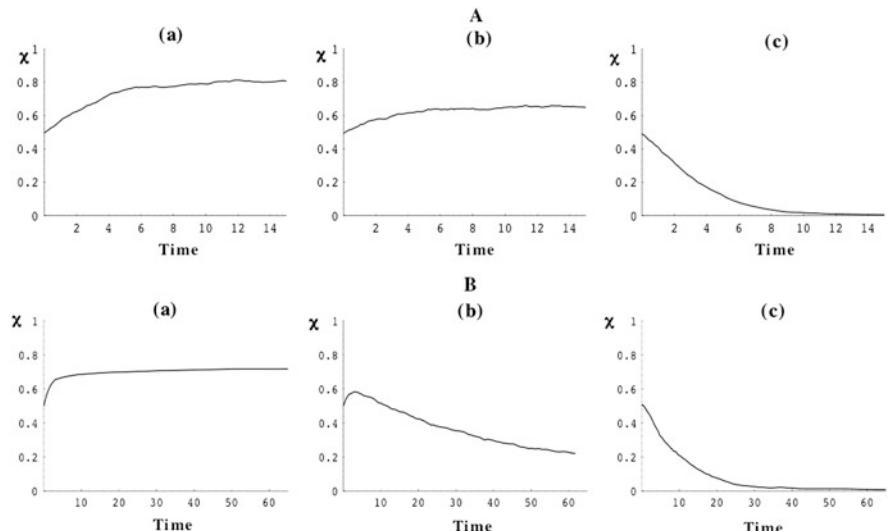


Fig. 2.8 Comparison of Monte Carlo simulations of the master equation for (A) high (0.08), and (B) low (0.006) connectivities per neuron. Initial conditions random and varying α (Redrawn from [59])

It will be seen that for intermediate values of the decay constant α there is a big difference between the mean-field macroscopic equations, and those resulting from the system-size expansion, which incorporate the effects of fluctuations. In particular the fluctuations kill any persistent activated state. The reader is referred to [59] for more details.

2.7 Another Look at Path Integrals

The final computation Ohira and I carried out was to derive a bosonic path integral. The results can be found in [58]. (At that time, 1993, I was not aware of the important paper by [61], on formulating such path integrals.) The resulting integral had the desired property that the Wiener measure was on a Euclidean domain, so, in principle, we could calculate statistics. Unfortunately it was not clear from the form of the path integral how to proceed, and we did not follow up the calculation. In retrospect this was another mistake on my part, and it was not until 2002 that another Physics graduate student, Michael Buice, who was very well-trained in the techniques of quantum field theory and the renormalization group, joined me and chose to work on the problem, that we started to make progress again. However Ohira and I had established the close connection with Reggeon Field Theory that I had earlier guessed at, and we had shown that the van Kampen system-size expansion produced equations that went beyond the mean-field macroscopic Wilson-Cowan equations. However, it must be said that I did not have these results clearly in my mind until Buice joined me, and we were able to greatly clarify and expand upon them.

2.7.1 Bosonic Path-Integrals for Neural Networks

Buice and I reviewed all my earlier efforts and we ended up with a much clearer and simpler account of the derivation of the bosonic path integrals for both two-state and three-state neural models, than I had formulated in 1986–1990. (In what follows I will confine this discussion to the two-state model.)

Our starting point was Eq. (2.47) for the bosonic master equation. This describes the stochastic evolution of neural activity, as specified by the configuration state vector

$$|\Omega\rangle = |v_1, v_2, \dots, v_r, \dots, v_N\rangle$$

from some initial configuration $|\Omega(0)\rangle$ to the current one $|\Omega(t)\rangle$. However, many observations indicate that the numbers of activated neurons at any instant are rather low, i.e. the distribution of active neurons is *sparse*. Thus most configurations are not far from the ground state $|q_1, q_2, \dots, q_r, \dots, q_N\rangle$. This suggested that we could

simplify the bosonic master equation by assuming that most of the time the network is in the ground state in which the fraction of cells in the quiescent state is unity, i.e, that the number operator

$$\sum_r \theta_r^\dagger \theta_r = 1 \quad (2.82)$$

which implies that $\theta_r^\dagger = \theta_r = 1$. This reduces Eq. (2.47) to the form

$$\frac{d}{dt} |P[t]\rangle = \sum_r [\alpha(\phi_r - \phi_r^\dagger \phi_r) + (\phi_r^\dagger - 1)f[s(\phi_r^\dagger \phi_r)]] |P[t]\rangle \quad (2.83)$$

At this point we introduced a shift in the representation of the number operator

$$\sum_r \phi_r^\dagger \phi_r \rightarrow \sum_r (\phi_r^\dagger + 1)\phi_r \quad (2.84)$$

which allows us to write all products of the form $\langle u | v \rangle$ in the form $\langle u | v' | 0 \rangle$, i.e. as a vacuum expectation. The effect of such a shifted representation is to simplify Eq. (2.83) still further, to the form

$$\frac{d}{dt} |P[t]\rangle = \sum_r [-\alpha \phi_r^\dagger \phi_r + \phi_r^\dagger f[s(\phi_r^\dagger \phi_r + \phi_r)]] |P[t]\rangle \quad (2.85)$$

Instead of representing this in terms of the spin coherent states, Eq. (2.60), which led to Eq. (2.54), the spin-coherent state master equation, we were now able to use instead the simpler Glauber states of Eq. (2.48), since the neural model is now effectively a one-state model, to obtain from the reduced bosonic master equation of Eq. (2.85), its coherent state representation

$$\frac{d}{dt} |P[t]\rangle = \sum_r [-\alpha \tilde{\varphi}_r \varphi_r + \tilde{\varphi}_r f[s(\tilde{\varphi}_r \varphi_r + \varphi_r)]] |P[t]\rangle \quad (2.86)$$

where we introduced a change of notation for the Glauber coherent states

$$\alpha_a \rightarrow \varphi, \quad \alpha_q \rightarrow \vartheta \quad (2.87)$$

and $\tilde{\varphi}, \tilde{\vartheta}$ are, respectively, the *adjoints* of φ, ϑ .

Thus the Glauber coherent state Hamiltonian for a sparsely firing two-state neural network is simply that of the one-state Hamiltonian

$$H_3(t) = \sum_r [\alpha \tilde{\varphi}_r \varphi_r - \tilde{\varphi}_r f[s(\tilde{\varphi}_r \varphi_r + \varphi_r)]] \quad (2.88)$$

and the solution of Eq. (2.86) is

$$|P[t]\rangle = \exp\left[-\int_0^t H_3(\tau)d\tau\right] |P[0]\rangle \quad (2.89)$$

We then followed the standard methodology to end up with an expression for the propagator $G[\Omega, t|\Omega, t']$ as

$$G[\Omega, t|\Omega, t'] = \langle \exp\left[-\int_{t'}^t d\tau L_3(\tilde{\varphi}_r, \varphi_r)\right] \rangle \quad (2.90)$$

with respect to the Wiener measure

$$\int \mathcal{D}\tilde{\varphi}_A(\tau) \mathcal{D}\varphi_A(\tau)$$

where

$$\mathcal{D}\varphi_A(\tau) \mathcal{D}\tilde{\varphi}_A(\tau) = \prod_r \mathcal{D}\varphi_r(\tau) \mathcal{D}\tilde{\varphi}_r(\tau) = \left(\frac{1}{\pi}\right)^n \lim_{n \rightarrow \infty} \prod_r \prod_{j=0}^n d\tilde{\varphi}_r(j) d\varphi_r(j) \quad (2.91)$$

and

$$L_3(\tilde{\varphi}_r, \varphi_r) = \sum_r \left[\tilde{\varphi}_r \frac{\partial}{\partial t} \varphi_r + H_3(\tilde{\varphi}_r, \varphi_r) \right] \quad (2.92)$$

is the neural network Lagrangian. The integral of the Lagrangian is referred to as the *action* $S(\tilde{\varphi}_r, \varphi_r)$, i.e.

$$S(\tilde{\varphi}_r, \varphi_r) = \int_{t'}^t d\tau L_3(\tilde{\varphi}_r, \varphi_r) \quad (2.93)$$

and the entire path integral solution of the master equation is written as the expectation over the Wiener measure as

$$\langle \exp[-S(\tilde{\varphi}_r, \varphi_r)] \rangle.$$

In addition Buice and I introduced the moment generating functional (MGF)

$$Z[\tilde{J}_r(t), J_r(t)] = \langle \exp[-S(\tilde{\varphi}_r, \varphi_r) + \tilde{J} \cdot \varphi + J \cdot \tilde{\varphi}] \rangle \quad (2.94)$$

where

$$\tilde{J} \cdot \varphi = \int dt \sum_r \tilde{J}_r(t) \varphi_r(t), \quad J \cdot \tilde{\varphi} = \int dt \sum_r J_r(t) \tilde{\varphi}_r(t) \quad (2.95)$$

By functional differentiation w.r.t. the auxiliary variables J and \tilde{J} and then setting them equal to zero, we can obtain all joint *moments* of the variables $\varphi(x, t)$ and $\tilde{\varphi}(x, t)$. Here then was the result that satisfied my efforts to obtain a statistical theory of large-scale neural activity. This was what Wiener had suggested I look for in 1962. It took 40 years! The details can be found in [12] and in [13, 14].

2.7.2 Observables of Neural Activity

Given this method of calculating neural statistics, we now confronted the problem of how to interpret our formalism. As I noted earlier, there are two differing interpretations of the variables representing neural activity. In the first interpretation v_r is interpreted to be the indicator of spiking activity, equal to 0 if the cell at site r is quiescent, equal to 1 if it is activated. In the second interpretation, v_r takes values in the interval $[0, 1]$ and indicates the fraction of N cells in a homogeneous population located at the r th lattice site that are activated at a given time t . Buice and I introduced another related possibility, that φ_r , the coherent state representation of v_r , is the number of activations of the cell at site r in a time bin of width Δ , i.e. of the spike *count* of the cell at the r th lattice site. As such it fluctuates, so that $\langle \varphi_r(t) \rangle$ is the mean spike count at the r th lattice site. The bosonic character of φ_r is not an issue in this interpretation. Since $\langle \varphi_r(t) \rangle$ is proportional to the mean current density $i_r(t)$ and therefore to the local field potential (LFP) $v_r(t)$, the fundamental observables of our formulation are directly related to spike counts obtained by binning, or to LFPs. In addition we can also compute the correlations $C(x - x', t - t') = \langle \varphi_r(t) \varphi_{r'}(t') \rangle$ and $D(x - x', t - t') = \langle v_r(t) v_{r'}(t') \rangle$ the properties of which provide important information about neocortical dynamics near the critical points of phase transitions.

2.7.3 The Effective Spiking Model

The full model action can also be written in the thermodynamic [$N \rightarrow \infty$] and continuum [$L \rightarrow 0$] approximations as

$$S(\tilde{\varphi}(x, t), \varphi(x, t)) = \int_0^t dt \left[\int d^d x [\tilde{\varphi}(x, t) \partial_t \varphi(x, t) + \alpha \tilde{\varphi}(x, t) \varphi(x, t) - \tilde{\varphi}(x, t) f[s(\tilde{\varphi}(x, t) \varphi(x, t) + \varphi(x, t))] \right] \quad (2.96)$$

where the input current is

$$s(\tilde{\varphi}(x, t) \varphi(x, t) + \varphi(x, t)) = \int d^d y w(x - y) (\tilde{\varphi}(y, t) \varphi(y, t) + \varphi(y, t)) + h(x, t) \quad (2.97)$$

where $w(x)$ is the distribution of weights, and $h(x, t)$ is an external stimulus. Buice and I referred to this action as that of the *effective spike* model. There was also an initial condition which could be incorporated into the action if necessary.

It is then just a matter of applying the standard techniques of quantum field theory to produce predictions from this action. The main technique is a form of perturbation theory adapted to computing moments from the path integral incorporating the effective spike action. In fact the MGF when the interactions are zero, $w(x) = 0$, can be expressed as

$$Z_0[\tilde{J}, J] = \exp \left[- \int dt dt' d^d x d^d x' \tilde{J}(x, t) \Delta(x, t; x', t') J_s(x', t') \right] \quad (2.98)$$

where $\Delta(x, t; x', t')$ is the solution of the linearized equation

$$\left(\frac{\partial}{\partial t} + \alpha \right) \Delta(x, t; x', t') = \delta(t - t') \delta^d(x - x') \quad (2.99)$$

i.e.

$$\Delta(x, t; x', t') = e^{-\alpha(t-t')} \Theta(t - t') \delta^d(x - x') \quad (2.100)$$

known as the Green's function or *propagator* of Eq. (2.99), and

$$J_s(x, t) = J(x, t) + f[h(x, t)] + p(x) \delta(t - t_0) \quad (2.101)$$

where $p(x)$ is the probability of a single spike occurring in the interval $(x, x+dx)$ at $t = t_0$, i.e., an initial condition equal to a Poisson distribution with mean rate $p(x)$.

$Z_0[\tilde{J}, J]$ is called the MGF for the *free* action $S_0(\tilde{\varphi}, \varphi)$ with $w = 0$. It follows immediately from this MGF that the first moment of the free, uncoupled neural network can be written as

$$\langle \varphi(x, t) \rangle = \frac{\delta}{\delta \tilde{J}(x, t)} Z_0[\tilde{J}, J] |_{\tilde{J}, J=0} = p(x) e^{-\alpha(t-t_0)} + \int_{t_0}^t dt' e^{-\alpha(t-t')} f[h(x, t')] \quad (2.102)$$

The quantity $\langle \varphi(x, t) \rangle$ is the mean spike count, analogous to n_r in our earlier analysis. Evidently it is the solution of the equation

$$\left(\frac{\partial}{\partial t} + \alpha \right) \langle \varphi(x, t) \rangle = f[h(x, t)] + p(x) \delta(t - t_0) \quad (2.103)$$

Similarly the quantity

$$\langle \varphi(x, t) \tilde{\varphi}(x', t') \rangle = \frac{\delta}{\delta \tilde{J}(x', t')} \frac{\delta}{\delta \tilde{J}(x, t)} Z_0[\tilde{J}, J] |_{\tilde{J}, J=0} = \Delta(x', t'; x, t) \quad (2.104)$$

is a solution of Eq. (2.99) and gives the linear response to a small change in activity at the location x .

In case $w \neq 0$ perturbation theory about the free action must be used to compute corrections to the expressions given above. All the techniques developed in quantum field theory, Feynman diagrams and renormalization Group Techniques, can be used to perform such computations. We obtained the general formula

$$\frac{\delta^n}{\delta \tilde{J}^n} \frac{\delta^m}{\delta J^m} Z[\tilde{J}, J] |_{\tilde{J}, J=0} = \sum_k \frac{(-1)^k}{k!} \langle [S_w(\tilde{\varphi}, \varphi)]^k \varphi^n \tilde{\varphi}^m \rangle \quad (2.105)$$

where $S_w(\tilde{\varphi}, \varphi) = S(\tilde{\varphi}, \varphi) - S_0(\tilde{\varphi}, \varphi)$ for the perturbative corrections to the moments.

2.7.4 Variational Techniques

One of the main reasons for introducing the path-integral approach pioneered by [77] and by [32] is that there is a close connection between the action $S(\tilde{\varphi}, \varphi)$ and certain *paths* from the initial to the final state. In fact the most probable path is obtained by assuming that S satisfies a form of the Euler-Lagrange equations of variational calculus. The result is

$$\left\langle \frac{\delta S(\varphi_a)}{\delta \varphi_a} \right\rangle = 0 \quad (2.106)$$

where $\varphi_a = \{\tilde{\varphi}, \varphi\}$. Using Eq. (2.96) gives

$$\partial_t \langle \varphi(x, t) \rangle = -\alpha \langle \varphi(x, t) \rangle + \langle f[w \star \varphi + h(x, t)] \rangle \quad (2.107)$$

where \star means spatial convolution, and it follows from the fact that $\tilde{\varphi}(x, t)$ is a representation of intrinsic neural fluctuations, that $\langle \tilde{\varphi}(x, t) \rangle = 0$.

Inspection of this equation indicates that depending on the nature of the firing rate function $f[x]$, the mean spike count $\langle \varphi(x, t) \rangle$ is coupled to many higher moments. The simplest assumption that breaks this coupling is to assume that all higher moments decompose into products of $\langle \varphi(x, t) \rangle$. The result is the mean-field or macroscopic equation

$$\partial_t \langle \varphi(x, t) \rangle = -\alpha \langle \varphi(x, t) \rangle + f[w \star \langle \varphi \rangle + h(x, t)] \quad (2.108)$$

This is a version of Eq. (2.27), the sparse firing rate Wilson-Cowan equation for a single homogeneous population of neurons.

The next level of approximation is the assumption that the higher moments are Gaussian and factor into products of the mean and the variance $\langle \varphi(x, t) \rangle$

and $\langle \varphi(x, t)\varphi(x', t') \rangle$, and the propagator $\langle \varphi(x, t)\tilde{\varphi}(x', t') \rangle$. Using the variational condition

$$\left\langle \varphi(x', t') \frac{\delta S(\varphi_a)}{\delta \varphi_a(x, t)} \right\rangle = 0 \quad (2.109)$$

together with the causality condition that any moments with a factor $\tilde{\varphi}(x, t)$, $t > t'$ multiplying $\varphi(x', t')$ must vanish, and the definitions $C[x, t; x', t'] = \langle \varphi(x, t)\varphi(x', t') \rangle - \langle \varphi(x, t) \rangle \langle \varphi(x', t') \rangle$ and $\langle \varphi(x, t) \rangle = a(x, t)$, gives the following equations

$$\begin{aligned} \partial_t a(x, t) &= -\alpha a(x, t) + f[s(x, t)] \\ &\quad + \int \int d^d x' d^d x'' f''[s(x, t)] w(x, x'') w(x, x') C[x', t : x'', t] \\ \partial_t \Delta(x, t; x', t') &= -\alpha \Delta(x, t; x', t') + \int d^d x'' f'[s(x, t)] w(x, x'') \Delta(x'', t; x', t') \\ &\quad + \delta(x - x') \delta(t - t') \\ -\partial_{t'} \Delta(x, t; x', t') &= -\alpha \Delta(x, t; x', t') + \int d^d x'' f'[s(x'', t)] w(x, x'') \Delta(x', t'; x'', t) \\ &\quad + \delta(x - x') \delta(t - t') \end{aligned} \quad (2.110)$$

and

$$\begin{aligned} \partial_t C[x, t; x', t'] &= -\alpha C[x, t; x', t'] + \int d^d x'' f'[s(x, t)] w(x, x'') C[x'', t; x, t] \\ &\quad + \int d^d x'' [f'[s(x, t)] w(x, x'') a(x'', t') + f'[s(x'', t)] w(x'', x) a(x, t)] \Delta(x', t'; x'', t) \end{aligned} \quad (2.111)$$

These equations extend the Wilson-Cowan equation for a single homogeneous neural population to include the effects of fluctuations and correlations beyond mean-field effects. Buice and I were able to show under what conditions this truncation of the hierarchy of moment equations is valid [13].

It turns out that the magnitude of the effects of fluctuations and correlations is connected with the mean-field dynamics described by the Wilson-Cowan equations. Recall that Wilson and I discovered that neural network dynamics has stable fixed points or attractors. At or near attractors fluctuations and correlations tend to be small, and in fact scale with $O(1/\sqrt{\kappa N})$ for fluctuations, and $O(1/\kappa N)$ for correlations, where N is the number of neurons in the local population, and κN is the number of contacts or synapses per neuron. The constant κ has been estimated to be about 0.025 [70] which translates to about 4,000 synapses for most neocortical cells, and $N \approx 150,000$ in a local neocortical population. Thus the Wilson-Cowan equations are a good model for neocortical dynamics near attractors. However there are also marginally stable fixed points which signal a transition from one attractor to

another. These are the critical points of phase transitions. Buice and I calculated the stability properties of the attractor in our single population model, and found that the eigenvalue equation is

$$\lambda = -\alpha + f' \cdot w(p) \quad (2.112)$$

where f' is the slope of the firing rate function at the fixed point, and $w(p)$ is the Fourier transform of the weighting function $w(x)$. Thus there is a critical point when $\alpha = \alpha_c = f' \cdot w(p)$. So if $\alpha > \alpha_c$, $\lambda < 0$ and the fixed point is an attractor, and the fluctuations and correlations decay, whereas as $\alpha \rightarrow \alpha_c$, $\lambda \rightarrow 0$ the fluctuations and correlations have an increasing effect, and the fixed point becomes a critical point.

Buice and I then derived the following condition for the existence of mean-field behavior, namely that the inequality

$$\left(\frac{w_2}{w_0}\right)^2 \gg \frac{|f''|w_0 A}{f'} L_d^{4-d}, \quad L_d = \frac{1}{\sqrt{2}} \sqrt{\frac{f' w_2}{\alpha - f' w_0}} = \sqrt{\frac{D}{\mu}} \quad (2.113)$$

is satisfied, where $\mu = \alpha - f' w_0$, $D = \frac{1}{2} f' w_2$ is the effective diffusion constant, $w_0 = \int w(x) dx$ and $w_2 = \int x^2 w(x) dx$ are, respectively, the zeroth and second moments of the weight function $w(x)$ and L_d is the effective diffusion length of the activity [13]. Thus only the bulk parameters of the neocortical model, f' and f'' which relate to the probability distribution of thresholds, and w_0 and w_2 which relate to the total number and variance of the weights, and α which is the decay rate of spiking activity, and finally the dimension d of the cortical slab, appear in the inequality.

2.7.5 Renormalizing the Action

These considerations can be made more precise by renormalizing the action using methods introduced by [78]. Buice and I applied these techniques to the path integral given in Eq. (2.96). The result is the renormalized Hamiltonian

$$S(\tilde{\varphi}, \varphi) = \int_0^t dt \int d^d x [\tilde{\varphi} \partial_t \varphi + \mu \tilde{\varphi} \varphi - D' \tilde{\varphi} \nabla^2 \varphi - g' \cdot (\tilde{\varphi} - \varphi) \cdot \tilde{\varphi} \varphi] \quad (2.114)$$

where μ , D' , and g' are renormalized constants. This result was the payoff for my 40 years of work since the renormalized action is that of Reggeon field theory, otherwise known as *directed percolation* or DP, equivalent to a *branching-aggregating* random walk. Thus we had successfully mapped the mathematics of stochastic neural network dynamics into that of a specific type of percolation process.

If we look again at Eq. (2.113) we can now interpret the dimension $d = 4$ as the *upper* critical dimension of the DP phase transition. This is the dimension at which the condition $\mu = 0$ no longer causes the inequality to be violated, thus guaranteeing mean field behavior for $d \geq 4$. In fact neural connectivity is quite dense, since there are about 4,000 contacts per cell, and in our continuum approximation it makes sense that the *effective* dimension $d \geq 4$. This implies that as an approximation, we can use the mean field scaling exponents of the DP phase transition, which are $\beta = 1, v_{\perp} = 1/2, v_{\parallel} = 1$, where β is related to the probability that an active cell at a given site decays, v_{\perp} and v_{\parallel} are related, respectively, to spatial and temporal correlation lengths of the activity at multiple sites. Given such values we can write down the approximate form of the renormalized propagators for subcritical, critical, and supercritical states of our neural network model, based on the calculations of [1, 2, 7] and [19]. The results are as follows: for the subcritical state, we obtain

$$G_R(x - x', t - t') \propto (t - t')^{-2} \exp\left[-\frac{(x - x')^2}{4(t - t')} - \mu(t - t')\right] \quad (2.115)$$

which is nothing more than simple diffusion with decay. The renormalized decay constant $\mu > 0$ in the subcritical case. Exactly at the critical point, where $\mu = 0$ the renormalized propagator is approximated by

$$G_R(x - x', t - t') \propto (t - t')^{-2} \exp\left[-\frac{(x - x')^2}{4(t - t')}\right] \quad (2.116)$$

which is just simple diffusion with $D' = 1$.

That this is the case can be seen by examining the form of the free propagator in momentum space, written as

$$G_0(p, t - t') = e^{-\alpha(t-t')} \Theta(t - t') \quad (2.117)$$

or more correctly the unrenormalized propagator $G_I(p, t - t')$ for the case when $w \neq 0$, i.e.

$$G_I(p, t - t') = e^{-[\alpha - f' w(p)](t-t')} \Theta(t - t') = e^{-[\mu + D p^2](t-t')} \Theta(t - t') \quad (2.118)$$

It is easy to show that Fourier transforming $G_I(p, t - t')$ gives the propagator in coordinate space, i.e.

$$G_I(x - x', t - t') \propto [D(t - t')]^{-2} \exp\left[-\frac{(x - x')^2}{4D(t - t')} - \mu(t - t')\right] \quad (2.119)$$

In the renormalized case, the effective diffusion constant $D \rightarrow D' = 1$ in case $d \geq 4$, so that the renormalized subcritical and critical propagators correspond to those obtained directly from $G_I(x - x', t - t')$.

The supercritical case, however, is different. The result is

$$G_R(x - x', t - t') \propto \mu^2 \Theta[\sqrt{\mu}(t - t') - |x - x'|] \quad (2.120)$$

This is a propagating front, effectively the surface of a hypersphere, with amplitude $\mu^2 \propto \langle \varphi \rangle^2$, and speed $\sqrt{\mu}$. Thus as any disturbance becomes more and more supercritical, its amplitude increases rapidly and its speed increases more slowly. This case is very interesting in the light of the early work of [16] on the responses of isolated neocortical slabs to current pulses, which exhibited fronts propagating at a velocity of approximately 15 cm/s.

More recently observations by [46] and [57] have indicated that the resting state of the neocortex has properties described quite precisely by the subcritical renormalized propagator, and that when stimulated, neocortical activity develops properties well described by the mean field Wilson-Cowan equations, in that long-range correlations characteristic of states just below or near criticality decay in favor of localized states that are only correlated with the stimulus.

2.7.6 Avalanches

In another series of experiments on isolated slabs lightly anaesthetized with ethyl-ether, [16] observed *bursts* of activity from a number of randomly occurring sites. Moreover, any variation in the level of anaesthetic, either up or down, abolished such bursts. Sixty years later these bursts were studied systematically by [8] using an 8×8 microelectrode array to record LFPs in isolated neocortical slabs. Their major finding was that spontaneous neural activity is close to being *scale-free*—the number of bursts of a given size (i.e. N , the number of activated electrodes per burst) is power function of the burst size n , such that

$$N(n) \propto n^{-\alpha}, \quad \alpha = \frac{3}{2} \quad (2.121)$$

Using the field theory Buice and I calculated the predicted value of α assuming that spontaneous neural activity is at a critical point. The result was

$$\alpha = \left[1 + \frac{\beta}{\nu(4+z)-\beta} \right] \quad (2.122)$$

where β , ν and z are critical exponents. Assuming that these exponents are those of mean-field DP, $\beta = 1$, $\nu = 1/2$, $z = 2$, so that $\alpha = \frac{3}{2}$. This is the exponent of mean-field *critical branching* [4] in which one active neuron gives rise to just one other active neuron. Thus as Beggs and Plenz had suggested, the critical branching process describes quite accurately spontaneous neural activity in cortical slices. In recognition of this Beggs and Plenz called the random bursts of neural activity

avalanches, thus making contact with the extensive literature on avalanches and criticality in phase transitions. What we had added is that such a branching process occurs in mean-field directed percolation.

2.7.7 Calculations Concerning Brain Rhythms and Spike Statistics

Buice and I also carried out some further calculations on neocortical properties that could be deduced from the field theory. Consider, for example, the Laplace transform of the renormalized subcritical propagator, which has the simple form

$$G_R(p, \omega) = \frac{1}{\mu + p^2 + i\omega} \quad (2.123)$$

and therefore its power spectrum is

$$G_R(p, \omega)G_R^*(p, \omega) = \frac{1}{(\mu + p^2)^2 + \omega^2} \quad (2.124)$$

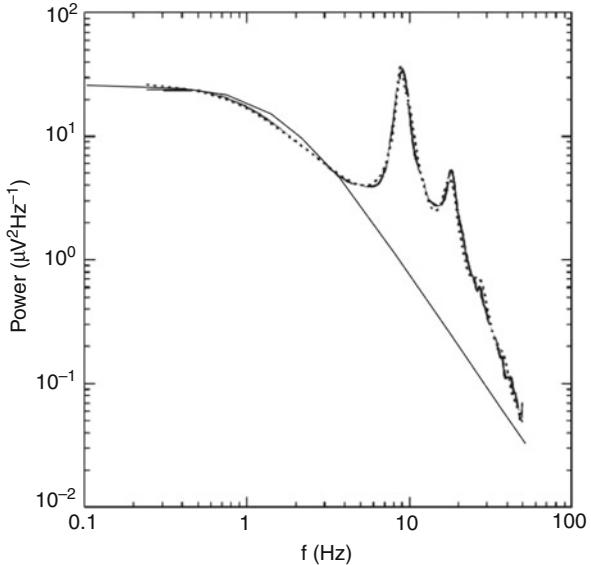
In the spatially homogeneous state, $p = 0$, this reduces to

$$|G_R(p, \omega)|^2 = \frac{1}{\mu^2 + \omega^2} \quad (2.125)$$

We plot the function $K |G_R(p, \omega)|^2$ with $K = 75$ and $\mu = 3$ on the same graph as that of the power spectrum of human EEG [65]. The result is shown in Fig. 2.9. It will be seen that the function $75/(3 + \omega^2)$ provides a good fit to the EEG *baseline*. This was an important result. It established that the EEG baseline spectrum was that of Brownian motion, with a $1/\omega^2$ tail. Of course sitting on top of the Brownian baseline are multiple peaks, corresponding to various brain rhythms, the most prominent of which are the *Alpha* (9 Hz) and *Beta* (18 Hz) rhythms of the occipital cortex. I will describe a model for the generation of such rhythms in a later section.

The tail ω^{-2} however suggested an interesting way to generate a tail equal to ω^{-1} for the power spectrum of spontaneous neural spiking activity recorded with microelectrodes. We based our analysis on a result of [44] who showed that a time series of particle counts generated by some branching process whose correlations correspond to a power spectrum with a $1/\omega^2$ tail, has itself a ω^{-1} tail. Since we had shown that the EEG spectrum, and the closely related local field potential (LFP) [14] had a ω^{-2} tail, here was a simple derivation of the $1/\omega$ tail of the power spectrum of spontaneous spiking activity. Interestingly, [26] derived such a spectrum by assuming that neural threshold fluctuations were Brownian motion! Here was a direct connection with our calculation that the neural field has a Brownian motion

Fig. 2.9 Fit of the function $75/3 + \omega^2$ to EEG data
(Reprinted from [14])



spectrum! The reader is referred to [14] for a fuller discussion of this, and many other results.

2.7.8 *Closed Moment Equations*

Buice graduated in 2007 and left to become a post-doctoral fellow at NIH with Carson Chow. We had managed to solve many of the problems I had gotten stuck on, and had ended up with the statistical field theory that I had aspired to several decades earlier. Buice and I continued to work on some further aspects of the theory, and were joined by Chow in an effort to find a more principled way to derive closed moment equations from the field theory. The results of this work are to be found in [15]. The equations we obtained are essentially those described in Eq. (2.111).

In what follows I describe the special case of a localized finite sized network of N neurons coupled in a spatially homogeneous all-to-all fashion, since it illustrates the essential features of the moment closure problem. The equations in this simplified case reduce to

$$\begin{aligned} \frac{d}{dt}a(t) &= -\alpha a(t) + f[s(t)] + \frac{1}{2}f''[s(t)]w_0^2C(t) \\ \frac{d}{dt}C(t) &= -2\mu C(t) + 2\frac{w_0}{N}f'[s(t)]a(t) \end{aligned} \quad (2.126)$$

where

$$a(t) = \frac{1}{N} \sum_i a_i(t), \quad C(t) = \frac{1}{N^2} \sum_{ij} C_{ij}(t), \quad s(t) = w_0 a(t) + h(t) \quad (2.127)$$

and $C_{ij}(t)$ is the pair correlation function for cells at the vertices i and j of the presumed neural lattice. Thus $a(t)$ is the mean activity in the network, and $C(t)$ is the covariance of the activity. The important feature of these equations is that they are *coupled* moment equations, so that the mean is affected by the covariance and vice-versa.

Equation (2.126) can be solved for C at the fixed point $d_t a(t) = d_t C(t) = 0$. The result is

$$C_0 = \frac{w_0}{\mu N} f'[s_0] a_0 \quad (2.128)$$

Thus $C_0 \propto (\mu N)^{-1}$. But it follows from Eq. (2.112) that $\mu = -\lambda_0$, the eigenvalue of the mean field equation

$$\frac{d}{dt} a(t) = -\alpha a(t) + f[s(t)]$$

in the spatially homogeneous case, so that $C_0 \propto (-\lambda_0 N)^{-1}$. So as long as the fixed point $\{a_0, C_0\}$ is *stable*, $-\lambda > 0$, $C(t)$ will become smaller with increasing numbers of neurons N , and its effects on the network dynamics can be discounted. This is a finite-size effect: for large enough N , the mean-field equation is a good model of the dynamics, any small effects of the correlations can be well represented by Eq. (2.126). Conversely, as $\lambda \rightarrow 0$ so that the fixed point approaches a critical point, μN will diminish and correlations will grow. Thus Eqs. (2.126) become singular at the critical point, and the correlations and intrinsic fluctuations dominate the dynamics. Thus the neighborhood of the critical point can be described as a *fluctuation dominated* region [18], and the renormalization group technique, or its mathematical equivalent—singular perturbation theory [20], must be used to analyze this region. This conclusion also holds for the general spatially inhomogeneous case.

Figure 2.10 shows a plot of some solutions of the spatially inhomogeneous network case, for Poisson initial conditions with $a_i(0) = 2$, $C_{ij}(0) = 0$ and $\alpha = 0.9$ where $\alpha = 1.0$ at the critical point. It will be seen that the mean field equations fail to capture the true behavior of the network. While the coupled moment equations do better, they do not give an accurate representation of the true behavior. Evidently $\alpha = 0.9$ is too close to criticality.

These results are consistent with those Ohira and I obtained in 1997 (see Figs. 2.7 and 2.8), using the system-size expansion of the master equation, in that the mean field prediction of a non-zero active state is incorrect. It became clear to me, and independently to Buice and Chow, sometime in 2008, that there was a close connection between the closed moment equations we had obtained, and equations

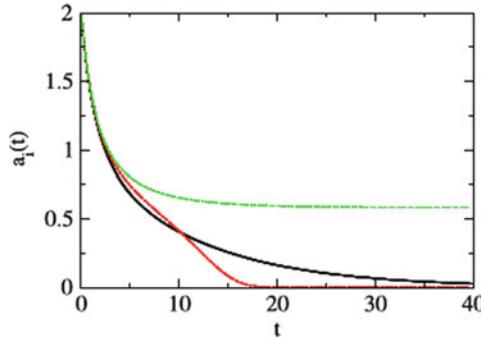


Fig. 2.10 $a_1(t)$ versus t for $\alpha = 0.9$, $N = 10$. Upper lines are solutions of mean-field theory, lower lines are solutions of spatially inhomogeneous equations, and black lines are expectation values of data from simulations of the underlying Markov process, carried out using the Gillespie algorithm [34] (Reprinted from [15])

that could be obtained using the van Kampen system size expansion of the master equation. I started to writeup some notes on this connection when Paul Bressloff, now at Oxford University, saw a preprint of the manuscript Buice, Chow, and I had submitted for publication, which we had placed in the public domain in Arxiv.org. Working with his customary lightning speed Bressloff wrote and published a paper on this topic [11] which formalized this connection, and greatly clarified, at least for me, a number of the issues I had grappled with over the years. In particular, the paper looked at the different ways in which one can develop the field theory by starting either from the spike count as a measure of neural activity, as Buice and I had done, versus using the fraction of activated cells in a population comprising N cells, as a measure of the activity, which I had used to formulate the original master equation in 1979. This paper made clear to me a result Buice had obtained when still my student, that what physicists call the one-loop correction to the free propagator of our theory, generates the same formulas as do the moment equations obtained from the system-size expansion of the master equation.

2.8 Stochastic Wilson-Cowan Equations

There remained the problem of using all this mathematical physics machinery to analyze more realistic models of the neocortex. For example, the models Ohira, Buice and I had developed did not distinguish between *excitatory*, E , and *inhibitory*, I , neurons, as did the mean-field Wilson-Cowan equations. It therefore remained a goal of mine to extend the field theory to cover such a distinguishing feature of the neocortex. By this time (2008), I had two other graduate students working with me, Edward Wallace (Mathematics) and Marc Benayoun (Pediatric Neurology), and a local collaborator Wim van Drongelen (Pediatrics), with whom I jointly supervised

Benayoun. After much animated discussion Wallace and Benayoun succeeded in convincing me to look again at the Fokker-Planck approach to modeling E/I networks, which I had more-or-less given up in favor of the field theory approach. The result was a master equation for the E/I case, and subsequently a Fokker-Planck equation derived from the master equation via the system-size expansion. The details can be found in [9].

2.8.1 The E/I Master Equation

We first introduced an extension of Eq. (2.29) for a single homogeneous population comprising M excitatory, and N inhibitory neurons, i.e. the master equation for $P_{m,n}(t)$, the probability of finding m active excitatory and n active inhibitory neurons at time t

$$\begin{aligned} \frac{dP_{m,n}(t)}{dt} = & \alpha_E [(m+1)P_{m+1,n} - mP_{m,n}] \\ & + (M-m+1)f_E[s_E(m-1, n)]P_{m-1,n} \\ & - (M-m)f_E[s_E(m, n)]P_{m,n} \\ & + \alpha_I [(n+1)P_{m,n+1} - nP_{m,n}] \\ & + (N-n+1)f_I[s_I(m, n-1)]P_{m,n-1} \\ & - (N-n)f_I[s_I(m, n)]P_{m,n} \end{aligned} \quad (2.129)$$

where

$$s_E(m, n) = w_{EE}m - w_{EI}n + h_E, \quad s_I(m, n) = w_{IE}m - w_{II}n + h_I \quad (2.130)$$

are the currents driving the neurons.

A straightforward generalization of this to the case of \mathcal{Q} such populations is

$$\begin{aligned} \frac{dP[\mathbf{m}, \mathbf{n}; t]}{dt} = & \alpha_E \sum_{r=1}^{\mathcal{Q}} \{((m_r + 1)P[\mathbf{m}_{r+}, \mathbf{n}; t] - m_r P[\mathbf{m}, \mathbf{n}; t] \\ & + (M_r - m_r + 1)f_E[s_E(m_r - 1, n_r)]P[\mathbf{m}_{r-}, \mathbf{n}; t] \\ & - (M_r - m_r)f_E[s_E(m_r, n_r)]P[\mathbf{m}, \mathbf{n}; t] \} \\ & + \alpha_I \sum_{r=1}^{\mathcal{Q}} \{((n_r + 1)P[\mathbf{m}, \mathbf{n}_+; t] - n_r P[\mathbf{m}, \mathbf{n}; t]) \\ & + (N_r - n_r + 1)f_I[s_I(m_r, n_r - 1)]P[\mathbf{m}, \mathbf{n}_-; t] \\ & - (N_r - n_r)f_I[s_I(m_r, n_r)]P[\mathbf{m}, \mathbf{n}; t] \} \end{aligned} \quad (2.131)$$

where

$$\begin{aligned} s_E(m_r, n_r) &= \sum_s w_{rs}^{EE} m_s - \sum_s w_{rs}^{EI} n_s + h_E \\ s_I(m_r, n_r) &= \sum_s w_{rs}^{IE} m_s - \sum_s w_{rs}^{II} n_s + h_I \end{aligned} \quad (2.132)$$

are the currents driving the neurons. Application of the van Kampen system-size expansion to this master equation generates, to lowest order, the Wilson-Cowan equations, and to the next order Fokker-Planck and Langevin equations for the fluctuations.

Let

$$\frac{m_r}{M_r} = E_r + \frac{\eta_r}{\sqrt{M_r}}, \quad \frac{n_r}{N_r} = I_r + \frac{\xi_r}{\sqrt{N_r}} \quad (2.133)$$

It then follows from the system-size expansion that to $O(1)$

$$\begin{aligned} \frac{dE_r}{dt} &= -\alpha_E E_r + (1 - E_r) f_E[w_r^{EE} \star E_r - w_r^{EI} \star I_r + h_r^E] \\ \frac{dI_r}{dt} &= -\alpha_I I_r + (1 - I_r) f_I[w_r^{IE} \star E_r - w_r^{II} \star I_r + h_r^I] \end{aligned} \quad (2.134)$$

i.e. the mean-field Wilson-Cowan equations for coupled E/I populations on a lattice.

Similarly, to $O(\frac{1}{\sqrt{N}})$ we obtain the coupled linear fluctuation or Langevin equations

$$\frac{d}{dt} \begin{pmatrix} \xi_r \\ \eta_r \end{pmatrix} = A \begin{pmatrix} \xi_r \\ \eta_r \end{pmatrix} + \begin{pmatrix} \sqrt{2\alpha_E E_{r,0}} & \cdot \\ \cdot & \sqrt{2\alpha_I I_{r,0}} \end{pmatrix} \begin{pmatrix} \chi_r \\ \zeta_r \end{pmatrix} \quad (2.135)$$

assuming $\xi_r(0) = \eta_r(0) = 0$, where the matrix A_r is the Jacobian of Eq. (2.134), B is the diagonal matrix

$$B_r = \begin{pmatrix} \sqrt{2\alpha_E E_{r,0}} & \cdot \\ \cdot & \sqrt{2\alpha_I I_{r,0}} \end{pmatrix} \quad (2.136)$$

χ_r and ζ_r are independent white noise variables, and $E_{r,0}$ and $I_{r,0}$ are the fixed point values of $E_r(t)$ and $I_r(t)$, respectively.

At $O(\frac{1}{N})$ we obtain lattice versions of Eq. (2.126) extended to deal with E/I populations.

In [9] we used this Langevin equation to analyze the effects of fluctuations on the dynamics of a single homogeneous coupled E/I neocortical population, in case the mean-field fixed point $\{E_0, I_0\}$ is a stable node. Equation (2.135) describes Gaussian fluctuations of the activity about such a node. We noticed in simulations that large

fluctuations tended to occur as the level of inhibition approaches that of excitation. This gives rise to the so-called *balanced* condition in which the net excitation per neuron is balanced by its net inhibition. In such a condition we can introduce the additional symmetries $w_{EE} = w_{IE} = w_E$, $w_{EI} = w_{II} = w_I$, $w_E = w_I$, and $M = N$, and the linear transformation

$$\Sigma = \frac{1}{2}(E + I), \quad \Delta = \frac{1}{2}(E - I) \quad (2.137)$$

so that the total input current is

$$s = w_0\Sigma + w_1\Delta + h \quad (2.138)$$

where $w_0 = w_E - w_I$, $w_1 = w_E + w_I$ whence in the balanced case $w_0 \ll w_1$.

(In my earlier work with Buice on the one population model, we had tacitly assumed that the net difference between excitation and inhibition was represented by w_0 the total synaptic weight per neuron, so that that $\mu = \alpha - f'w_0 \rightarrow \alpha$ as $w_0 \rightarrow 0$ would give a stable node for positive decay constant α .)

Equation (2.134), in the homogeneous case, now transforms to the simpler form:

$$\begin{aligned} \frac{d\Sigma}{dt} &= -\alpha\Sigma + (1 - \Sigma)f[w_0\Sigma + w_1\Delta + h] \\ \frac{d\Delta}{dt} &= -\Delta(\alpha + f[w_0\Sigma + w_1\Delta + h]) \end{aligned} \quad (2.139)$$

which has a fixed point, the stable node $\{\Sigma_0, 0\}$. Evidently Σ_0 is left unchanged by variations in w_1 , since $\Delta_0 = 0$.

Similarly Eq. (2.140) takes the form

$$\frac{d}{dt} \begin{pmatrix} \xi_\Sigma \\ \eta_\Delta \end{pmatrix} = \begin{pmatrix} -\lambda_1 & w_{FF} \\ \cdot & -\lambda_2 \end{pmatrix} \begin{pmatrix} \xi_\Sigma \\ \eta_\Delta \end{pmatrix} + \sqrt{\alpha\Sigma_0} \begin{pmatrix} \chi_\Sigma \\ \zeta_\Delta \end{pmatrix} \quad (2.140)$$

where $\lambda_1 = \alpha + f[s_0] + (1 - \Sigma_0)w_0f'[s_0]$, $\lambda_2 = \alpha + f[s_0]$, and $w_{FF} = (1 - \Sigma_0)w_1f'[s_0]$. The matrix

$$\begin{pmatrix} -\lambda_1 & w_{FF} \\ \cdot & -\lambda_2 \end{pmatrix} \quad (2.141)$$

has eigenvalues $-\lambda_1$ and λ_2 . It is easy to see that w_0 small and positive (w_E slightly larger than w_I) implies that the eigenvalues are small and negative, so that the fixed point $\{\Sigma_0, 0\}$ is a weakly stable node. The effect of expression w_{FF} is referred to as *hidden feedforward* [56], hidden because the linear transformation of Eq. (2.137) is required to make it explicit, feedforward because fluctuations in Δ affect Σ but not vice-versa. The net effect of such terms is that small changes in η_Δ lead to larger changes in ξ_Σ , the magnitude of which increases with w_{FF} . This effect is

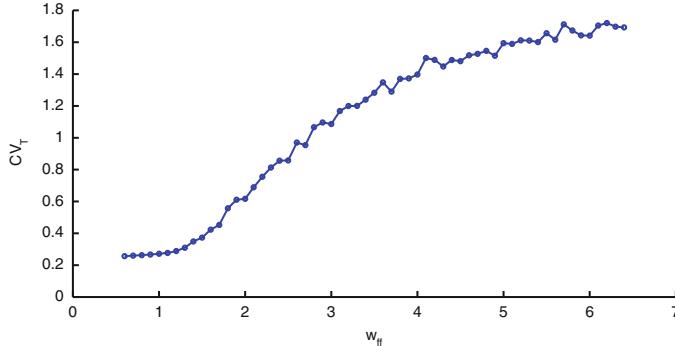


Fig. 2.11 Coefficient of variation (CV) plotted against feedforward weight w_{FF} , from simulations with other parameters fixed, $w_0 = 0.2$, $h = 0.001$ and $N = 800$. The time bin width is $T = 1 \text{ ms}$. Note that the feedforward weight depends linearly on w_1 (Reprinted from [9])

referred to as *balanced amplification* [56]. Evidently in a population subject to small fluctuations, large w_{FF} and small w_0 can generate large sustained fluctuations of ξ_{Σ} .

We can express the details of this effect more precisely by calculating the stationary value of $\text{var}[\tilde{\Sigma}]$ the variance of the total activity of the *E/I* population of N neurons as

$$\text{var}[\tilde{\Sigma}] = \frac{\text{var}[\xi_{\Sigma}]}{N} = \frac{\alpha \Sigma_0}{2N\lambda_1} \left[1 + \frac{w_{\text{FF}}^2}{\lambda_2(\lambda_2 + \lambda_1)} \right] \quad (2.142)$$

Evidently the fluctuations of $\tilde{\Sigma}$ increase with w_{FF} , and also increase as $\lambda_1, \lambda_2 \rightarrow 0$, so that $\text{var}[\tilde{\Sigma}] \propto (\lambda_1 N)^{-1}$, which is similar to the result of Eq. (2.128) for the one population case.

Figure 2.11 shows the result of a simulation in which the coefficient of variation of time-binned spike counts, a measure of $\text{var}[\tilde{\Sigma}] / \langle \tilde{\Sigma} \rangle$, as a function of w_{FF} .

It will be seen that the CV increases monotonically with w_{FF} . Thus the amplitude of fluctuations is determined in large part by large w_{FF} , and small eigenvalues. It is clear that these are the conditions which generate neural avalanches. To support this deduction, we carried out simulations of a network of 800 sparsely connected neurons, given the same balanced amplification conditions as in the previous cases. Figure 2.12 shows the results.

It will be seen that avalanches occur with an approximate power law distribution even in the spatially inhomogeneous sparse case, and even when the eigenvalues are small but non-zero negative, i.e. even when the fixed point $\{\Sigma_0, 0\}$ is a weakly stable node. Thus it may be the case that criticality is not a necessary condition for the appearance of approximate power laws in the statistics of spontaneous neural activity.

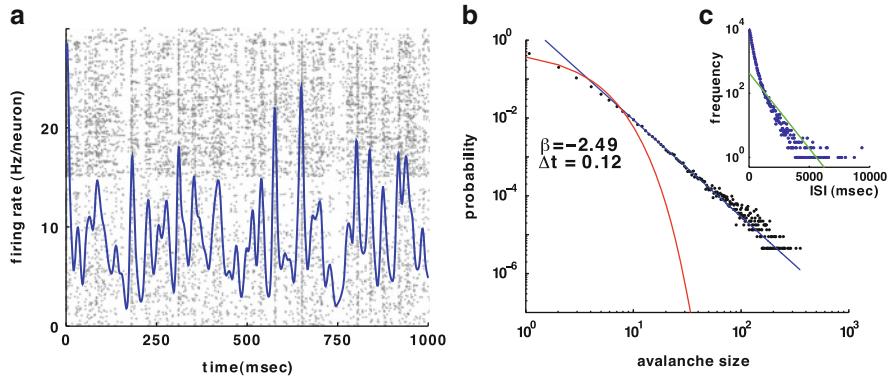


Fig. 2.12 Results from an excitatory and inhibitory network with $N = 800$, with 17 % connectivity. (a): Raster plot and mean firing rate (solid plot). (b): Avalanche size distribution. Lower curve: Poisson statistics. Blue line: Power law statistics with slope $\beta \approx -2.5$. (c): Inter-spike-interval distribution. Exponential statistics (Reprinted from [9])

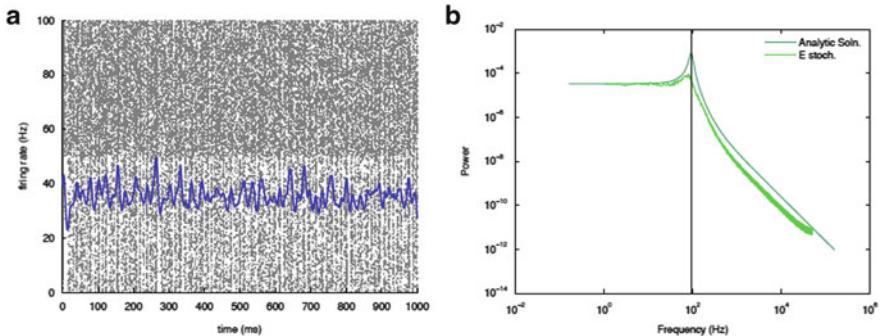


Fig. 2.13 Results from an excitatory and inhibitory network with $N = 800$, $h_E = 0.325$, $w_{EE} = w_{IE} = 1$, $w_{EI} = 1.6$, $w_{II} = 1$. (a): Raster plot and mean firing rate (solid plot) of the network. Lower half excitatory neurons, upper half inhibitory neurons. (b): Normalized power spectrum of the excitatory population in simulation (lower curve) and from Eq. (2.135) (upper curve). The peak occurs at a frequency $\omega_0 \approx 100$ Hz (Reprinted from [76])

2.8.2 Work in Progress

It is evident that there are many new problems to investigate. Among them are two problems concerning further effects of fluctuations and correlations on neural dynamics near weakly stable nodes and foci. Thus from Eq. (2.135) we can calculate the power spectrum of the total excitatory activity in a homogeneous E/I population in case the fixed point $\{E_0, I_0\}$ is a weakly stable focus [17, 76]. Figure 2.13 shows a plot of this spectrum. It will be seen that unlike the one population model, there is a pronounced peak in the spectrum on top of the

Brownian motion baseline. This raises the possibility that fluctuation driven finite size effects may play a role in the origin of brain oscillations or rhythms.

In fact it can also be shown that this property is present in the more general case of multiple populations of size N_r , $r = 1, 2, \dots, \Omega$, in which case there are multiple peaks with amplitudes proportional to $\text{RI}(\lambda_r)^2 N_r$.

As a final example we note another piece of work in progress. Consider Eqs. (2.126) and (2.127) containing $O(\frac{1}{N})$ terms, in which pair correlations have an effect. One can show that there exists a fixed point in such equations which is a weakly stable focus. Thus even a single homogeneous population of neurons can generate oscillations if the effects of pair correlations are non-negligible. In such a case, we can linearize these equations about such a fixed point to obtain solutions of the eigenvalue equation for its stability.

The result (Buice and Cowan In Preparation) is that

$$\lambda_{\pm} = -\frac{3}{2}\mu \pm \frac{1}{2} \left[\mu^2 + \frac{4f''w_0^3}{N}(f' + f''[s_0 - h]) \right]^{\frac{1}{2}} \quad (2.143)$$

where $\mu = \alpha - f'w_0$. These eigenvalues are complex and represent weakly damped oscillations. In the limit $\alpha \rightarrow \alpha_c = f'w_0$, corresponding to $\mu \rightarrow 0$, this formula reduces to the purely imaginary complex conjugate eigenvalues of a Hopf bifurcation to a limit cycle, i.e.

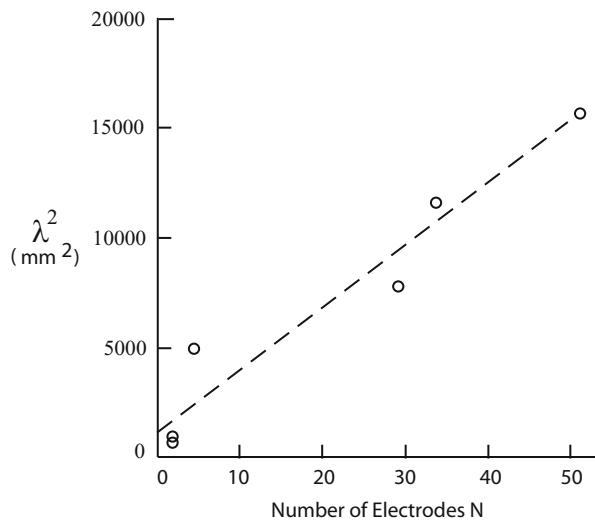
$$\lambda_{\pm} \rightarrow \lambda_{c,\pm} = \pm \frac{w_0^{3/2}}{\sqrt{N}} i \left[|f''|(f' + f''a_0w_0) \right]^{\frac{1}{2}} \quad (2.144)$$

Interestingly this is exactly the result we obtained in 2006 (unpublished) by calculating one-loop corrections to the propagator of the one population model, which we now know is equivalent to the system size expansion of the master equation. This result indicated something that may prove to be important, namely the *frequency* of the resulting oscillations scales inversely with \sqrt{N} , provided $f'' < 0$. This is a property that is not seen in the lower order fluctuation analysis. If we interpret such oscillations to be associated with some kind of periodic travelling front or wave-like neocortical disturbance, then we can infer that the *period* of such a travelling disturbance scales with \sqrt{N} . There is in fact some data which relates the wavelength of synchronous oscillations to \sqrt{N} . Figure 2.14 shows, for example, the relation between the wavelength λ of synchronous epileptiform oscillations as measured with surface EEG electrodes, and the number of electrodes N involved.

It will be seen that λ^2 is linearly proportional to N , and therefore the period τ of the oscillation (assuming simple harmonic waves) is linearly proportional to \sqrt{N} . The details can be seen in [14].

It seems plausible that fluctuations and correlations of neocortical activity near criticality play an important role in the generation of the various oscillations seen in such activity.

Fig. 2.14 Wavelength λ of synchronous EEG as a function of electrode number N . W. VanDongen (Personal Communication 2007) (Reprinted from [14])



2.9 Concluding Remarks

*The past is a foreign country:
they do things differently there.*
L.P. Hartley: *The Go-Between* (1953)

Evidently there is a great difference between the state of knowledge about neural networks in 1945, and what we now know today. Even when I started to think seriously about biological neural networks in 1962, apart from the earlier papers of Uttley and Beurle, there was not too much in the literature to provide guidance. Today the situation is different. Neuroscience has made great strides since 1945, and the modeling of biological neural networks is beginning to make much better contact with the experimental literature.

My initial idea about an analogy between neural networks and Lotka-Volterra systems seems to have stood the test of time. Currently there is a substantial amount of work on what has been called Stochastic Lattice Lotka-Volterra Systems (SLLVS), both using field theory [53, 54] and the van Kampen system-size expansion [51, 52], that closely parallels the work on neural networks described here. As we have shown, the two methodologies are closely related. However the system-size expansion breaks down at criticality, whereas the renormalization group method enables field theory to work even at criticality, and leads to the prediction of a universal phase transition in neural network dynamics, the DP phase transition, known to be equivalent to a branching and aggregating process. This in turn leads to the interpretation of the critical branching found in neural avalanches by Beggs and Plenz, as mean-field directed percolation, and to the use of the renormalized subcritical, critical, and supercritical propagators of DP as models of large-scale neural activity. Finally work on neural avalanches and oscillations using the system

size expansion has made it clear that in the stochastic Wilson-Cowan equations, proximity to a critical point either through a weakly stable node or focus is necessary for their generation. Both avalanches and such oscillations can therefore be seen as fluctuation driven finite-sized effects.

It seems clear to me that many problems involving the large-scale dynamics of neocortical activity can now be analyzed using these methodologies. My job is nearly done!

Acknowledgements I would like to acknowledge with thanks the financial support from ONR on two separate contracts, and that from the Grant Foundation and the James S. McDonnell Foundation, which greatly facilitated the development of this work. I will always appreciate the many contributions to this work made by my Postdoctoral Fellow and later colleague Hugh Wilson in the early days of this work, and the more recent contributions of my graduate students Toru Ohira, Michael Buice, Edward Wallace and Marc Benyoun, and of my collaborators Paul Bressloff and Wim van Drongelen, and of Nigel Goldenfeld and his graduate student Tom Butler at the University of Illinois, Urbana-Champaign, not to speak of the many other colleagues at the University of Chicago who, over the years since 1967, have provided much stimulation and constructive criticism.

References

1. Abarbanel, H., Broznan, J.: Structure of the vacuum singularity in Reggeon field theory. *Phys. Lett. B* **48**, 345–348 (1974)
2. Abarbanel, H., Broznan, J., Schwimmer, A., Sugar, R.: Intercept of the Pomeron. *Phys. Rev. D* **14**(2), 632–646 (1976)
3. Abarbanel, H., Broznan, J., Sugar, R., White, A.: Reggeon field theory: formulation and use. *Phys. Rep.* **21**(3), 119–182 (1975)
4. Alstrom, P.: Mean-field exponents for self-organized critical phenomena. *Phys. Rev. A* **38**(9), 4905–4906 (1988)
5. Amari, S.I.: A method of statistical neurodynamics. *Kybernetik* **14**, 201–215 (1974)
6. Amari, S.I., Yoshida, K., Kanatani, K.: A mathematical foundation for statistical neurodynamics. *SIAM J. Appl. Math.* **33**, 95–126 (1977)
7. Amati, D., Marchesini, G., Ciafoloni, M., Parisi, G.: Expanding disc as a dynamical vacuum instability in Reggeon field theory. *Nucl. Phys. B* **114**, 483–504 (1976)
8. Beggs, J., Plenz, D.: Neuronal avalanches in neocortical circuits. *J. Neurosci.* **23**(35), 11167–11177 (2003)
9. Benyoun, M., Cowan, J., van Drongelen, W., Wallace, E.: Avalanches in a stochastic model of spiking neurons. *PLoS Comput. Biol.* **6**(7), e1000846 (2010)
10. Beurle, R.L.: Properties of a mass of cells capable of regenerating pulses. *Philos. Trans. R. Soc. Lond. B* **240**(669), 55–94 (1956)
11. Bressloff, P.: Stochastic neural field theory and the system-size expansion. *SIAM J. Appl. Math.* **70**(5), 1488–1521 (2009)
12. Buice, M.: Neural networks, field theory, directed percolation, and critical branching. Ph.D. thesis, University of Chicago (2005)
13. Buice, M.A., Cowan, J.D.: Field theoretic approach to fluctuation effects for neural networks. *Phys. Rev. E* **75**, 051919 (2007)
14. Buice, M., Cowan, J.: Statistical mechanics of the neocortex. *Prog. Biophys. Theor. Biol.* **99**(2,3), 53–86 (2009)

15. Buice, M.A., Cowan, J.D., Chow, C.C.: Systematic fluctuation expansion for neural network activity equations. *Neural Comput.* **22**(2), 377–426 (2010)
16. Burns, B.D.: Some properties of isolated cerebral cortex in the unanaesthetized cat. *J. Physiol.* **112**, 156–175 (1951)
17. Butler, T., Benayoun, M., Wallace, E., van Drongelen, W., Goldenfeld, N., Cowan, J.: Evolutionary constraints on visual cortex architecture from the dynamics of hallucinations. *PNAS* **109**(2), 606–609 (2012)
18. Cai, D., Tao, L., Shelley, M., McLaughlin, D.: An effective kinetic representation of fluctuation-driven neuronal networks with application to simple and complex cells in visual cortex. *Proc. Nat. Acad. Sci. USA* **101**(20), 7757–7762 (2004)
19. Cardy, J., Sugar, R.: Directed percolation and Reggeon field theory. *J. Phys. A* **13**, L423–L427 (1980)
20. Chen, L.Y., Goldenfeld, N., Oono, Y.: Renormalization group and singular perturbations: multiple scales, boundary layers, and reductive perturbation theory. *Phys. Rev. E* **54**(1), 376–394 (1996)
21. Cowan, J.: A mathematical theory of central nervous activity. Ph.D. thesis, Imperial College of Science and Technology, University of London (1967)
22. Cowan, J.: Statistical mechanics of nervous nets. In: Caianiello E.R. (ed.) *Neural Networks*, pp. 181–188. Springer, Berlin/Heidelberg/New York (1968)
23. Cowan, J.: A statistical mechanics of nervous activity. In: Gerstenhaber, M. (ed.) *Some Mathematical Questions in Biology*, vol. 2. American Mathematical Society, Providence (1971)
24. Cowan, J.D.: Stochastic neurodynamics. In: Touretzsky, D., Lippman, R., Moody, J. (eds.) *Advances in Neural Information Processing Systems*, vol. 3, pp. 62–68. Morgan Kaufmann, San Mateo (1991)
25. Cragg, B., Temperley, H.: The organization of neurones: a cooperative analogy. *Electroencephalogr. Clin. Neurophysiol.* **6**, 85–92 (1954)
26. Davidsen, J., Schuster, H.: Simple model for $1/f^\alpha$ noise. *Phys. Rev. E* **65**(2), 026120–026123 (2002)
27. Destexhe, A., Sejnowski, T.: The Wilson-Cowan model, 36 years later. *Biol. Cybern.* **101**(1), 1–2 (2009)
28. Dickman, R.: Nonequilibrium lattice models: series analysis of steady states. *J. Stat. Phys.* **55**(5–6), 997–1026 (1989)
29. Doi, M.: Second quantization representation for classical many-particle system. *J. Phys. A Math. Gen.* **9**(9), 1465–1477 (1976)
30. Doi, M.: Stochastic theory of diffusion controlled reaction. *J. Phys. A Math. Gen.* **9**(9), 1479–1495 (1976)
31. Fetter, A., Walecka, J.: *Quantum Theory of Many-Particle Systems*. McGraw-Hill, New York (1971)
32. Feynman, R.P.: Space–time approach to non–relativistic quantum mechanics. *Rev. Mod. Phys.* **20**(2), 367–387 (1948)
33. Georgi, H.: *Lie Algebras in Particle Physics*. Benjamin Books, Menlo Park (1982)
34. Gillespie, D.: The chemical Langevin equation. *J. Chem. Phys.* **113**(1), 297–306 (2000)
35. Glauber, R.: Photon correlations. *Phys. Rev. Lett.* **10**(3), 84–86 (1963)
36. Grassberger, P., de la Torre, A.: Reggeon field theory (Schlögle's first model) on a lattice: Monte carlo calculations of critical behaviour. *Ann. Phys.* **122**, 373–396 (1979)
37. Grassberger, P., Scheunert, M.: Fock–space methods for identical classical objects. *Fortschr. Phys.* **28**, 547–578 (1980)
38. Gribov, V.: A Reggeon diagram technique. *Sov. Phys. JETP* **26**, 414 (1968)
39. Hecht, K.: *The Vector Coherent State Method*. Springer, New York (1987)
40. Hopfield, J.: Neural networks and physical systems with emergent collective computational properties. *Proc. Nat. Acad. Sci.* **79**, 2554–2558 (1982)
41. Hoppensteadt, F., Izhikevich, E.: *Weakly Connected Neural Networks*. Springer, New York (1997)

42. Kerner, E.: A statistical mechanics of interacting species. *Bull. Math. Biophys.* **19**, 121–146 (1957)
43. Kirkwood, J.: The statistical mechanical theory of transport processes i: general theory. *J. Chem. Phys.* **14**(2), 180–201 (1946)
44. Kobayashi, T.: f^{-1} series generated by using the branching process model. *J. Phys. A* **22**, L931–L938 (1989)
45. Kramers, H.: Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* **7**(4), 284–304 (1940)
46. Lampl, I., Reichova, I., Ferster, D.: Synchronous membrane potential fluctuations in neurons of the cat visual cortex. *Neuron* **22**, 361–374 (1999)
47. Liggett, T.: *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*. Springer, New York (1985)
48. Little, W.: The existence of persistent states in the brain. *Math. Biosci.* **19**, 101–120 (1974)
49. Lotka, A.: *Elements of Physical Biology*. Williams and Wilkins Co., Baltimore (1925)
50. McCulloch, W., Pitts, W.: A logical calculus of ideas immanent in nervous tissue. *Bull. Math. Biophys.* **5**, 115–133 (1943)
51. McKane, A., Newman, T.: Stochastic models in population biology and their deterministic analogs. *Phys. Rev. E* **70**(4), 41902 (2004)
52. McKane, A., Newman, T.: Predator-prey cycles from resonant amplification of demographic stochasticity. *Phys. Rev. Lett.* **94**(21), 218102 (2005)
53. Mobilia, M., Georgiev, I., Täuber, U.: Fluctuations and correlations in lattice models of predator-prey interactions. *Phys. Rev. E* **73**(4), 04093–04096 (2006)
54. Mobilia, M., Georgiev, I., Täuber, U.: Phase transitions and spatio-temporal fluctuations in stochastic Lotka-Volterra models. *J. Stat. Phys.* **123**(6), 1061–1097 (2006)
55. Moyal, J.: Stochastic processes and statistical physics. *J. R. Stat. Soc. B* **11**(2), 150–210 (1949)
56. Murphy, B., Miller, K.: Balanced amplification: a new mechanism of selective amplification of neural activity patterns. *Neuron* **61**(4), 635–648 (2009)
57. Nauhaus, I., Busse, L., Carandini, M., Ringach, D.: Stimulus contrast modulates functional connectivity in visual cortex. *Nat. Neurosci.* **12**(1), 70–76 (2009)
58. Ohira, T.: A master equation approach to stochastic neurodynamics. Ph.D. thesis, University of Chicago (1993)
59. Ohira, T., Cowan, J.: Stochastic neurodynamics and the system size expansion. In: Ellacort, S., Anderson, I. (eds.) *Proceedings of the First International Conference on the Mathematics of Neural Networks*, pp. 290–294. Kluwer, Boston (1997)
60. Pearl, R.: The growth of populations. *Q. Rev. Biol.* **2**, 532–548 (1927)
61. Peliti, L.: Path integral approach to birth-death processes on a lattice. *J. Phys.* **46**, 1469–1483 (1985)
62. Perelomov, A.: *Generalized Coherent States and Their Applications*. Springer, New York (1986)
63. Rose, H.: Renormalised kinetic theory for nonequilibrium many-particle classical systems. *J. Stat. Phys.* **20**(4), 415–447 (1980)
64. Rosenblatt, F.: The perceptron, a probabilistic model for information storage and organization in the brain. *Psychol. Rev.* **62**, 386–408 (1958)
65. Rowe, D., Robinson, P., Rennie, C.: Estimation of neurophysiological parameters from the waking EEG using a biophysical model of brain dynamics. *J. Theor. Biol.* **231**, 413–433 (2004)
66. Rumelhart, D., Hinton, G., Williams, R.: Learning internal representations by error propagation. In: Rumelhart, D., McClelland, J. (eds.) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. I: foundations*. MIT, Cambridge (1986)
67. Schulman, L.: *Techniques and Applications of Path Integration*. Wiley, New York (1981)
68. Schwinger, J.: On angular momentum. In: Biedenharn, L., van Dam, H. (eds.) *Quantum Theory of Angular Momentum*, pp. 229–279. Academic, New York (1965)
69. Smith, D., Davidson, C.: Maintained activity in neural nets. *J. ACM* **9**(2), 268–279 (1962)

70. Stevens, C.: How cortical interconnectedness varies with network size. *Neural Comput.* **1**, 473–479 (1989)
71. Uttley, A.: The probability of neural connexions. *Proc. R. Soc. Lond. B* **144**(915), 229–240 (1955)
72. Uttley, A.: A theory of the mechanism of learning based on the computation of conditional probabilities. In: Proceedings of the 1st International Congress on Cybernetics, Namur, pp. 830–856. Gauthier-Villars, Paris (1956)
73. Van Kampen, N.: Stochastic Processes in Physics and Chemistry. North Holland, Amsterdam/New York (1981)
74. Verhulst, P.: Recherches mathématiques sur la loi d'accroissement de la population. *Mem. Acad. R. Belg.* **18**, 1–38 (1845)
75. Volterra, V.: Leçons sur la Théorie Mathématique de la Lutte pour la Vie. Gauthier-Villars, Paris (1931)
76. Wallace, E., Benayoun, M., van Drongelen, W., Cowan, J.: Emergent oscillations in networks of stochastic spiking neurons. *PLoS ONE* **6**(5), e14804 (2011)
77. Wiener, N.: Nonlinear Problems in Random Theory. MIT, Cambridge (1958)
78. Wilson, K.: Renormalization group and critical phenomena. I. Renormalization group and the Kadanoff scaling picture. *Phys. Rev. B* **4**(9), 3174–3183 (1971)
79. Wilson, H., Cowan, J.: Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.* **12**, 1–22 (1972)
80. Wilson, H., Cowan, J.: A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* **13**, 55–80 (1973)

Chapter 3

Heaviside World: Excitation and Self-Organization of Neural Fields

Shun-ichi Amari

Abstract Mathematical treatments of the dynamics of neural fields become much simpler when the Heaviside function is used as an activation function. This is because the dynamics of an excited or active region reduce to the dynamics of the boundary. We call this regime the Heaviside world. Here, we visit the Heaviside world and briefly review bump dynamics in the 1D, 1D two-layer, and 2D cases. We further review the dynamics of forming topological maps by self-organization. The Heaviside world is useful for studying the learning or self-organization equation of receptive fields. The stability analysis shows the formation of a continuous map or the emergence of a block structure responsible for columnar microstructures. The stability of the Kohonen map is also discussed.

3.1 Introduction

The dynamics of excitations in a neural field, first proposed by Wilson and Cowan [18], are described by nonlinear partial integro-differential equations. The dynamics include rich phenomena, but sophisticated mathematical techniques are required for solving the equations. Amari [3] analyzed dynamical behavior rigorously by using the Heaviside activation function and showed the existence and stability of a bump solution as well as a traveling bump solution. The Heaviside activation function, instead of a general sigmoid function, makes it possible to analyze the dynamics and to obtain explicit solutions. This framework is called the Heaviside world.

The dynamics of excitation patterns can be reduced to much simpler dynamics of the boundaries of an excitation region [3] in the Heaviside world. In a one-dimensional field, the boundaries of a simple excitation pattern consist of two points, and hence, their dynamics can be described by ordinary differential equations.

S. Amari (✉)

RIKEN Brain Science Institute, Hirosawa 2-1, Wako-shi, Saitama 351-0198, Japan

e-mail: amari@brain.riken.jp

The boundaries consist of a closed curve in the 2D case. The dynamical behavior of a curve is not so simple but it is still much simpler than the original field equation. See Coombes, Schmit and Bojak [7] and Bressloff [5] for detailed mathematical techniques in the 2D case.

The present Chapter will briefly review the results of pattern dynamics in the Heaviside world. It further demonstrates the dynamics of learning (self-organization) in a neural field, and it elucidates the mechanism of formation of a topological map to fit the environmental information. A model for forming a topological map was proposed by Willshaw and von der Malsburg [16], and it was analyzed by Takeuchi and Amari [15] (see also Amari [4]). Kohonen [11] proposed an engineering model for forming a topological map. We show that the Heaviside world works even in this situation and is applicable to a Kohonen-type map, as was studied in Kurata [12].

3.2 Dynamics of Excitation in a Homogeneous Neural Field

3.2.1 1D 1-Layer Field

We begin with a simplest case of a 1D neural field X with one layer. Let x be a position coordinate of the field. The dynamics are described by

$$\tau \frac{\partial u(x, t)}{\partial t} = -u(x, t) + \int w(x - x') f[u(x', t)] dx' + s(x, t). \quad (3.1)$$

Here, $u(x, t)$ is the average membrane potential of neurons at a position x at time t , $w(x, x')$ is the synaptic connection weight from a position x' to x , $f(u)$ is the activation function such that $z = f(u)$ is the output of neurons at x and $s(x, t)$ is the external stimuli applied to x at t . A threshold is included in the term of external stimuli. The Heaviside world assumes that the activation function is the Heaviside function:

$$f(u) = \begin{cases} 1, & u > 0, \\ 0, & u \leq 0. \end{cases} \quad (3.2)$$

We further assume that the field is homogeneous and the connections are symmetric:

$$w(x, x') = w(|x - x'|). \quad (3.3)$$

First, we shall study the case in which the external stimuli $s(x, t)$ is a constant.

Given $u(x, t)$, the active region of X is defined by

$$A(t) = \{x | u(x, t) > 0\}. \quad (3.4)$$

The equation is then rewritten as

$$\tau \frac{\partial u(x, t)}{\partial t} = -u + \int_A w(x - x') dx' + s. \quad (3.5)$$

It is not easy to solve the equation even in this simplified case. We assume a bump solution such that $u(x, t)$ is positive only in an interval $[x_1, x_2]$; that is, the active or excited region is

$$A(t) = [x_1(t), x_2(t)]. \quad (3.6)$$

Note that the boundary points of $A(t)$ satisfy

$$u\{x_i(t), t\} = 0, \quad i = 1, 2. \quad (3.7)$$

By differentiating this with respect to t , we get

$$\alpha_i \frac{dx_i(t)}{dt} + \frac{\partial u(x_i, t)}{\partial t} = 0, \quad (3.8)$$

where

$$\alpha_i = \frac{\partial u(x_i, t)}{\partial x}. \quad (3.9)$$

Therefore, the dynamics of the boundaries are described by

$$\tau \frac{dx_i}{dt} = -\frac{1}{\alpha_i} \left(\int_A w(x_i - x') dx' + s \right) \quad (3.10)$$

Let us define

$$W(x) = \int_0^x w(x') dx'. \quad (3.11)$$

Thus, we find

$$\int_A w(x_i - x') dx' = \int_{x_1}^{x_2} w(x_i - x') dx' \quad (3.12)$$

$$= W(x_2 - x_1), \quad (3.13)$$

so that

$$\frac{\partial u(x_i, t)}{\partial t} = \frac{1}{\tau} \{W(x_2 - x_1) + s\}. \quad (3.14)$$

The equations of the boundaries have a simple expression:

$$\frac{dx_i(t)}{dt} = -\frac{1}{\tau\alpha_i} \{W(x_2 - x_1) + s\}. \quad (3.15)$$

Here, α_1 and α_2 are the slopes of the waveform $u(x, t)$ at x_1 and x_2 , and hence,

$$\alpha_1 > 0, \quad \alpha_2 < 0. \quad (3.16)$$

Since α_i are variables depending on the waveform, (3.15) is not a closed expression of the boundaries $x_i(t)$. However, the Heaviside world tells us lots of information on its dynamics.

An equilibrium solution, if it exists, satisfies

$$W(x_2 - x_1) + s = 0. \quad (3.17)$$

Note that the width of the active region is

$$a(t) = x_2(t) - x_1(t). \quad (3.18)$$

The dynamics have a simple form,

$$\frac{da(t)}{dt} = \frac{1}{\tau\alpha} \{W(a) + s\} \quad (3.19)$$

where

$$\frac{1}{\alpha} = \frac{1}{\alpha_1} - \frac{1}{\alpha_2} > 0. \quad (3.20)$$

The equilibrium solution a satisfies

$$W(a) + s = 0. \quad (3.21)$$

Moreover, by considering the variational equation of (3.19), we see that it is stable when and only when

$$w(a) = W'(a) < 0. \quad (3.22)$$

The waveform of a stable bump solution is explicitly obtained from (3.5) as

$$u(x) = \int_0^a w(x - x') dx' + s = W(x) + W(a - x) + s. \quad (3.23)$$

Obviously, when $u(x)$ is a stable bump solution, $u(x - c)$ is also a stable bump solution for any constant c . Hence, stable solutions form a one-dimensional set of solutions $\{u(x - c)\}$, which is a line attractor [14].

We studied the case in which s is constant. When s is not constant but has a spatial distribution $s(x)$, a bump moves toward its maximum. Let us assume that an external stimulus $s(x)$ is suddenly applied and consider how the stable bump moves. The velocity of a bump is described by the equation of motion of its center,

$$\frac{1}{2} \frac{d}{dt} (x_1 + x_2) \approx \frac{1}{\tau \alpha_1} \{s(x_2) - s(x_1)\}, \quad (3.24)$$

which depends on the intensity of external stimuli at $s(x_1)$ and $s(x_2)$. The tracking ability of a bump has been analyzed with detail in another solvable model [10, 19], where the Hermite world is used instead of the Heaviside world.

3.2.2 1D Field with Two Layers

We can easily study the equation of a field with two layers, one excitatory and one inhibitory. The equations are

$$\begin{aligned} \tau_E \frac{\partial u_E(x, t)}{\partial \tau} &= -u_E(x, t) + w_{EE} * f[u_E] \\ &\quad - w_{EI} * f[u_I] + s_E, \end{aligned} \quad (3.25)$$

$$\tau_I \frac{\partial u_I(x, t)}{\partial t} = -u_I(x, t) + w_{IE} * f[u_E] - w_{II} * f[u_I] + s_I, \quad (3.26)$$

where $u_E(x, t)$ and $u_I(x, t)$ are the potentials of excitatory and inhibitory layers, respectively, and the connection weights are $w_{EE}(x)$, $w_{EI}(x)$, $w_{IE}(x)$ and $w_{II}(x)$, depending on the originating layer and terminating layer of excitation and inhibition, with a convolution operator $*$, such as

$$w_{EE} * f[u_E] = \int w_{EE}(x - x') f[u_E(x')] dx'. \quad (3.27)$$

We can analyze the dynamics of the boundary points of active regions in the excitatory and inhibitory layers in a similar manner to obtain a bump solution.

We shall first consider a uniform oscillatory solution in the Heaviside world. The uniform solutions, $u_E(t)$ and $u_I(t)$, do not depend on the position x , so that the equations are simple ordinary differential equations

$$\tau_E \frac{du_E(t)}{dt} = -u_E + W_{EE}f(u_E) - W_{EI}f(u_I) + s_E, \quad (3.28)$$

$$\tau_I \frac{du_I(t)}{dt} = -u_I + W_{IE}f(u_E) + s_I, \quad (3.29)$$

where W_{EE} etc. are, for example,

$$W_{EE} = \int_{-\infty}^{\infty} W_{EE}(x) dx. \quad (3.30)$$

We put $W_{II} = 0$ for simplicity's sake. The equations show the population dynamics of excitatory and inhibitory neuron pools proposed by Amari [1, 2] and Wilson and Cowan [18] as a model of neural oscillators. The state space is $\mathbf{u} = (u_E, u_I)$. When we use the Heaviside world, we have piecewise linear equations, linear in each quadrant of \mathbf{u} , determined by $u_E \geq 0$, and $u_I \geq 0$.

The equations can be written as

$$\tau \frac{d\mathbf{u}}{dt} = -\mathbf{u} + \bar{\mathbf{u}}_k, \quad k = I, II, III, IV \quad (3.31)$$

in the k -th quadrant, where

$$\bar{\mathbf{u}}_I = \mathbf{s} + \begin{pmatrix} W_{EE} - W_{EI} \\ W_{IE} \end{pmatrix}, \quad I : u_E > 0, u_I > 0, \quad (3.32)$$

$$\bar{\mathbf{u}}_{II} = \mathbf{s} + \begin{pmatrix} -W_{EI} \\ 0 \end{pmatrix}, \quad II : u_E < 0, u_I > 0, \quad (3.33)$$

$$\bar{\mathbf{u}}_{III} = \mathbf{s}, \quad III : u_E < 0, u_I < 0, \quad (3.34)$$

$$\bar{\mathbf{u}}_{IV} = \mathbf{s} + \begin{pmatrix} W_{EE} \\ W_{IE} \end{pmatrix}, \quad IV : u_E > 0, u_I < 0, \quad (3.35)$$

where

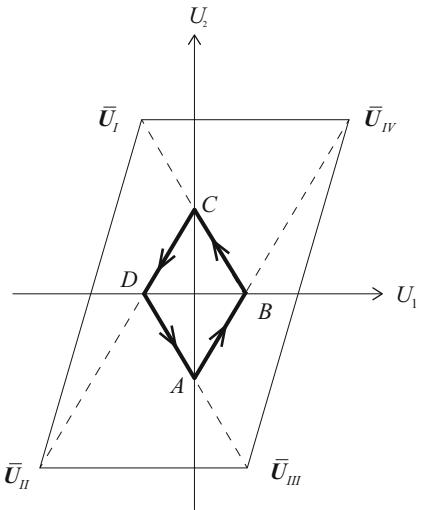
$$\mathbf{s} = \begin{pmatrix} s_E \\ s_I \end{pmatrix}. \quad (3.36)$$

In the k -th quadrant, the dynamical flow is linear converging to $\bar{\mathbf{u}}_k$, but $\bar{\mathbf{u}}_k$ is not necessarily in the k -th quadrant (see Fig. 3.1). The direction of the flow changes when $\mathbf{u}(t)$ intersects the coordinate axes and enters another quadrant. The dynamical behaviors depend on the positions of $\bar{\mathbf{u}}_k$. They are monostable, bistable, and oscillatory, as studied in Amari [1, 2] and Wilson and Cowan [18]. We show the existence and stability of an oscillatory solution in the Heaviside world.

Theorem 1. *A stable oscillation exists when $W_{EE} < W_{EI}$ for an adequate \mathbf{s} such that the parallelepiped $\bar{\mathbf{u}}_I \bar{\mathbf{u}}_{II} \bar{\mathbf{u}}_{III} \bar{\mathbf{u}}_{IV}$ encircles the origin.*

The existence and stability are clearly shown in Fig. 3.1. In this case $\bar{\mathbf{u}}_k$ does not exist in the k -th quadrant and there are no equilibrium states. The oscillatory solution is hence globally stable.

Fig. 3.1 Limit cycle ABCD of oscillation in the two-layers 1D field



In addition to a stationary bump solution, a moving bump solution exists in this field, where such a solution does not exist in a one-layer field. By using moving coordinates, we get

$$\xi = x - ct, \quad (3.37)$$

where c is the velocity of the bump, we have an explicit solution of a moving bump when certain conditions are satisfied. It was analyzed in Amari [3] so that we will not describe it here. The existence of a breathing solution shows the richness of the solutions to these equations. The Heaviside world makes it easier to analyze bump solutions [8, 9].

3.2.3 2D Field of Neural Excitation

The equation of a one layer 2D field is described as

$$\tau \frac{\partial u(\mathbf{x}, t)}{\partial t} = -u(\mathbf{x}, t) + \int w(\mathbf{x} - \mathbf{x}') f[u(\mathbf{x}', t)] d\mathbf{x}' + s, \quad (3.38)$$

where $\mathbf{x} = (x_1, x_2)$ is the coordinates of the field and $w(\mathbf{x})$ is a radially symmetric connection function. Let $A(t)$ be an active region in the Heaviside world on which $u(\mathbf{x}, t) > 0$. Let \mathbf{x}_A be a point on the boundary of A . It satisfies

$$u(\mathbf{x}_A, t) = 0. \quad (3.39)$$

Let us denote the gradient of a waveform by

$$\boldsymbol{\alpha} = \frac{\partial u(\mathbf{x}, t)}{\partial \mathbf{x}}. \quad (3.40)$$

Then, by differentiating (3.39), we arrive at an equation which describes the motion of the boundary of the excited region:

$$\boldsymbol{\alpha} \cdot \frac{d\mathbf{x}_A}{dt} = -\frac{\partial u(\mathbf{x}_A, t)}{\partial t} = -\frac{1}{\tau} \left[\int_A w(\mathbf{x}_A - \mathbf{x}') d\mathbf{x}' + s \right]. \quad (3.41)$$

The equilibrium solution satisfies

$$\int_A w(\mathbf{x}_A - \mathbf{x}') d\mathbf{x}' + s = 0. \quad (3.42)$$

The equilibrium solution having an active region A is written as

$$u(\mathbf{x}) = \int_A w(\mathbf{x} - \mathbf{x}') d\mathbf{x}' + s. \quad (3.43)$$

When $w(\mathbf{x})$ is a radially symmetric function, the radially symmetric equilibrium solution of radius a satisfies

$$u(\mathbf{x}) > 0, \quad |\mathbf{x}| < a, \quad (3.44)$$

$$u(\mathbf{x}) < 0, \quad |\mathbf{x}| > a. \quad (3.45)$$

The equilibrium radius a is obtained from

$$\tilde{W}(a) + s = 0, \quad (3.46)$$

where

$$\tilde{W}(a) = \int_0^{2a} 2rw(r) \cos^{-1} \frac{r}{2a} dr. \quad (3.47)$$

However, its stability condition is not easy to determine, because the variational equation for stability has freedom of deformation of A not only in the radial direction, but also of the shape of the boundary circle. The stability was first analyzed in Amari, a Japanese book. See Bressloff and Coombes [6] for later developments. We use polar coordinates (r, θ) to write down the variation of the excited region. Let us write

$$u(r, \theta, t) = u(r, \theta) + \varepsilon v(\theta, t), \quad (3.48)$$

where $u(r, \theta)$ is the equilibrium solution, ε is a small constant and $v(\theta, t)$ denotes the shape of the variation. We use the Fourier expansion of $v(\theta, t)$. The stability condition is given by

$$\lambda_n = \frac{1}{c} \int_0^{2\pi} w \left(2a \left| \cos \frac{\theta}{2} \right| \right) \cos n\theta d\theta - 1 < 0, \quad n = 0, 2, 3, \dots . \quad (3.49)$$

where

$$c = \int_0^{2\pi} w \left(2a \left| \cos \frac{\theta}{2} \right| \right) \cos \theta d\theta. \quad (3.50)$$

We can extend the idea to a two-layer field having rich dynamical phenomena such as multiple bumps, spiral waves, breathing waves and others. See Chap. 7 in this book. Lu, Sato, and Amari [13] studied traveling bumps and their collisions in a 2D field by simulation. However, the analysis is difficult even in the Heaviside world. See Coombes, Schmidt and Bojak [7] for further developments. Wu and colleagues [10, 19] use another technique of total inhibition, which uses a Gaussian approximation of the bump's shape and its Hermite expansion. This leads us to another world called the Hermite world.

3.3 Self-Organization of Neural Fields

3.3.1 Field Model of Self-Organization

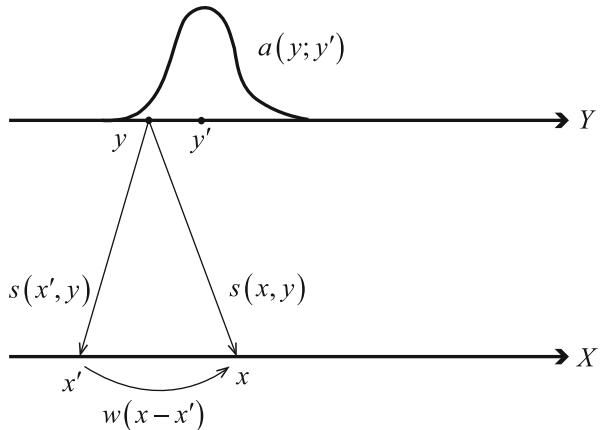
We shall study the self-organization of a 1D homogeneous neural field X which receives inputs from another 1D neural field Y . Fields X and Y have position coordinates x and y , respectively. The input signal to X is assumed to be a bump solution of Y . The neural field X receives stimuli from a bump signal $a(y)$ concentrated around a position y , and they induce a bump solution $u(x)$ of X centered at x . This establishes a correspondence of the positions y and x of two fields Y and X (Fig. 3.2). When Y is a retinal field responsible for the external light stimuli and X is a visual cortex, both being 2D, the correspondence is called a retinotopic map. The map is generated by self-organization of neural fields.

We assume that a bump of Y centered at y' is composed of the activities of neurons at y ,

$$a(y; y') = a(y - y'), \quad (3.51)$$

where $a(y)$ is a unimodal waveform of a bump solution. The activities $a(y; y')$ of Y stimulate neurons of X . The connection weight from the position y of Y to the position x of X is written as $s(x, y)$. We also assume that neurons at x receive an inhibitory input of constant intensity a_0 with a synaptic weight $s_0(x)$.

Fig. 3.2 Self-organization of field X by receiving stimuli from field Y



Hence, the total amount of input stimuli given to the neurons at x caused by $a(y; y')$ is written as

$$S(x, y') = \int s(x, y)a(y; y') dy - s_0(x)a_0. \quad (3.52)$$

The dynamics of excitation in field X is described as follows: Given an input $a(y; y')$, the neurons at x calculate the inputs $S(x, y')$ by using (3.52) and $u(x, t)$ changes subject to the dynamics,

$$\tau \frac{\partial u(x, t)}{\partial t} = -u(x, t) + w * f[u] + S(x, y'). \quad (3.53)$$

The inputs $S(x, y')$ depend on the connection weights $s(x, y)$ and $s_0(x)$. The connection weights are modified in the process of neural activation. This modification is learning or self-organization. We assume a Hebb type of learning rule: Synaptic weight $s(x, y)$ increases in proportion to the input $a(y; y')$ when the neuron at x fires and decays with a small time constant. We also assume that a Hebbian rule applies to an inhibitory synapse. For neurons at x , the learning rule is written as

$$\tau' \frac{\partial s(x, y)}{\partial t} = -s(x, y) + cf[u(x, t)]a(y; y'), \quad (3.54)$$

$$\tau' \frac{\partial s_0(x)}{\partial t} = -s_0(x) + c'f[u(x, t)]a_0. \quad (3.55)$$

Here, τ' is a time constant which is much larger than that of the neural excitation, and c and c' are different constants. We also assume that the inhibitory input a_0 is always constant.

A bump excitation of Y randomly appears around the position y' . Let $p(y')$ be the probability density of a bump appearing at y' . A bump $a(y; y')$ continues for a

duration which is sufficiently large compared with τ for forming a stable bump in X , but sufficiently small compared with the time constant τ' of learning. Hence, we use the adiabatic approximation and consider the following equation of learning,

$$\tau' \frac{\partial s(x, y)}{\partial t} = -s(x, y) + cf[U(x, y')]a(y; y'), \quad (3.56)$$

$$\tau' \frac{\partial s_0(x)}{\partial t} = -s_0(x) + c'f[U(x, y')]a_0, \quad (3.57)$$

where the current excitation $u(x, t)$ is replaced by the stationary state $U(x, y')$ given rise to by input $a(y; y')$,

$$U(x, y') = \int w(x - x') f[U(x', y')] dx' + S(x, y'). \quad (3.58)$$

Considering that τ' is large, before $s(x, y)$ and $s_0(x)$ change substantially, a number of bumps $a(y; y')$ at various y' are randomly chosen. Hence, we further use an averaging approximation for stochastic choices of y' . Let $\langle \cdot \rangle$ denote the average over all possible input bumps $a(y, y')$:

$$\langle f[U(x, y')]a(y; y') \rangle = \int p(y') f[U(x, y')]a(y; y') dy'. \quad (3.59)$$

Accordingly, we get the fundamental equations of learning,

$$\tau' \frac{\partial s(x, y)}{\partial t} = -s(x, y) + c\langle f[U(x, y')]a(y; y') \rangle, \quad (3.60)$$

$$\tau' \frac{\partial s_0(x)}{\partial t} = -s_0(x) + c'\langle f[U(x, y')] \rangle a_0. \quad (3.61)$$

We thus have two important field quantities $U(x, y')$ and $S(x, y')$, both of which depend on $s(x, y)$ and $s_0(x)$ and are hence modified by learning. By differentiating (3.52) and substituting (3.56) and (3.57), we obtain the dynamic equation describing the change of $S(x, y)$,

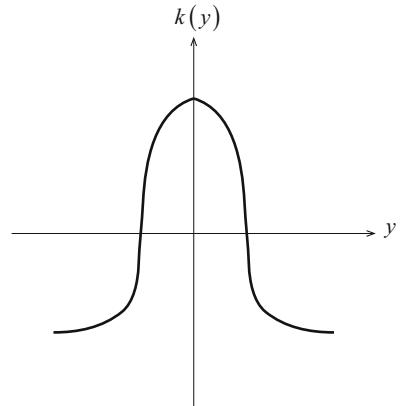
$$\tau' \frac{\partial S(x, y; t)}{\partial t} = -S(x, y; t) + \int k(y - y') f[U(x, y')] p(y') dy', \quad (3.62)$$

where we put

$$k(y - y') = c \int a(y'' - y) a(y'' - y') dy'' - c'a_0. \quad (3.63)$$

This term shows how two stimuli centered at y and y' overlap. The topology of Y is represented in it (see Fig. 3.3).

Fig. 3.3 Shape of $k(y)$ defined in (3.63)



Equation (3.62) can be written as

$$\tau' \frac{\partial S(x, y')}{\partial t} = -S + k \circ f[U], \quad (3.64)$$

where \circ is another convolution operator defined by

$$k \circ f[U] = \int p(y') k(y - y') f[U(x, y')] dy'. \quad (3.65)$$

3.3.2 Dynamics of the Receptive Field

An equilibrium solution $U(x, y')$ is determined by (3.58), depending on $s(x, y')$ and $s_0(x)$. Let A be a region on $X \times Y$ such that

$$A = \{(x, y) \mid U(x, y) > 0\}. \quad (3.66)$$

The receptive field $R(x)$ of a neuron at x is a region of Y such that

$$U(x, y) > 0, \quad y \in R(x). \quad (3.67)$$

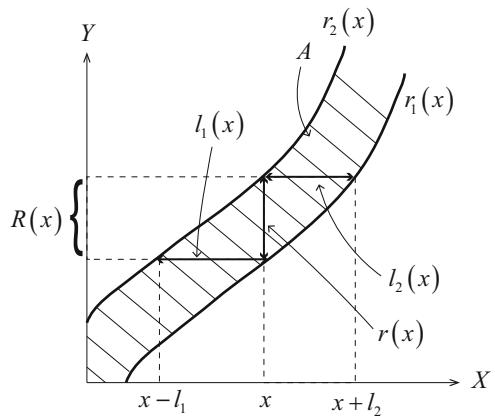
That is, neurons at x are excited by an input bump around y . We assume that it is an interval,

$$R(x) = [r_1(x), r_2(x)], \quad (3.68)$$

so that A is bounded by two lines $y = r_1(x)$ and $y = r_2(x)$ (see Fig. 3.4). The size of the receptive field is

$$r(x) = r_2(x) - r_1(x). \quad (3.69)$$

Fig. 3.4 Boundary lines of the active region A in $X-Y$ plane, which represents receptive fields



In contrast, when a bump around y is used, the active region is

$$A(y) = [x_1(y), x_2(y)], \quad (3.70)$$

where

$$x_1(y) = r_2^{-1}(y), \quad x_2(y) = r_1^{-1}(y). \quad (3.71)$$

The length of the active region of X is

$$\bar{x}(y) = r_2^{-1}(y) - r_1^{-1}(y). \quad (3.72)$$

We use the following notations: For an input bump at around $y = r_1(x)$, the excited region is $[x - l_1, x]$ and for an input bump at around $y = r_2(x)$, the excited region is $[x, x + l_2]$. This implies

$$l_1 = x - r_2^{-1}\{r_1(x)\}, \quad (3.73)$$

$$l_2 = r_1^{-1}\{r_2(x)\} - x. \quad (3.74)$$

The equilibrium $U(x, y)$ of (3.58) changes as a result of the change in $S(x, y; t)$ or $s(x, y; t)$ and $s_0(x; t)$. By differentiating (3.58) and using (3.64), we get

$$\tau' \frac{\partial U(x, y; t)}{\partial t} = -S + \tau' \frac{\partial}{\partial t} w * f[U] + k \circ f[U] \quad (3.75)$$

$$= -U + w * f[U] + \tau' \frac{\partial}{\partial t} w * f[U] + k \circ f[U]. \quad (3.76)$$

The dynamical equation governing changes in the receptive field $R(x)$ is gotten by observing the boundary of A , that is $r_1(x)$ and $r_2(x)$. We use

$$\frac{\partial U(x, r_i, t)}{\partial t} \frac{\partial r_i(x, t)}{\partial t} + \frac{\partial U(x, r_i, t)}{\partial t} = 0 \quad (3.77)$$

to derive the equations of $\partial r_i(x, t)/dt$ in the Heaviside world.

Theorem 2. *The dynamics describing the boundaries of the receptive field is given by*

$$\tau'(\alpha_1 + \beta_1) \frac{\partial r_1(x, t)}{\partial t} - \tau' \beta_1 \frac{\partial r_2(x - l_1, t)}{\partial t} = -W(l_1) - K(r) - s, \quad (3.78)$$

$$\tau'(\alpha_2 - \beta_2) \frac{\partial r_2(x, t)}{\partial t} + \tau' \beta_2 \frac{\partial r_1(x + l_2, t)}{\partial t} = -W(l_2) - K(r) - s, \quad (3.79)$$

where $l_i = l_i(x, t)$, $r = r(x, t)$ and

$$W(l) = \int_0^l w(x) dx, \quad (3.80)$$

$$K(r) = \int_0^r k(y) dy, \quad (3.81)$$

$$\alpha_i = \frac{\partial U(x, r_i, t)}{\partial y}, \quad i = 1, 2 \quad (3.82)$$

$$\beta_1 = -w(l_1) \left/ \frac{\partial r_2(x - l_1, t)}{\partial x} \right., \quad (3.83)$$

$$\beta_2 = -w(l_2) \left/ \frac{\partial r_1(x + l_2, t)}{\partial x} \right.. \quad (3.84)$$

Proof. To evaluate $\partial U/\partial t$ at $y = r_1$, we first calculate $K \circ f[U]$ at $r_1(x)$. We easily have

$$\int k(r_1 - y') f[U(x, y')] dy' = \int_{r_1}^{r_2} k(r_1 - y') dy' = K(r_2 - r_1), \quad (3.85)$$

where we have used $K(r) = K(-r)$ and we have assumed that the distribution of input stimuli of Y is uniform, $p(y') = 1$, by normalizing the length of Y equal to 1. Similarly, we have

$$\int w(x - x') f[U(x', r_1)] dx' = \int_{x-l_1}^x w(x - x') = W(l_1). \quad (3.86)$$

Hence,

$$\frac{\partial}{\partial t} w * f[U] = w(l_1) \frac{\partial l_1}{\partial t} \quad (3.87)$$

at $y = r_1$. Therefore,

$$\tau' \frac{\partial U(x, r_1)}{\partial t} = \tau' w(l_1) \frac{\partial l_1}{\partial t} + W(l_1) + K(r) + s. \quad (3.88)$$

We need to evaluate $w(l_1) \frac{\partial l_1}{\partial t}$. By differentiating

$$r_2(x - l_1) = r_1(x) \quad (3.89)$$

with respect to t , we get

$$\frac{\partial r_1(x)}{\partial t} = \frac{\partial r_2(x - l_1)}{\partial t} - \frac{\partial r_2(x - l_1)}{\partial x} \frac{\partial l_1(x)}{\partial t}. \quad (3.90)$$

We substitute $\frac{\partial l_1}{\partial t}$ obtained from (3.90) in (3.88), and we finally get (3.78). Calculations at $y = r_2$ yield (3.79).

3.3.3 Equilibrium Solution of Learning

The equilibrium solutions \bar{U} and \bar{S} of the dynamics of learning are derived by putting $\partial S / \partial t = 0$ in (3.62). They satisfy the following equations

$$\bar{U}(x, y) = w * f[\bar{U}] + \bar{S}, \quad (3.91)$$

$$\bar{S}(x, y) = k \circ f[\bar{U}], \quad (3.92)$$

In order to understand the formation of a topological map from Y to X and its stability, we consider a simple situation: We eliminate the boundary conditions by assuming that both Y and X are rings, where we normalize the lengths of the rings equal to 1, $L_X = L_Y = 1$.

We then search for an equilibrium solution of (3.78) and (3.79). Since the equilibrium solution satisfies

$$W(l_1) + K(r) + s = 0, \quad (3.93)$$

$$W(l_2) + K(r) + s = 0, \quad (3.94)$$

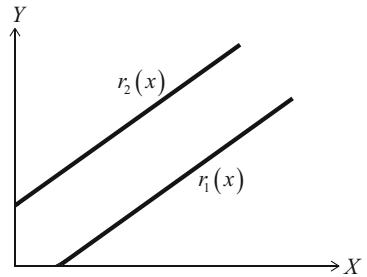
it is easy to see that

$$l_1(x) = l_2(x) = l(x). \quad (3.95)$$

As can be seen from Fig. 3.4, we get

$$l(x) = l\{x + l(x)\}, \quad \text{mod } 1. \quad (3.96)$$

Fig. 3.5 Continuous map from Y to X



When $l(x)$ is continuous, it has only a constant solution (see Takeuchi and Amari [15]),

$$l(x) = \bar{l}. \quad (3.97)$$

This implies

$$r(x) = \bar{r} \quad (3.98)$$

and hence $r_2(x)$ is a shift of $r_1(x)$, $r_2(x) = r_1(x) + \bar{r}$.

The solution

$$r_1(x) = x + c, \quad (3.99)$$

$$r_2(x) = x + c + \bar{r} \quad (3.100)$$

is an equilibrium for any constant c . This gives a natural correspondence between Y and X (Fig. 3.5). Now, let us study its stability.

However, first, there is a delicate problem as to whether other equilibrium solutions exist or not. When \bar{l} is a rational number,

$$\bar{l} = \frac{m}{n}, \quad (3.101)$$

we find that

$$r_1(x) = x + c + g(x) \quad (3.102)$$

$$r_2(x) = x + c + \bar{r} + g(x) \quad (3.103)$$

is also an equilibrium solution, where $g(x)$ is a periodic function with a period $1/n$. This is a rippled solution, but we may disregard the ripple when n is large.

3.3.4 Stability of the Equilibrium Solution

We shall study the stability of the simple equilibrium solution (3.99), (3.100) in order to check if the continuous topological mapping is stable. We put $c = 0$ and perturb the solution as

$$r_1(x, t) = x + \varepsilon v_1(x, t) \quad (3.104)$$

$$r_2(x, t) = x + \bar{r} + \varepsilon v_2(x, t) \quad (3.105)$$

where ε is a small constant. The variational equation is

$$\begin{aligned} \tau'(\alpha + \beta) \frac{\partial v_1(x, t)}{\partial t} - \tau' \beta \frac{\partial v_2(x - \bar{l}, t)}{\partial t} &= \beta \{v_2(x - \bar{l}, t) - v_1(x, t)\} \\ &\quad - k(\bar{r}) \{v_2(x, t) - v_1(x, t)\}, \end{aligned} \quad (3.106)$$

$$\begin{aligned} \tau'(\alpha + \beta) \frac{\partial v_2(x, t)}{\partial t} - \tau' \beta \frac{\partial v_1(x + \bar{l}, t)}{\partial t} &= -\beta \{v_2(x, t) - v_1(x + \bar{l}, t)\} \\ &\quad + k(\bar{r}) \{v_2(x, t) - v_1(x, t)\}, \end{aligned} \quad (3.107)$$

where

$$\alpha = \frac{\partial \bar{U}(x, r_1(x))}{\partial y}, \quad \beta = -w(\bar{l}) \left/ \frac{dr_1(x)}{dt} \right.. \quad (3.108)$$

We expand v_1 and v_2 in a Fourier series,

$$v_1(x, t) = \sum V_1(n, t) \exp \{i 2n\pi x\}, \quad (3.109)$$

$$v_2(x, t) = \sum V_2(n, t) \exp \{i 2n\pi x\}. \quad (3.110)$$

The variational equation then separates for every n , giving ordinary differential equations for each $n = 0, 1, \dots$:

$$\tau' A_n \frac{d}{dt} \begin{bmatrix} V_1(n, t) \\ V_2(n, t) \end{bmatrix} = B_n \begin{bmatrix} V_2 \\ V_1 \end{bmatrix}, \quad n = 0, 1, \dots, \quad (3.111)$$

where

$$A_n = \begin{bmatrix} \alpha + \beta & -\beta \bar{z}_n \\ -\beta z_n & \alpha + \beta \end{bmatrix}, \quad B_n = \begin{bmatrix} k - \beta & \beta \bar{z}_n - k \\ \beta z_n - k & k - \beta \end{bmatrix}, \quad (3.112)$$

Fig. 3.6 Stability of continuous solutions:
(a) stable, **(b)** unstable

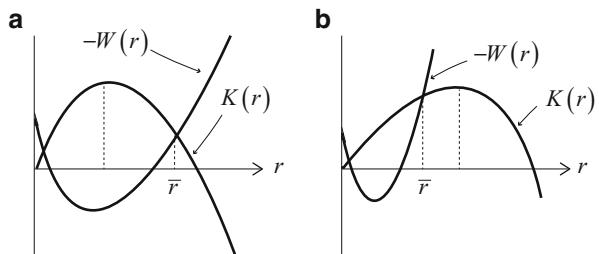
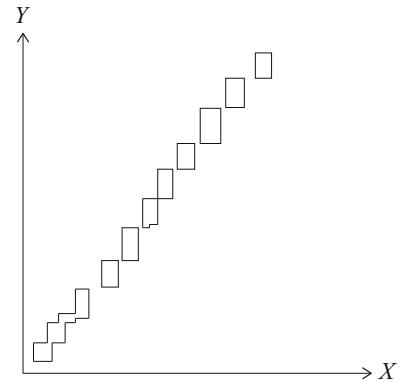


Fig. 3.7 Simulation result of block formation in unstable case



$B_n = \exp\{i2n\pi\bar{l}\}$. The stability depends on the eigenvalues λ :

$$\det |\lambda A_n - B_n| = 0. \quad (3.113)$$

We shall omit the detailed derivation, since they are rather technical (see Takeuchi and Amari [15]).

Theorem 3. *The equilibrium solution is stable when and only when*

$$k(\bar{r}) < 0, \quad w(\bar{l}) < 0. \quad (3.114)$$

Roughly speaking, the continuous map is stable when the length of the receptive field \bar{r} is wider than the length \bar{l} of the active region of the input field, and is unstable otherwise. (See Fig. 3.6a for the stable and Fig. 3.6b for unstable solutions.) We can also analyze the stability of a rippled solution, but the result is the same.

The variational analysis does not tell what will happen in the unstable case. Computer simulations show that both X and Y are divided into discrete blocks, and there exists a mapping from a block of Y to a block of X (see Fig. 3.7). The topology of X and Y is preserved in the discretized sense. When an input bump is in a block of Y , then all the neurons in the corresponding block of X are excited. This might explain the mechanism behind the formation of microscopic columnar structures observed in the cerebrum. This mechanism is important when

Y has more dimensions than X . For example, let us consider the case where Y is a set of stimuli given to the retina. We assume that a bar with an arbitrary orientation is presented at any position on the retina. The set of stimuli is three-dimensional, having two dimensions corresponding to the positions of a bar and one dimension corresponding to the orientation of the bar. In this case, $Y = \mathbf{R}^2 \times S^1$. Such stimuli are mapped to the visual cortex $X = \mathbf{R}^2$. It is known that X decomposes into an aggregate of blocks called a column. The position of a bar is mapped to the position of blocks of X in a discretized manner and keeps the topography. There is a microstructure inside a block such that the orientation of the bar is continuously mapped inside a columnar block. This is the wisdom of nature expressed through evolution. Our theory might explain it.

When $p(y')$ is not uniform, some part of Y is stimulated more frequently than the other part. It is plausible that a frequently stimulated part of Y has a finer representation occupying a larger part of X . This effect is analyzed in [4].

3.3.5 Kohonen Map

Kohonen [11] proposed a neural mechanism of self-organization, which generates a topological map from the space of input signals to a neural field. It is an engineering model simplifying the Willshaw and Malsburg model [17] such that the dynamics of neural excitation are omitted. This mechanism is applicable to various engineering problems and is known as Kohonen's SOM (self-organizing map).

The input signal field Y gives a vector-valued output $\mathbf{a}(y')$, when position y' is activated. In the previous case, this is a bump $a(y; y')$ in Y . Here, we regard this as a vector $\mathbf{a}(y')$ whose components are $a(y; y')$, $y \in Y$. In our previous case, Y is 1-dimensional and the activation vector is $\mathbf{a}(y') = a(y; y')$. We regard the distribution $a(y; y')$ over $y \in Y$ as a vector $\mathbf{a}(y')$. The output layer X is a neural field, typically 2-dimensional. A neuron at position x has a connection weight vector $s(x)$. When it receives an input signal $\mathbf{a}(y')$, it calculates the inner product of $s(x)$ and $\mathbf{a}(y')$, obtaining

$$u(x, y') = s(x) \cdot \mathbf{a}(y') = \int s(x, y)a(y; y') dy, \quad (3.115)$$

where we have put $s(x) = s(x, y)$. The neurons of X are not recurrently connected and no dynamics of excitation take place in X in the Kohonen model. Instead, the activation $f[u(x)]$ of the neural field is decided by a simple rule stated in the following. The neuron at position x is said to be the winner when it has the highest value of $u(x, y')$ when a stimulus $\mathbf{a}(y')$ is applied. The winner neuron \bar{x} corresponding to input $\mathbf{a}(y')$ is hence defined by

$$\bar{x}(\mathbf{a}) = \arg \max_x u(x, \mathbf{a}(y')). \quad (3.116)$$

When the absolute values of vector \mathbf{a} and weight s are normalized to 1, the winner is the neuron whose weight $s(x)$ is closer to the input pattern vector \mathbf{a} .

We then define a neighborhood of a neuron x by using a distance function $d(x, x')$. The neighborhood $N(x)$ of x is a set of positions x' given by

$$N(x) = \{x' \mid d(x, x') \leq c\} \quad (3.117)$$

for a constant c . When \bar{x} is the winner for input \mathbf{a} , the neurons in the neighborhood of \bar{x} are excited. Hence, we have

$$f[u(x, y')] = \begin{cases} 1, & x \in N(\bar{x}), \\ 0, & \text{otherwise.} \end{cases} \quad (3.118)$$

This is the Heaviside world, although the excited neurons are determined by this simple rule, not by recurrent dynamics.

The connection weight vector $s(x)$ of an excited neuron $x \in N(\bar{x})$ changes according to a Hebb-like rule: The weight vector $s(x)$ moves toward the input \mathbf{a} , and then, the normalization takes place. Hence, the rule is written as

$$s(x) \rightarrow \frac{s(x) + \varepsilon \mathbf{a}}{|s(x) + \varepsilon \mathbf{a}|} \approx \{1 - \varepsilon |\mathbf{a} \cdot s|\} s + \varepsilon \mathbf{a} \quad (3.119)$$

for small constant ε , when x belongs to the neighborhood $N\{\bar{x}(\mathbf{a})\}$. The weight vectors outside the neighborhood do not change.

We use the continuous time version of learning and rewrite the dynamics of $s(x)$ in the differential equation,

$$\tau' \frac{ds(x)}{dt} = -s + \varepsilon \langle \{\mathbf{a}(y') - \mathbf{a}(y') \cdot s(x)\} s(x) \rangle. \quad (3.120)$$

Here, we use the adiabatic assumption such that the time constant τ' of learning is sufficiently small compared with the duration in which a randomly chosen stimulus $\mathbf{a}(y')$ is applied.

Let us see if a topological map between Y and X is formed stably by the Kohonen SOM mechanism. K. Kurata analyzed this problem in his doctoral dissertation and in a Japanese paper [12]. He proved that, when Y and X are one-dimensional, the continuous topological map $y = r(x)$ is always neutrally stable. He also analyzed the case when Y is one-dimensional and X is two-dimensional. Then, self-organization embeds Y as a curve in X . He showed that there are three possibilities depending on the sizes of the abscissa and ordinate of X .

Consider the map

$$Y \rightarrow X : x_1(y) = y, \quad (3.121)$$

$$x_2(y) = c. \quad (3.122)$$

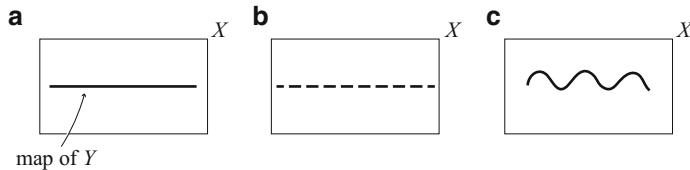


Fig. 3.8 Stability of Kohonen map: (a) 1-D Y is stably embedded, (b) unstable and block structure emerges, (c) unstable and wave emerges

This is an equilibrium solution (Fig. 3.8a). It is stable or unstable, depending on the parameters. There are two types of instability. In one case, it is unstable in the vertical direction, so that a wave pattern emerges (see Fig. 3.8c). The second case is that it is unstable in the horizontal direction, so that the block structure emerges (see Fig. 3.8b). This is an interesting but regrettably not well known result.

3.4 Conclusions

The study of dynamics of a neural field is currently a ‘hot’ research topic. However, its mathematical treatment is difficult. When one approximates the activation function by the Heaviside function, the dynamics can be dramatically simplified and one can find exact results in some cases. This is the Heaviside world, in which the dynamics of an active region can be analyzed in terms of the dynamics of its boundary. The Heaviside world recaptures previous important results including the formation of a topological (topographic) map in the cortex.

The Heaviside world plays an important role in the analysis of topographic map formation. We analyzed a one-dimensional mapping, and proved that a topological map is unstable under a certain condition. In such a case, a block structure or micro-columnar structure emerges. This analysis may account for the formation of columnar microstructures in the visual cortex.

References

1. Amari, S.: Characteristics of randomly connected threshold-element networks and network systems. Proc. IEEE **59**(1), 35–47 (1971)
2. Amari, S.: Characteristics of random nets of analog neuron-like elements. IEEE Trans. Syst. Man Cybern. (SMC) **2**(5), 643–657 (1972). (Also Vemri, V. (ed.): Artificial Neural Networks Theoretical Concepts. IEEE Computer Society (1988))
3. Amari, S.: Dynamics of pattern formation in lateral-inhibition type neural fields. Biol. Cybern. **27**, 77–87 (1977)
4. Amari, S.: Topographic organization of nerve fields. Bull. Math. Biol. **42**, 339–364 (1980)
5. Bressloff, P.C.: Spatiotemporal dynamics of continuum neural fields. J. Phys. A **45**, 1–109 (2012)

6. Bressloff, P.C., Coombes, S.: Neural ‘bubble’ dynamics revisited. *Cogn. Comput.* **5**, 281–294 (2013)
7. Coombes, S., Schmidt, H., Bojak, I.: Interface dynamics in planar neural field models. *J. Math. Neurosci.* **2**, 9 (2012)
8. Folias, S.E.: Nonlinear analysis of breathing pulses in synaptically coupled excitable neural activity. *J. Comput. Neurosci.* **11**, 121–134 (2011)
9. Folias, S.E., Bressloff, P.C.: Breathing pulses in an excitatory neural network. *SIAM J. Appl. Dyn. Syst.* **3**, 378–407 (1974)
10. Fung, C.A., Wong, M., Wang, H., Wu, S.: Dynamical synapses enhance neural information processing: Gracefulness, accuracy and mobility. *Neural Comput.* **24**, 1147–1185 (2012)
11. Kohonen, T.: Self-organized formation of topologically correct feature maps. *Biol. Cybern.* **43**, 59–69 (1982)
12. Kurata, K.: Formation of information representation by self-organization. In: Japanese Association of Industrial Technology (eds.) *Basics of Neurocomputing*, chap. 5. Japanese Association of Industrial Technology, Tokyo (1989). in Japanese
13. Lu, Y., Sato, Y., Amari, S.: Travelling bumps and their collisions in a two-dimensional neural field. *Neural Comput.* **23**, 1248–1260 (2011)
14. Seung, H.S.: Continuous attractors and oculomotor control. *Neural Netw.* **11**, 1253–58 (1998)
15. Takeuchi, A., Amari, S.: Formation of topographic maps and columnar microstructures. *Biol. Cybern.* **35**, 63–72 (1979)
16. von der Malsburg, C.: Self-organization of orientation sensitivity and columns in the visual cortex. *Kybernetik* **14**, 85–100 (1973)
17. Willshaw, D.J., von der Malsburg, C.: How patterned neural connections can be set up by self-organization. *Proc. R. Soc. B* **194**, 431–445 (1976)
18. Wilson, H.R., Cowan, J.D.: Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.* **12**, 1–24 (1972)
19. Wu, S., Amari, S.: Computing with continuous attractors: Stability and on-line aspects. *Neural Comput.* **17**, 2215–2239 (2005)

Chapter 4

Spatiotemporal Pattern Formation in Neural Fields with Linear Adaptation

G. Bard Ermentrout, Stefanos E. Folias, and Zachary P. Kilpatrick

Abstract We study spatiotemporal patterns of activity that emerge in neural fields in the presence of linear adaptation. Using an amplitude equation approach, we show that bifurcations from the homogeneous rest state can lead to a wide variety of stationary and propagating patterns on one- and two-dimensional periodic domains, particularly in the case of lateral-inhibitory synaptic weights. Other typical solutions are stationary and traveling localized activity bumps; however, we observe exotic time-periodic localized patterns as well. Using linear stability analysis that perturbs about stationary and traveling bump solutions, we study conditions for the activity to lock to a stationary or traveling external input on both periodic and infinite one-dimensional spatial domains. Hopf and saddle-node bifurcations can signify the boundary beyond which stationary or traveling bumps fail to lock to external inputs. Just beyond a Hopf bifurcation point, activity bumps often begin to oscillate, becoming *breather* or *slosher* solutions.

4.1 Introduction

Neural fields that include local negative feedback have proven very useful in qualitatively describing the propagation of experimentally observed neural activity [26, 39]. Disinhibited *in vitro* cortical slices can support traveling pulses and spiral

G.B. Ermentrout (✉)

Department of Mathematics, University of Pittsburgh, Pittsburgh, PA, USA

e-mail: bard@pitt.edu

S.E. Folias

Department of Mathematics & Statistics, University of Alaska Anchorage, Anchorage, AK, USA
e-mail: sfolias@uua.alaska.edu

Z.P. Kilpatrick

Department of Mathematics, University of Houston, Houston, TX, USA
e-mail: zpkilpat@math.uh.edu

waves [27, 53], suggesting that some process other than inhibition must curtail large-scale neural excitations. A common candidate for this negative feedback is spike frequency adaptation, a cellular process that brings neurons back to their resting voltage after periods of high activity [2, 48]. Often, adaptation is modeled as an additional subtractive variable in the activity equation of a spatially extended neural field [26, 38, 39]. Pinto, in his PhD dissertation with Ermentrout, explored how linear adaptation leads to the formation of traveling pulses [38]. Both singular perturbation theory and the Heaviside formalism of Amari (see Chap. 3 and [1]) were used to analyze an excitatory network on the infinite spatial domain [38, 39]. At the same time, Hansel and Sompolinsky showed adaptation leads to traveling pulses (traveling bumps) in a neural field on the ring domain [26]. In the absence of adaptation, excitatory neural fields generate stable traveling fronts [21, 25]. For weak adaptation, the model still supports fronts which undergo a symmetry breaking bifurcation, leading to bidirectional front propagation at a critical value of the adaptation rate [6]. In fact, adaptive neural fields generate a rich variety of spatiotemporal dynamics like stimulus-induced breathers [7], spiral waves (see Chap. 5 and [27]), drifting spots (see Chap. 7), multipulse solutions [52], and self-sustained oscillations [46]. Coombes and Owen have implemented a related model, employing nonlinear adaptation, that is shown to generate breathers, traveling bumps, and more exotic solutions [11]. However, it has been shown that great care must be taken when performing stability analysis of such a model [29]. Thus, we restrict the contents of this chapter to analyzing models with linear adaptation.

We review a variety of results concerning bifurcations that arise in spatially extended neural fields when an auxiliary variable representing linear adaptation is included [13, 23, 25, 31]. In particular, we study the dynamics of the system of non-local integro-differential equations [10, 26, 35, 39]

$$\tau \frac{\partial u(x, t)}{\partial t} = -u(x, t) - \beta v(x, t) + \int_D w(x - y) F(u(y, t)) dy + I(x, t), \quad (4.1a)$$

$$\frac{1}{\alpha} \frac{\partial v(x, t)}{\partial t} = u(x, t) - v(x, t). \quad (4.1b)$$

The variable $u(x, t)$ represents the total synaptic input arriving at location $x \in D$ in the network at time t . We fix time units by setting $\tau = 1$ without loss of generality. The convolution term represents the effects of recurrent synaptic interactions, and $w(x - y) = w(y - x)$ is a reflection-symmetric, distance-dependent synaptic weight function encoding the strength of connections between location y and x . The nonlinearity F is a transfer function that converts the synaptic inputs to an output firing rate. Local negative feedback $v(x, t)$ represents the effects of spike frequency adaptation [2, 26, 39, 48], occurring at rate α with strength β . Finally, $I(x, t)$ represents an external spatiotemporal input. In Sect. 4.2, we begin by analyzing bifurcations from the rest state on one- and two-dimensional periodic domains in the absence of any input ($I(x, t) \equiv 0$) with the use of amplitude equations. We show that a lateral-inhibitory synaptic weight organizes activity of the network into

a wide variety of stationary and propagating spatiotemporal patterns. In Sect. 4.3, we study the processing of external inputs on a ring domain ($D = (-\pi, \pi)$). Since adaptation can lead to spontaneous propagation of activity, inputs must move at a speed that is close to the natural wavespeed of the network to be well tracked by its activity. Finally, in Sect. 4.4, we study bifurcations of stationary and traveling bumps in a network on the infinite spatial domain ($D = (-\infty, \infty)$). Both natural and stimulus-induced bump solutions are analyzed. Depending on whether the synaptic weight function is purely excitatory or lateral-inhibitory, either spatial mode of a stimulus-locked bump can destabilize in a Hopf bifurcation, leading to a *breather* or a *slosher*. Conditions for the locking of traveling bumps to moving inputs are discussed as well.

4.2 Bifurcations from the Homogeneous State

The simplest type of analysis that can be done with continuum neural field models is to study bifurcations from the homogeneous state. As in [13], we focus on the one-dimensional ring model, and then make some comments about the dynamics of systems in two space dimensions with periodic boundary conditions. Here, our domain is either the ring ($D = (-\pi, \pi)$) or the square ($D = (-\pi, \pi) \times (-\pi, \pi)$) with periodic boundary conditions. With some abuse of notation, x is either a scalar or a two-dimensional vector. The function $w(x)$ is periodic in its coordinates and furthermore, we assume that it is symmetric in one-dimension and isotropic in two-dimensions. Translation invariance and periodicity assures us that $\int_D w(x-y)dy = W_0$. A constant steady state has the form

$$u(x, t) = \bar{u}, \quad \text{where} \quad (1 + \beta)\bar{u} = W_0 F(\bar{u}).$$

Since F is monotonically increasing with $F(-\infty) = 0$ and $F(+\infty) = 1$, we are guaranteed at least one root. To simplify the analysis further, we assume that $F(u) = k(f(u) - f(0))/f'(0)$ with $f(u) = 1/(1 + \exp(-r(u - u_{th}))$) as in [13]. Note that $F(0) = 0$ and $F'(0) = k$ which serves as our bifurcation parameter. With this assumption, $\bar{u} = \bar{v} = 0$ is the homogeneous rest state.

To study the stability, we linearize, letting $u(x, t) = \bar{u} + q(x, t)$ and $v(x, t) = \bar{u} + p(x, y)$ so that to linear order in $q(x, t), p(x, t)$ we have

$$\begin{aligned} \frac{\partial q}{\partial t} &= -q(x, t) + k \int_{\Omega} w(x-y)q(y, t) dy - \beta p(x, t) \\ \frac{\partial p}{\partial t} &= \alpha(-p(x, t) + q(x, t)). \end{aligned} \tag{4.2}$$

Because $w(x)$ is translational invariant and the domain is periodic, solutions to the linearized equations have the form $\exp(\lambda t) \exp(in \cdot x)$ where in one-dimension n is

an integer and in two-dimensions, it is a pair of integers, (n_1, n_2) . Let $m = |n|$ be the magnitude of this vector (scalar) and let

$$W(m) := \int_{\Omega} w(y) e^{-in \cdot y} dy.$$

(The isotropy of w guarantees that the integral depends only on the magnitude of n .) We then see that λ must satisfy

$$\lambda \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix} = \begin{pmatrix} -1 + k W_m & -\beta \\ \alpha & -\alpha \end{pmatrix} \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix}, \quad (4.3)$$

where $(\chi_1, \chi_2)^T$ is a constant eigenvector.

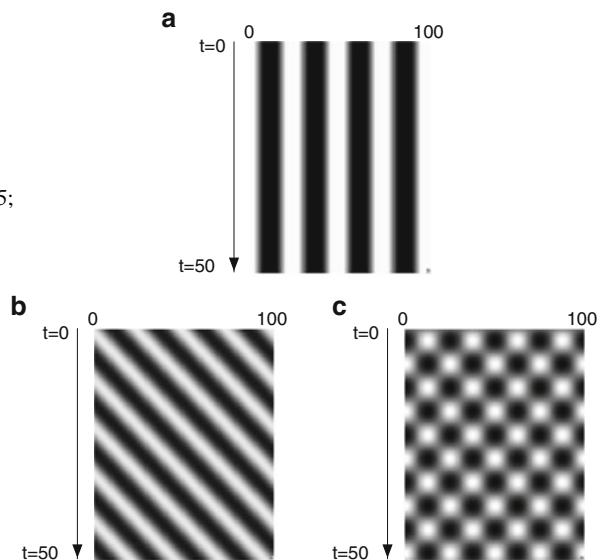
There are several cases with which to contend, and we now describe them. The easiest parameter to vary in this system is the sensitivity, k (This is the slope of F at the equilibrium point). The trace of this matrix is $\mathcal{T}(m) := -(1 + \alpha) + k W(m)$ and the determinant is $\mathcal{D}(m) := \alpha[1 + \beta - k W(m)]$. Note that $W(0) = W_0$ and $W(m) \rightarrow 0$ as $m \rightarrow \infty$. The uniform state is linearly stable if and only if $\mathcal{T}(m) < 0$ and $\mathcal{D}(m) > 0$ for all m . If $W(m) < 0$, then both stability conditions hold, so, consider the sets $k_m^{\mathcal{T}} = (1 + \alpha)/W(m)$ and $k_m^{\mathcal{D}} = (1 + \beta)/W(m)$ which represent critical values of k where the trace and determinant vanish respectively. We are interested in the minimum of these sets over all values of m where $W(m) > 0$. Let n denote the critical wavenumber at which $W(m)$ is maximal. It is clear that if $\alpha > \beta$ then the determinant vanishes at a lower value of k than the trace does and *vice versa*. That is, there is a critical ratio $R = \beta/\alpha$ such that if $R > 1$, then the trace is critical (and there is a Hopf bifurcation) while if $R < 1$, the determinant is critical (and there is a stationary bifurcation). The ratio R is the product of the strength and the time constant of the adaptation. If the adaptation is weak and fast, there is a steady state bifurcation, while if it is large and slow, there is a Hopf bifurcation. Curtu and Ermentrout [13] studied the special case where R is close to 1. At $R = 1$, there is a double zero eigenvalue at the critical wavenumber m and thus a Takens-Bogdanov bifurcation. For the rest of this section, let m^* denote the value of $|n|$ at which $W(m)$ is maximal. We also assume that $W(m^*) > 0$. For one dimension, $n = \pm m^*$ and in two spatial dimensions, at criticality, $n = (n_1, n_2)$ where $m^* = \sqrt{n_1^2 + n_2^2}$. For concreteness and illustration of the results, we use $f(u) = 1/(1 + \exp(-r(u - u_{th})))$ with two free parameters that set the shape of f and thus F . We remark that (i) if $u_{th} = 0$, then $F''(0) = 0$ and (ii) for a range of u_{th} surrounding 0, $F'''(0) < 0$. We also use $w(x) = Aa^{p/2} \exp(-ax^2) - Bb^{p/2} \exp(-bx^2)$ (where p is the dimension of the domain) and note that $W(m) = \pi(A \exp(-m^2/a) - B \exp(-m^2/b))$. With $A = 5, a = 0.125, B = 4, b = 0.005$, this kernel has a fairly narrow Mexican hat profile.

Fig. 4.1 Space-time plots of the solutions to (4.1). (a)

Stationary stripes for
 $k = 0.24$, $\beta = 0$, $\alpha = 0.1$,
 $u_{th} = 0.05$, and $r = 0.25$;

(b) Traveling waves with
parameters as in (a), but
 $\beta = 0.25$, $k = 0.26$, $u_{th} = 0.05$;

(c) Standing waves with
parameters as in A, B, but
 $u_{th} = 0.3$



4.2.1 One Spatial Dimension

4.2.1.1 Zero Eigenvalue

In the case of $R < 1$, the bifurcation is at a zero eigenvalue and we expect a spatial pattern that has the form $u(x, t) = z \exp(im^*x) + \text{c.c}$ (here c.c means complex conjugates) and

$$z_t = z(a(k - k_c) - b|z|^2)$$

where a and b are complicated, but readily computed, functions of w , $F''(0)^2$, and $F'''(0)$. Both a, b are real, $a > 0$, and for our choice of w and F , we have $b > 0$. The non-zero solution to this equation is $z = Ae^{-i\Theta}$ where $A^2 = a(k - k_c)/b$ and Θ is an arbitrary constant corresponding to a phase-shift of the periodic pattern. The solution exists as long as $k \geq k_c$ (since a, b are positive) and, furthermore, the solution is stable. Thus as we increase k , we expect to see a spatially periodic pattern emerge that has the form

$$u(x) = \sqrt{a(k - k_c)/b} \cos(m^*x + \Theta) + O(k - k_c).$$

Figure 4.1a shows a simulation of Eq. (4.1) where we have discretized the one-dimensional ring into 100 units. In this case $W(m)$ takes its maximum at $m^* = 4$, so as expected, we see a stationary pattern consisting of four peaks. In the case where $m^* = 1$, (which occurs for sufficiently broad inhibition) these spatially periodic patterns are interpreted as localized activity for tuning in a directionally based neural

system [26, 54]. This single stationary ‘‘bump’’ can be perturbed and pinned with external stimuli as we see in subsequent sections of this chapter.

4.2.1.2 Imaginary Eigenvalues

When $R > 1$ (strong or slow adaptation), then the trace vanishes at a lower critical k than the determinant. Let $m^* > 0$ be the critical wavenumber and $i\omega$ be the imaginary eigenvalue. Then $u(x, t)$ has the form

$$u(x, t) = z(t)e^{i(\omega t - m^* x)} + w(t)e^{i(\omega t + m^* x)} + \text{c.c}$$

where

$$\begin{aligned} z' &= z[(a_1 + ia_2)(k - k_c) - (b_1 + ib_2)|z|^2 - (c_1 + ic_2)|w|^2] \\ w' &= w[(a_1 + ia_2)(k - k_c) - (b_1 + ib_2)|w|^2 - (c_1 + ic_2)|z|^2]. \end{aligned} \quad (4.4)$$

These coefficients can be computed for (4.1) (and, indeed, for a variant of the equations, [13] computes them explicitly) and they depend only on $F''(0)^2$, $F'''(0)$, $W(2m)$, $W(m)$, α , and β . In particular, with our choice of $f(u)$ and for u_{th} not large, $b_1, c_1 > 0$. There are three distinct types of nontrivial solutions: $(z, w) \in \{(Z, 0), (0, Z), (Y, Y)\}$, where:

$$\begin{aligned} Z &= Ae^{i\Omega t}, & Y &= Be^{i\Xi t}, \\ \Omega &= (a_2 - a_1 b_2/b_1)(k - k_c), & \Xi &= (a_2 - a_1(b_2 + c_2)/(b_1 + c_1))(k - k_c), \\ A^2 &= (a_1/b_1)(k - k_c), & B^2 &= (a_1/(b_1 + c_1))(k - k_c). \end{aligned}$$

Solutions of the form $(Z, 0), (0, Z)$ correspond to traveling wavetrains with opposite velocities and those of the form (Y, Y) correspond to standing time-periodic waves. To see this, we note that the solutions have the form

$$u(x, t) = \Re\{ze^{i(\omega t + m^* x)} + we^{i(\omega t - m^* x)}\},$$

so that for the solution, $(Z, 0)$, we get

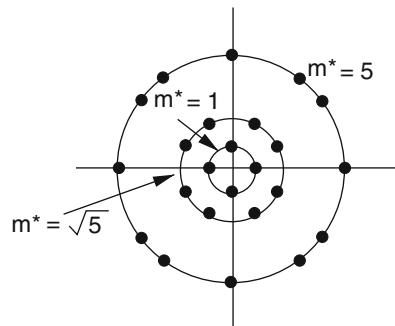
$$u(x, t) = A \cos((\omega + \Omega)t + m^* x),$$

while for the (Y, Y) case

$$u(x, t) = B \cos((\omega + \Xi)t) \cos(m^* x).$$

The traveling (standing) waves are stable if and only if $c_1 > b_1$ (resp. $c_1 < b_1$) and, importantly, if $F''(0)$ is zero or close to zero (that is, $u_{th} \approx 0$), then $c_1 > b_1$ no matter what you choose for the other parameters. Thus, for u_{th} small, we expect to

Fig. 4.2 Three different cases of critical wavenumbers in the square lattice. The critical wavenumbers are (from out to in), $\{(\pm 1, 0), (0, \pm 1)\}$, $\{(\pm 2, 1), (\pm 2, -1), (\pm 1, 2), (\pm 1, -2)\}$ and $\{(\pm 3, 4), (\pm 3, -4), (\pm 4, 3), (\pm 4, -3), (\pm 5, 0), (0, \pm 5)\}$



see only stable traveling waves. Figure 4.1b, c shows simulations of (4.1) for two different choices of u_{th} ; near zero, the result is traveling waves, while for $u_{th} = 0.3$, standing waves emerge. Choosing the interaction kernel, $w(x)$, so that $m^* = 1$, leads to a single traveling pulse or bump of activity which, itself, can be entrained and perturbed by external stimuli (see the next sections).

4.2.2 Two Spatial Dimensions

While most of the focus in this chapter is on one space dimension, the theory of pattern formation is much richer in two-dimensions and Eq.(4.1) provides an excellent example of the variety of patterns. The isotropy of the weight matrix implies that the eigensolutions to the linear convolution equation (4.2) have the form $\exp(i n \cdot x)$. In two dimensions, n is a two-vector of integers. We then obtain exactly the same formula for the determinant and the trace as in one-dimension, however, $m = |n|$ in this case so that there are at least two distinct eigenvectors and their complex conjugates and there are often many more. Figure 4.2 illustrates three cases where $m^* = 1, \sqrt{5}, 5$ corresponding to 4, 8, and 12 different pairs (n_1, n_2) . We treat and numerically illustrate several possibilities by discretizing (4.1) on a 50×50 array. Our choice of $w(x)$ gives a maximum at $m^* = 2$ which is the simplest case.

4.2.2.1 Zero Eigenvalue

The simplest possible case in two dimensions has only four distinct wave vectors (inner circle in Fig.4.2). For example, if $m^* = 2$, then $n \in \{(2, 0), (0, 2), (-2, 0), (0, -2)\}$. (Note that in cases where there are only four vectors, the critical waves have either of the two forms: $(k, 0), (0, k), (-k, 0), (0, -k)$ or $(k, k), (k, -k), (-k, -k), (-k, k)$.) If we write $x = (x_1, x_2)$, then, $u(x, t)$ has the form $u(x_1, x_2, t) = z_1 \exp(i 2x_1) + z_2 \exp(i 2x_2) + \text{c.c}$ and

$$\begin{aligned} z'_1 &= z_1(a(k - k_c) - b|z_1|^2 - c|z_2|^2), \\ z'_2 &= z_2(a(k - k_c) - b|z_2|^2 - c|z_1|^2), \end{aligned} \quad (4.5)$$

where as in the one-dimensional case, b, c depend on $F''(0)^2, F'''(0)$. All coefficients are real and can be computed. They are all positive for our choices of $F(u)$. We let $z_j = A_j e^{i\theta_j}$ and we then find that

$$\begin{aligned} A'_1 &= A_1(a(k - k_c) - bA_1^2 - cA_2^2), \\ A'_2 &= A_2(a(k - k_c) - bA_2^2 - cA_1^2). \end{aligned}$$

It is an elementary calculation to show that there are three types of solutions, $(z_1, z_2) = \{(r_1, 0), (0, r_1), (r_2, r_2)\}$ where $r_1^2 = a(k - k_c)/b, r_2^2 = a(k - k_c)/(b + c)$. For this example, the first two solutions correspond to vertical and horizontal stripes respectively and the third solution represents a spotted or checkerboard pattern. Stripes (spots) are stable if and only if $b < c$ (resp. $b > c$) [16]. As in the traveling/standing wave case above, if $F''(0)$ is zero ($u_{th} = 0$), then, $c > b$ and there are only stable stripes [16]. The resulting stationary patterns look identical to those in Fig. 4.3a, b without the implied motion. (To get stationary patterns, choose, e.g., $\beta = 0, r = 3$, and $u_{th} = 0$ for stripes or $u_{th} = 0.3$ for spots.)

This case (of two real amplitude equations) is the simplest case. The critical wave vector can be more complicated, for example, if $m^* = \sqrt{5}$, then, $n \in \{(1, 2), (1, -2), (2, 1), (2, -1), (-1, -2), (-1, 2), (-2, -1), (-2, 1)\}$ for which there are eight eigenvectors and the solution has the form

$$u(x, t) = \sum_{j=1}^4 z_j(t) e^{in_j \cdot x} + \text{c.c.}$$

where $n_j = (1, 2), \dots$ and z_j satisfy the four independent amplitude equations

$$\begin{aligned} z'_1 &= z_1(a(k - k_c) - b|z_1|^2 - c|z_2|^2 - d|z_3|^2 - e|z_4|^2), \\ z'_2 &= z_2(a(k - k_c) - b|z_2|^2 - c|z_1|^2 - d|z_4|^2 - e|z_3|^2), \\ z'_3 &= z_3(a(k - k_c) - b|z_3|^2 - c|z_4|^2 - d|z_1|^2 - e|z_2|^2), \\ z'_4 &= z_4(a(k - k_c) - b|z_4|^2 - c|z_3|^2 - d|z_2|^2 - e|z_1|^2). \end{aligned}$$

As in Eqs. (4.5), since a, \dots, e are all real coefficients, this model can be reduced to the analysis of a four dimensional real system. Dionne et al. [15] derive and analyze this case (among many others). In the context of neural fields, Tass [49] and Ermentrout [17] provide stability conditions for the equilibria, all of which consist of z_j taking on values of some $A \neq 0$ or 0. For example, the pure patterns $z_1 = A, z_2, z_3, z_4 = 0$ are stable if and only if $a < \{b, c, d\}$, there are also pairwise mixed solutions (checkerboards) of the form $z_1 = z_2 = A', z_3 = z_4 = 0$, etc., and fully

nonzero solutions, $z_1 = z_2 = z_3 = z_4 = A''$ which are stable if $a > \{d + c - b, d + b - c, b + c - d\}$. We remark that the triplet solutions $z_j = z_k = z_l = A'''$ are never stable and that if $F''(0) = 0$, then only stripes (one z_j , nonzero).

In two spatial dimensions, $m^* = 1$ can correspond to a single bump of activity which has been used to model hippocampal place cells [28]. For narrower inhibition, the more complex patterns describe the onset of geometric visual hallucinations [5, 18, 49, 50]. Simple geometric hallucinations take the form of spirals, pinwheels, bullseyes, mosaics, and honeycombs [33]. When transformed from the retinocentric coordinates of the eyeball to the coordinates of the visual cortex, these patterns take the form of simple geometric planforms such as rolls, hexagons, squares, etc. [45]. Thus, spontaneous bifurcations to patterned activity form a natural model for the simple visual patterns seen when the visual system is perturbed by hallucinogens, flicker [43] or other excitation. (See [3] for a comprehensive review.)

4.2.2.2 Imaginary Eigenvalues

The case of imaginary eigenvalues on a square lattice is quite complicated and only partially analyzed. Tass [50] has studied this case extensively when there are no even terms in the nonlinear equations (corresponding to $u_t = 0$ in our model). Silber and Knobloch [47] provide a comprehensive and extremely readable analysis of case where there are four critical wavenumbers.

Let us first consider the four dimensional case and take as a specific example: $n \in \{(2, 0), (0, 2), (-2, 0), (0, -2)\}$. In this case, the firing rate has the form:

$$u(x, t) = z_1 e^{i2x_1+i\omega t} + z_2 e^{i2x_2+i\omega t} + z_3 e^{-i2x_1+i\omega t} + z_4 e^{-i2x_2+i\omega t} + \text{c.c.}$$

The complex amplitudes z_j satisfy normal form equations ([47], equation 5.3):

$$\begin{aligned} z'_1 &= z_1[a(k - k_c) - b|z_1|^2 - cN_1 - dN_2] - e\bar{z}_3z_2z_4 \\ z'_2 &= z_2[a(k - k_c) - b|z_2|^2 - cN_2 - dN_1] - e\bar{z}_4z_1z_3 \\ z'_3 &= z_3[a(k - k_c) - b|z_3|^2 - cN_1 - dN_2] - e\bar{z}_1z_2z_4 \\ z'_4 &= z_4[a(k - k_c) - b|z_4|^2 - cN_2 - dN_1] - e\bar{z}_2z_1z_3 \end{aligned} \quad (4.6)$$

where $N_1 = |z_1|^2 + |z_3|^2$ and $N_2 = |z_2|^2 + |z_4|^2$. Here, a, \dots, e are all complex numbers; a depends only on the linearized equation, while b, \dots, e depend on $F''(0)^2, F'''(0)$ and $w(x)$. For the case of no quadratic nonlinearities ($u_t = 0$), $b = c = d = e$. There are many qualitatively different solutions to this system which correspond to interesting patterns. Silber and Knobloch [47] describe each of them as well as their conditions for stability. Travelling roll patterns (TR) consist of either horizontal or vertical traveling waves that are constant along one direction. They correspond to solutions to Eq. (4.6) where exactly one $z_j \neq 0$. Standing rolls correspond to $z_1 = z_3 \neq 0, z_2 = z_4 = 0$. (Note, the contrary case with $z_1 = z_3 = 0$

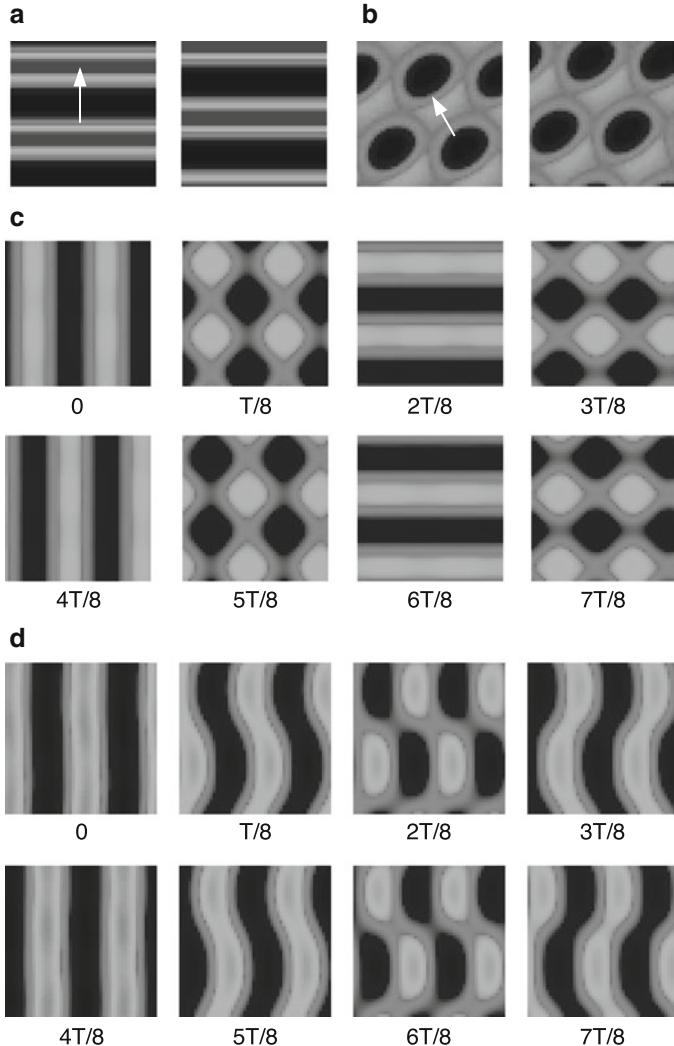


Fig. 4.3 Two-dimensional time-periodic patterns with period T in (4.1) for $\beta = 0.25, \alpha = 0.1$: **(a)** $k = 0.1, r = 3, u_{th} = 0$; **(b)** $k = 0.09, r = 5, u_{th} = 0.3$; **(c)** $k = 0.085, r = 3, u_{th} = 0$; **(d)** $k = 0.09, r = 3, u_{th} = 0$

and $z_2 = z_4 \neq 0$ are also standing rolls.) Traveling squares or spots correspond to $z_1 = z_2 \neq 0$ and $z_3 = z_4 = 0$. Standing squares (a blinking checkerboard pattern) correspond to $z_1 = z_2 = z_3 = z_4 \neq 0$. A very interesting pattern that we see is the alternating roll pattern where horizontal blinking stripes switch to vertical blinking stripes. These correspond to solutions of the form $z_1 = -iz_2 = z_3 = -iz_4 \neq 0$. Figure 4.3 illustrates the results of simulations of Eq. (4.1) on the square doubly periodic domain in the case where $m^* = 2$. Thus, all the patterns show two spatial

cycles along the principle directions. In the simulations illustrated in the figure, we change u_{th}, r which affect the values of $F''(0), F'''(0)$ and thus the values of the coefficients of the normal form, (4.6). The relative sizes of these coefficients determine both the amplitude and the stability of the patterns. Figure 4.3a shows the TR solutions for $u_{th} = 0$ (which makes $F''(0)$ vanish), while panel b shows a traveling spot pattern. Neither of these patterns can be simultaneously stable. However, there can be other patterns that stably coexist. Figure 4.3c illustrates the “alternating roll” pattern in which there is a switch from vertical to horizontal standing roles. Figure 4.3d shows a pattern that combines a standing roll (alternating vertical stripes) with a checkerboard pattern in between.

Dawes [14] has partially analyzed the more complicated case in which there are 8 critical wave vectors, for example $m^* = \sqrt{5}$ in Fig. 4.2. All of the patterns we described above are also found as solutions to his amplitude equations. In some specific cases, he finds evidence of chaotic behavior. Thus, even near the bifurcation, we can expect the possibility of complex spatiotemporal dynamics in models like present equations. Tass [50] also considers this case, but only when the quadratic terms (e.g., $F''(0)$) are zero. Obviously, there is a great reduction in the complexity of the patterns and the resulting possibilities are restricted. The $m^* = 5$ case has, to our knowledge, not yet been analyzed.

4.2.3 Summary of Pattern Formation

On a periodic one-dimensional domain, Eq. (4.1) can undergo a variety of bifurcations from the homogeneous state and these can be reduced via the construction of normal forms to one or two ordinary differential equations for the complex amplitudes. These bifurcations are generic in the sense that you can expect them to happen as you vary a *single* parameter. If you have the freedom to vary several parameters, then it is possible to arrange them so that multiple instabilities occur at the same time. For example [19] looked at the Wilson-Cowan neural field equations when $W(m) = W(m+1)$ with corresponding imaginary eigenvalues (a double Hopf bifurcation). More recently, [13] studied (4.1) near $R = 1$. When $R = 1$, recall that both the trace and the determinant vanish at the critical wave number and critical sensitivity k . Thus, there is a Bogdanov-Takens bifurcation. The normal form is more complicated in this case; however for (4.1), the only solutions that were found were the stationary periodic patterns, standing waves, and traveling waves.

In two spatial dimensions, the dynamics is considerably richer due the symmetry of the square allowing for many critical wave vectors to become unstable simultaneously. The richness increases with the size of the critical wavenumber m^* . As a ballpark estimate, the critical wavenumber is proportional to the ratio of the domain size and the spatial scale of the connectivity function $w(x)$. Thus, for, say, global inhibition, the critical wavenumber is close to 1 and the possible patterns are very simple. We remark that by estimating the spatial frequency of visual hallucinations, it is possible to estimate the characteristic length scale in visual cortex [5].

4.3 Response to Inputs in the Ring Network

We now consider the effects of linear adaptation in the ring model [13, 26] in the presence of external inputs. We show that adaptation usually degrades the ability of the network to track input locations. We consider the domain $D = (-\pi, \pi)$ and take w to be the cosine function [26]

$$w(x - y) = \cos(x - y), \quad (4.7)$$

so $w(x - y) \geq 0$ when $|x - y| \leq \pi/2$. Networks with lateral-inhibitory synaptic weights like (4.7) are known to sustain stable stationary bumps [1, 4, 8, 26]. Many of our calculations are demonstrated in the case that the firing rate function f is the Heaviside step function (see Chaps. 3, 5, 7 and [1, 4, 8, 39]).

$$F(u) \equiv H(u - \theta) = \begin{cases} 1 : x > \theta, \\ 0 : x < \theta. \end{cases} \quad (4.8)$$

We consider both stationary and propagating inputs with the simple functional form

$$I(x, t) = I_0 \cos(x - c_0 t), \quad (4.9)$$

so they are unimodal in x . We study the variety of bifurcations that can arise in the system (4.1) due to the inclusion of adaptation and inputs.

For vanishing adaptation ($\beta \rightarrow 0$), we find stable stationary bumps. For sufficiently strong adaptation, the input-free ($I_0 = 0$) network (4.1) supports traveling bumps (pulses). The network locks to moving inputs as long as their speed is sufficiently close to that of naturally arising traveling bumps. Otherwise, activity periodically slips off of the stimulus or sloshes about the vicinity of the stimulus location. Previously, Hansel and Sompolinsky [26] studied many of these results, and recently [31] reinterpreted many of these findings in the context of hallucinogen-related visual pathologies.

4.3.1 Existence of Stationary Bumps

First, we study existence of stationary bump solutions in the presence of stationary inputs ($I(x, t) \equiv I(x)$). Assuming stationary solutions $(u(x, t), v(x, t)) = (U(x), V(x))$ to (4.1) generates the single equation

$$(1 + \beta)U(x) = \int_{-\pi}^{\pi} w(x - y)F(U(y))dy + I(x). \quad (4.10)$$

For a cosine weight kernel (4.7), we can exploit the trigonometric identity

$$\cos(x - y) = \cos y \cos x + \sin y \sin x, \quad (4.11)$$

and consider the cosine input (4.9), which we take to be stationary ($c_0 = 0$). This suggests looking for even-symmetric solutions

$$U(x) = \left(A + \frac{I_0}{1 + \beta} \right) \cos x, \quad (4.12)$$

so that the amplitude of (4.12) is specified by the implicit equation

$$A = \frac{1}{1 + \beta} \int_{-\pi}^{\pi} \cos y F((A + (1 + \beta)^{-1} I_0) \cos y) dy. \quad (4.13)$$

For a Heaviside firing rate function (4.8), we can simplify the implicit equation (4.13), using the fact that (4.12) is unimodal and symmetric so that $U(x) > \theta$ for $x \in (-a, a)$ for solutions $A > 0$. First of all, this means that the profile of $U(x)$ crosses through threshold θ at two distinct points [1, 4, 8]

$$U(\pm a) = [A + (1 + \beta)^{-1} I_0] \cos a = \theta \Rightarrow a = \cos^{-1} \left[\frac{(1 + \beta)\theta}{(1 + \beta)A + I_0} \right]. \quad (4.14)$$

The threshold condition (4.14) converts the integral equation (4.13) to

$$A = \frac{1}{1 + \beta} \int_{-a}^a \cos y dy = \frac{2}{1 + \beta} \sqrt{1 - \frac{(1 + \beta)^2 \theta^2}{((1 + \beta)A + I_0)^2}}, \quad (4.15)$$

which can be converted to a quartic equation and solved analytically [30].

In the limit of no input $I_0 \rightarrow 0$, the amplitude of the bump is given by the pair of real roots of (4.15)

$$A_{\pm} = \frac{\sqrt{1 + (1 + \beta)\theta} \pm \sqrt{1 - (1 + \beta)\theta}}{1 + \beta}, \quad (4.16)$$

so there are two bump solutions. As is usually found in lateral inhibitory neural fields, the wide bump (+) is stable and the narrow bump (−) is unstable in the limit of vanishing adaptation ($\beta \rightarrow 0$) [1, 4, 12, 40]. At a critical β , the wide bump undergoes a drift instability leading to a traveling bump.

4.3.2 Linear Stability of Stationary Bumps

We now compute stability of the bump (4.12) by studying the evolution of small, smooth, separable perturbations. By plugging $u = U(x) + \psi(x)e^{\lambda t}$ and $v = V(x) +$

$\phi(x)e^{\lambda t}$ (where $|\psi(x)| \ll 1$ and $|\phi(x)| \ll 1$) into (4.1), Taylor expanding, and truncating to first order we find the linear system

$$(\lambda + 1)\psi(x) = -\beta\phi(x) + \int_{-\pi}^{\pi} w(x-y)F'(U(y))\psi(y)dy, \quad (4.17)$$

$$(\lambda + \alpha)\phi(x) = \alpha\psi(x). \quad (4.18)$$

For the cosine weight function (4.7), we apply the identity (4.11) and substitute (4.18) into (4.17) to yield the single equation

$$\mathcal{Q}(\lambda)\psi(x) = (\lambda + \alpha)(\mathcal{A}\cos x + \mathcal{B}\sin x) \quad (4.19)$$

where $\mathcal{Q}(\lambda) = (\lambda + \alpha)(\lambda + 1) + \alpha\beta$ and

$$\mathcal{A} = \int_{-\pi}^{\pi} \cos x F'(U(x))\psi(x)dx, \quad \mathcal{B} = \int_{-\pi}^{\pi} \sin x F'(U(x))\psi(x)dx. \quad (4.20)$$

We can then plug (4.19) into the system of equations (4.20) and simplify to yield

$$\mathcal{Q}(\lambda)\mathcal{A} = (\lambda + \alpha) \left(\int_{-\pi}^{\pi} F'(U(x))dx - \frac{(1 + \beta)^2 A}{(1 + \beta)A + I_0} \right) \mathcal{A}, \quad (4.21)$$

$$\mathcal{Q}(\lambda)\mathcal{B} = \frac{(\lambda + \alpha)(1 + \beta)^2 A}{(1 + \beta)A + I_0} \mathcal{B}, \quad (4.22)$$

where we have used the fact that integrating (4.13) by parts yields

$$A = \frac{A + (1 + \beta)^{-1}I_0}{1 + \beta} \int_{-\pi}^{\pi} \sin^2 x F'((A + (1 + \beta)^{-1}I_0) \cos x)dx,$$

as well as the fact that the off-diagonal terms vanish, since their integrands are odd. This means that the eigenvalues determining the linear stability of the bump (4.12) are of two classes: (a) those of even perturbations so $\psi(x) = \cos x$ and (b) those of odd perturbations where $\psi(x) = \sin x$. We primarily study eigenvalues associated with odd perturbations, given by the quadratic equation

$$\lambda^2 + [1 + \alpha - (1 + \beta)\Omega]\lambda + \alpha(1 + \beta)(1 - \Omega) = 0, \quad \Omega = \frac{(1 + \beta)A}{(1 + \beta)A + I_0}. \quad (4.23)$$

We can use (4.23) to study two bifurcations of stationary bumps in the system (4.1). First, we show a drift instability arises in the input-free ($I_0 = 0$) network, leading to a pitchfork bifurcation whose resultant attracting solutions are traveling bumps [10, 13, 26, 35, 39]. Second, we show that in the input-driven system ($I_0 > 0$), an

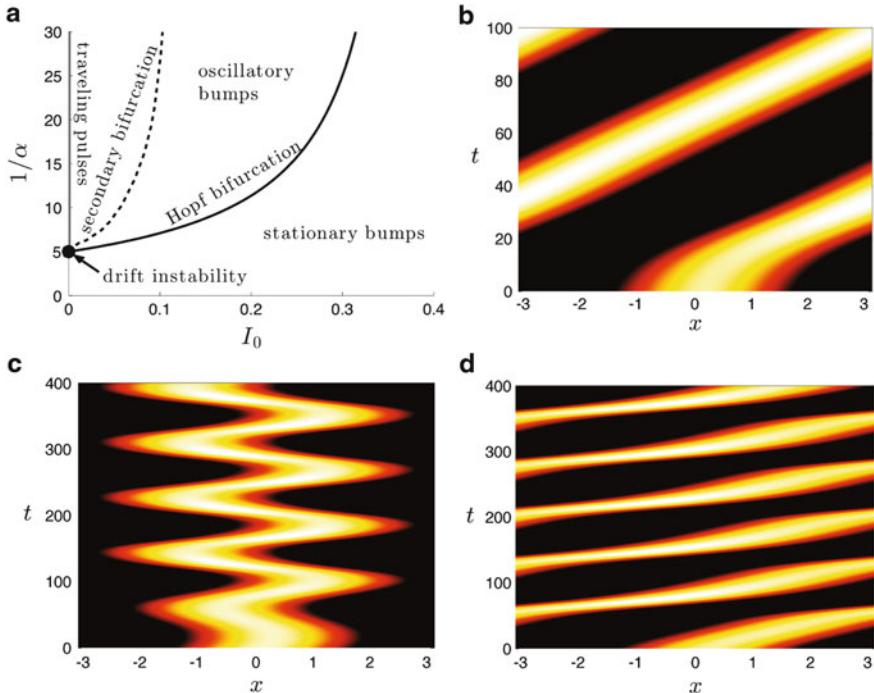


Fig. 4.4 (a) Partition of (I_0, α^{-1}) parameter space into different dynamical behaviors of the bump solution (4.12) for Heaviside firing rate (4.8). Numerical simulation of the (b) drift instability of the bump (4.12) in the case of no input ($I_0 = 0$); (c) sloshing oscillatory instability in the case of input $I_0 = 0.1$; and (d) translation variant propagation in the case of weak input $I_0 = 0.05$. Other parameters are $\theta = 0.5$, $\alpha = 0.1$, and $\beta = 0.2$

oscillatory instability arises where the edges of the ‘‘slosh’’ periodically. This is a Hopf bifurcation, which also persists for moving inputs ($c_0 > 0$).

In the limit of no input ($I_0 \rightarrow 0$), $\Omega \rightarrow 1$, so (4.23) reduces to

$$\lambda^2 + [\alpha - \beta]\lambda = 0. \quad (4.24)$$

There is always a zero eigenvalue, due to the translation symmetry of the input-free network [1, 40]. Fixing adaptation strength β , we can decrease the rate α from infinity to find the other eigenvalue crosses zero when $\alpha = \beta$. We mark this point in our partition of parameter space into different dynamical behaviors in Fig. 4.4a. This non-oscillatory instability results in a traveling bump, indicated by the associated shift eigenfunction ($\sin x$). Traveling pulses can propagate in either direction, so the full system (4.1) undergoes a pitchfork bifurcation. We demonstrate the instability resulting in a traveling bump in Fig. 4.4b.

We could also ensure that instabilities associated with even perturbations ($\cos x$) of the bump (4.12) do not occur prior to this loss of instability of the odd

perturbation. For brevity, we omit this calculation. Numerical simulations (as shown in Fig. 4.4b) verify odd perturbations are the first to destabilize. Therefore, we would always expect that as α is decreased from infinity, the first instability that arises is associated with odd perturbations of the bump, leading to a drift instability and thus a traveling bump solution (see Fig. 4.4).

For nonzero input ($I_0 > 0$), the primary bifurcation of the stable (wide) stationary bump solution is shown to be oscillatory. To identify the location of this Hopf bifurcation, we plug the ansatz $\lambda = i\omega$ into (4.23) to find

$$-\omega^2 + i[(1 + \alpha) - (1 + \beta)\Omega]\lambda + \alpha(1 + \beta)(1 - \Omega) = 0. \quad (4.25)$$

Equating real and imaginary parts of (4.25), we find a Hopf bifurcation occurs when

$$\alpha_H = (1 + \beta)\Omega - 1, \quad (4.26)$$

with onset frequency

$$\omega_H = \sqrt{\alpha(1 + \beta)(1 - \Omega)}. \quad (4.27)$$

Since $\Omega \in (0, 1)$ when $I_0 > 0$, we know that $\omega_H > 0$ for all parameter values we consider. Therefore, there is never an instability with purely real eigenvalues associated with odd perturbations, in the case of nonzero input. We show the curve of Hopf bifurcations in (I_0, α^{-1}) parameter space in Fig. 4.4a as well as a simulation of the resulting oscillatory solution in Fig. 4.4c. Studies of input-driven excitatory networks reveal it is the even mode that destabilizes into oscillations, yielding reflection symmetric breathers [22, 23]. Here, due to the lateral inhibitory kernel, the odd eigenmode destabilizes, leading to sloshing breathers [22, 42]. As in the case of the drift instability, we should ensure that instabilities associated with even perturbations do not arise prior to the Hopf bifurcation. We have ensured this for the calculations of Fig. 4.4 but do not show this explicitly here.

Finally, we note a secondary bifurcation which leads to dynamics that evolves as a propagating pattern with varying width (see Fig. 4.4d). Essentially, the “sloshing” bump breaks free from the attraction of the pinning stimulus and begins to propagate. As it passes over the location of the stimulus, it expands. Such secondary bifurcations have been observed in adaptive neural fields on infinite spatial domains too [23]. While we cannot develop a linear theory for this bifurcation, we can determine the location of this bifurcation numerically.

4.3.3 Existence of Traveling Bumps

Our linear stability analysis of stationary bumps predicts the existence of traveling bumps for substantially slow and strong adaptation. We can also show that when

a moving input is introduced, the system tends to lock to it if it has speed commensurate with that of the natural wave. Converting to a wave coordinate frame $\xi = x - c_0 t$ where we choose the stimulus speed c_0 , we can study traveling wave solutions $(u(x, t), v(x, t)) = (U(\xi), V(\xi))$ of (4.1) with the second order differential equation [23]

$$-c_0^2 U''(\xi) + c_0(1 + \alpha)U'(\xi) - \alpha(1 + \beta)U(\xi) = G(\xi) \quad (4.28)$$

where

$$G(\xi) = \left(c \frac{d}{d\xi} - \alpha \right) \left[\int_{-\pi}^{\pi} w(\xi - y) F(U(y)) dy + I(\xi + \Delta_I) \right], \quad (4.29)$$

and Δ_I specifies the spatial shift between the moving input and the pulse that tracks it. In the case of a cosine weight kernel (4.7) and input (4.9), we can apply the identity (4.11)–(4.29) so we may write Eq. (4.28) as

$$-c_0^2 U''(\xi) + c_0(1 + \alpha)U'(\xi) - \alpha(1 + \beta)U(\xi) = \mathcal{C} \cos \xi + \mathcal{S} \sin \xi, \quad (4.30)$$

where

$$\mathcal{C} = \int_{-\pi}^{\pi} \cos x \left[c_0 F'(U(x))U'(x) - \alpha F(U(x)) \right] dx - I_0(\alpha \cos \Delta_I + c_0 \sin \Delta_I), \quad (4.31)$$

$$\mathcal{S} = \int_{-\pi}^{\pi} \sin x \left[c_0 F'(U(x))U'(x) - \alpha F(U(x)) \right] dx + I_0(\alpha \sin \Delta_I - c_0 \cos \Delta_I). \quad (4.32)$$

By treating \mathcal{C} and \mathcal{S} as constants, it is straightforward to solve the second order differential equation (4.30) to find

$$U(\xi) = \frac{(c_0^2 - \alpha - \alpha\beta)[\mathcal{C} \cos \xi + \mathcal{S} \sin \xi] + c_0(1 + \alpha)[\mathcal{C} \sin \xi - \mathcal{S} \cos \xi]}{(c_0^2 - \alpha(1 + \beta))^2 + c_0^2(1 + \alpha)^2}. \quad (4.33)$$

In the case of a Heaviside firing rate function (4.8), we can evaluate the integral terms of \mathcal{C} and \mathcal{S} directly. First, we break the translation symmetry of the system by fixing the threshold crossing points, $U(\pi) = U(\pi - \Delta) = \theta$. This specifies the input shift parameter Δ_I as well. We also require that the superthreshold region $U(\xi) > \theta$ when $x \in (\pi - \Delta, \pi)$ and $U(\xi) < \theta$ otherwise. This yields

$$\mathcal{C} = \alpha \sin \Delta + c_0(1 - \cos \Delta) - I_0(\alpha \cos \Delta_I + c_0 \sin \Delta_I), \quad (4.34)$$

$$\mathcal{S} = c_0 \sin \Delta - \alpha(1 - \cos \Delta) + I_0(\alpha \sin \Delta_I - c_0 \cos \Delta_I). \quad (4.35)$$

Plugging this into (4.33) and imposing threshold conditions, we have the system

$$\frac{\mathcal{X}_1[\sin \Delta - I_0 \cos \Delta_I] - \mathcal{X}_2[1 - \cos \Delta - I_0 \sin \Delta_I]}{(c_0^2 - \alpha(1 + \beta))^2 + c_0^2(1 + \alpha)^2} = \theta, \quad (4.36)$$

$$\frac{\mathcal{X}_1[\sin \Delta - I_0 \cos(\Delta - \Delta_I)] + \mathcal{X}_2[1 - \cos \Delta - I_0 \sin(\Delta - \Delta_I)]}{(c_0^2 - \alpha(1 + \beta))^2 + c_0^2(1 + \alpha)^2} = \theta, \quad (4.37)$$

where $\mathcal{X}_1 = c_0^2 + \alpha^2(1 + \beta)$ and $\mathcal{X}_2 = c_0^3 + c_0\alpha^2 - c_0\alpha\beta$, which we could solve the numerically (see [31]).

In the limit of no input ($I_0 \rightarrow 0$), we can treat $c = c_0$ as an unknown parameter. By taking the difference of (4.37) and (4.36) in this limit, we see that we can compute the speed of natural waves by studying solutions of

$$c^3 + c\alpha^2 - c\alpha\beta = 0, \quad (4.38)$$

a cubic equation providing up to three possible speeds for a traveling bump solution. The trivial $c = 0$ solution is the limiting case of stationary bump solutions that we have already studied and is unstable when $\alpha < \beta$. In line with our bump stability predictions, for $\alpha \leq \beta$, we have the two additional solutions $c_{\pm} = \pm\sqrt{\alpha\beta - \alpha^2}$, which provides a right-moving (+) and left-moving (-) traveling bump solution. The pulse widths are then given applying the expression (4.38) into (4.36) and (4.37) and taking their mean to find $\sin \Delta = (1 + \alpha)\theta$. Thus, we can expect to find four traveling bump solutions, two with each speed, that have widths $\Delta_s = \pi - \sin^{-1}[\theta(1 + \alpha)]$ and $\Delta_u = \sin^{-1}[\theta(1 + \alpha)]$. We can find, using linear stability analysis, that the two traveling bumps associated with the width Δ_s are stable [35, 39].

4.3.4 Linear Stability of Traveling Bumps

To analyze the linear stability of stimulus-locked traveling bumps (4.33), we study the evolution of small, smooth, separable perturbations to $(U(\xi), V(\xi))$. To find this, we plug the expansions $u(x, t) = U(\xi) + \psi(\xi)e^{\lambda t}$ and $v(x, t) = V(\xi) + \phi(\xi)e^{\lambda t}$ (where $|\psi(\xi)| \ll 1$ and $|\phi(\xi)| \ll 1$) and truncate to first order to find the linear equation [10, 25, 56]

$$-c_0\psi'(\xi) + (\lambda + 1)\psi(\xi) = -\beta\phi(\xi) + \int_{-\pi}^{\pi} w(\xi - y)F'(U(y))\psi(y)dy, \quad (4.39)$$

$$-c_0\phi'(\xi) + (\lambda + \alpha)\phi(\xi) = \alpha\psi(\xi). \quad (4.40)$$

For the cosine weight function (4.7), we can apply the identity (4.11), so that upon converting the system to a second order differential equation, we

$$-c_0^2\psi'' + c(2\lambda + 1 + \alpha)\psi' - [(\lambda + 1)(\lambda + \alpha) + \alpha\beta]\psi = \mathcal{A}\cos\xi + \mathcal{B}\sin\xi, \quad (4.41)$$

where

$$\mathcal{A} = -(\lambda + \alpha) \int_{-\pi}^{\pi} \cos \xi F'(U(\xi)) \psi(\xi) d\xi + c_0 \int_{-\pi}^{\pi} \sin \xi F'(U(\xi)) \psi(\xi) d\xi, \quad (4.42)$$

$$\mathcal{B} = -c_0 \int_{-\pi}^{\pi} \cos \xi F'(U(\xi)) \psi(\xi) d\xi - (\lambda + \alpha) \int_{-\pi}^{\pi} \sin \xi F'(U(\xi)) \psi(\xi) d\xi. \quad (4.43)$$

Employing periodic boundary conditions $\psi(-\pi) = \psi(\pi)$ and $\psi'(-\pi) = \psi'(\pi)$ and treating \mathcal{A} and \mathcal{B} as constants, it is then straightforward to solve (4.41) to find

$$\psi(\xi) = \frac{\mathcal{P}_2 \mathcal{A} - \mathcal{P}_1 \mathcal{B}}{\mathcal{D}_p} \cos \xi + \frac{\mathcal{P}_1 \mathcal{A} + \mathcal{P}_2 \mathcal{B}}{\mathcal{D}_p} \sin \xi. \quad (4.44)$$

where $\mathcal{P}_1 = c_0(2\lambda + 1 + \alpha)$, $\mathcal{P}_2 = c_0^2 - [(\lambda + 1)(\lambda + \alpha) + \alpha\beta]$, and $\mathcal{D}_p = \mathcal{P}_1^2 + \mathcal{P}_2^2$. We can then use self-consistency to determine the constants \mathcal{A} and \mathcal{B} , which implicitly depend upon ψ itself. In the case that the firing rate function is a Heaviside (4.8), we can reduce this to a pointwise dependence, so that

$$\mathcal{A} = \frac{c_0 \sin \Delta \psi(\pi - \Delta)}{|U'(\pi - \Delta)|} + (\lambda + \alpha) \left[\frac{\psi(\pi)}{|U'(\pi)|} + \frac{\cos \Delta \psi(\pi - \Delta)}{|U'(\pi - \Delta)|} \right], \quad (4.45)$$

$$\mathcal{B} = c_0 \left[\frac{\psi(\pi)}{|U'(\pi)|} + \frac{\cos \Delta \psi(\pi - \Delta)}{|U'(\pi - \Delta)|} \right] - \frac{(\lambda + \alpha) \sin \Delta \psi(\pi - \Delta)}{|U'(\pi - \Delta)|}, \quad (4.46)$$

and we can write the solution

$$\psi(\xi) = \frac{\mathcal{C}_1 \cos \xi + \mathcal{S}_1 \sin \xi}{\mathcal{D}_p} \frac{\psi(\pi)}{|U'(\pi)|} + \frac{\mathcal{C}_2 \cos \xi + \mathcal{S}_2 \sin \xi}{\mathcal{D}_p} \frac{\psi(\pi - \Delta)}{|U'(\pi - \Delta)|},$$

where

$$\begin{aligned} \mathcal{C}_1 &= \mathcal{P}_2(\lambda + \alpha) - \mathcal{P}_1 c_0, & \mathcal{S}_1 &= \mathcal{P}_1(\lambda + \alpha) + \mathcal{P}_2 c_0, \\ \mathcal{C}_2 &= \mathcal{P}_1((\lambda + \alpha) \sin \Delta - c_0 \cos \Delta) + \mathcal{P}_2(c_0 \sin \Delta + (\lambda + \alpha) \cos \Delta), \\ \mathcal{S}_2 &= \mathcal{P}_1((\lambda + \alpha) \cos \Delta + c_0 \sin \Delta) + \mathcal{P}_2(c_0 \cos \Delta - (\lambda + \alpha) \sin \Delta). \end{aligned}$$

Applying self consistency, we have a 2×2 eigenvalue problem $\Psi = \mathcal{A}_p \Psi$, where

$$\Psi = \begin{pmatrix} \psi(\pi) \\ \psi(\pi - \Delta) \end{pmatrix}, \quad \mathcal{A}_p = \begin{pmatrix} \mathcal{A}_{\pi\pi} & \mathcal{A}_{\pi\Delta} \\ \mathcal{A}_{\Delta\pi} & \mathcal{A}_{\Delta\Delta} \end{pmatrix}, \quad (4.47)$$

with

$$\mathcal{A}_{\pi\pi} = -\frac{\mathcal{C}_1}{\mathcal{D}_p |U'(\pi)|}, \quad \mathcal{A}_{\pi\Delta} = -\frac{\mathcal{C}_2}{\mathcal{D}_p |U'(\pi - \Delta)|},$$

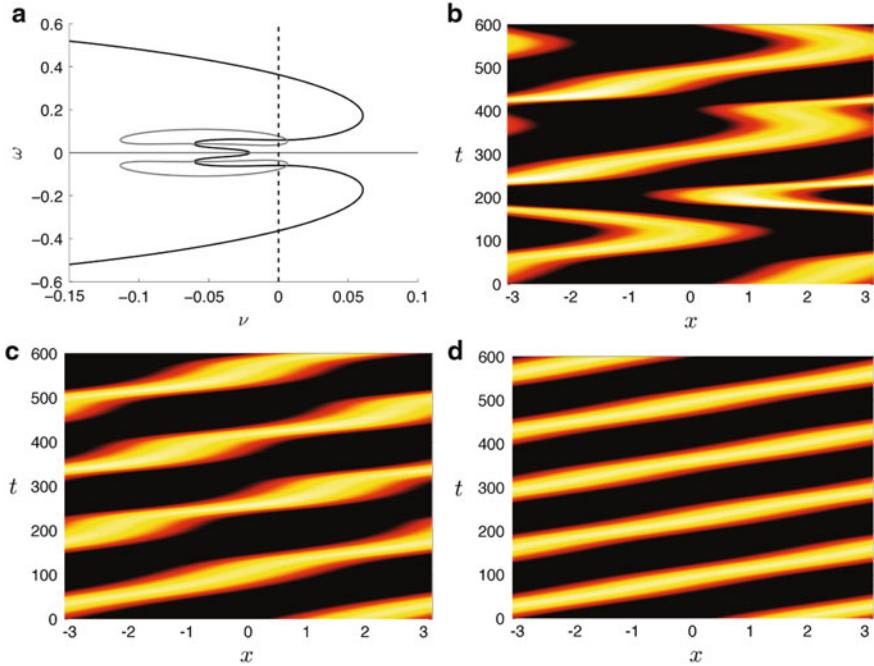


Fig. 4.5 Sloshing instability of stimulus-locked traveling bumps (4.33) in adaptive neural field (4.1) with Heaviside firing rate (4.8). **(a)** Dependence of stimulus locked pulse width Δ on stimulus speed c_0 , calculated using the implicit equations (4.36) and (4.37). **(a)** Zeros of the Evans function $\mathcal{E}(\lambda) = \det(\mathcal{A}_p - I)$, with (4.47), occur at the crossings of the zero contours of $\text{Re}\mathcal{E}(\lambda)$ (black) and $\text{Im}\mathcal{E}(\lambda)$ (grey). Presented here for stimulus speed $c_0 = 0.042$, just beyond the Hopf bifurcation at $c_H \approx 0.046$. Breathing instability occurs in numerical simulations for **(b)** $c_0 = 0.036$ and **(c)** $c_0 = 0.042$. **(d)** When stimulus speed $c_0 = 0.047$ is sufficiently fast, stable traveling bumps lock. Other parameters are $\theta = 0.5$, $\alpha = 0.05$, $\beta = 0.2$, and $I_0 = 0.1$

$$\mathcal{A}_{\Delta\pi} = \frac{\mathcal{S}_1 \sin \Delta - \mathcal{C}_1 \cos \Delta}{\mathcal{D}_p |U'(\pi)|}, \quad \mathcal{A}_{\Delta\Delta} = \frac{\mathcal{S}_2 \sin \Delta - \mathcal{C}_2 \cos \Delta}{\mathcal{D}_p |U'(\pi - \Delta)|}.$$

Then, applying the approach of previous stability analyses of traveling waves in neural fields [10, 25, 56], we examine nontrivial solutions of $\Psi = \mathcal{A}_p \Psi$ so that $\mathcal{E}(\lambda) = 0$, where $\mathcal{E}(\lambda) = \det(\mathcal{A}_p - I)$ is called the Evans function of the traveling bump solution (4.33). Since no other parts of the spectrum contribute to instabilities in this case, the traveling bump is linearly stable as long as $\text{Re } \lambda < 0$ for all λ such that $\mathcal{E}(\lambda) = 0$. We can find the zeros of the Evans function by following the approach of [10, 25] and writing $\lambda = \nu + i\omega$ and plotting the zero contours of $\text{Re } \mathcal{E}(\lambda)$ and $\text{Im } \mathcal{E}(\lambda)$ in the (ν, ω) -plane. The Evans function is zero where the lines intersect.

We present examples of this analysis in Fig. 4.5. As shown, we can use the implicit equations (4.36) and (4.37) to compute the width of a stimulus-locked pulse as it depends upon the speed of the input in the case of a Heaviside firing rate

function (4.8). In parameter regime we show, there are two pulses for each parameter value, either both are unstable or one is stable. As the speed of stimuli is decreased, a stable traveling bump undergoes a Hopf bifurcation. For sufficiently fast stimuli, a stable traveling bump can lock to the stimulus, as shown in Fig. 4.5d. However, for sufficiently, slow stimuli, the speed of natural traveling bumps of the stimulus free network is too fast to track the stimuli. Therefore, an oscillatory instability results. We plot the zeros of the Evans functions associated with this instability in Fig. 4.5a. The sloshing pulses that result are pictured in Fig. 4.5b, c. Note that, as was shown in [31], it is possible for pulses to destabilize due to stimuli being too fast. In this context, such an instability occurs through a saddle-node bifurcation, rather than a Hopf.

4.4 Activity Bumps on the Infinite Line

We consider neural field (4.1) with a Heaviside firing rate function $F(u) = H(u - \theta)$ with firing threshold θ where $u(x, t)$ and $v(x, t)$ are defined along the infinite line with $u(x, t), v(x, t) \rightarrow 0$ as $x \rightarrow \pm\infty$. The even-symmetric synaptic weight function w is assumed to be either excitatory ($w(x) > 0$) or of Mexican hat form (locally positive, laterally negative) satisfying $\int_{-\infty}^{\infty} w(y) dy < \infty$. We consider *stationary* activity bumps in Sect. 4.4.1 and *traveling* activity bumps in Sect. 4.4.2 and examine the cases of (i) bumps generated intrinsically by the network with no input ($I(x, t) = 0$) and (ii) bumps induced by a localized, excitatory input inhomogeneity ($I(x, t) > 0$) which can be either stationary ($I(x)$) or traveling ($I(x - ct)$) with constant speed c . The input is assumed to have an even-symmetric, Gaussian-like profile satisfying $I(x) \rightarrow 0$ as $x \rightarrow \pm\infty$.

4.4.1 Natural and Stimulus-Induced Stationary Activity Bumps

Existence of Stationary Bumps. An equilibrium solution of (4.1) is expressed as $(u(x, t), v(x, t))^T = (U_o(x), V_o(x))^T$ and satisfies $V_o(x) = U_o(x)$ and

$$(1 + \beta)U_o(x) = \int_{-\infty}^{\infty} w(x - y) H(U_o(y) - \theta) dy + I(x). \quad (4.48)$$

We follow the approach of Amari [1] to use the Heaviside firing rate and make the ansatz of an even-symmetric stationary bump $U_o(x)$ that is centered about $x = 0$, is superthreshold $U_o(x) > \theta$ for $x \in (-a, a)$, satisfies $U_o(\pm a) = \theta$, and is subthreshold otherwise with $U_o(x) \rightarrow 0$ as $x \rightarrow \pm\infty$ (see Fig. 4.6). That the stationary bump is centered about $x = 0$ is by choice both (i) in the case of no input ($I(x) = 0$) due to translation symmetry of the bump and (ii) in the presence

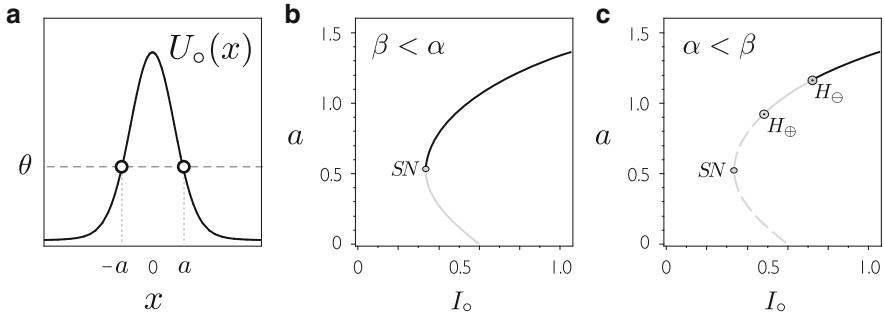


Fig. 4.6 (a) Stationary bump profile $U_o(x)$ with halfwidth a . Bifurcation curves satisfying (4.50) and illustrating the dependence of a on the bifurcation parameter I_o are shown in (b) for $\beta < \alpha$ and in (c) for $\alpha < \beta$. Black (gray) denote stability (instability) of the stationary bump. SN denotes a saddle-node bifurcation and H_{\oplus} and H_{\ominus} denote Hopf bifurcations with respect to the sum mode Ω_+ and difference mode Ω_- , respectively. Parameters are $\bar{w}_e = 1$, $\sigma_e = 1$, $\bar{w}_i = 0.4$, $\sigma_i = 2$, $\theta = 0.3$, $\alpha = 0.025$, $\beta = 1$, $\sigma = 1.2$ (Figure adapted from Folias [22])

of a stationary input ($I(x) \neq 0$) where the stationary bump and the input share the same center, which is set to be $x = 0$. The profile $U_o(x)$ of the stationary bump can then be expressed as

$$(1 + \beta)U_o(x) = \int_{-a}^a w(x - y) dy + I(x) = [W(x + a) - W(x - a) + I(x)] \quad (4.49)$$

where $W(x) = \int_0^x w(y) dy$. The bump halfwidth a is then determined by requiring (4.49) to satisfy the *threshold conditions* $U_o(\pm a) = \theta$ which, by even symmetry, result in

$$W(2a) + I(a) = (1 + \beta)\theta. \quad (4.50)$$

This determines the existence of the stationary bump if all assumptions are satisfied. Condition (4.50) was solved numerically in Fig. 4.6 where w and I were taken to be

$$w(x) = \frac{\bar{w}_e}{\sqrt{\pi}\sigma_e} e^{-(x/\sigma_e)^2} - \frac{\bar{w}_i}{\sqrt{\pi}\sigma_i} e^{-(x/\sigma_i)^2}, \quad I(x) = I_o e^{-(x/\sigma)^2}. \quad (4.51)$$

Existence Results for Stationary Bumps for General w and Gaussian-like I .

CASE I: *No Input* ($I(x) = 0$). For an excitatory weight function ($w(x) > 0$), stationary bumps exist and satisfy (4.50) when parameters permit ($0 < \theta < \lim_{x \rightarrow \infty} W(x)$); however, they are always linearly unstable [22, 23, 39]. The case of a Mexican hat weight function w is an extension of the Amari neural field [1] with the existence equation containing an extra factor due to adaptation ($W(2a) = \theta(1 + \beta)$); however, the

dynamics of the adaptation variable v additionally governs the stability of the stationary bump [22]. In particular, if $\alpha < \beta$, stationary bumps are always unstable. Stable bumps in the scalar model of Amari can extend to this model only for $\alpha > \beta$, and a stable bump for $\alpha > \beta$ destabilizes as α decreases through $\alpha = \beta$ leading to a drift instability [22] that gives rise to traveling bumps.

CASE II: *Localized Excitatory Input ($I(x) > 0$)*. A variety of bifurcation scenarios can occur [22, 23], and, importantly, stationary bumps can emerge in a saddle-node bifurcation for strong inputs in parameter regimes where stationary bumps do not exist for weak or zero input as shown in Fig. 4.6. When stationary bumps exist for $\alpha > \beta$, the stability of a bump is determined directly by the geometry of the bifurcation curves [22, 23] (e.g., see Fig. 4.6). As α decreases through $\alpha = \beta$, a Hopf bifurcation point emerges from a saddle-node bifurcation point (associated with the sum mode Ω_+) and destabilizes a segment of a branch of stable bumps for $\alpha < \beta$. Generally, Hopf bifurcations occur with respect to either of two spatial modes Ω_\pm (discussed later), and their relative positions (denoted by H_\oplus and H_\ominus , respectively, on the bifurcation curves in Fig. 4.6) can switch depending on parameters [22].

Stability of Stationary Bumps. By setting $u(x, t) = U_o(x, t) + \tilde{\varphi}(x, t)$ and $v(x, t) = V_o(x, t) + \tilde{\psi}(x, t)$, we study the evolution of small perturbations $(\tilde{\varphi}, \tilde{\psi})^\top$ in a Taylor expansion of (4.1) about the stationary bump $(U_o, V_o)^\top$. To first order in $(\tilde{\varphi}, \tilde{\psi})^\top$, the perturbations are governed by the linearization

$$\begin{aligned} \partial_t \tilde{\varphi} &= -\tilde{\varphi} - \beta \tilde{\psi} + \int_{-\infty}^{\infty} w(x-y) H'(U_o(y) - \theta) \tilde{\varphi}(y, t) dy, \\ \frac{1}{\alpha} \partial_t \tilde{\psi} &= +\tilde{\varphi} - \tilde{\psi}. \end{aligned} \quad (4.52)$$

Separating variables, we set $\tilde{\varphi}(x, t) = e^{\lambda t} \varphi(x)$ and $\tilde{\psi}(x, t) = e^{\lambda t} \psi(x)$ in (4.52) where $(\varphi, \psi)^\top \in C_u^1(\mathbb{R}, \mathbb{C}^2)$ denoting uniformly continuously differentiable vector-valued functions $\mathbf{u} : \mathbb{R} \rightarrow \mathbb{C}^2$. This leads to the spectral problem for λ and $(\varphi, \psi)^\top$

$$M \begin{pmatrix} \varphi \\ \psi \end{pmatrix} = \lambda \begin{pmatrix} \varphi \\ \psi \end{pmatrix}, \quad M \begin{pmatrix} \varphi \\ \psi \end{pmatrix} = \begin{bmatrix} -1 & -\beta \\ \alpha & -\alpha \end{bmatrix} \begin{pmatrix} \varphi \\ \psi \end{pmatrix} + \begin{pmatrix} \mathcal{N} \varphi \\ 0 \end{pmatrix}, \quad (4.53)$$

where $\mathcal{N} \varphi(x) = \int_{-\infty}^{\infty} w(x-y) H'(U_o(y) - \theta) \varphi(y) dy$. The essential spectrum lies in the left-half complex plane and plays no role in instability [22, 23]. To calculate the point spectrum, define $\rho(\lambda) = \lambda + 1 + \frac{\alpha\beta}{\lambda+\alpha}$ and reduce (4.53) to $\psi(x) = (\frac{\alpha}{\lambda+\alpha})\varphi(x)$ and

$$\rho(\lambda)\varphi(x) = \frac{w(x-a)}{|U'_o(+a)|} \varphi(a) + \frac{w(x+a)}{|U'_o(-a)|} \varphi(-a). \quad (4.54)$$

Setting $x = \pm a$ in (4.54) yields a compatibility condition for the values of $\varphi(\pm a)$

$$\begin{pmatrix} \Lambda_o - \rho(\lambda) I \\ \varphi(-a) \end{pmatrix} \begin{pmatrix} \varphi(+a) \\ \varphi(-a) \end{pmatrix} = 0, \quad \Lambda_o = \frac{1}{|U'_o(a)|} \begin{bmatrix} w(0) & w(2a) \\ w(2a) & w(0) \end{bmatrix}.$$

Consequently, nontrivial solutions of (4.53) exist when $\det(\Lambda_o - \rho(\lambda)I) = 0$, thereby identifying eigenvalues λ . The point spectrum comprises two pairs of eigenvalues λ_{\pm}^+ , λ_{\pm}^- and eigenfunctions \mathbf{v}_{\pm}^+ , \mathbf{v}_{\pm}^- defining two characteristic spatial modes [22, 23]:

Sum mode: eigenvalues λ_{\pm}^+ and eigenfunctions $\mathbf{v}_{\pm}^+(x) = \Omega_+(x)(\lambda_{\pm}^+ + \alpha, \alpha)^T$,

$$\lambda_{\pm}^+(a) = -\frac{1}{2}\Upsilon_{\pm} \pm \frac{1}{2}\sqrt{\Upsilon_{\pm}^2 - 4\Gamma_{\pm}}, \quad \Omega_+(x) = w(x-a) + w(x+a),$$

Difference Mode: eigenvalues λ_{\pm}^- and eigenfunctions $\mathbf{v}_{\pm}^-(x) = \Omega_-(x)(\lambda_{\pm}^- + \alpha, \alpha)^T$,

$$\lambda_{\pm}^-(a) = -\frac{1}{2}\Upsilon_{\pm} \pm \frac{1}{2}\sqrt{\Upsilon_{\pm}^2 - 4\Gamma_{\pm}}, \quad \Omega_-(x) = w(x-a) - w(x+a),$$

where $\Omega_+(x)$ is *even*-symmetric, $\Omega_-(x)$ is *odd*-symmetric, and Υ_{\pm} , Γ_{\pm} are given by

$$\Upsilon_{\pm}(a) = 1 + \alpha - \frac{\Omega_{\pm}(a)}{|U'_o(a)|}, \quad \Gamma_{\pm}(a) = \alpha \left[1 + \beta - \frac{\Omega_{\pm}(a)}{|U'_o(a)|} \right].$$

Stability Results for Stationary Bumps for General w and Gaussian-like I .

CASE I: *No Input* ($I(x) = 0$) [22, 23, 25, 41]. Since $|U'_o(a)| = \Omega_o(a)/(1 + \beta)$, $\Gamma_o(a) = 0$ and we can redefine the eigenvalues λ_{\pm}^- as $\lambda_+^- \equiv 0$ and $\lambda_-^- = \beta - \alpha$. The persistent 0-eigenvalue $\lambda_+^- \equiv 0$ corresponds to the translation invariance of the stationary bump and is associated with an eigenfunction in the difference mode Ω_- . The other eigenfunction in the difference mode (associated with λ_-^-) is stable for $\beta < \alpha$ and unstable for $\alpha < \beta$. Thus, for $\alpha < \beta$, a stationary bump is always linearly unstable. For $\beta < \alpha$, a stationary bump can be linearly stable for a Mexican hat weight function (if $w(2a) < 0$) but is always unstable for an excitatory weight function ($w(x) > 0$) [22]. For $\beta < \alpha$, it is not possible for a stable stationary bump to undergo a Hopf bifurcation and, as β is increased through α , a stable stationary bump undergoes a drift instability due to eigenvalue λ_-^- increasing through 0 [22]. Interestingly, a multibump solution in (4.1) on a two-dimensional domain is capable of undergoing a bifurcation to a rotating traveling multibump solution [37].

CASE II: *Localized Excitatory Input* ($I(x) > 0$) [7, 22, 23]. The input inhomogeneity ($I(x) \neq 0$) breaks translation symmetry and $\lambda_+^- \neq 0$ generically. A stationary bump is linearly stable when $\operatorname{Re} \lambda_{\pm}^+, \operatorname{Re} \lambda_{\pm}^- < 0$ which reduce to the conditions

$$\frac{\Omega_+(a)}{|U'_o(a)|} < 1 + \beta \quad \text{if} \quad \alpha > \beta, \quad \frac{\Omega_\pm(a)}{|U'_o(a)|} < 1 + \alpha \quad \text{if} \quad \alpha < \beta.$$

For any stationary bump $U_o(x)$, (4.49) implies $(1 + \beta)|U'_o(a)| = w(0) - w(2a) + |I'(a)|$. Consequently, the stability conditions translate, in terms of the gradient $|I'(a)|$, to

$$\begin{aligned} \alpha > \beta : \quad & |I'(a)| > D_{SN}(a) \equiv 2w(2a), \\ \alpha < \beta : \quad & |I'(a)| > D_H(a) \equiv \begin{cases} \left(\frac{\beta-\alpha}{1+\alpha}\right)\Omega_+(a) + 2w(2a), & w(2a) \geq 0, \\ \left(\frac{\beta-\alpha}{1+\alpha}\right)\Omega_-(a), & w(2a) < 0. \end{cases} \end{aligned}$$

$|I'(a)| = D_{SN}(a)$ denotes a saddle-node bifurcation point and $|I'(a)| = D_H(a)$ denotes a Hopf bifurcation where a pair of complex eigenvalues associated with one of the two spatial modes Ω_\pm crosses into the right-half plane. If $w(2a) > 0$ at the Hopf bifurcation point, the sum mode Ω_+ destabilizes and gives rise to a *breather*—a time-periodic, localized bump-like solution that expands and contracts. If $w(2a) < 0$ at the Hopf bifurcation point, the difference mode Ω_- destabilizes and gives rise to a *slosher*—a time-periodic localized solution that instead sloshes side-to-side as shown in Fig. 4.7. Sloshers were also found to occur in [26]. Nonlinear analysis of the Hopf bifurcation reveals that, to first order, the breather and slosher are time-periodic modulations of the stationary bump $U_o(x)$ based upon the even and odd geometry of the sum and difference modes, respectively [22]. The bifurcation can be super/subcritical, which can be determined from the normal form or amplitude equation derived in [22]. Stimulus-induced breathers can undergo further transitions, e.g., period doubling, and can also exhibit mode-locking between breathing and emission of traveling bumps (when supported by the network) [23, 25]. Alternatively, stationary and breathing *fronts* occur in the case of step function input inhomogeneities $I(x)$ [6, 7]. Hopf bifurcation of radially symmetric stationary bumps extends to (4.1) on two-dimensional domains, leading to a variety of localized time-periodic solutions including nonradially symmetric oscillatory structures [23, 24].

4.4.2 Natural and Stimulus-Locked Traveling Activity Bumps

Existence of Traveling Bumps. We simultaneously consider the two cases of *natural* traveling bumps ($I(x, t) = 0$) and *stimulus-locked* traveling bumps which are locked to a stimulus $I(x, t) = I(x - ct)$ traveling with constant speed c . Natural traveling bumps in neural field (4.1) on the infinite line $D = (-\infty, \infty)$ were first considered in [38, 39] and can occur in the absence of an input or in a region of the

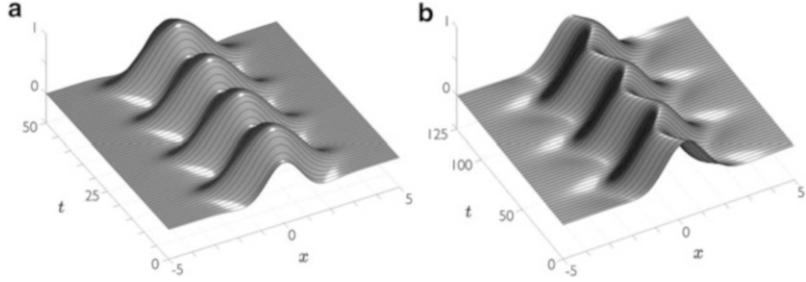


Fig. 4.7 Destabilization of spatial modes $\Omega_+(x)$ and $\Omega_-(x)$, as the bifurcation parameter I_o is varied through a Hopf bifurcation, can give rise to a stable *breather* or *slosher*, respectively, depending on the relative position of the bifurcation points for each spatial mode (e.g., H_\oplus and H_\ominus , Fig. 4.6c). (a) a plot of $u(x, t)$ exhibiting a breather arising from destabilization of the sum mode $\Omega_+(x)$ for $I_o = 1.9, \bar{w}_i = 0, \beta = 2.75, \alpha = 0.1, \theta = 0.375$. (b) a plot of $u(x, t)$ exhibiting a slosher arising from destabilization of the difference mode $\Omega_-(x)$ for $I_o = 1.5, \bar{w}_i = 0.4, \sigma_i = 2, \beta = 2.6, \alpha = 0.01, \theta = 0.35$. Common parameters: $\sigma = 1.2, \bar{w}_e = 1, \sigma_e = 1$

neural medium where an input is effectively zero [23, 25]. An important distinction between the two cases is that the natural traveling bump in the absence of the input is translationally invariant and we have stability with respect to a family of translates, whereas in the stimulus-locked case there is a fixed position of the bump relative to the input.

Assume $u(x, t) = U(x - ct)$ and $v(x, t) = V(x - ct)$ in (4.1) and, in traveling wave coordinates $\xi = x - ct$, make the assumption that the activity $U(\xi)$ is superthreshold $U(\xi) > \theta$ for $\xi \in (\xi_1, \xi_2)$, satisfies $U(\xi_{1,2}) = \theta$, and is subthreshold otherwise with $U(\xi) \rightarrow 0$ as $\xi \rightarrow \pm\infty$. Consequently, the profile of the bump satisfies

$$\begin{aligned} -c U_\xi &= -U - \beta V + \int_{-\infty}^{\infty} w(\xi - \eta) H(U(\eta) - \theta) d\eta + I(\xi), \\ -\frac{c}{\alpha} V_\xi &= +U - V. \end{aligned} \quad (4.55)$$

Variation of parameters [25, 55] can be used to solve (4.55) to construct the profile $(U_c, V_c)^T$ of the traveling bump which can be expressed as [25]

$$\begin{aligned} U_c(\xi) &= (1 - \mu_-) \mathcal{M}_+(\xi) - (1 - \mu_+) \mathcal{M}_-(\xi) \\ V_c(\xi) &= -\alpha [\mathcal{M}_+(\xi) - \mathcal{M}_-(\xi)]. \end{aligned}$$

where $m(\xi) = W(\xi - \xi_1) - W(\xi - \xi_2) + I(\xi)$,

$$\mathcal{M}_\pm(\xi) = \frac{1}{c(\mu_+ - \mu_-)} \int_\xi^\infty e^{\frac{\mu_\pm}{c}(\xi - \eta)} m(\eta) d\eta, \quad \mu_\pm = \frac{1}{2}(1 + \alpha \pm \sqrt{(1 - \alpha)^2 - 4\alpha\beta}),$$

and $0 < \Re \mu_- \leq \Re \mu_+$. Since $m(\xi)$ is dependent upon ξ_1, ξ_2 , the threshold conditions $U_c(\xi_i) = \theta$, where $i = 1, 2$ and $\xi_1 < \xi_2$, determine the position of

the two threshold crossings of the bump relative to the position of the input $I(\xi)$. This results in consistency conditions for the existence of a *stimulus-locked* traveling bump:

$$\begin{aligned}\theta &= (1 - \mu_-)\mathcal{M}_+(\xi_1) - (1 - \mu_+)\mathcal{M}_-(\xi_1), \\ \theta &= (1 - \mu_-)\mathcal{M}_+(\xi_2) - (1 - \mu_+)\mathcal{M}_-(\xi_2).\end{aligned}$$

These determine the existence of a traveling bump (provided the profile satisfies the assumed threshold conditions) and also include the case of *natural waves* ($I_\circ = 0$) with the difference being that the conditions instead determine the width $a = \xi_2 - \xi_1$ of any translate of the bump and its wave speed c which is selected by the network. Note that existence equations for the traveling bump in (4.55) can also be derived using a second order ODE formulation [23, 39] or an integral formulation [9].

Existence Conditions for a Positive, Exponential w and Gaussian I . For explicit calculations in this section, w and I are taken to be

$$w(x) = \frac{\bar{w}_e}{2\sigma_e} e^{-|x|/\sigma_e}, \quad I(x - ct) = I_\circ e^{-(|x-ct|/\sigma)^2}. \quad (4.56)$$

CASE I: *Natural Traveling Bump* ($I(\xi) = 0$) with Speed c [9, 23, 25, 39, 41]. In the absence of an input, translation invariance of the bump allows the simplification $(\xi_1, \xi_2) = (0, a)$ where the wave speed c and bump width a are naturally selected by the network according to the following threshold conditions [25]

$$\theta = J_+(-a), \quad \theta = K(-a), \quad (4.57)$$

where $K(\zeta) = J_-(\zeta) - H_+(\zeta) + H_-(\zeta)$, and, for w given in (4.56),

$$J_\pm(\zeta) = \frac{(\alpha \pm c)(1 - e^\zeta)}{2(c \pm \mu_+)(c \pm \mu_-)}, \quad H_\pm(\zeta) = \frac{c^2(1 - \mu_\mp)(1 - e^{\frac{\mu_\pm}{c}\zeta})}{\mu_\pm(c^2 - \mu_\pm^2)(\mu_+ - \mu_-)}. \quad (4.58)$$

Note that $(c \pm \mu_+)(c \pm \mu_-) = c^2 \pm c(1 + \alpha) + \alpha(1 + \beta)$. Existence equations (4.57) were solved numerically in Fig. 4.8b indicating two branches of traveling bumps for small α . The wide, faster bump is found to be stable and the narrow, slower bump is unstable.

Detailed analyses of the existence of natural traveling bumps can be found in [41, 52], including the case where the homogeneous state has complex eigenvalues [52]. A singular perturbation construction for the pulse was carried out for smooth firing rate functions F in [39]. For moderate values of β traveling fronts occur in (4.1) and were shown to undergo a *front bifurcation* in the form of a cusp bifurcation with respect to the wave speed of the front [6].

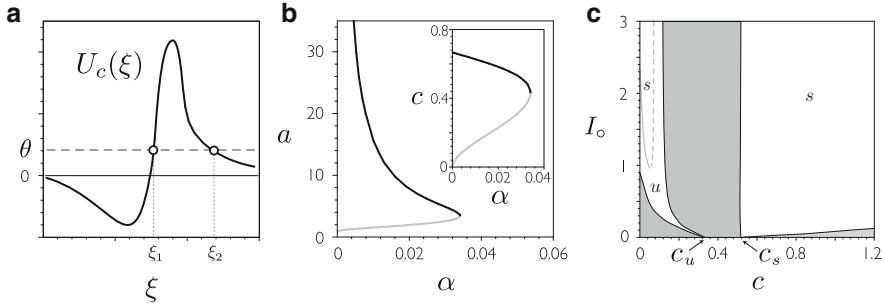


Fig. 4.8 (a) Traveling bump profile. (b) Bifurcation curves for *natural* traveling bumps ($I_o = 0$) in (4.1) in the (α, a) -plane and (α, c) -plane. The stable branch (black) of wide, fast bumps and the unstable branch (gray) of narrow, slow bumps annihilate in a saddle-node bifurcation at a critical value $\alpha_c \approx 0.341$. (c) Regions of existence (white) of the *stimulus-locked* traveling bumps in the (c, I_o) -plane for fixed $\sigma = 1, \alpha = 0.03$. The left and right regions form tongues that issue from the unstable c_u and stable c_s natural traveling bumps, respectively. The curve of Hopf bifurcations within the left tongue is shown in gray, above which the bump is stable (s) and below which it is unstable (u). Stable traveling breathers bifurcate from the left branch (solid gray) Hopf curve, and stationary bumps correspond to the intersection of the tongue and the line $c = 0$. When bumps and breathers are unstable or do not exist, there is mode-locking between breathing and the emission of natural traveling bumps. Parameters in (b) and (c): $\theta = 0.3, \beta = 2.5, \bar{w}_e = \sigma_e = 1, \bar{w}_i = 0$ (Figure adapted from Folias and Bressloff [25])

CASE II: *Stimulus-Locked Traveling Bump ($I(\xi) \neq 0$) with Speed c* [25]. The wave and stimulus speeds c are identical, and the threshold conditions for (ξ_1, ξ_2) are [25]

$$\begin{aligned} \theta &= K(\xi_1 - \xi_2) + T_+(\xi_1) - T_-(\xi_1), \\ \theta &= J_+(\xi_1 - \xi_2) + T_+(\xi_2) - T_-(\xi_2), \end{aligned} \quad (4.59)$$

where K, J_+ are given in (4.58) and T_{\pm} arises from the input and is given by

$$T_{\pm}(\xi) = \frac{\sqrt{\pi} \sigma I_o}{2c} \left(\frac{1 - \mu_{\mp}}{\mu_+ - \mu_-} \right) \exp \left(\frac{\mu_{\pm}\xi}{c} + \left[\frac{\mu_{\pm}\sigma}{2c} \right]^2 \right) \operatorname{erfc} \left(\frac{\xi}{\sigma} + \frac{\mu_{\pm}\sigma}{2c} \right),$$

with $\operatorname{erfc}(z)$ denoting the complementary error function. Equation (4.59) can be solved numerically to determine the regions of existence of stimulus-locked traveling bumps as both the speed c and amplitude I_o are varied (assuming $U_c(\xi)$ satisfies the threshold assumptions). This allows us to connect the stationary bumps to natural traveling bumps via stimulus-locked traveling bumps as shown in Fig. 4.8. This analysis for stimulus-locked fronts was carried out in [6] and an extension of stimulus-locked bumps for a general smooth firing rate function F was studied in [20].

Stability of Traveling Bumps. By setting $u = U_c + \tilde{\varphi}$ and $v = V_c + \tilde{\psi}$, we study the evolution of small perturbations $(\tilde{\varphi}, \tilde{\psi})^T$ in the linearization of (4.1) about the traveling bump $(U_c, V_c)^T$ which, in traveling wave coordinates, are governed by

$$\begin{aligned}\partial_t \tilde{\varphi} &= c \partial_\xi \tilde{\varphi} - \tilde{\varphi} - \beta \tilde{\psi} + \int_{-\infty}^{\infty} w(\xi - \eta) H'(U_c(\eta) - \theta) \tilde{\varphi}(\eta, t) d\eta, \\ \partial_t \tilde{\psi} &= c \partial_\xi \tilde{\psi} + \alpha \tilde{\varphi} - \alpha \tilde{\psi}.\end{aligned}\quad (4.60)$$

Separating variables by setting $\tilde{\varphi}(\xi, t) = e^{\lambda t} \varphi(\xi)$ and $\tilde{\psi}(\xi, t) = e^{\lambda t} \psi(\xi)$ in (4.60), where $(\varphi, \psi)^T \in C_u^1(\mathbb{R}, \mathbb{C}^2)$, leads to the spectral problem for λ and $(\varphi, \psi)^T$

$$(L + \mathcal{N}) \begin{pmatrix} \varphi \\ \psi \end{pmatrix} = \lambda \begin{pmatrix} \varphi \\ \psi \end{pmatrix} \quad (4.61)$$

where

$$L = c \frac{\partial}{\partial \xi} - \begin{bmatrix} 1 & \beta \\ -\alpha & \alpha \end{bmatrix}, \quad \mathcal{N} \begin{pmatrix} \varphi \\ \psi \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \left[\frac{w(\xi - \xi_1)}{|U'_c(\xi_1)|} \varphi(\xi_1) + \frac{w(\xi - \xi_2)}{|U'_c(\xi_2)|} \varphi(\xi_2) \right].$$

The essential spectrum lies within the set $D = \{z : \Re z \in [-\Re \mu_+, -\Re \mu_-]\}$, where $\Re \mu_\pm > 0$, inducing no instability [25, 41, 55]. Stability is then determined by elements of the point spectrum that lie in the region $R = \{z : \Re z > -\Re \mu_-\}$ which can be calculated using an Evans function. In particular, we determine a condition for $(L + \mathcal{N} - \lambda I)$ to have a bounded inverse. The Evans function $\mathcal{E}(\lambda)$ subsequently arises from the condition that $(L + \mathcal{N} - \lambda I)$ is not invertible and $(L + \mathcal{N} - \lambda I) = 0$ has nontrivial solutions. We set $\mathbf{u} = (\varphi, \psi)^T$ and use variation of parameters [25, 55] to construct a bounded inverse for $(L + \mathcal{N} - \lambda I)$ based on the integral kernel

$$\mathbf{M}(\xi, \eta, \lambda) = \frac{1}{c\beta(\mu_+ - \mu_-)} \left[\Phi_+(\xi) | \Phi_-(\xi) \right] \left[\Psi_+(\eta) | \Psi_-(\eta) \right]^T \quad (4.62)$$

where $[\mathbf{A}! \mathbf{B}]$ denotes the matrix with column vectors \mathbf{A} and \mathbf{B} , respectively, and

$$\Phi_\pm(\xi) = \begin{pmatrix} \beta \\ \mu_\pm - 1 \end{pmatrix} e^{\left(\frac{\lambda+\mu_\pm}{c}\right)\xi}, \quad \Psi_\pm(\xi) = \pm \begin{pmatrix} 1 - \mu_\mp \\ \beta \end{pmatrix} e^{-\left(\frac{\lambda+\mu_\pm}{c}\right)\xi}.$$

For $\Re(\lambda) > -\mu_-$, we can express $(L + \mathcal{N} - \lambda I)\mathbf{u} = -\mathbf{f}$, where $\mathbf{f} = (f_1, f_2)^T$, as

$$\mathbf{u}(\xi) - \int_\xi^\infty \mathbf{M}(\xi, \eta, \lambda) \mathcal{N} \mathbf{u}(\eta) d\eta = \int_\xi^\infty \mathbf{M}(\xi, \eta, \lambda) \mathbf{f}(\eta) d\eta. \quad (4.63)$$

From (4.63), ψ is calculated in terms of $\varphi(\xi_1)$, $\varphi(\xi_2)$, \mathcal{F}_2 , and λ, φ are determined by

$$\varphi(\xi) - \Lambda_1(\lambda, \xi)\varphi(\xi_1) - \Lambda_2(\lambda, \xi)\varphi(\xi_2) = \mathcal{F}_1(\xi) \quad (4.64)$$

where M_{11} denotes the $(1,1)$ entry of M in (4.62) and $i = 1, 2$ in the expression below

$$\Lambda_i(\lambda, \xi) = \int_{\xi}^{\infty} M_{11}(\xi, \eta, \lambda) \frac{w(\eta - \xi_i)}{|U'_c(\xi_i)|} d\eta, \quad \begin{pmatrix} \mathcal{F}_1(\xi) \\ \mathcal{F}_2(\xi) \end{pmatrix} = \int_{\xi}^{\infty} M(\xi, \eta, \lambda) \mathbf{f}(\eta) d\eta.$$

By the Hölder inequality, Λ_i and $\mathcal{F}_{1,2}$ are bounded for all $\xi \in \mathbb{R}$ and $\mathbf{f} \in C_u^0(\mathbb{R}, \mathbb{C}^2)$. A compatibility condition that determines the values of $\varphi(\xi_1)$ and $\varphi(\xi_2)$ is produced by substituting $\xi = \xi_1$ and $\xi = \xi_2$ into (4.64) to obtain the matrix equation

$$(I - \Lambda(\lambda)) \begin{pmatrix} \varphi(\xi_1) \\ \varphi(\xi_2) \end{pmatrix} = \begin{pmatrix} \mathcal{F}_1(\xi_1) \\ \mathcal{F}_1(\xi_2) \end{pmatrix}, \quad \Lambda(\lambda) = \begin{bmatrix} \Lambda_1(\lambda, \xi_1) & \Lambda_2(\lambda, \xi_1) \\ \Lambda_1(\lambda, \xi_2) & \Lambda_2(\lambda, \xi_2) \end{bmatrix}$$

which has a unique solution if and only if $\det(I - \Lambda(\lambda)) \neq 0$, resulting in a bounded inverse $(L + \mathcal{N} - \lambda I)^{-1}$ defined on all of $C_u^0(\mathbb{R}, \mathbb{C}^2)$. Conversely, we cannot invert the operator for λ such that $\det(I - \Lambda(\lambda)) = 0$, in which case $(L + \mathcal{N} - \lambda)\mathbf{u} = \mathbf{0}$ has nontrivial solutions corresponding to eigenvalues λ and eigenfunctions $(\varphi, \psi)^\top$ in the point spectrum. Thus, for $\Re(\lambda) > -\mu_-$, we can express the *Evans function* as

$$\mathcal{E}(\lambda) = \det(I - \Lambda(\lambda)), \quad \Re(\lambda) > -\mu_-, \quad (4.65)$$

which has eigenvalues λ given by its zero set.

Evans Function for an Exponential Weight w and Gaussian-like Input I . The following gives an explicit construction of the Evans function for natural ($I_o = 0$) and stimulus-locked ($I_o > 0$) traveling bumps in (4.1) with a Heaviside firing rate function, exponential weight distribution and Gaussian input given in (4.56). For natural traveling bumps ($I_o = 0$), by translation invariance we set $(\xi_1, \xi_2) = (0, a)$.

For $\Re(\lambda) > -\mu_-$, the Evans function $\mathcal{E}(\lambda)$ is given by [25]

$$\mathcal{E}(\lambda) = \left[1 - \frac{\Theta_+(\lambda)}{|U'_c(\xi_1)|} \right] \left[1 - \frac{\Theta_+(\lambda)}{|U'_c(\xi_2)|} \right] - \frac{\Theta_+(\lambda) \mathcal{E}(\lambda, \xi_1 - \xi_2)}{|U'_c(\xi_1) U'_c(\xi_2)|},$$

where

$$\Gamma_{\pm}(\lambda) = \frac{(1 - \mu_{\mp})c}{(\mu_+ - \mu_-)(c^2 - (\lambda + \mu_{\pm})^2)}, \quad \Theta_{\pm}(\lambda) = \frac{\lambda + \alpha \pm c}{2(\lambda + \mu_+ \pm c)(\lambda + \mu_- \pm c)},$$

$$\mathcal{E}(\lambda, \xi) = \Theta_-(\lambda) e^{2\xi} + \Gamma_+(\lambda) e^{\left[\frac{\lambda + \mu_+ + c}{c}\right]\xi} - \Gamma_-(\lambda) e^{\left[\frac{\lambda + \mu_- + c}{c}\right]\xi}.$$

Since the zero set of the Evans function $\mathcal{E}(\lambda)$ comprises solutions of a transcendental equation, the eigenvalues λ can be determined numerically by finding the intersection points of the zero sets of the real and imaginary parts of the Evans function which was used to determine the stability results in Fig. 4.8. Hopf bifurcations, identified by complex conjugate eigenvalues crossing the imaginary axis, can give rise to traveling breathers or mode-locking between breathing and the emission of natural traveling bumps [25].

For various treatments of the stability of natural traveling bumps and Evans functions in (4.1) see [4, 10, 25, 41, 44, 55], and a comparison between different approaches is found in [44]. Zhang developed the Evans function and analyzed the stability of traveling bumps in the singularly perturbed case $0 < \alpha \ll 1$ [55]. Finally, on one-dimensional domains, traveling multibump waves were studied in [52], and traveling waves have been extended to the case of inhomogeneous synaptic coupling in [32] and asymmetric coupling [51]. On two-dimensional domains, circular waves/target patterns [23], spiral waves [34, 52], traveling and rotating multibumps [37], and the collision of traveling bumps [36] have also been examined.

References

1. Amari, S.: Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.* **27**(2), 77–87 (1977)
2. Benda, J., Herz, A.V.M.: A universal model for spike-frequency adaptation. *Neural Comput.* **15**(11), 2523–2564 (2003)
3. Billcock, V.A., Tsou, B.H.: Elementary visual hallucinations and their relationships to neural pattern-forming mechanisms. *Psychol. Bull.* **138**(4), 744–774 (2012)
4. Bressloff, P.C.: Spatiotemporal dynamics of continuum neural fields. *J. Phys. A: Math. Theor.* **45**(3), 033,001 (2012)
5. Bressloff, P.C., Cowan, J.D., Golubitsky, M., Thomas, P.J., Wiener, M.C.: Geometric visual hallucinations, Euclidean symmetry and the functional architecture of striate cortex. *Philos. Trans. R. Soc. B.* **356**(1407), 299–330 (2001)
6. Bressloff, P.C., Folias, S.E.: Front bifurcations in an excitatory neural network. *SIAM J. Appl. Math.* **65**, 131–151 (2004)
7. Bressloff, P.C., Folias, S.E., Prat, A., Li, Y.X.: Oscillatory waves in inhomogeneous neural media. *Phys. Rev. Lett.* **91**(17), 178,101 (2003)
8. Coombes, S.: Waves, bumps, and patterns in neural field theories. *Biol. Cybern.* **93**(2), 91–108 (2005)
9. Coombes, S., Lord, G.J., Owen, M.R.: Waves and bumps in neuronal networks with axo-dendritic synaptic interactions. *Phys. D* **178**, 219–241 (2003)
10. Coombes, S., Owen, M.R.: Evans functions for integral neural field equations with Heaviside firing rate function. *SIAM J. Appl. Dyn. Syst.* **3**(4), 574–600 (2004)
11. Coombes, S., Owen, M.R.: Bumps, breathers, and waves in a neural network with spike frequency adaptation. *Phys. Rev. Lett.* **94**(14), 148,102 (2005)
12. Coombes, S., Schmidt, H., Bojak, I.: Interface dynamics in planar neural field models. *J. Math. Neurosci.* **2**, 9 (2012)
13. Curtu, R., Ermentrout, B.: Pattern formation in a network of excitatory and inhibitory cells with adaptation. *SIAM J. Appl. Dyn. Syst.* **3**, 191–231 (2004)
14. Dawes, J.H.P.: Hopf bifurcation on a square superlattice. *Nonlinearity* **14**, 491 (2001)

15. Dionne, B., Silber, M., Skeldon, A.C.: Stability results for steady, spatially periodic planforms. *Nonlinearity* **10**, 321 (1997)
16. Ermentrout, B.: Stripes or spots? Nonlinear effects in bifurcation of reaction-diffusion equations on the square. *Proc. R. Soc. Lond. Ser. A: Math. Phys. Sci.* **434**(1891), 413–417 (1991)
17. Ermentrout, B.: Neural networks as spatio-temporal pattern-forming systems. *Rep. Prog. Phys.* **61**, 353–430 (1998)
18. Ermentrout, G.B., Cowan, J.D.: A mathematical theory of visual hallucination patterns. *Biol. Cybern.* **34**(3), 137–150 (1979)
19. Ermentrout, G.B., Cowan, J.D.: Secondary bifurcation in neuronal nets. *SIAM J. Appl. Math.* **39**(2), 323–340 (1980)
20. Ermentrout, G.B., Jalics, J.Z., Rubin, J.E.: Stimulus-driven traveling solutions in continuum neuronal models with a general smooth firing rate function. *SIAM J. Appl. Math.* **70**(8), 3039–3064 (2010)
21. Ermentrout, G.B., McLeod, J.B.: Existence and uniqueness of travelling waves for a neural network. *Proc. R. Soc. Edinb.* **123A**, 461–478 (1993)
22. Folias, S.E.: Nonlinear analysis of breathing pulses in a synaptically coupled neural network. *SIAM J. Appl. Dyn. Syst.* **10**, 744–787 (2011)
23. Folias, S.E., Bressloff, P.C.: Breathing pulses in an excitatory neural network. *SIAM J. Appl. Dyn. Syst.* **3**(3), 378–407 (2004)
24. Folias, S.E., Bressloff, P.C.: Breathers in two-dimensional neural media. *Phys. Rev. Lett.* **95**(20), 208,107 (2005)
25. Folias, S.E., Bressloff, P.C.: Stimulus-locked traveling waves and breathers in an excitatory neural network. *SIAM J. Appl. Math.* **65**(6), 2067–2092 (2005)
26. Hansel, D., Sompolinsky, H.: Modeling feature selectivity in local cortical circuits. In: Koch, C., Segev, I. (eds.) *Methods in Neuronal Modeling: From Ions to Networks*, Chap. 13, pp. 499–567. MIT, Cambridge (1998)
27. Huang, X., Troy, W.C., Yang, Q., Ma, H., Laing, C.R., Schiff, S.J., Wu, J.Y.: Spiral waves in disinhibited mammalian neocortex. *J. Neurosci.* **24**(44), 9897–9902 (2004)
28. Itskov, V., Curto, C., Pastalkova, E., Buzsáki, G.: Cell assembly sequences arising from spike threshold adaptation keep track of time in the hippocampus. *J. Neurosci.* **31**(8), 2828–2834 (2011)
29. Kilpatrick, Z.P., Bressloff, P.C.: Stability of bumps in piecewise smooth neural fields with nonlinear adaptation. *Phys. D* **239**(12), 1048–1060 (2010)
30. Kilpatrick, Z.P., Ermentrout, B.: Wandering bumps in stochastic neural fields. arXiv (nlin.PS/1205.3072v1) (2012)
31. Kilpatrick, Z.P., Ermentrout, G.B.: Hallucinogen persisting perception disorder in neuronal networks with adaptation. *J. Comput. Neurosci.* **32**(1), 25–53 (2012)
32. Kilpatrick, Z.P., Folias, S.E., Bressloff, P.C.: Traveling pulses and wave propagation failure in inhomogeneous neural media. *SIAM J. Appl. Dyn. Syst.* **7**(1), 161–185 (2008)
33. Klüver, H.: *Mescal and the Mechanisms of Hallucinations*. University of Chicago Press, Chicago (1966)
34. Laing, C.R.: Spiral waves in nonlocal equations. *SIAM J. Appl. Dyn. Syst.* **4**(3), 588–606 (2005)
35. Laing, C.R., Longtin, A.: Noise-induced stabilization of bumps in systems with long-range spatial coupling. *Phys. D* **160**(3–4), 149–172 (2001)
36. Lu, Y., Sato, Y., Amari, S.I.: Traveling bumps and their collisions in a two-dimensional neural field. *Neural Comput.* **23**(5), 1248–60 (2011)
37. Owen, M.R., Laing, C.R., Coombes, S.: Bumps and rings in a two-dimensional neural field: splitting and rotational instabilities. *New J. Phys.* **9**, 378 (2007)
38. Pinto, D.J.: Computational, experimental, and analytical explorations of neuronal circuits in the cerebral cortex. Ph.D. thesis, Department of Mathematics, University of Pittsburgh, Pittsburgh (1997)
39. Pinto, D.J., Ermentrout, G.B.: Spatially structured activity in synaptically coupled neuronal networks: I. Traveling fronts and pulses. *SIAM J. Appl. Math.* **62**(1), 206–225 (2001)

40. Pinto, D.J., Ermentrout, G.B.: Spatially structured activity in synaptically coupled neuronal networks: II. Lateral inhibition and standing pulses. *SIAM J. Appl. Math.* **62**(1), 226–243 (2001)
41. Pinto, D.J., Jackson, R.K., Wayne, C.E.: Existence and stability of traveling pulses in a continuous neuronal network. *SIAM J. Appl. Dyn. Syst.* **4**, 954–984 (2005)
42. Rankin, J., Tlapale, E., Veltz, R., Faugeras, O., Kornprobst, P.: Bifurcation analysis applied to a model of motion integration with a multistable stimulus. Inria: Research report (2011)
43. Rule, M., Stoffregen, M., Ermentrout, B.: A model for the origin and properties of flicker-induced geometric phosphenes. *PLoS Comput. Biol.* **7**(9), e1002158 (2011)
44. Sandstede, B.: Evans functions and nonlinear stability of traveling waves in neuronal network models. *Int. J. Bifur. Chaos Appl. Sci. Eng.* **17**, 2693–2704 (2007)
45. Schwartz, E.: Spatial mapping in the primate sensory projection: analytic structure and relevance to projection. *Biol. Cybern.* **25**, 181–194 (1977)
46. Shusterman, V., Troy, W.C.: From baseline to epileptiform activity: a path to synchronized rhythmicity in large-scale neural networks. *Phys. Rev. E* **77**(6), 061,911 (2008)
47. Silber, M., Knobloch, E.: Hopf bifurcation on a square lattice. *Nonlinearity* **4**, 1063 (1991)
48. Stocker, M., Krause, M., Pedarzani, P.: An apamin-sensitive Ca^{2+} -activated K^+ current in hippocampal pyramidal neurons. *Proc. Natl. Acad. Sci. USA* **96**(8), 4662–4667 (1999)
49. Tass, P.: Cortical pattern formation during visual hallucinations. *J. Biol. Phys.* **21**(3), 177–210 (1995)
50. Tass, P.: Oscillatory cortical activity during visual hallucinations. *J. Biol. Phys.* **23**(1), 21–66 (1997)
51. Troy, W.C.: Traveling waves and synchrony in an excitable large-scale neuronal network with asymmetric connections. *SIAM J. Appl. Dyn. Syst.* **7**(4), 1247–1282 (2008)
52. Troy, W.C., Shusterman, V.: Patterns and features of families of traveling waves in large-scale neuronal networks. *SIAM J. Appl. Dyn. Syst.* **6**(1), 263–292 (2007)
53. Wu, J.Y.: Propagating waves of activity in the neocortex: what they are, what they do. *Neuroscientist* **14**(5), 487–502 (2008)
54. Zhang, K.: Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. *J. Neurosci.* **16**(6), 2112–2126 (1996)
55. Zhang, L.: On stability of traveling wave solutions in synaptically coupled neuronal networks. *Differ. Integral Equ.* **16**, 513–536 (2003)
56. Zhang, L.: Existence, uniqueness and exponential stability of traveling wave solutions of some integral differential equations arising from neuronal networks. *J. Differ. Equ.* **197**(1), 162–196 (2004)

Chapter 5

PDE Methods for Two-Dimensional Neural Fields

Carlo R. Laing

Abstract We consider neural field models in both one and two spatial dimensions and show how for some coupling functions they can be transformed into equivalent partial differential equations (PDEs). In one dimension we find snaking families of spatially-localised solutions, very similar to those found in reversible fourth-order ordinary differential equations. In two dimensions we analyse spatially-localised bump and ring solutions and show how they can be unstable with respect to perturbations which break rotational symmetry, thus leading to the formation of complex patterns. Finally, we consider spiral waves in a system with purely positive coupling and a second slow variable. These waves are solutions of a PDE in two spatial dimensions, and by numerically following these solutions as parameters are varied, we can determine regions of parameter space in which stable spiral waves exist.

5.1 Introduction

Neural field models are generally considered to date back to the 1970s [1, 41], although several earlier papers consider similar equations [4, 25]. These types of equations were originally formulated as models for the dynamics of macroscopic activity patterns in the cortex, on a much larger spatial scale than that of a single neuron. They have been used to model phenomena such as short-term memory [36], the head direction system [43], visual hallucinations [19, 20], and EEG rhythms [7].

C.R. Laing (✉)

Institute of Natural and Mathematical Sciences, Massey University, Private Bag 102-904 NSMC, Auckland, New Zealand

e-mail: c.r.laing@massey.ac.nz

Perhaps the simplest formulation of such a model in one spatial dimension is

$$\frac{\partial u(x, t)}{\partial t} = -u(x, t) + \int_{-\infty}^{\infty} w(x - y) f(u(y, t)) dy \quad (5.1)$$

where

- w is symmetric, i.e. $w(-x) = w(x)$,
- $\lim_{x \rightarrow \infty} w(x) = 0$,
- $\int_{-\infty}^{\infty} w(x) dx < \infty$,
- $w(x)$ is continuous,

and f is a non-decreasing function with $\lim_{u \rightarrow -\infty} f(u) = 0$ and $\lim_{u \rightarrow \infty} f(u) = 1$ [12, 36]. The physical interpretation of this type of model is that $u(x, t)$ is the average voltage of a large group of neurons at position $x \in \mathbb{R}$ and time t , and $f(u(x, t))$ is their firing rate, normalised to have a maximum of 1. The function $w(x)$ describes how neurons a distance x apart affect one another. Typical forms of this function are purely positive [6], ‘‘Mexican hat’’ [19, 26] (positive for small x and negative for large x) and decaying oscillatory [18, 36]. To find the influence of neurons at position y on those at position x we evaluate $f(u(y, t))$ and weight it by $w(x - y)$. The influence of all neurons is thus the integral over y of $w(x - y) f(u(y, t))$. In the absence of inputs from other parts of the network, u decays exponentially to a steady state, which we define to be zero. Equation (5.1) is a nonlocal differential equation, with the nonlocal term arising from the biological reality that we are modelling. Typically, researchers are interested in either ‘‘bump’’ solutions of (5.1), for which $f(u(x)) > 0$ only on a finite number of finite, disjoint intervals, or front solutions which connect a region of high activity to one of zero activity [12] (see Chap. 7). Note that this type of model is invariant with respect to spatial translations, which is reflected in the fact that w appears as a function of relative position only (i.e. $x - y$), not the actual values of x and y .

The function f is normally thought of as a sigmoid (although other functions are sometimes considered [26]), and in the limit of infinite steepness it becomes the Heaviside step function [12, 36]. In this case stationary solutions are easily constructed since to evaluate the integral in (5.1) we just integrate $w(x - y)$ over the interval(s) of y where $f(u(y, t)) = 1$. The stability of these solutions can be determined by linearising (5.1) about them and using the fact that the derivative of the Heaviside function is the Dirac delta function [6, 40].

When f is not a Heaviside, constructing stationary solutions becomes more difficult and we generally have to do so numerically. A stationary solution of (5.1) satisfies

$$u(x) = \int_{-\infty}^{\infty} w(x - y) f(u(y)) dy. \quad (5.2)$$

In all but Sect. 5.4 of this chapter we consider only spatially-localised solutions, i.e. ones for which u and all of its relevant spatial derivatives decay to zero as

$|x| \rightarrow \infty$. Generally speaking, integral equations such as (5.2) are not studied in as much detail as differential equations. As a result more methods for analysis—and software packages—exist for the numerical solution of differential equations, as opposed to integral equations. For these reasons we consider rewriting (5.2) as a differential equation for the function $u(x)$. The key to doing so is to recognise that the integral in (5.2) is a convolution. This observation provides several equivalent ways of converting (5.2) into a differential equation.

The first method involves recalling that the Fourier transform of the convolution of two functions is the product of their Fourier transforms. Thus, denoting by $F[u](k)$ the Fourier transform of $u(x)$, where k is the transform variable, Fourier transforming (5.2) gives

$$F[u](k) = F[w](k) \times F[f(u)](k) \quad (5.3)$$

where “ \times ” indicates normal multiplication. Suppose that the Fourier transform of w was a rational function of k^2 , i.e.

$$F[w](k) = \frac{P(k^2)}{Q(k^2)} \quad (5.4)$$

where P and Q are polynomials. Multiplying both sides of (5.3) by $Q(k^2)$ we obtain

$$Q(k^2) \times F[u](k) = P(k^2) \times F[f(u)](k) \quad (5.5)$$

Recalling that if the Fourier transform of $u(x)$ is $F[u](k)$, then the Fourier transform of $u''(x)$ is $-k^2 F[u](k)$, the Fourier transform of $u'''(x)$ is $k^4 F[u](k)$ and so on, where the primes indicate spatial derivatives, taking the inverse Fourier transform of (5.5) gives

$$D_1 u(x) = D_2 f(u(x)) \quad (5.6)$$

where D_1 and D_2 are linear differential operators involving only even derivatives associated with Q and P respectively [32]. As an example, consider the decaying oscillatory coupling function

$$w(x) = e^{-b|x|} (b \sin |x| + \cos x) \quad (5.7)$$

where b is a parameter (plotted in Fig. 5.1 (left) for $b = 0.5$), which has the Fourier transform

$$\frac{4b(b^2 + 1)}{k^4 + 2(b^2 - 1)k^2 + (b^2 + 1)^2}. \quad (5.8)$$

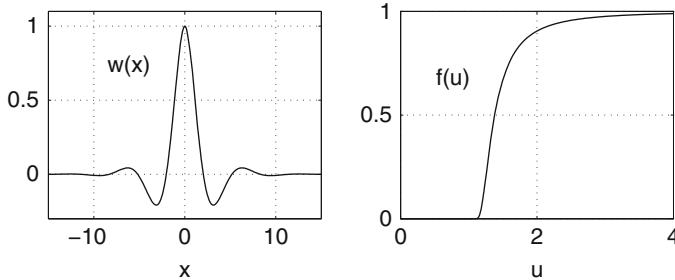


Fig. 5.1 Left: $w(x)$ given by (5.7) when $b = 0.5$. Right: $f(u)$ given by (5.21) for $\kappa = 0.1, h = 1$

For this example D_2 is just the constant $4b(b^2 + 1)$ and

$$D_1 = \frac{d^4}{dx^4} - 2(b^2 - 1)\frac{d^2}{dx^2} + (b^2 + 1)^2 \quad (5.9)$$

and thus (for this choice of w) Eq. (5.2) can be written

$$\frac{d^4u}{dx^4} - 2(b^2 - 1)\frac{d^2u}{dx^2} + (b^2 + 1)^2u = 4b(b^2 + 1)f(u(x)). \quad (5.10)$$

Our decision to consider only spatially-localised solutions validates the use of Fourier transforms and gives the boundary conditions for (5.10), namely

$$\lim_{x \rightarrow \pm\infty} (u, u', u'', u''') = (0, 0, 0, 0). \quad (5.11)$$

The other method for converting (5.2) into a differential equation is to recall that the solution of an inhomogeneous linear differential equation can be formally written as the convolution of the Green's function of the linear differential operator together with the appropriate boundary conditions, and the function on the right hand side (RHS) of the differential equation. Thus if w was such a Green's function, we could recognise (5.2) as being the solution of a linear differential equation with $f(u)$ as its RHS.

Using the example above one can show that the Green's function of the operator (5.9) with boundary conditions (5.11), i.e. the solution of

$$\frac{d^4w}{dx^4} - 2(b^2 - 1)\frac{d^2w}{dx^2} + (b^2 + 1)^2w = \delta(x) \quad (5.12)$$

satisfying (5.11), where δ is the Dirac delta function, is

$$w(x) = \frac{e^{-b|x|}(b \sin |x| + \cos x)}{4b(b^2 + 1)} \quad (5.13)$$

and thus the solution of (5.10)–(5.11) is (5.2). This second method, of recognising that the coupling function w is the Green's function of a linear differential operator, is perhaps less easy to generalise, so we concentrate mostly on the first method in this chapter. An important point to note is that the Fourier transform method applies equally well to (5.1), i.e. the full time-dependent problem. Using the function (5.7) and keeping the time derivative one can convert (5.1) to

$$\left[\frac{\partial^4}{\partial x^4} - 2(b^2 - 1) \frac{\partial^2}{\partial x^2} + (b^2 + 1)^2 \right] \left(u(x, t) + \frac{\partial u(x, t)}{\partial t} \right) = 4b(b^2 + 1)f(u(x, t)) \quad (5.14)$$

Clearly stationary solutions of (5.14) satisfy (5.10), but keeping the time dependence in (5.14) enables us to determine the stability of these stationary solutions via linearisation about them.

Note that the Fourier transform of $(1/2)e^{-|x|}$ is $1/(1 + k^2)$, and thus for this coupling function (5.2) is equivalent to

$$\left(1 - \frac{\partial^2}{\partial x^2} \right) u = f(u) \quad (5.15)$$

Also, the Fourier transform of the “wizard hat” $w(x) = (1/4)(1 - |x|)e^{-|x|}$ is $k^2/(1 + k^2)^2$, giving the differential equation [12]

$$\left(1 - \frac{\partial^2}{\partial x^2} \right)^2 u = -\frac{\partial^2}{\partial x^2} f(u) \quad (5.16)$$

and thus a variety of commonly used connectivity functions are amenable to this type of transformation. (See also [26] for another example.)

The model (5.1) assumes that information about activity at position y propagates instantaneously to position x , but a more realistic model could include a distance-dependent delay:

$$\frac{\partial u(x, t)}{\partial t} = -u(x, t) + \int_{-\infty}^{\infty} w(x - y)f\left(u\left(y, t - \frac{|x - y|}{v}\right)\right) dy \quad (5.17)$$

where $v > 0$ is the velocity of propagation of information. Equation (5.17) can be written

$$\frac{\partial u(x, t)}{\partial t} = -u(x, t) + \psi(x, t) \quad (5.18)$$

where

$$\psi(x, t) \equiv \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(x - y, t - s)f(u(y, s))dy ds \quad (5.19)$$

and $K(x, t) = w(x)\delta(t - |x|/v)$ [12, 33]. Recognising that both integrals in (5.19) are convolutions, and making the choice $w(x) = (1/2)e^{-|x|}$, one can take Fourier transforms in both space and time and convert (5.19) to

$$\left(\frac{\partial^2}{\partial t^2} + 2v \frac{\partial}{\partial t} + v^2 - v^2 \frac{\partial^2}{\partial x^2} \right) \psi(x, t) = \left(v^2 + v \frac{\partial}{\partial t} \right) f(u(x, t)) \quad (5.20)$$

This equation was first derived by [30], and these authors may well have been the first to use Fourier transforms to convert neural field models to PDEs. We will not consider delays here, but see [15] for a recent approach in two spatial dimensions.

5.2 Results in One Spatial Dimension

We now present some results of the analysis of (5.14), similar to those in [36]. From now on we make the specific choice of the firing rate function

$$f(u) = e^{-\kappa/(u-h)^2} H(u-h) \quad (5.21)$$

where $\kappa > 0$ and $h \in \mathbb{R}$ are parameters, and H is the Heaviside step function. The function (5.21) for typical parameter values is shown in Fig. 5.1 (right). Note that if $h > 0$ then $f(0) = 0$.

We start with a few comments regarding Eqs. (5.10) and (5.11). Firstly, Eq. (5.10) is reversible under the involution $(u, u', u'', u''') \mapsto (u, -u', u'', -u''')$ [18]. Secondly, spatially-localised solutions of (5.10) can be regarded as homoclinic orbits to the origin, i.e. orbits for which u and all of its derivatives tend to zero as $x \rightarrow \pm\infty$. Linearising (5.10) about the origin one finds that it has eigenvalues $b \pm i$ and $-b \pm i$, i.e. the fixed point at the origin is a bifocus [34], and thus the homoclinic orbits spiral into and out of the origin. Thirdly, Eq. (5.10) is Hamiltonian, and homoclinic orbits to the origin satisfy the first integral

$$u' u''' - \frac{(u'')^2}{2} - (b^2 - 1)(u')^2 + (b^2 + 1)^2 Q(u) = 0 \quad (5.22)$$

where

$$Q(u) \equiv \int_0^u \left(s - \frac{4bf(s)}{b^2 + 1} \right) ds \quad (5.23)$$

This Hamiltonian nature can be exploited to understand the solutions of (5.10)–(5.11) and the bifurcations they undergo as parameters are varied [18]. See for example [11] for more details on homoclinic orbits in reversible systems.

We are interested in stationary spatially-localised solutions of (5.14), and how they vary as parameters are varied. Figure 5.2 shows the result of following such

Fig. 5.2 Spatially-localised steady states of (5.14) as a function of h . The vertical axis is the maximum over the domain of $u(x)$. Solid curves indicate stable while dashed indicate unstable. The solutions at points A, B and C are shown in Fig. 5.3. Other parameters are $b = 0.25, \kappa = 0.1$

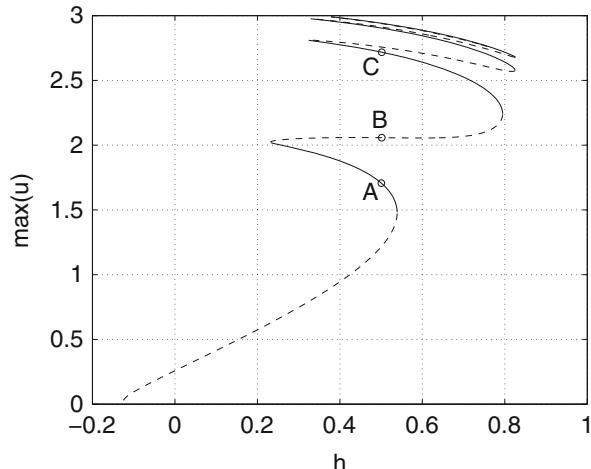
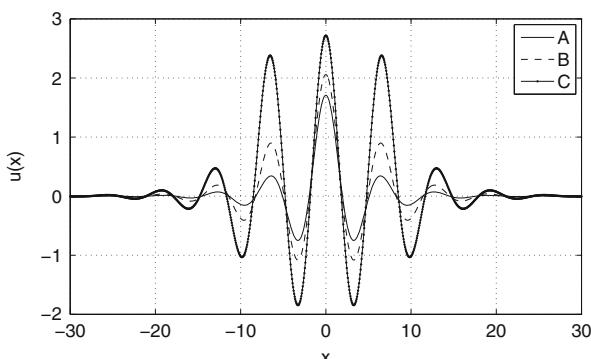


Fig. 5.3 Spatially-localised steady states of (5.14) at the three points marked A, B and C in Fig. 5.2. Other parameters are $b = 0.25, \kappa = 0.1$



solutions as the parameter h (firing threshold) is varied. As was found in [18, 35] the family of solutions forms a “snake” with successively more large amplitude oscillations added to the solution as one moves from one branch of the snake to the next in the direction of increasing $\max(u)$. (Note that b , not h , was varied in [18, 35].) Similar snakes of homoclinic orbits have been found in other reversible systems of fourth-order differential equations [10, 42], and Faye et al. [21] very recently analysed snaking behaviour in a model of the form (5.1).

Figure 5.3 shows three solutions from the family shown in Fig. 5.2, all at $h = 0.5$. Solutions at A and C are stable, and are referred to as “1-bump” and “3-bump” solutions, respectively, since they have 1 and 3, respectively, regions for which $u > h$. The solution at B is an unstable 3-bump solution. Stability of solutions was determined by linearising (5.14) about them. The curve in Fig. 5.2 shows N -bump solutions which are symmetric about the origin, where N is odd. A similar curve exists for N even (not shown) and asymmetric solutions also exist [17]. In summary, spatially-localised solutions of (5.10) are generic and form families which are connected in a snake-like fashion which can be uncovered as parameters are

varied. For more details on (5.10)–(5.11) the reader is referred to [36]. We next consider the generalisation of neural field models to two spatial dimensions and again investigate spatially-localised solutions.

5.3 Two Dimensional Bumps and Rings

Neural field equations are easily generalised to two spatial dimensions, and the simplest are of the form

$$\frac{\partial u(\mathbf{x}, t)}{\partial t} = -u(\mathbf{x}, t) + \int_{\mathbb{R}^2} w(|\mathbf{x} - \mathbf{y}|) f(u(\mathbf{y}, t)) d\mathbf{y} \quad (5.24)$$

where $\mathbf{x} \in \mathbb{R}^2$ and w and f have their previous meanings. Note that w is a function of the scalar distance between points \mathbf{x} and \mathbf{y} . Spatially-localised solutions of equations of the form (5.24) have only recently been analysed in any depth [9, 16, 22–24, 29, 35, 38]. The study of such solutions is harder than in one spatial dimension for the following reasons:

- Their analytical construction involves integrals over subsets of the plane rather than over intervals.
- The determination of the stability of, say, a circular stationary solution is more difficult because perturbations which break the rotational symmetry must be considered.
- Numerical studies require vastly more mesh points in a discretisation of the domain.

However, the use of the techniques presented in Sect. 5.1 has been fruitful for the construction and analysis of such solutions. One important point to note is that the techniques cannot be applied directly when the function w is one of the commonly used ones mentioned above. For example, if $w(x) = e^{-x} - Me^{-mx}$ (of Mexican-hat type when $0 < M < 1$ and $0 < m < 1$) then its Fourier transform is

$$F[w](|\mathbf{k}|) = \frac{1}{(1 + |\mathbf{k}|^2)^{3/2}} - \frac{Mm}{(m^2 + |\mathbf{k}|^2)^{3/2}} \quad (5.25)$$

where $\mathbf{k} \in \mathbb{R}^2$ is the transform variable. Rearranging and then taking the inverse Fourier transform one faces the question as to what a differential equation containing an operator like $(1 - \nabla^2)^{3/2}$ actually means [15]. One way around this is to expand a term like $(1 + |\mathbf{k}|^2)^{3/2}$ around $|\mathbf{k}| = 0$ as $1 + (3/2)|\mathbf{k}|^2 + O(|\mathbf{k}|^4)$ and keep only the first few terms, thus (after inverse transforming) giving one a PDE. This is known as the long wavelength approximation [37]; see [15] for a discussion.

A more fruitful approach is to realise that neural field models are qualitative only, and we can gain insight from models in which the functions w and f are qualitatively correct. Thus we have some freedom in our choice of these functions.

The approach of Laing and co-workers [32, 35, 36] was to use this freedom to choose not w , but its Fourier transform. If the Fourier transform of w is chosen so that the Fourier transform of (5.24) can be rearranged and then inverse transformed to give a simple differential equation, and the resulting function w is qualitatively correct (i.e. has the same general properties as connectivity functions of interest) then one can make much progress.

As an example, consider the case when

$$F[w](|\mathbf{k}|) = \frac{A}{B + (|\mathbf{k}|^2 - M)^2} \quad (5.26)$$

where A , B and M are parameters [35]. Taking the Fourier transform of (5.24), using (5.26), and rearranging, one obtains

$$\left\{ |\mathbf{k}|^4 - 2M|\mathbf{k}|^2 + B + M^2 \right\} F \left[u + \frac{\partial u}{\partial t} \right] (\mathbf{k}) = AF[f(u)](\mathbf{k}) \quad (5.27)$$

and upon taking the inverse Fourier transform one obtains the differential equation

$$[\nabla^4 + 2M\nabla^2 + B + M^2] \left(u + \frac{\partial u}{\partial t} \right) = Af(u) \quad (5.28)$$

The function w is then defined as the inverse Fourier transform of its Fourier transform, i.e.

$$w(x) = A \int_0^\infty \frac{s J_0(xs)}{B + (s^2 - M)^2} ds \quad (5.29)$$

where J_0 is the Bessel function of the first kind of order 0 [35]. ($w(x)$ is the Hankel transform of order 0 of $F[w]$.) Figure 5.4 shows a plot of $w(x)$ for parameter values $M = 1$, $A = 0.4$, $B = 0.1$. We see that it is of a physiologically-plausible form, qualitatively similar to that shown in Fig. 5.1 (left). We have thus formally transformed (5.24) into the PDE (5.28).

As a start we consider spatially-localised and rotationally-invariant solutions of (5.28), which satisfy

$$\begin{aligned} & \left[\frac{\partial^4}{\partial r^4} + \frac{2}{r} \frac{\partial^3}{\partial r^3} - \frac{1}{r^2} \frac{\partial^2}{\partial r^2} + \frac{1}{r^3} \frac{\partial}{\partial r} + 2M \left(\frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} \right) + (B + M^2) \right] \left(u + \frac{\partial u}{\partial t} \right) \\ &= Af(u) \end{aligned} \quad (5.30)$$

with

$$\frac{\partial u}{\partial r} \Big|_{r=0} = \frac{\partial^3 u}{\partial r^3} \Big|_{r=0} = 0 \quad \text{and} \quad \lim_{r \rightarrow \infty} \left(u, \frac{\partial u}{\partial r}, \frac{\partial^2 u}{\partial r^2}, \frac{\partial^3 u}{\partial r^3} \right) = (0, 0, 0, 0) \quad (5.31)$$

Fig. 5.4 The function $w(x)$ defined by (5.29) for parameter values $M = 1, A = 0.4, B = 0.1$

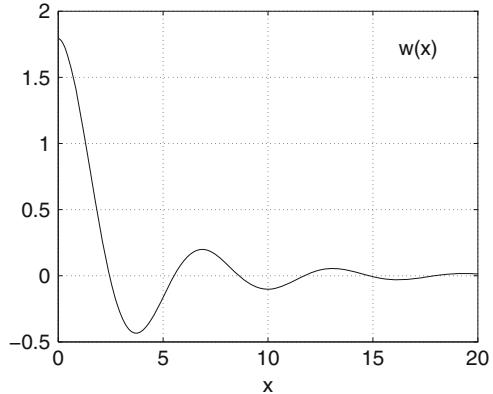
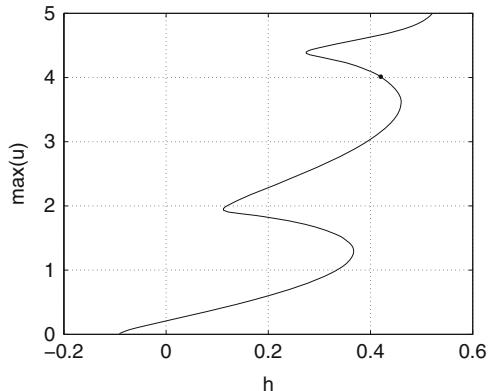


Fig. 5.5 Solutions of (5.30)–(5.31) with $u(0) > 0$ and $u''(0) < 0$ as a function of h . Other parameter values: $\kappa = 0.05, M = 1, A = 0.4, B = 0.1$. The solution $\bar{u}(r)$ at the point indicated by the circle is shown in Fig. 5.6 (left)



where u is now a function of radius r and time t only. We can numerically find and then follow stationary solutions of (5.30)–(5.31) as parameters are varied. For example, Fig. 5.5 shows the effects of varying h for solutions with $u(0) > 0$ and $u''(0) < 0$. We see a snaking curve similar to that in Fig. 5.2, and as we move up the snake, on each successive branch the solution gains one more large amplitude oscillation.

For any particular solution, $\bar{u}(r)$ on the curve in Fig. 5.5 one can find its stability by linearising (5.28) about it. To do this we write

$$u(r, \theta, t) = \bar{u}(r) + \epsilon v(t, r) \cos(m\theta) \quad (5.32)$$

where $0 < \epsilon \ll 1$ and $m \geq 0$ is an integer, the azimuthal index. We choose this form of perturbation in order to find solutions which break the circular symmetry of the system. Substituting (5.32) into (5.28) and keeping only first order terms in ϵ we obtain

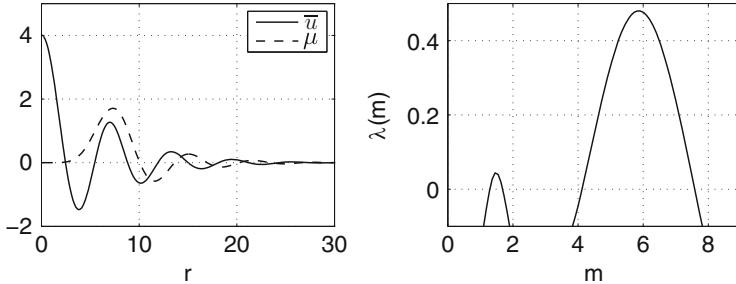


Fig. 5.6 *Left:* the solid curve shows $\bar{u}(r)$ at the point indicated by the circle in Fig. 5.5. The dashed curve shows the eigenfunction $\mu(r)$ corresponding to $\lambda(6)$. *Right:* $\lambda(m)$ for the solution shown solid in the left panel. The integer with largest λ is $N = 6$

$$\left[\frac{\partial^4}{\partial r^4} + \frac{2}{r} \frac{\partial^3}{\partial r^3} + \left(\frac{2Mr^2 - 2m^2 - 1}{r^2} \right) \frac{\partial^2}{\partial r^2} + \left(\frac{2m^2 + 1 + 2Mr^2}{r^3} \right) \frac{\partial}{\partial r} + \frac{m^4 - 4m^2 + (B + M^2)r^4 - 2Mm^2r^2}{r^4} \right] \left(v + \frac{\partial v}{\partial t} \right) = Af'(\bar{u})v \quad (5.33)$$

Since this equation is linear in v we expect solutions of the form $v(r, t) \sim \mu(r)e^{\lambda t}$ as $t \rightarrow \infty$, where λ is the most positive eigenvalue associated with the stability of \bar{u} (which we assume to be real) and $\mu(r)$ is the corresponding eigenfunction.

Thus to determine the stability of a circularly-symmetric solution with radial profile $\bar{u}(r)$, we solve (5.33) for integer $m \geq 0$ and determine $\lambda(m)$. If N is the integer for which $\lambda(N)$ is largest, and $\lambda(N) > 0$, then this circularly-symmetric solution will be unstable with respect to perturbations with D_N symmetry, and the radial location of the growing perturbation will be given by $\mu(r)$.

For example, consider the solution shown solid in the left panel of Fig. 5.6. This solution exists at $h = 0.42$, so in terms of active regions (where $u > h$) this solution corresponds to a central circular bump with a ring surrounding it. Calculating $\lambda(m)$ for this solution we obtain the curve in Fig. 5.6 (right). (We do not need to be restricted to integer m for the calculation.) We see that for this solution $N = 6$, and thus we expect a circularly-symmetric solution of (5.28) with radial profile given by $\bar{u}(r)$ to be unstable at these parameter values, and most unstable with respect to perturbations with D_6 symmetry. The eigenfunction $\mu(r)$ corresponding to $\lambda(6)$ is shown dashed in Fig. 5.6 (left). It is spatially-localised around the ring at $r \approx 7$, so we expect the instability to appear here.

Figure 5.7 shows the result of simulating (5.28) with an initial condition formed by rotating the radial profile in Fig. 5.6 (left) through a full circle in the angular direction, and then adding a small random perturbation to u at each grid point. The initial condition is shown in the left panel and the final state (which is stable) is shown in the right panel. We see the formation of six bumps at the location of the first ring, as expected. This analysis has thus successfully predicted the appearance

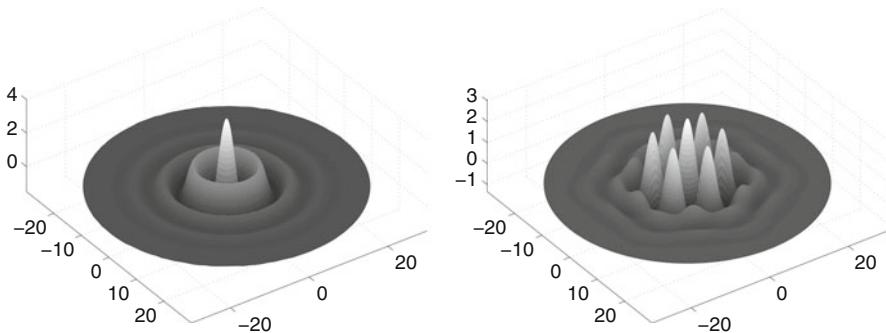


Fig. 5.7 A simulation of (5.28) with initial condition corresponding to $\bar{u}(r)$ in Fig. 5.6. *Left:* initial condition. *Right:* stable final state. $u(r, \theta)$ is plotted vertically

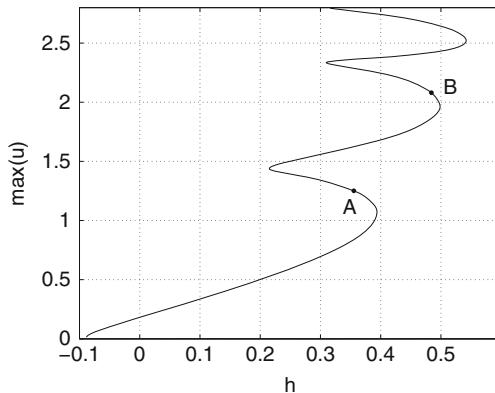


Fig. 5.8 Solutions of (5.30)–(5.31) with $u(0) < 0$ and $u''(0) > 0$ as a function of h . Other parameter values: $\kappa = 0.05$, $M = 1$, $A = 0.4$, $B = 0.1$. The solutions $\bar{u}(r)$ at the points A and B are shown in Fig. 5.9 (left) and Fig. 5.11 (left), respectively

of a stable “7-bump” solution from the initial condition shown in Fig. 5.7 (left). (We used a regular grid in polar coordinates, with domain radius 30, using 200 points in the radial direction and 140 in the angular. The spatial derivatives in (5.28) were approximated using second-order accurate finite differences.)

We can also consider stationary solutions of (5.30)–(5.31) for which $u(0) < 0$ and $u''(0) > 0$, i.e. which have a “hole” in the centre. Following these solutions as h is varied we obtain Fig. 5.8. As in Fig. 5.5 we see a snake of solutions, with successive branches having one more large amplitude oscillation. We will consider the stability of two solutions on the curve in Fig. 5.8; first, the solution at point A, shown in the left panel of Fig. 5.9. This solution corresponds to one with just a single ring of active neurons. Calculating $\lambda(m)$ for this solution we obtain the curve in Fig. 5.9 (right), and we see that a circularly-symmetric solution of (5.28) with radial profile given by this $\bar{u}(r)$ will be most unstable with respect to perturbations with

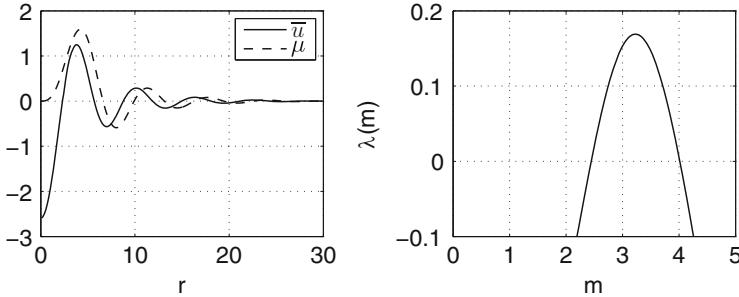


Fig. 5.9 Left: the solid curve shows $\bar{u}(r)$ at the point indicated by the point A in Fig. 5.8. The dashed curve shows the eigenfunction $\mu(r)$ corresponding to $\lambda(3)$. Right: $\lambda(m)$ for the solution shown solid in the left panel. The integer with largest λ is $N = 3$

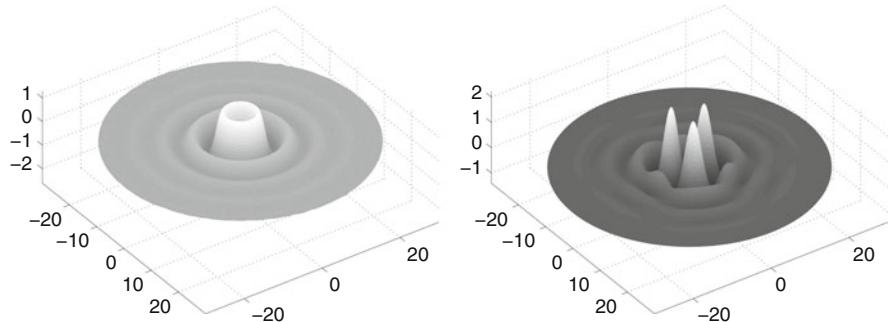


Fig. 5.10 A simulation of (5.28) with initial condition corresponding to $\bar{u}(r)$ in Fig. 5.9. Left: initial condition. Right: stable final state. $u(r, \theta)$ is plotted vertically

D_3 symmetry. The eigenfunction $\mu(r)$ corresponding to $N = 3$ is shown dashed in Fig. 5.9 (left), and it is localised at the first maximum of $\bar{u}(r)$.

Figure 5.10 shows the result of simulating (5.28) with an initial condition formed by rotating the radial profile in Fig. 5.9 (left) through a full circle in the angular direction, and then adding a small random perturbation to u at each grid point. The initial condition is shown in the left panel and the final state (which is stable) is shown in the right panel. We see the formation of three bumps at the first ring, as expected.

Now consider the solution at point B in Fig. 5.8. This solution, shown in Fig. 5.11 (left) corresponds to one with two active rings. An analysis of its stability is shown in Fig. 5.11 (right) and we see that it is most unstable with respect to perturbations with D_9 symmetry, and that these should appear at the outer ring. Figure 5.12 shows the result of simulating (5.28) with an initial condition formed by rotating the radial profile in Fig. 5.11 (left) through a full circle in the angular direction, and then adding a small random perturbation to u at each grid point. The initial condition

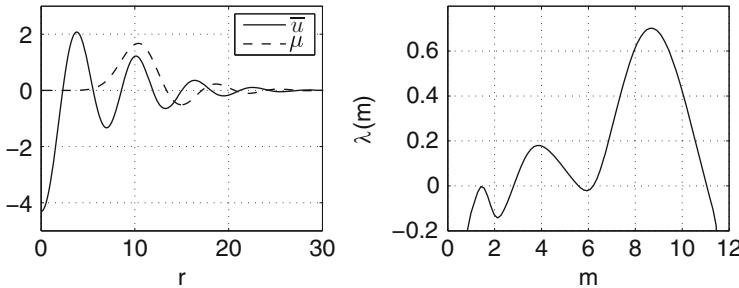


Fig. 5.11 Left: the solid curve shows $\bar{u}(r)$ at the point indicated by the point B in Fig. 5.8. The dashed curve shows the eigenfunction $\mu(r)$ corresponding to $\lambda(9)$. Right: $\lambda(m)$ for the solution shown solid in the left panel

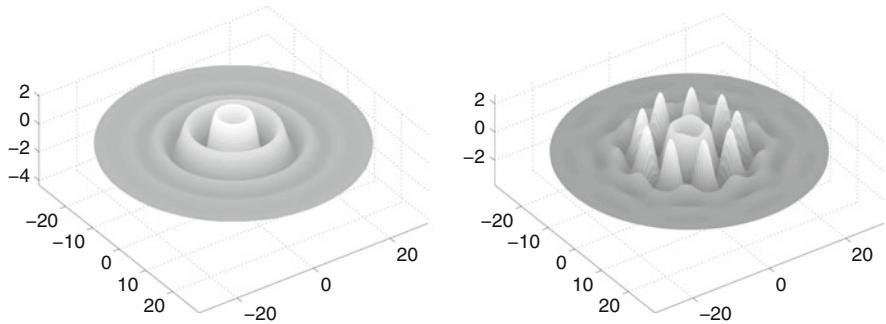


Fig. 5.12 A simulation of (5.28) with initial condition corresponding to $\bar{u}(r)$ in Fig. 5.11. Left: initial condition. Right: stable final state. $u(r, \theta)$ is plotted vertically

is shown in the left panel and the final state (which is stable) is shown in the right panel. We see the formation of nine bumps at the second ring, as expected.

In summary we have shown how to analyse the stability of rotationally-symmetric solutions of the neural field equation (5.24), where w is given by (5.29), via transformation to a PDE. Notice that for all functions \bar{u} shown in the left panels of Figs. 5.6, 5.9 and 5.11, $\lambda(0) < 0$, i.e. these are *stable* solutions of (5.30). However, they are *unstable* with respect to some perturbations which break their rotational invariance. The stable states for all three examples considered consist of a small number of spatially-localised active regions.

Similar results to those presented in this section were obtained subsequently by [38] using a Heaviside firing rate function, which allowed for the construction of an Evans function to determine stability of localised patterns. These authors also showed that the presence of a second, slow variable could cause a *rotational* instability of a pattern like that in Fig. 5.10 (right), resulting in it rotating at a constant speed. Very recently, instabilities of rotationally-symmetric solutions

were addressed by considering the dynamics of the interface dividing regions of high activity from those with low activity, again using the Heaviside firing rate function [14] (and Coombes chapter). Several other authors have also recently investigated symmetry breaking bifurcations of spatially-localised bumps [9, 16]. We now consider solutions of two-dimensional neural field equations which are not spatially-localised, specifically, spiral waves.

5.4 Spiral Waves

The function w used in the previous section was of the decaying oscillatory type (Fig. 5.4). Another form of coupling of interest is purely positive, i.e. excitatory. However, without some form of negative feedback, activity in a neural system with purely excitatory coupling will typically spread over the whole domain. With the inclusion of some form of slow negative feedback such as spike frequency adaptation [13] or synaptic depression [31], travelling pulses of activity are possible [1, 12, 19]. In two spatial dimensions the analogue of a travelling pulse is a spiral wave [2, 3], which we now study. Let us consider the system

$$\frac{\partial u(\mathbf{x}, t)}{\partial t} = -u(\mathbf{x}, t) + B \int_{\Omega} w(|\mathbf{x} - \mathbf{y}|) F(u(\mathbf{y}, t)) d\mathbf{y} - a(\mathbf{x}, t) \quad (5.34)$$

$$\tau \frac{\partial a(\mathbf{x}, t)}{\partial t} = Au(\mathbf{x}, t) - a(\mathbf{x}, t) \quad (5.35)$$

where $\Omega \subset \mathbb{R}^2$ which, in practice, we choose to be a disk, and the firing rate function is

$$F(u) = \frac{1}{1 + e^{-\beta(u-h)}}. \quad (5.36)$$

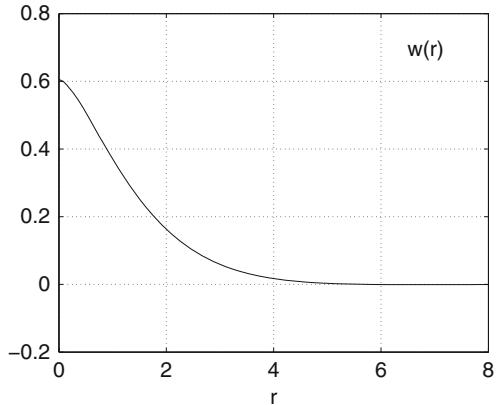
where h and β are parameters. This system is very similar to that in [24] and is the two-dimensional version of that considered in [23, 39]. If we choose the coupling function to be

$$w(r) = \int_0^\infty \frac{s J_0(rs)}{s^4 + s^2 + 1} ds \quad (5.37)$$

then, using the same ideas as above (and ignoring the fact that we are not dealing with spatially-localised solutions) (5.34) is equivalent to

$$[\nabla^4 - \nabla^2 + 1] \left(\frac{\partial u(\mathbf{x}, t)}{\partial t} + u(\mathbf{x}, t) + a(\mathbf{x}, t) \right) = BF(u(\mathbf{x}, t)) \quad (5.38)$$

Fig. 5.13 The function $w(r)$ defined by (5.37)



We choose boundary conditions

$$u(R, \theta, t) = \frac{\partial^2 u(r, \theta, t)}{\partial r^2} \Big|_{r=R} = 0 \quad (5.39)$$

for all θ and t , where R is radius of the circular domain and we have written u in polar coordinates. The two differences between the system considered here and that in [32] are that here we use the firing rate function F (Eq. (5.36)), which is non-zero everywhere (the function f (Eq. (5.21)) was used in [32]), and the boundary conditions given in (5.39) are different from those in [32].

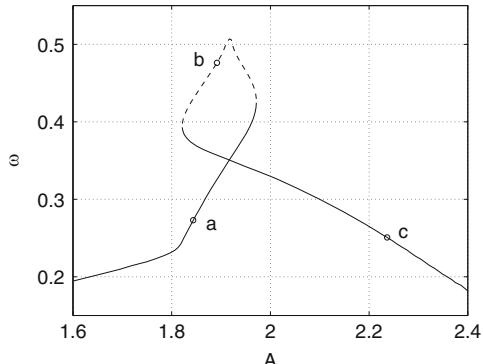
The function $w(r)$ defined by (5.37) is shown in Fig. 5.13 and we see that it is positive and decays monotonically as $r \rightarrow \infty$. For a variety of parameters, the system (5.34)–(5.35) supports a rigidly-rotating spiral wave on a circular domain. To find and study such a wave we recognise that rigidly-rotating patterns on a circular domain can be “frozen” by moving to a coordinate frame rotating at the same speed as the pattern [2, 3, 5]. These rigidly rotating patterns satisfy the time-independent equations

$$[\nabla^4 - \nabla^2 + 1] \left(-\omega \frac{\partial u}{\partial \theta} + u + a \right) = BF(u) \quad (5.40)$$

$$-\omega \tau \frac{\partial a}{\partial \theta} = Au - a \quad (5.41)$$

where ω is the rotation speed of the pattern and θ is the angular variable in polar coordinates. Rigidly rotating spiral waves are then solutions of (5.40)–(5.41), together with a scalar “pinning” equation [2, 32] which allows us to determine ω as well as u and a . In practice, one solves (5.41) to obtain a as a function of u and substitutes into (5.40), giving the single equation for u

Fig. 5.14 ω as a function of A for spiral wave solutions of (5.40)–(5.41). Solid curves are stable, dashed unstable. The spiral wave at points marked “a”, “b” and “c” are shown in Fig. 5.15. Other parameters are $h = 0.6$, $\beta = 20$, $\tau = 3$, $B = 3.5$. The domain has radius 35



$$[\nabla^4 - \nabla^2 + 1] \left(1 - \omega \frac{\partial}{\partial \theta} + A \left[1 - \omega \tau \frac{\partial}{\partial \theta} \right]^{-1} \right) u = BF(u) \quad (5.42)$$

Having found a solution \bar{u} of (5.42) its stability can be determined by linearising (5.34)–(5.35) about (\bar{u}, \bar{a}) , where

$$\left(1 - \omega \tau \frac{\partial}{\partial \theta} \right) \bar{a} = A \bar{u} \quad (5.43)$$

As we have done in previous sections, we can numerically follow solutions of (5.42) as parameters are varied, determining their stability.

In Fig. 5.14 we show ω as a function of A and also indicate the stability of solutions. Interestingly, there is a region of bistability for moderate values of A . Typical solutions (of both u and a) at three different points on the curve are shown in Fig. 5.15. In agreement with the results in [32] we see that as A (the strength of the negative feedback) is decreased, more of the domain becomes active, and as A is increased, less of the domain is active. The results of varying h (the threshold of the firing rate function) are shown in Fig. 5.16. We obtain results quite similar to those in Fig. 5.14—as h is decreased, more of the domain becomes active, and vice versa, and we also have a region of bistability. Figure 5.17 shows the result of varying τ : for large τ the spiral is unstable. The bifurcations seen in Figs. 5.14, 5.16 and 5.17 are all generic saddle-node bifurcations. In principle they could be followed as two parameters are varied, thus mapping out regions of parameter space in which stable spiral waves exist.

We conclude this section by noting that spiral waves have been observed in simulations which include synaptic depression rather than spike frequency adaptation [8, 31], and also seen experimentally in brain slice preparations [27, 28].

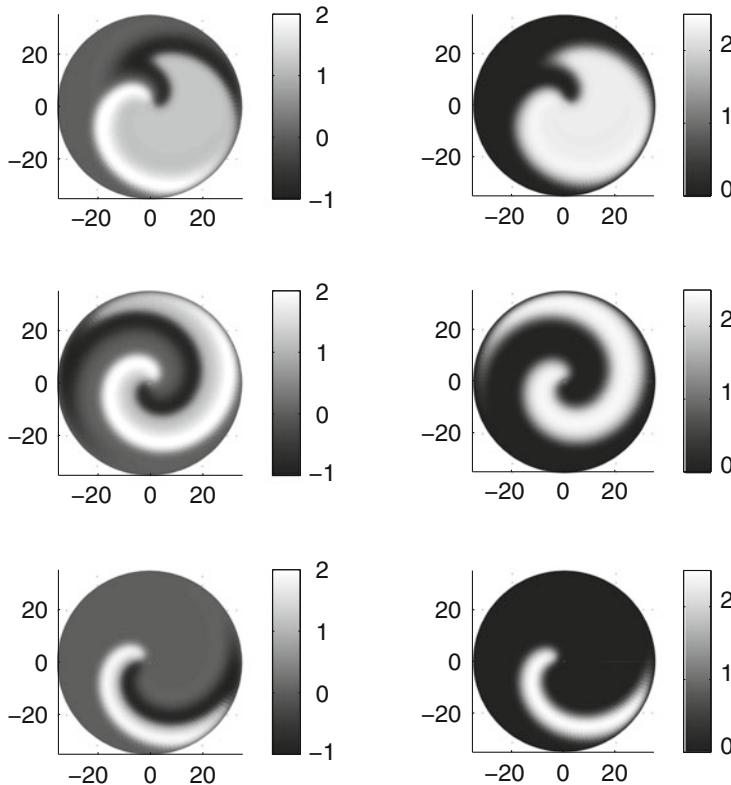


Fig. 5.15 Solutions of (5.40)–(5.41) at the three points marked in Fig. 5.14. The *left column* shows u and the *right column* shows a . The *top, middle and bottom rows* correspond to points “a”, “b” and “c”, respectively

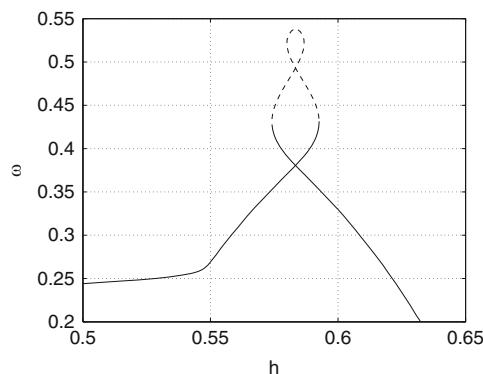


Fig. 5.16 ω as a function of h for spiral wave solutions of (5.40)–(5.41). *Solid curves* are stable, *dashed* unstable. Other parameters are $A = 2, \beta = 20, \tau = 3, B = 3.5$. The domain has radius 35

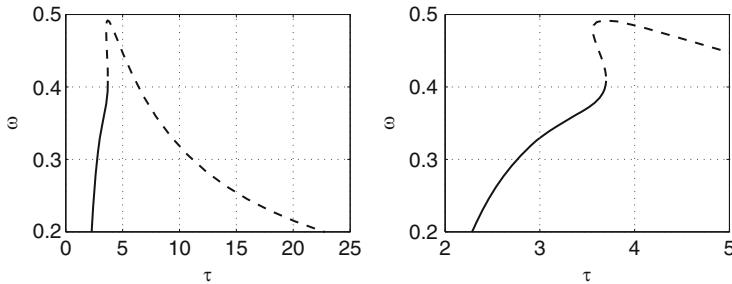


Fig. 5.17 ω as a function of τ for spiral wave solutions of (5.40)–(5.41). Solid curves are stable, dashed unstable. The right panel is an enlargement of the left. Other parameters are $A = 2$, $\beta = 20$, $B = 3.5$, $h = 0.6$. The domain has radius 35

5.5 Conclusion

This chapter has summarised some of the results from [32, 35, 36], in which neural field equations in one and two spatial dimensions were studied by being converted into PDEs via a Fourier transform in space. In two spatial dimensions we showed how to investigate the instabilities of spatially-localised “bumps” and rings of activity, and also how to study spiral waves. An important technique used was the numerical continuation of solutions of large systems of coupled, nonlinear, algebraic equations defined by the discretisation of PDEs. Since the work summarised here was first published a number of other authors have used some of the techniques presented here to further investigate neural field models [9, 15, 21, 26, 31, 33].

References

1. Amari, S.: Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.* **27**(2), 77–87 (1977)
2. Bär, M., Bangia, A., Kevrekidis, I.: Bifurcation and stability analysis of rotating chemical spirals in circular domains: boundary-induced meandering and stabilization. *Phys. Rev. E* **67**(5), 056,126 (2003)
3. Barkley, D.: Linear stability analysis of rotating spiral waves in excitable media. *Phys. Rev. Lett.* **68**(13), 2090–2093 (1992)
4. Beurle, R.L.: Properties of a mass of cells capable of regenerating pulses. *Philos. Trans. R. Soc. B: Biol. Sci.* **240**(669), 55–94 (1956)
5. Beyn, W., Thümmler, V.: Freezing solutions of equivariant evolution equations. *SIAM J. Appl. Dyn. Syst.* **3**(2), 85–116 (2004)
6. Blomquist, P., Wyller, J., Einevoll, G.: Localized activity patterns in two-population neuronal networks. *Phys. D: Nonlinear Phenom.* **206**(3–4), 180–212 (2005)
7. Bojak, I., Liley, D.T.J.: Modeling the effects of anesthesia on the electroencephalogram. *Phys. Rev. E* **71**, 041,902 (2005). doi:10.1103/PhysRevE.71.041902. <http://link.aps.org/doi/10.1103/PhysRevE.71.041902>

8. Bressloff, P.C.: Spatiotemporal dynamics of continuum neural fields. *J. Phys. A: Math. Theor.* **45**(3), 033,001 (2012). <http://stacks.iop.org/1751-8121/45/i=3/a=033001>
9. Bressloff, P.C., Kilpatrick, Z.P.: Two-dimensional bumps in piecewise smooth neural fields with synaptic depression. *SIAM J. Appl. Math.* **71**(2), 379–408 (2011). doi:10.1137/100799423. <http://link.aip.org/link/?SMM/71/379/1>
10. Burke, J., Knobloch, E.: Homoclinic snaking: structure and stability. *Chaos* **17**(3), 037,102 (2007). doi:10.1063/1.2746816. <http://link.aip.org/link/?CHA/17/037102/1>
11. Champneys, A.: Homoclinic orbits in reversible systems and their applications in mechanics, fluids and optics. *Phys. D: Nonlinear Phenom.* **112**(1–2), 158–186 (1998)
12. Coombes, S.: Waves, bumps, and patterns in neural field theories. *Biol. Cybern.* **93**(2), 91–108 (2005)
13. Coombes, S., Owen, M.: Bumps, breathers, and waves in a neural network with spike frequency adaptation. *Phys. Rev. Lett.* **94**(14), 148,102 (2005)
14. Coombes, S., Schmidt, H., Bojak, I.: Interface dynamics in planar neural field models. *J. Math. Neurosci.* **2**(1), 1–27 (2012)
15. Coombes, S., Venkov, N., Shiu, L., Bojak, I., Liley, D., Laing, C.: Modeling electrocortical activity through improved local approximations of integral neural field equations. *Phys. Rev. E* **76**(5), 051,901 (2007)
16. Doubovinski, K., Herrmann, J.: Stability of localized patterns in neural fields. *Neural comput.* **21**(4), 1125–1144 (2009)
17. Elvin, A.: Pattern formation in a neural field model. Ph.D. thesis, Massey University, New Zealand (2008)
18. Elvin, A., Laing, C., McLachlan, R., Roberts, M.: Exploiting the Hamiltonian structure of a neural field model. *Phys. D: Nonlinear Phenom.* **239**(9), 537–546 (2010)
19. Ermentrout, B.: Neural networks as spatio-temporal pattern-forming systems. *Rep. Prog. Phys.* **61**, 353 (1998)
20. Ermentrout, G.B., Cowan, J.D.: A mathematical theory of visual hallucination patterns. *Biol. Cybern.* **34**, 137–150 (1979)
21. Faye, G., Rankin, J., Chossat, P.: Localized states in an unbounded neural field equation with smooth firing rate function: a multi-parameter analysis. *J. Math. Biol.* **66**(6), 1303–1338 (2013)
22. Folias, S.E.: Nonlinear analysis of breathing pulses in a synaptically coupled neural network. *SIAM J. Appl. Dyn. Syst.* **10**, 744–787 (2011)
23. Folias, S., Bressloff, P.: Breathing pulses in an excitatory neural network. *SIAM J. Appl. Dyn. Syst.* **3**(3), 378–407 (2004)
24. Folias, S.E., Bressloff, P.C.: Breathers in two-dimensional neural media. *Phys. Rev. Lett.* **95**, 208,107 (2005). doi:10.1103/PhysRevLett.95.208107. <http://link.aps.org/doi/10.1103/PhysRevLett.95.208107>
25. Griffith, J.: A field theory of neural nets: I: derivation of field equations. *Bull. Math. Biol.* **25**, 111–120 (1963)
26. Guo, Y., Chow, C.C.: Existence and stability of standing pulses in neural networks: I. Existence. *SIAM J. Appl. Dyn. Syst.* **4**, 217–248 (2005). doi:10.1137/040609471. <http://link.aip.org/link/?SJA/4/217/1>
27. Huang, X., Troy, W., Yang, Q., Ma, H., Laing, C., Schiff, S., Wu, J.: Spiral waves in disinhibited mammalian neocortex. *J. Neurosci.* **24**(44), 9897 (2004)
28. Huang, X., Xu, W., Liang, J., Takagaki, K., Gao, X., Wu, J.Y.: Spiral wave dynamics in neocortex. *Neuron* **68**(5), 978–990 (2010)
29. Hutt, A., Rougier, N.: Activity spread and breathers induced by finite transmission speeds in two-dimensional neural fields. *Phys. Rev. E* **82**, 055,701 (2010). doi:10.1103/PhysRevE.82.055701. <http://link.aps.org/doi/10.1103/PhysRevE.82.055701>
30. Jirsa, V.K., Haken, H.: Field theory of electromagnetic brain activity. *Phys. Rev. Lett.* **77**, 960–963 (1996). doi:10.1103/PhysRevLett.77.960. <http://link.aps.org/doi/10.1103/PhysRevLett.77.960>
31. Kilpatrick, Z., Bressloff, P.: Spatially structured oscillations in a two-dimensional excitatory neuronal network with synaptic depression. *J. Comput. Neurosci.* **28**, 193–209 (2010)

32. Laing, C.: Spiral waves in nonlocal equations. *SIAM J. Appl. Dyn. Syst.* **4**(3), 588–606 (2005)
33. Laing, C., Coombes, S.: The importance of different timings of excitatory and inhibitory pathways in neural field models. *Netw. Comput. Neural Syst.* **17**(2), 151–172 (2006)
34. Laing, C., Glendinning, P.: Bifocal homoclinic bifurcations. *Phys. D: Nonlinear Phenom.* **102**(1–2), 1–14 (1997)
35. Laing, C., Troy, W.: PDE methods for nonlocal models. *SIAM J. Appl. Dyn. Syst.* **2**(3), 487–516 (2003)
36. Laing, C., Troy, W., Gutkin, B., Ermentrout, G.: Multiple bumps in a neuronal model of working memory. *SIAM J. Appl. Math.* **63**, 62 (2002)
37. Liley, D., Cadusch, P., Dafilis, M.: A spatially continuous mean field theory of electrocortical activity. *Netw. Comput. Neural Syst.* **13**(1), 67–113 (2002)
38. Owen, M., Laing, C., Coombes, S.: Bumps and rings in a two-dimensional neural field: splitting and rotational instabilities. *New J. Phys.* **9**, 378 (2007)
39. Pinto, D., Ermentrout, G.: Spatially structured activity in synaptically coupled neuronal networks: I. Traveling fronts and pulses. *SIAM J. Appl. Math.* **62**(1), 206–225 (2001)
40. Pinto, D., Ermentrout, G.: Spatially structured activity in synaptically coupled neuronal networks: II. Lateral inhibition and standing pulses. *SIAM J. Appl. Math.* **62**(1), 226–243 (2001)
41. Wilson, H., Cowan, J.: A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Biol. Cybern.* **13**(2), 55–80 (1973)
42. Woods, P., Champneys, A.: Heteroclinic tangles and homoclinic snaking in the unfolding of a degenerate reversible Hamiltonian-Hopf bifurcation. *Phys. D: Nonlinear Phenom.* **129**(3–4), 147–170 (1999). doi:10.1016/S0167-2789(98)00309-1. <http://www.sciencedirect.com/science/article/pii/S0167278998003091>
43. Zhang, K.: Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. *J. Neurosci.* **16**(6), 2112–2126 (1996)

Chapter 6

Numerical Simulation Scheme of One- and Two Dimensional Neural Fields Involving Space-Dependent Delays

Axel Hutt and Nicolas Rougier

Abstract Neural Fields describe the spatiotemporal dynamics of neural populations involving spatial axonal connections between neurons. These neuronal connections are delayed due to the finite axonal transmission speeds along the fibers inducing a distance-dependent delay between two spatial locations. The numerical simulation in 1-dimensional neural fields is numerically demanding but may be performed in a reasonable run time by implementing standard numerical techniques. However 2-dimensional neural fields demand a more sophisticated numerical technique to simulate solutions in a reasonable time. The work presented shows a recently developed numerical iteration scheme that allows to speed up standard implementations by a factor 10–20. Applications to some pattern forming systems illustrate the power of the technique.

6.1 Introduction

Finite transmission speed in physical systems has attracted research for decades. Previous work on heat diffusion has shown experimentally that the transmission speed (also called propagation speed in the literature) is finite in certain media [14, 16]. These results do not show accordance to classical diffusion theory implying infinite transmission speed. To cope with this problem theoretically, Cattaneo was

A. Hutt (✉)

Neurosys, Inria CR Grand Est – Nancy, Villers-les-Nancy, France
e-mail: axel.hutt@inria.fr

N. Rougier

Mnemosyne, Inria Bordeaux Sud-Ouest, Talence, France

LaBRI, UMR 5800 CNRS, Bordeaux University, Bordeaux, France

Institute of Neurodegenerative Diseases, UMR 5293, Bordeaux, France

one of the first to insert delayterms into the diffusion equation to achieve a finite transmission speed [5].

Recently, an integral model has been proposed which takes into account a finite transmission speed as a space-dependent retardation [10]. It was shown that the Cattaneo-equation can be derived from this model. This model is well-established in computational neuroscience and known as the neural field model. It describes the activity evolution in a neural population involving finite transmission speed along axonal fibres. The neural field model has been shown to model successfully neural activity known from experiments [3, 9]. In the recent decades, neural fields have been studied analytically and numerically in one and two spatial dimensions [15, 19], while previous studies considered finite axonal transmission speeds in one-dimensional models only [1, 2, 6, 11]. To our best knowledge, only few previous studies considered analytically and numerically finite transmission speeds in two-dimensional neural fields. The current work presents a recently developed method [13] to reveal finite transmission speed effects in two-dimensional systems.

The subsequent paragraphs derive a novel fast numerical scheme to simulate the corresponding evolution equations in one and then in some detail in two spatial dimensions. Stimulus-induced activity propagation in two spatial dimensions is studied numerically to illustrate the delayed activity spread. We find numerically transmission delay-induced breathers.

The underlying model considers a one-dimensional line Ω with length l or a two-dimensional rectangle spatial domain Ω with side length l , in both cases assuming periodic boundary conditions. In addition, the center of the coordinate system is chosen to be the center of the domain in the following. Then the neural population activity $V(\mathbf{x}, t)$, i.e. the mean membrane potential, at spatial location $\mathbf{x} \in \Omega$ and time t obeys the evolution equation

$$\tau \frac{\partial}{\partial t} V(\mathbf{x}, t) = I(\mathbf{x}, t) - V(\mathbf{x}, t) + \int_{\Omega} d^n y K(|\mathbf{x} - \mathbf{y}|) S \left[V \left(\mathbf{y}, t - \frac{|\mathbf{x} - \mathbf{y}|}{c} \right) \right] \quad (6.1)$$

with $n = 1$ or $n = 2$, the synaptic time constant τ , the external stimulus $I(\mathbf{x}, t)$, the finite axonal transmission speed c and the nonlinear transfer function S . Moreover, the spatial interaction is non-local and is given by the spatial synaptic connectivity kernel $K(|\mathbf{x} - \mathbf{y}|)$, that depends on the distance between two spatial locations \mathbf{x} and \mathbf{y} only.

6.2 The Novel Principle

For notational simplicity, let us consider a one-dimensional spatial domain. Then the integral on the right hand side of Eq. (6.1) can be re-written as

$$\begin{aligned}
& \int_{\Omega} dz K(|x - z|) S \left[V \left(z, t - \frac{|x - z|}{c} \right) \right] \\
&= \int_{\Omega} dz \int_{-\infty}^{\infty} dt' \delta \left(\frac{|x - z|}{c} - (t - t') \right) K(|x - z|) S[V(z, t')] \\
&= \int_{\Omega} dz \int_{-\infty}^{\infty} dt' L(x - z, t - t') S[V(z, t')] \\
&= \int_{\Omega} dz \int_0^{\tau_{max}} d\tau L(x - z, \tau) S[V(z, t - \tau)],
\end{aligned}$$

with $\tau_{max} = l/c$. This shows that the introduction of the space-time kernel $L(x, t) = K(x)\delta(|x|/c - t)$ allows us to write the single space integral as two integrals: one spatial convolution and one integral over delays. To understand the logic of the computation, let us discretize the time and space by $t \rightarrow t_n = n\Delta t$, $x \rightarrow x_m = m\Delta x$ with $n \in \mathcal{N}_0$, $m \in \mathcal{Z}_0$, $|m| < M/2$ and $l = M\Delta x$. This implies that the speed c also takes discrete values. We obtain

$$\begin{aligned}
L(x, t) \rightarrow L(x_m, t_n) &\sim K(m\Delta x)\delta((\Delta x/c)|m| - n\Delta t) \\
&= K_m \delta \left(\Delta t \left(\frac{\Delta x}{c\Delta t} |m| - n \right) \right) \\
&= K_m \delta(\Delta t(r|m| - n)) \\
&= K_m \delta(|m|\Delta t(r - n/|m|)) \\
&\sim K_m \delta_{|m|, n/r}
\end{aligned}$$

with $K_m = K(m\Delta x)$, $r = \Delta x/(c\Delta t)$. The last equation shows that $L(x_m, t_n) \sim K_{\pm n/r} \neq 0$ only if r is rational number with $r = n/m$ and thus

$$c = \frac{\Delta x}{\Delta t} \frac{1}{r} = \frac{\Delta x}{\Delta t} \frac{m}{n}$$

is discrete.

6.3 The Numerical Implementation in Two Spatial Dimensions

To investigate the activity propagation in detail, we derive a novel iteration scheme for the numerical integration of (6.1) for $n = 2$. Since the integral over space in (6.1) is not a convolution in the presence of a finite transmission speed c , one can not apply directly fast numerical algorithms such as the Discrete Fast Fourier transform (DFT) to calculate the integral. Hence the numerical integration of (6.1) is very time

consuming with standard quadrature techniques. For instance, with a discretization of the spatial domain by N^2 grid intervals and applying the Gaussian quadrature rule for the spatial integral, it would be necessary to compute N^4 elements in each time step which is very time-consuming in the case of a good spatial resolution. The work [13] proposes a fast numerical method that is based on the DFT and resembles the Ritz-Galerkin method well-established to solve partial differential equations.

As in the previous section, the integral in (6.1) reads

$$\begin{aligned} A(\mathbf{x}, t) &= \int_{\Omega} d^2y K(|\mathbf{x} - \mathbf{y}|) S \left[V \left(\mathbf{y}, t - \frac{|\mathbf{x} - \mathbf{y}|}{c} \right) \right] \\ &= \int_{\Omega} d^2y \int_{-\infty}^{\infty} dt' \delta \left(\frac{|\mathbf{x} - \mathbf{y}|}{c} - (t - t') \right) K(|\mathbf{x} - \mathbf{y}|) S[V(\mathbf{y}, t')] \\ &= \int_{\Omega} d^2y \int_{-\infty}^{\infty} dt' L(\mathbf{x} - \mathbf{y}, t - t') S[V(\mathbf{y}, t')] \end{aligned} \quad (6.2)$$

$$= \int_{\Omega} d^2y \int_0^{\tau_{max}} d\tau L(\mathbf{x} - \mathbf{y}, \tau) S[V(\mathbf{y}, t - \tau)], \quad (6.3)$$

with $\tau_{max} = l/\sqrt{2}c$, the novel spatial delay kernel $L(\mathbf{x}, \tau) = K(\mathbf{x})\delta(|\mathbf{x}|/c - \tau)$ and the delta-distribution $\delta(\cdot)$. These simple calculations show that A may be written as a two-dimensional spatial convolution, but with a new delayed spatio-temporal kernel L that now considers the past activity. The form (6.2) has been used previously to study spatio-temporal instabilities in one- and two-dimensional neural fields [19].

The new delay kernel L is independent of time t and is computed on the delay interval only. Hence it represents the contribution of the current and past activity to the current activity at time t . In addition A implies multiple delays and the corresponding delay distribution function depends strongly on the spatial kernel K . In other words, axonal transmission speeds represent a delay distribution as found before in other contexts [7, 12].

Figure 6.1 illustrates the construction of the kernel: given the kernel function K in space (Fig. 6.1, left), $L(\mathbf{x}, \tau)$ is generated by cutting out a ring of radius $c\tau$ (Fig. 6.1, right hand side). In a continuous spatial domain these rings are infinitely thin, whereas a spatially discretized domain yields finite ring widths, see the paragraphs below for more details.

Now let us derive the rules to compute A numerically. The periodic boundary conditions implied lead to discrete wave vectors $\mathbf{k}_{mn} = (k_m, k_n)$ with $k_p = 2\pi p/l$, $p \in \mathbb{Z}_0$. The Fourier series of $V(\mathbf{x}, t)$ reads

$$V(x, y, t) = \frac{1}{l} \sum_{m,n=-\infty}^{\infty} \tilde{V}_{mn}(t) e^{i(k_m x + k_n y)} \quad (6.4)$$

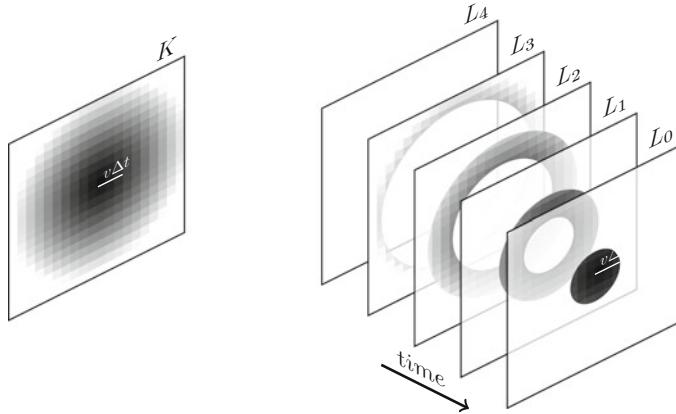


Fig. 6.1 The construction of the delay-kernel $L(\mathbf{x}, \tau)$. Assuming a spatial kernel (left image), L exhibits rings with radius $c\tau$ (images on the right for different delay times) which is the interaction distance of the system at the delay time τ

with the Fourier vector component $\tilde{V}_{mn}(t) = \tilde{V}(k_m, k_n, t)$ and the spatial Fourier transform

$$\tilde{V}_{mn}(t) = \frac{1}{l} \int_{-l/2}^{l/2} dx \int_{-l/2}^{l/2} dy V(x, y, t) e^{-i(k_m x + k_n y)}. \quad (6.5)$$

Inserting (6.4) into (6.3) and applying (6.5) leads to

$$A(x, y, t) = \sum_{m,n=-\infty}^{\infty} e^{i(k_m x + k_n y)} \int_0^{\tau_{max}} dT \tilde{L}_{mn}(T) \tilde{S}_{mn}(t - T), \quad (6.6)$$

with the spatial Fourier transforms of $L(\mathbf{x}, t)$ $\tilde{L}_{mn}(t)$, $\tilde{S}_{mn}(t)$ and the nonlinear functional $S[V(\mathbf{x}, t)]$. Moreover

$$\tilde{L}_{mn}(T) = \frac{c^2}{l} \int_{-l/2c}^{l/2c} \int_{-l/2c}^{l/2c} \delta(|\boldsymbol{\tau}| - T) K(|c\boldsymbol{\tau}|) e^{-i c \mathbf{k}_{mn} \boldsymbol{\tau}} d^2 \boldsymbol{\tau}. \quad (6.7)$$

After obtaining $A(x, y, t)$ in the Fourier space for a continuous spatial domain, now we discretize the spatial domain to gain the final numerical scheme. To this end, Ω is discretized in a regular spatial grid of $N \times N$ elements with grid interval $\Delta x = l/N$. Hence $x \rightarrow x_n = n\Delta x$, $n = -N/2, \dots, N/2 - 1$. By virtue of this discretization, we can approximate (6.6) and (6.7) by applying the rectangular rule $\int_a^b f(x) dx \approx \Delta x \sum_{n=-N/2}^{N/2-1} f(x_n)$. The error is $E < (b-a)\Delta x^2 f''(\xi)/24$, $a < \xi < b$ for twice-differentiable functions f , i.e. the rectangular rule is a good

approximation for smooth functions and large enough N . The same holds true for the discretization of the time integral and Eq. (6.7) reads

$$\tilde{L}_{mn}(T_u) = \frac{l}{N^2 \Delta t} \sum_{k,p=-N/2}^{N/2} L_{kp}(T_u) e^{-i 2\pi(km+np)/N} \quad (6.8)$$

with the discrete version of the delay kernel L

$$L_{kp}(T_u) = \delta(\Delta\tau \sqrt{k^2 + p^2}, T_u) K(|\mathbf{x}_{kp}|).$$

The symbol $\delta(\cdot, \cdot)$ is identical to the Kronecker symbol and is introduced for notational convenience. By virtue of the isotropy of the spatial interactions, in addition we find the simple relation

$$L_{kp}(T_u) = K(uc\Delta t).$$

In other words the width of the rings in Fig. 6.1 is $c\Delta t$. In these latter calculations, we introduced the time discretization $\tau_{kp} = (k, p)\Delta\tau$, $\Delta\tau = \Delta x/c$, $T \rightarrow T_u = u\Delta t$ and $\delta(\tau - T) \rightarrow \delta_{nu}/\Delta t$ for $\tau \rightarrow \tau_n$.

Although the relation $L_{kp}(T_u) = K(uc\Delta t)$ seems to be quite intuitive and elucidates the discrete ring structure of L , it does not give directly the condition which grid point (k, p) belongs to which delay ring. This condition may be read off the Kronecker symbol: u is an integer number and hence $\delta(\Delta\tau \sqrt{k^2 + p^2}, T_u) = 1$ if $[\Delta\tau \sqrt{k^2 + p^2}/\Delta t] = u$ with the integer operation $[a]$ that cuts off the decimal numbers. Consequently the grid points (k, p) that contribute to the delay time T_u obey

$$u \leq \frac{\Delta x}{c\Delta t} \sqrt{k^2 + p^2} < u + 1, \quad u = 0, 1, 2, \dots, u_{max}$$

with $u_{max} = [\tau_{max}/\Delta t]$, i.e. they lie in a ring with inner and outer radius $(c\Delta t/\Delta x)u$ and $(c\Delta t/\Delta x)(u + 1)$, respectively.

Moreover, the definition of $L_{kp}(T_u)$ allows us to derive some conditions on the numerical parameters. The ring width in Fig. 6.1 is $\Delta r = c\Delta t/\Delta x$ which is the number of spatial grid intervals. Hence the maximum radius of a ring is $l/\sqrt{2}\Delta x$ and hence the maximum transmission speed that can be implemented is $c_{max} = l/\sqrt{2}\Delta t$. Since $c_{max} \rightarrow \infty$ for $\Delta t \rightarrow 0$, the transmission speed $c > c_{max}$ in the discrete scheme is equivalent to an infinite transmission speed in the analytical original model and the finiteness of c_{max} results from the time discretization. Moreover, $c \rightarrow c_{max}$ yields $\tau_{max} \rightarrow 0$, i.e. the transmission delay vanishes.

We add that the maximum wave number is $k_{max} = 2\pi/\Delta x$ and, by the definition of Δx , the number of Fourier modes is limited to N .

Combining the latter results now Eq. (6.6) reads

$$A(x_r, y_s, t_v) = \sum_{m,n=-N/2}^{N/2-1} e^{i2\pi(mr+ns)/N} \times \sum_{u=0}^{u_{max}-1} \tilde{L}_{mn}(T_u) \tilde{S}_{mn}(t_v - T_u). \quad (6.9)$$

With the standard definition of the two-dimensional Discrete Fourier Transform

$$DFT[A]_{kp} = \sum_{n,m} A_{nm} e^{-i2\pi(nk+mp)/N}, \quad n, m \in [-N/2; N/2 - 1]$$

and its inverse (IDFT) correspondingly, we find finally

$$A(\mathbf{x}, t_v) = \frac{l^2}{N^2} IDFT \left[\sum_{u=0}^{u_{max}-1} DFT [L(T_u)] \times DFT [S(t_v - T_u)] \right]. \quad (6.10)$$

Some numerical implementations of the DFT assume that the index n runs in the interval $[0; N - 1]$. In this case, Eq. (6.10) is also valid but $DFT[A]_{kp}$ is modulated by a factor $e^{-i\pi(k+p)} = (-1)^{k+p}$.

In practice, $DFT[L(T_u)]$ is computed once for all T_u in the beginning of the simulation since it does not depend on the system activity. Moreover, for $N = 2^n$, $n \in \mathcal{N}$, the discrete Fourier transform may be implemented numerically by a Fast Fourier transform, that speeds up the numerical computation dramatically. This possible algorithm choice represents the major advantage of the proposed method compared to other non-convolution methods.

The discrete version of A can be applied to any explicit or implicit numerical integration scheme. For instance, the numerical Euler scheme stipulates

$$\begin{aligned} \tilde{V}_{mn}(t_i + \Delta t) &= \tilde{V}_{mn}(t_i) + \frac{\Delta t}{\tau} (I_{mn}(t_i) - \tilde{V}_{mn}(t_i)) \\ &\quad + \frac{\Delta t}{\tau} \frac{L^3}{N^4} \sum_{u=0}^d DFT [L(t_u)]_{mn} \times DFT [S(t_v - t_u)]_{mn} \end{aligned} \quad (6.11)$$

where $I_{mn}(t)$ is the Fourier transform of the input $I(\mathbf{x}, t)$ and one obtains $V(\mathbf{x}, t_v)$ by applying Eq. (6.4).

In the following, we study analytically and numerically the response to an external stimulus. At first, let us consider a small input. Then the response is linear about the systems' stationary state. Since we are interested in responses that approach the stationary state after removal of the stimulus, it is necessary to ensure the linear stability of the stationary state.

The stationary state of Eq. (6.1) constant in space and time implies $V_0 = \kappa S[V_0] + I_0$ for a constant input I_0 with the kernel norm $\kappa = \int_{\Omega} K(\mathbf{x}) d^2x$.

Considering small additional external inputs $\bar{I}(\mathbf{x}, t) = I(\mathbf{x}, t) - I_0$, small deviations $u(\mathbf{x}, t) = V(\mathbf{x}, t) - V_0$ from this stationary state obey

$$\frac{du(\mathbf{x}, t)}{dt} = -u(\mathbf{x}, t) + \bar{I}(\mathbf{x}, t) + s' \int_{\Omega} K(\mathbf{x} - \mathbf{y}) u(\mathbf{y}, t - |\mathbf{x} - \mathbf{y}|/c) dy^2. \quad (6.12)$$

with $s' = \delta S[V]/\delta V$, $V = V_0$. Now expanding $u(\mathbf{x}, t)$ into a spatial Fourier series according to Eq. (6.4) and applying a temporal Laplace transform to each Fourier mode amplitude, we find the characteristic equation

$$\lambda + 1 = \int_{\Omega} K(\mathbf{x}) e^{i\mathbf{k}\mathbf{x} - \lambda|\mathbf{x}|/c} d^2x \quad (6.13)$$

with the wave vector $\mathbf{k} = (k_m, k_n)^t$ and the Lyapunov exponent $\lambda \in \mathcal{C}$. The stationary state V_0 is linearly stable if $Re(\lambda) < 0$.

Now let us consider the spatio-temporal response of the system involving the spatially periodic interactions

$$K(\mathbf{x}) = K_o \sum_{i=0}^2 \cos(\mathbf{k}_i \mathbf{x}) \exp(-|\mathbf{x}|/\sigma)$$

with

$$\mathbf{k}_i = k_c (\cos(\phi_i), \sin(\phi_i))^t, \quad \phi_i = i\pi/3.$$

This kernel reflects spatial hexagonal connections which have been found, e.g., in layer 2/3 of the visual cortex in monkeys [17]. Stimulating the stable system by a small external input in the presence of the finite transmission speed c elucidates the transmission delay effect on the linear response. This delay effect has attracted some attention in previous studies on the activity propagation in the visual cortex [4, 18]. For the given kernel, the characteristic equation (6.13) reads

$$\lambda + 1 = \sum_{i=0}^2 f_+(\lambda, \phi_i) + f_-(\lambda, \phi_i)$$

with

$$f_{\pm}(\lambda, \phi_i) = 1/\sqrt{(1/\sigma + \lambda/c)^2 + k^2 + k_c^2 \pm 2kk_c \cos(\phi_i - \theta)}^3$$

and $\mathbf{k} = k(\cos(\theta), \sin(\theta))^t$. The numerical simulation applies parameters which guarantee the stability of the stationary state.

Figure 6.2 shows snapshots of the simulated spatio-temporal response of the system about a stable stationary state applying the numerical scheme (6.11).

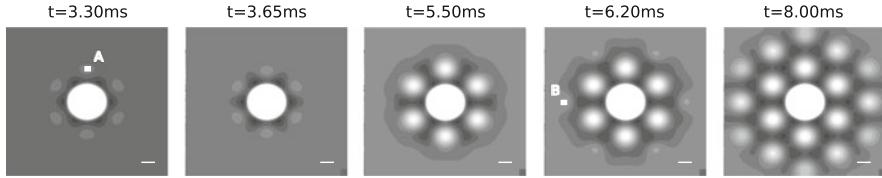


Fig. 6.2 Spatio-temporal response activity to the external stimulus $I(\mathbf{x}, t) = I_0 + e^{-\mathbf{x}^2/\sigma_I^2}$ for the spatial connectivity function $K(\mathbf{x})$ given by the numerical simulation of Eq. (6.1). Used (dimensionless) parameters are $K_o = 0.1, c = 10, l = 10, k_c = 10\pi/l, \sigma = 10, \sigma_I = 0.2, N = 512, \tau = 1, \Delta t = 0.005$. Moreover, $I_0 = 2.0, S[V] = 2/(1 + \exp(-5.5(V - 3)))$ and $V_0 = 2.00083$. The initial values are chosen to $V(\mathbf{x}, \theta) = V_0$ for the delay interval $-l/\sqrt{2c} \leq \theta \leq 0$. Introducing the temporal and spatial scale $\tau = 10\text{ ms}$ and $\lambda = 1.0\text{ mm}$, the results reflect the spatio-temporal activity with transmission speed $c = 1.0\text{ m/s}$ and the domain length $l = 10\text{ mm}$, which are realistic values for layer 2/3 in visual cortex. Then the points A and B are located at a distance $d_A = 2.1\text{ mm}$ and $d_B = 3.8\text{ mm}$ from the stimulus onset location at the origin, respectively. The bar in the plots is 0.83 mm long

We observe the lateral activity propagation starting from the stimulus location in the domain centre. The spreading activity reveals the maxima of axonal connections close to previous experimental findings [17]. To validate the numerical results, we take a closer look at two single spatial locations, denoted A and B in Fig. 6.2 at distance d_A and d_B from the stimulus location in the center. Before stimulus onset, they show the stationary activity constant in time. After stimulus onset, it takes the activity some time to propagate from the stimulus location to these distant points, e.g. the transmission delays $d_A/c = 3.3\text{ ms}$ and $d_B/c = 6.2\text{ ms}$. Figure 6.2 shows that the activity reaches the locations A and B about these times for the first time as expected. This finding validates the numerical algorithm proposed above.

We investigate whether the transmission delay induces oscillatory instabilities in the presence of external input. The following brief numerical study is motivated by previous theoretical studies on breathers [8]. In that study, the authors computed analytically conditions for Hopf-bifurcations from stimulus-induced stable standing bumps in a neural model involving spike rate adaption. The presence of the spike rate adaption permits the occurrence of the Hopf-bifurcation. The corresponding control parameter of these instability studies is the magnitude of the applied external stimulus. In contrast, the present model does not consider spike rate adaption to gain a Hopf-bifurcation, but consider transmission delays. We decreases the axonal transmission speed from large speeds, i.e. increases the transmission delay, to evoke a delay-induced Hopf-bifurcation while keeping the other parameters constant.

Let us assume an anisotropic Gaussian stimulus

$$I(\mathbf{x}, t) = I_0 e^{-\mathbf{x}' \Sigma^{-1} \mathbf{x}/2}$$

with the 2×2 diagonal variance matrix Σ^{-1} with $\Sigma_{ii}^{-1} = 1/\sigma_i^2, i = 1, 2$. Moreover the spatial kernel $K(\mathbf{x})$ represents locally excitatory and laterally inhibitory connections and the transfer function is the Heaviside function $S[V] = H[V - V_{\text{thresh}}]$.

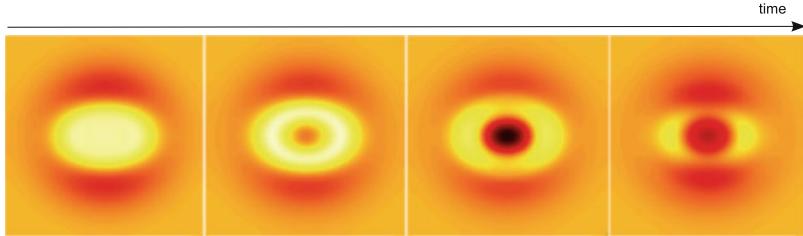


Fig. 6.3 One cycle of a transmission delay-induced breathers evoked by an anisotropic external stimulus. The spatial connectivity function is chosen to $K(r) = 10 \exp(-r/3)/(18\pi) - 14 \exp(-r/7)/(98\pi)$ and the input magnitude and variances are $I_0 = 10$ and $\Sigma_{11}^{-1} = 3$, $\Sigma_{22}^{-1} = 5$, respectively. Other parameters are $c = 100$, $l = 30$, $N = 512$, $\tau = 1$, $\Delta t = 0.05$ and $V_{\text{thresh}} = 0.005$. The initial values are chosen to $V(\mathbf{x}, \theta) = 0$ for the delay interval $-L/\sqrt{2c} \leq \theta \leq 0$

The numerical computation of Eq. (6.1) applying the numerical scheme (6.11) yields delay-induced breathers in two dimensions. Figure 6.3 shows the temporal sequence of a single oscillation cycle. To our best knowledge such delay-induced breathers in two dimensions have not been found before.

6.4 Conclusion

We have motivated briefly a one-dimensional numerical method to integrate a spatial integral involving finite transmission speeds. Moreover we have derived analytically and validated numerically in detail a novel numerical scheme for two-dimensional neural fields involving transmission delay that includes a convolution structure and hence allows the implementation of fast numerical algorithms, such as Fast Fourier transform. We have demonstrated numerically a transmission delay-induced breather [13]. To facilitate future applications of the algorithm, the implementation code for both numerical examples is made available for download.¹ We point out that the method can be easily extended to higher dimensions.

In future research, the transmission delay will play an important role in the understanding of fast activity propagations whose time scales are close to the transmission delay, e.g. in the presence of ultra-fast pulses and/or at spatial nanometer scales. An open source Graphical User interface written in Python for a user-friendly application of the method proposed will be available soon.²

¹<http://www.loria.fr/~rougier/coding/index.html>

²NeuralFieldSimulator: <https://gforge.inria.fr/projects/nfsimulator/>

References

1. Atay, F.M., Hutt, A.: Stability and bifurcations in neural fields with finite propagation speed and general connectivity. *SIAM J. Appl. Math.* **65**(2), 644–666 (2005)
2. Atay, F.M., Hutt, A.: Neural fields with distributed transmission speeds and constant feedback delays. *SIAM J. Appl. Dyn. Syst.* **5**(4), 670–698 (2006)
3. Bressloff, P.C., Cowan, J.D., Golubitsky, M., Thomas, P.J., Wiener, M.C.: What geometric visual hallucinations tell us about the visual cortex. *Neural Comput.* **14**, 473–491 (2002)
4. Bringuier, V., Chavane, F., Glaeser, L., Fregnac, Y.: Horizontal propagation of visual activity in the synaptic integration field of area 17 neurons. *Science* **283**, 695–699 (1999)
5. Cattaneo, C.: A form of heat conduction equation which eliminates the paradox of instantaneous propagation. *Comptes Rendues* **247**, 431–433 (1958)
6. Coombes, S.: Waves, bumps and patterns in neural field theories. *Biol. Cybern.* **93**, 91–108 (2005)
7. Faye, G., Faugeras, O.: Some theoretical and numerical results for delayed neural field equations. *Phys. D* **239**, 561–578 (2010)
8. Folias, S.E., Bressloff, P.C.: Breathers in two-dimensional excitable neural media. *Phys. Rev. Lett.* **95**, 208107 (2005)
9. Huang, X., Troy, W.C., Schiff, S.J., Yang, Q., Ma, H., Laing, C.R., Wu, J.Y.: Spiral waves in disinhibited mammalian neocortex. *J. Neurosci.* **24**(44), 9897–9902 (2004)
10. Hutt, A.: Generalization of the reaction-diffusion, Swift-Hohenberg, and Kuramoto-Sivashinsky equations and effects of finite propagation speeds. *Phys. Rev. E* **75**, 026214 (2007)
11. Hutt, A., Bestehorn, M., Wennekers, T.: Pattern formation in intracortical neuronal fields. *Netw. Comput. Neural Syst.* **14**, 351–368 (2003)
12. Hutt, A., Frank, T.D.: Critical fluctuations and 1/f-activity of neural fields involving transmission delays. *Acta Phys. Pol. A* **108**(6), 1021 (2005)
13. Hutt, A., Rougier, N.: Activity spread and breathers induced by finite transmission speeds in two-dimensional neural fields. *Phys. Rev. E* **82**, R055701 (2010)
14. Klossika, J.J., Gratzke, U., Vicanek, M., Simon, G.: Importance of a finite speed of heat propagation in metals irradiated by femtosecond laser pulses. *Phys. Rev. B* **54**(15), 10277–10279 (1996)
15. Laing, C.R.: Spiral waves in nonlocal equations. *SIAM J. Appl. Dyn. Syst.* **4**(3), 588–606 (2005)
16. Lazzaro, E., Wilhelmsson, H.: Fast heat pulse propagation in hot plasmas. *Phys. Plasmas* **5**(4), 2830–2835 (1998)
17. Lund, J.S., Angelucci, A., Bressloff, P.C.: Anatomical substrates for functional columns in macaque monkey primary visual cortex. *Cereb. Cortex* **13**, 15–24 (2003)
18. Schwabe, L., Obermayer, K., Angelucci, A., Bressloff, P.C.: The role of feedback in shaping the extra-classical receptive field of cortical neurons: a recurrent network model. *J. Neurosci.* **26**, 9117–9126 (2006)
19. Venkov, N.A., Coombes, S., Matthews, P.C.: Dynamic instabilities in scalar neural field equations with space-dependent delays. *Phys. D* **232**, 1–15 (2007)

Chapter 7

Spots: Breathing, Drifting and Scattering in a Neural Field Model

Stephen Coombes, Helmut Schmidt, and Daniele Avitabile

Abstract Two dimensional neural field models with short range excitation and long range inhibition can exhibit localised solutions in the form of spots. Moreover, with the inclusion of a spike frequency adaptation current, these models can also support breathers and travelling spots. In this chapter we show how to analyse the properties of spots in a neural field model with linear spike frequency adaptation. For a Heaviside firing rate function we use an interface description to derive a set of four nonlinear ordinary differential equations to describe the width of a spot, and show how a stationary solution can undergo a Hopf instability leading to a branch of periodic solutions (breathers). For smooth firing rate functions we develop numerical codes for the evolution of the full space-time model and perform a numerical bifurcation analysis of radially symmetric solutions. An amplitude equation for analysing breathing behaviour in the vicinity of the bifurcation point is determined. The condition for a drift instability is also derived and a center manifold reduction is used to describe a slowly moving spot in the vicinity of this bifurcation. This analysis is extended to cover the case of two slowly moving spots, and establishes that these will reflect from each other in a head-on collision.

7.1 Introduction

Given the well known laminar structure of real cortical tissue it is natural to formulate neural field models in two spatial dimensions. For models with short range excitation and long range inhibition these have long been known to support localised solutions in the form of spots (commonly called bumps in one dimensional models). They are of particular interest to the neuroscience community since spatially localised bumps of activity have been linked to working memory (the

S. Coombes (✉) • H. Schmidt • D. Avitabile
School of Mathematical Sciences, University of Nottingham, Nottingham, NG7 2RD, UK.
e-mail: stephen.coombes@nottingham.ac.uk

temporary storage of information within the brain) in prefrontal cortex [17]. Perhaps not surprisingly their initial mathematical study was limited to solutions of one-dimensional models, and see [5, 12] for a review. With a further reduction in model complexity obtained by considering an effective single population model, obviating the need to distinguish between excitatory and inhibitory neuronal populations, Amari [1] showed for a Heaviside firing rate that bump solutions come in pairs, and that it is only the wider of the two that is stable. It was a surprisingly long time before Pinto and Ermentrout [24] demonstrated that a fuller treatment of inhibitory dynamics would allow a dynamic (Hopf) instability that could actually destabilise a wide bump. Blomquist et al. [3], further showed that this could lead to the formation of a stable breathing (spatially localised time periodic) solution. However, it is now known that inhibitory feedback is not the only way to generate dynamic instabilities of localised states, and a number of other physiological mechanisms are also possible. These include localised drive to the tissue [16], threshold accommodation (whereby the firing threshold is itself dynamic, mimicking a refractory mechanism) [6], synaptic depression [19], and spike frequency adaptation [8]. In comparison to their one dimensional counterparts, spots and their instabilities in two dimensions have received far less attention. Notable exceptions to this include the work of Laing and Troy [21] (focusing on numerical bifurcation analysis for smooth firing rates), Folias and Bressloff [15, 16] (focusing on localised drive), Owen et al. [23] (using an Evans function analysis to probe instabilities), and Coombes et al. [8] (using an interface approach). These last three pieces of work all focus on the Heaviside firing rate function.

In this chapter we develop new results for the description of spots in a two dimensional neural field model with spike frequency adaptation with both Heaviside and smooth firing rate choices. The techniques we develop are quite generic and may also be adapted to treat the other physiological mechanisms mentioned above for the generation of dynamic spot instabilities. We focus on a planar single population model that can be written as an integro-differential equation of the form

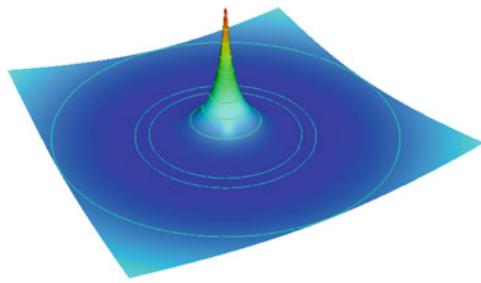
$$\frac{1}{\alpha} \frac{\partial u(\mathbf{r}, t)}{\partial t} = -u(\mathbf{r}, t) + \int_{\mathbb{R}^2} d\mathbf{r}' w(|\mathbf{r} - \mathbf{r}'|) f(u(\mathbf{r}', t)) - ga(\mathbf{r}, t), \quad (7.1)$$

$$\frac{\partial a(\mathbf{r}, t)}{\partial t} = -a(\mathbf{r}, t) + u(\mathbf{r}, t), \quad (7.2)$$

where $\mathbf{r} = (x_1, x_2) \in \mathbb{R}^2$ and $t \in \mathbb{R}^+$. Here the variable u represents synaptic activity and the kernel w represents anatomical connectivity. In real cortical tissues there are an abundance of metabolic processes whose combined effect is to modulate neuronal response. It is convenient to think of these processes in terms of local feedback mechanisms that modulate synaptic currents, described by the field a . Here, $g \in \mathbb{R}$ is the strength of the negative feedback. We shall take the firing rate to be a sigmoidal function, such as

$$f(u) = \frac{1}{1 + e^{-\beta(u-h)}}, \quad (7.3)$$

Fig. 7.1 A plot of the anatomical connectivity function describing short range excitation and long range inhibition. The form of $w(\mathbf{r})$, as expressed in (7.4), for $A_1 = 1 = \alpha_1$, $A_2 = -3/4$, and $\alpha_2 = 1/4$, generates a two dimensional wizard hat function



where $\beta > 0$ controls the steepness of the sigmoid around the threshold value h . Throughout the rest of this paper we shall work with the radially symmetric choice $w(\mathbf{r}) = w(r)$, with $r = |\mathbf{r}|$. To allow for some explicit calculations (though many of the results we develop do not require such a choice), we shall use the representation

$$w(r) = \sum_{i=1}^N A_i K_0(\alpha_i r), \quad A_i \in \mathbb{R}, \quad \alpha_i > 0, \quad (7.4)$$

where $K_\nu(x)$ is the modified Bessel function of the second kind, of order ν . For an appropriate combination of coefficients A_i and α_i this can generate an anatomical connectivity describing short range excitation and long range inhibition, with a wizard hat shape, as shown in Fig. 7.1.

In Sect. 7.2 we focus on a Heaviside firing rate and show how to extend the Amari interface approach to treat the spike frequency adaptation term. We develop a reduced description of a spot in terms of a set of four coupled nonlinear ordinary differential equations (ODEs). We solve these numerically, to find a narrow and wide branch of spot solutions that annihilate in a saddle-node bifurcation (with increasing threshold). The branch of wide spots is found to support a Hopf instability to a stable breathing solution. We move away from the Heaviside case in Sect. 7.3, and develop an equivalent partial differential equation (PDE) model that allows for straight-forward numerical implementation. We use this to probe radially symmetric solutions for models with sigmoidal firing rates, and not only confirm the results of our Heaviside analysis but show how results vary as one moves away from the limit of steep sigmoids. In Sect. 7.4 we develop an amplitude equation for analysing breathing behaviour (for a smooth firing rate function) in the vicinity of the bifurcation point. The condition for a drift instability, which describes the transition of a stationary spot to a (non-circular) travelling spot, is derived in Sect. 7.5. In Sect. 7.6 a center manifold reduction is used to describe a slowly moving spot in the vicinity of this bifurcation, and extended to cover the case of two slowly moving spots in Sect. 7.7. Interestingly the coupled ODE model for the spot pair can be analysed to show that these will reflect from each other in a head-on collision. Finally, in Sect. 7.8 we discuss natural extensions of the work in this Chapter.

7.2 Heaviside Firing Rate and Interface Dynamics

In the limit $\beta \rightarrow \infty$ the firing rate function (7.3) approximates a Heaviside function $H(u - h)$, and it is possible to explicitly construct localised states. Here we show how to extend the standard Amari interface approach to treat linear spike frequency adaptation. For simplicity we shall focus on radially symmetric solutions. A more general framework for describing the evolution of spreading interfaces that lack such a symmetry has recently been developed in [8]. We shall not pursue this here.

First let us rewrite the pair of Eqs. (7.1) and (7.2) in the form of a second order differential system:

$$u_{tt} + (1 + \alpha)u_t + \alpha(1 + g)u = \alpha(\psi + \psi_t), \quad (7.5)$$

where $\psi(\mathbf{r}, t) = \int_{\mathcal{B}(\mathbf{r}', t)} d\mathbf{r}' w(|\mathbf{r} - \mathbf{r}'|)$, and $\mathcal{B}(\mathbf{r}, t) = \{\mathbf{r}' | u(\mathbf{r}', t) \geq h\}$. For radially symmetric spot solutions of radius $R(t)$ that intersect the threshold exactly once (so that the active region is a single, simply connected region) we have that $(u(\mathbf{r}, t), \psi(\mathbf{r}, t)) = (u(r, t), \psi(r, t))$ with

$$\psi(r, t) = \int_0^{2\pi} d\theta \int_0^{R(t)} w\left(\sqrt{r^2 + r'^2 - 2rr' \cos \theta}\right) r' dr'. \quad (7.6)$$

Here $R(t)$ is defined by the level set condition $u(R(t), t) = h$. Differentiating this with respect to time gives an equation for the velocity of the spot interface in the form

$$\frac{dR}{dt} = -\left. \frac{\partial u(r, t)/\partial t}{\partial u(r, t)/\partial r} \right|_{r=R}, \quad (7.7)$$

Using (7.5) we may derive ODEs for $v = \partial u(r, t)/\partial t|_{r=R}$ and $z = \partial u(r, t)/\partial r|_{r=R}$ as

$$\frac{dv}{dt} + (1 + \alpha)v + \alpha(1 + g)h = \alpha [\psi + \psi_t]|_{r=R}, \quad (7.8)$$

$$\frac{d^2z}{dt^2} + (1 + \alpha)\frac{dz}{dt} + \alpha(1 + g)z = \alpha \left. \frac{\partial}{\partial r} [\psi + \psi_t] \right|_{r=R}, \quad (7.9)$$

where we have assumed that R is slowly evolving so that $dz/dt = dR/dt z_R + z_t \approx z_t$. Hence, we may generate a system of four nonlinear ODEs for (R, v, z, y) to describe the evolution of the (radially symmetric) spot:

$$\frac{dR}{dt} = -\frac{v}{z}, \quad (7.10)$$

$$\frac{dv}{dt} = -(1 + \alpha)v - \alpha(1 + g)h + \alpha\Psi(R) - \alpha R v \Phi(R)/z, \quad (7.11)$$

$$\frac{dz}{dt} = y, \quad (7.12)$$

$$\frac{dy}{dt} = -(1 + \alpha)y - \alpha(1 + g)z + \alpha\Upsilon(R) - \alpha R^2 v \Omega(R)/z, \quad (7.13)$$

where

$$\Psi(R) = \int_0^{2\pi} d\theta \int_0^R w(\sqrt{R^2 + r^2 - 2Rr \cos \theta}) r dr, \quad (7.14)$$

$$\Phi(R) = \int_0^{2\pi} d\theta w(\mathcal{R}(\theta)), \quad \mathcal{R}(\theta) = R\sqrt{2(1 - \cos \theta)}, \quad (7.15)$$

$$\Upsilon(R) = \int_0^{2\pi} d\theta \int_0^R \frac{w'(\sqrt{R^2 + r^2 - 2Rr \cos \theta})}{\sqrt{R^2 + r^2 - 2Rr \cos \theta}} (R - r \cos \theta) r dr, \quad (7.16)$$

$$\Omega(R) = \int_0^{2\pi} d\theta \frac{w'(\mathcal{R}(\theta))}{\mathcal{R}(\theta)} (1 - \cos \theta). \quad (7.17)$$

For the anatomical connectivity function (7.4) then we have explicitly (using the summation properties of Bessel functions and Graf's formula) that

$$\Psi(R) = 2\pi \sum_{i=1}^N A_i \left\{ \frac{1}{\alpha_i^2} - \frac{R}{\alpha_i} K_1(\alpha_i R) I_0(\alpha_i R) \right\}, \quad (7.18)$$

$$\Phi(R) = 2\pi \sum_{i=1}^N A_i K_0(\alpha_i R) I_0(\alpha_i R), \quad (7.19)$$

$$\Upsilon(R) = \Psi'(R) - R\Phi(R) = -2\pi R \sum_{i=1}^N A_i K_1(\alpha_i R) I_1(\alpha_i R), \quad (7.20)$$

$$\Omega(R) = \Phi'(R)/(2R) = 2\pi \sum_{i=1}^N A_i \left\{ \frac{1}{2R^2} - \frac{\alpha_i}{R} K_1(\alpha_i R) I_0(\alpha_i R) \right\}. \quad (7.21)$$

Here $I_\nu(x)$ is the modified Bessel function of the first kind.

Steady states of (7.10)–(7.13) are given by $(R, v, z, y) = (R, 0, \Upsilon(R)/(1 + g), 0)$, where R is a solution of $\Psi(R) = (1 + g)h$. We can reconstruct the spatial profile of the stationary spot from knowledge of its radius by using $q(r) = \psi(r)/(1 + g)$, where $\psi(r)$ is given by (7.6) with $R(t) = R$. This can be explicitly calculated as

$$\psi(r) = 2\pi R \sum_{i=1}^N A_i L_i(r), \quad (7.22)$$

where

$$L_i(r) = \begin{cases} \frac{1}{\alpha_i} I_1(\alpha_i R) K_0(\alpha_i r) & r \geq R \\ \frac{1}{\alpha_i^2 R} - \frac{1}{\alpha_i} I_0(\alpha_i r) K_1(\alpha_i R) & r < R \end{cases}. \quad (7.23)$$

Here we restrict our attention to only those values of R that satisfy our original hypothesis, namely that the spot is a simply-connected active region.

The Jacobian of the system (7.10)–(7.13) around the steady state has one pair of eigenvalues given by

$$\lambda = \frac{-(1 + \alpha) \pm \sqrt{(1 - \alpha)^2 - 4\alpha g}}{2}, \quad (7.24)$$

and another pair that satisfy $\lambda^2 - \lambda J + K = 0$, where

$$J = -(1 + \alpha) - \alpha(1 + g)R \frac{\Phi(R)}{\Upsilon(R)}, \quad K = \alpha(1 + g) \frac{\Psi'(R)}{\Upsilon(R)}. \quad (7.25)$$

A Hopf bifurcation is possible when $J = 0$ and $K > 0$. Using (7.20) we see that this can happen when

$$R \frac{\Phi(R)}{\Upsilon(R)} = -\frac{(1 + \alpha)}{\alpha(1 + g)}, \quad \alpha g > 1. \quad (7.26)$$

which recovers a result in [8] obtained via a different method. The frequency of oscillation at the Hopf bifurcation is given by $\sqrt{\alpha g - 1}$.

The system of ODEs (7.10)–(7.13) is solved numerically with AUTO-07P [9] to generate the bifurcation diagram in Fig. 7.2 for a wizard hat function given by (7.4) with $N = 2$, $A_1 = (2\pi)^{-1}$, $A_2 = -(2\pi\gamma)^{-1}$, $\alpha_1 = 1$, $\alpha_2 = \mu$, with $\gamma, \mu > 0$. As expected we see a branch of wide spots and narrow spots that annihilate in a saddle-node bifurcation with increasing h . The inclusion of spike-frequency adaptation now allows for a pair of Hopf bifurcations on the wide branch of solutions, marked as HB_1 and HB_2 , determined by Eq. (7.26). Further numerical exploration shows that $HB_{1,2}$ give rise to a branch of periodic orbits, describing radially-symmetric breathers, whose period increases to infinity with decreasing h . This gives rise to homoclinic orbits, and we denote the associated homoclinic bifurcation points by $M_{1,2}$. Only the branch of breathers emanating from HB_2 is stable and it undergoes a sequence of period-doubling bifurcations in a small region of parameter space, as shown in the inset of Fig. 7.2, where h varies between 0.03609 and 0.03611.

In Fig. 7.3 we repeat the continuation for $\alpha = 1.25$. As expected, the branch of spots is the same as Fig. 7.2, but its stability properties change: the Hopf bifurcations $HB_{1,2}$ shift along the branch and are now connected in parameter space.

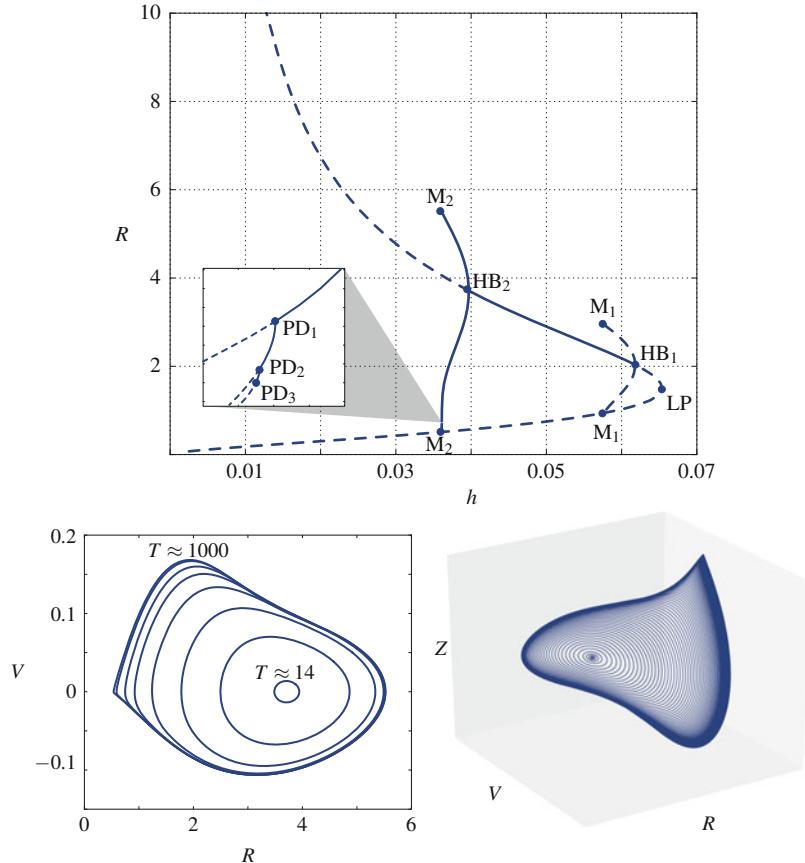


Fig. 7.2 Bifurcation diagram for spot solutions with a Heaviside firing rate function. *Solid (dashed) lines* are stable (unstable). Branches of periodic orbits originate from Hopf bifurcations of stationary spots. *Top:* branches of periodic solutions terminate at global homoclinic bifurcations M₁ and M₂. Solutions emanating from HB₂ are initially stable and destabilise close to M₂, where the branch undergoes a sequence of period-doubling bifurcations in a small region of parameter space (the *inset* shows h between 0.03609 and 0.03611). *Bottom:* stable periodic solutions between HB₂ and M₂, with increasingly high periods, projected on the (V, R) -plane and the (V, R, Z) -space. Parameters are $\gamma = 4$, $\mu = 0.5$, $g = 1$, and $\alpha = 1.2$

Moreover, stable breathers emanating from HB₂ destabilise at a limit point rather than a global bifurcation. On the same plot, we show points along the branch where stationary spots become unstable with respect to azimuthal instabilities with D_m symmetry ($m = 2, \dots, 8$). Such critical points satisfy the equation (see [8] and also Chap. 3)

$$\frac{R}{|\Psi'(R)|} \int_0^{2\pi} d\theta \cos(m\theta) w(\mathcal{R}(\theta)) = \frac{\sum_{i=1}^N A_i K_m(\alpha_i R) I_m(\alpha_i R)}{\sum_{i=1}^N A_i K_1(\alpha_i R) I_1(\alpha_i R)} = 1, \quad (7.27)$$

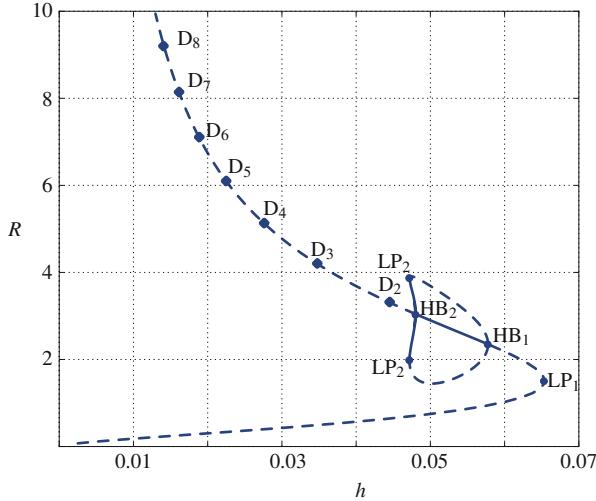


Fig. 7.3 Bifurcation diagram for spot solutions with a Heaviside firing rate function. Parameters as in Fig. 7.2 with the exception of α , which is set to 1.25. The branch of steady states coincides with the one in Fig. 7.2, as expected, but its stability depends on α : in this case HB_1 and HB_2 are connected in parameter space and stable breathers destabilise at a limit point. We also show points along the branch where spots become unstable to planar perturbations with D_2, \dots, D_8 symmetry

for $m = 2, 3, \dots$, where $\Psi(R)$ and $\mathcal{R}(\theta)$ are defined in (7.14) and (7.15). We point out that, following the branch for $\alpha = 1.25$ from bottom to top, the Hopf bifurcation HB_2 occurs before the D_2 instability, while the situation is reversed for $\alpha = 1.2$ (not shown). In Sect. 7.5 we further show that a solution with a drift instability, leading to a travelling spot, can occur as g increases through $1/\alpha$ (and note that the condition $\alpha g > 1$ is necessary for a breathing instability). A further weakly nonlinear analysis would be necessary to understand the competition between drifting and breathing at $g = 1/\alpha$.

The proximity of the limit point LP_1 to the Hopf bifurcation HB_1 (similarly to what was found for the case $\alpha = 1.2$ in Fig. 7.2, where we also have the global bifurcation M_1) suggests the possibility of a Bogdanov-Takens codimension 2 bifurcation. This scenario is present in a similar model (with localised drive) [15] and is confirmed by the two-parameter continuations plotted in Fig. 7.4.

7.3 Equivalent PDE Model and Numerical Bifurcation Analysis

In the previous section we derived a set of ODEs describing localised radially-symmetric solutions to (7.1) and (7.2), in the case when the firing rate f is a Heaviside function. It is interesting to compute similar solutions when f is a smooth

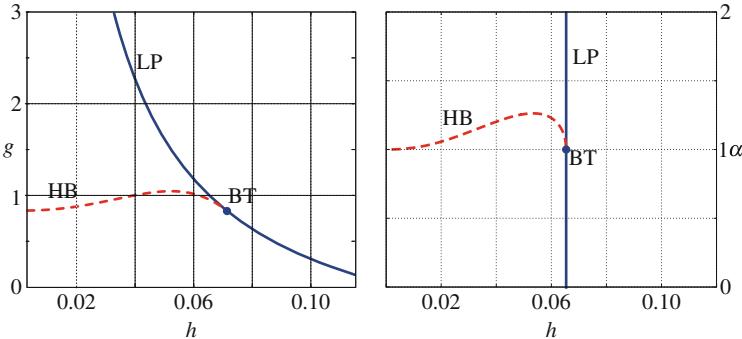


Fig. 7.4 Two-parameter continuations of the saddle-node bifurcation (*solid*), and Hopf bifurcations (*dashed*) of the stationary spots. Figure 7.2 corresponds to slices at $g = 1$ (in the *left panel*) and $\alpha = 1.2$ (in the *right panel*). Figure 7.3 corresponds to a slice $\alpha = 1.25$ (in the *right panel*). The curves meet at a Bogdanov-Takens bifurcation. Curves of homoclinic bifurcations $M_{1,2}$ are not shown

sigmoidal function, such as the one in (7.3). In this section, we follow ideas in [7, 13, 21] (and see also Chap. 5) that allow us to reformulate the nonlocal model defined by (7.1), (7.2), and (7.4), as a local model more suitable for direct numerical simulation and numerical continuation.

We begin by writing time-independent solutions of (7.1) and (7.2) as $(u(\mathbf{r}, t), a(\mathbf{r}, t)) = (q(\mathbf{r}), q(\mathbf{r}))$, where

$$q(\mathbf{r}) = \frac{1}{1+g} \int_{\mathbb{R}^2} d\mathbf{r}' w(|\mathbf{r} - \mathbf{r}'|) f(q(\mathbf{r}')). \quad (7.28)$$

A spot is a radially-symmetric solution of (7.28) such that $q(\mathbf{r}) = q(r)$. We shall denote the (vector) spot solution by $S(r) = (q(r), q(r))$. Other localised solutions with dihedral symmetries of the regular polygon are also expected, and may arise via bifurcations of spots. This has been established for the case of a Heaviside firing rate [8], though we will not pursue this further here.

By re-arranging and taking a two dimensional Fourier transform of (7.28) we obtain

$$(1+g)\hat{q}(k) = \hat{w}(k) \widehat{f \circ q}(k), \quad \hat{w}(k) = \sum_{i=1}^N \frac{2\pi A_i}{\alpha_i^2 + k^2}. \quad (7.29)$$

By taking the inverse Fourier transform of (7.29), we obtain a nonlinear PDE of the form

$$0 = -(1+g)\mathcal{L}_1 q + \mathcal{L}_2 f(q), \quad (7.30)$$

where \mathcal{L}_1 and \mathcal{L}_2 are linear differential operators containing even spatial derivatives up to order $2N$ and $2N - 2$ respectively (using the result that the inverse Fourier transform of k^2 is $-\nabla^2$). The PDE formulation (7.30) is formally equivalent to (7.28).

In a similar way, it is possible to study the stability of stationary states with respect to radial perturbations. System (7.1) and (7.2) is written as a PDE

$$\begin{aligned}\mathcal{L}_1(\alpha^{-1}\partial_t u + u + ga) &= \mathcal{L}_2 f(u), \\ a_t &= u - a.\end{aligned}\tag{7.31}$$

Linear stability to radial perturbations can be inferred by posing $u(r, t) = q(r) + Q(r)e^{\lambda t}$, $a(r, t) = q(r) + A(r)e^{\lambda t}$ and linearising $\mathcal{L}_2 f(u)$ around $q(r)$, giving the generalised eigenvalue problem

$$\rho \mathcal{L}_1 Q = \mathcal{L}_2(f'(q)Q), \quad \rho = 1 + \frac{\lambda}{\alpha} + \frac{g}{1 + \lambda}.\tag{7.32}$$

We can then compute stationary solutions of system (7.1) and (7.2) and their spectra by prescribing suitable boundary conditions for (7.30) and (7.32) and discretising the operators \mathcal{L}_1 and \mathcal{L}_2 : for steady states, we use Newton's iterations to solve (7.30); a few eigenvalues are then computed (without inverting \mathcal{L}_1) applying Arnoldi iterations to (7.32). In passing, we note that it is also possible to study stability of radial solutions with respect to perturbations with dihedral symmetry, albeit this requires an amendment of the generalised eigenvalue problem (7.32).

Let us consider, for illustrative purposes, the wizard hat kernel used in Sect. 7.2 so that:

$$\mathcal{L}_1 = \Gamma_1 - \Gamma_2 \Delta_r + \Delta_r^2, \quad \mathcal{L}_2 = \Gamma_3 - \Gamma_4 \Delta_r,\tag{7.33}$$

where $\Delta_r = \partial_{rr} + r^{-1}\partial_r$ is the Laplacian operator expressed in radial coordinates and

$$\Gamma_1 = \mu^2, \quad \Gamma_2 = 1 + \mu^2, \quad \Gamma_3 = \mu^2 - \frac{1}{\gamma}, \quad \Gamma_4 = 1 - \frac{1}{\gamma}.\tag{7.34}$$

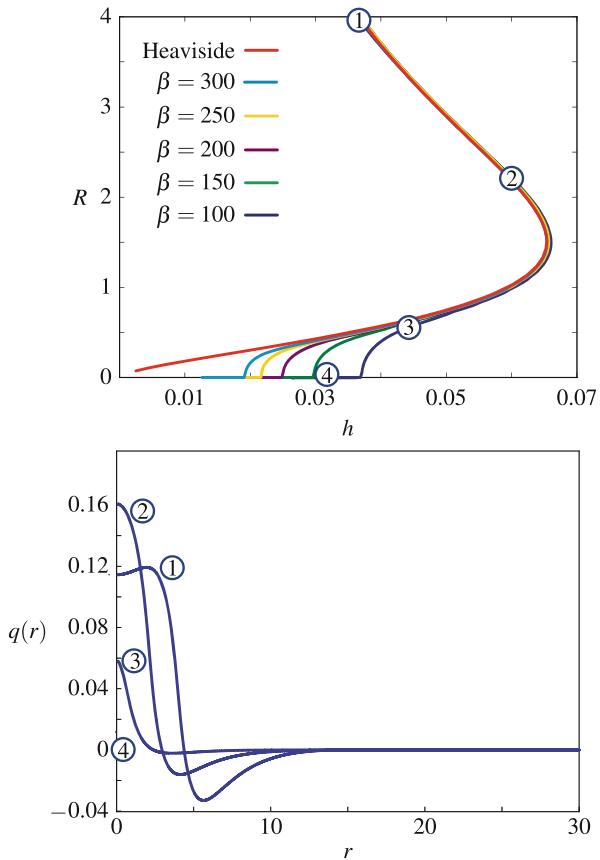
We then solve the nonlinear boundary-value problem

$$0 = -(1 + g)(\Gamma_1 - \Gamma_2 \Delta_r + \Delta_r^2)q + (\Gamma_3 - \Gamma_4 \Delta_r)f(q; h, \beta), \quad r \in (0, L)\tag{7.35}$$

$$0 = \partial_r q(0) = \partial_{rrr} q(0), \quad 0 = \partial_r q(L) = \partial_{rrr} q(L),\tag{7.36}$$

whose solutions can be continued in any of the control parameters μ, γ, h, β and g . We point out that, while equilibria of (7.1) and (7.2) do not depend upon α , this parameter influences their stability. The boundary-value problem (7.35) and (7.36) features no-flux boundary conditions at $r = L$, as in [20], but other choices are also possible: alternatively, one could pose $q(L) = \partial_{rr} q(L) = 0$.

Fig. 7.5 Localised spots found as solutions to the boundary-value problem (7.35). *Top:* branches of spots for various values of the sigmoidal firing-rate steepness, β ; on the vertical axis we plot a measure of the active region, that is, R is the largest number such that $q(R) = h$; the Heaviside branch is obtained by continuing equilibria of the system (7.10)–(7.13). *Bottom:* representative solutions along the branch $\beta = 100$. Control parameters: $\gamma = 4$, $\mu = 0.5$, $g = 1$, $L = 30$



We discretised (7.35) via second order centred finite differences with 3,200 grid points, on a domain with $L = 30$ and implemented a numerical continuation code written in MATLAB. Stability computations are performed using MATLAB's in-built function `eigs`. In Fig. 7.5 we show a set of h -continuations for several values of the parameter β . These bifurcation diagrams are plotted in terms of the largest coordinate $R \in [0, L]$ for which $q(R) = h$; if $q(r) < h$ for all $r \in [0, L]$, as in pattern 4 of Fig. 7.5, we set $R = 0$. This allows us to compare steady states of (7.1) and (7.2), for increasingly high values of the sigmoidal firing rate steepness parameter β , with steady states of (7.10)–(7.13), which correspond to the Heaviside limit. Our numerical results show that the PDE formulation recovers the Heaviside case when $\beta \rightarrow \infty$, as confirmed also from the stability computations, contained in Fig. 7.6. Small-amplitude localised spots are unstable (bottom part of the branches) for our choice of control parameters, namely $\gamma = 4$, $\mu = 0.5$, $g = 1$, $L = 30$ and $\alpha = 1.2$. The branch features a saddle-node bifurcation and two Hopf bifurcations, which delimit a region of stable stationary spots. We also plot the spectra corresponding to the Hopf bifurcations, noting that the

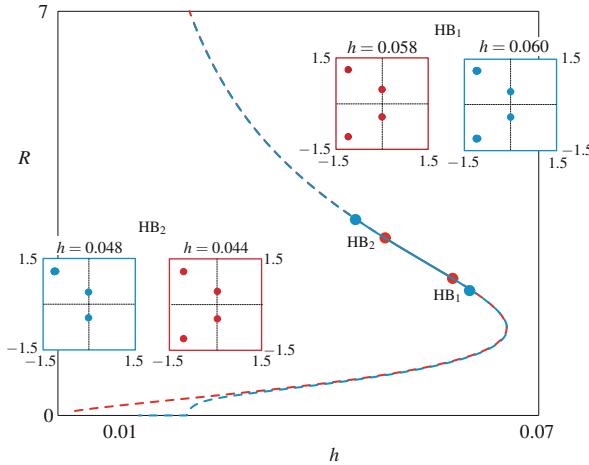


Fig. 7.6 Stability of radial spots for two solution branches of Fig. 7.5, namely Heaviside (red) and $\beta = 300$ (light blue). The bottom part of the branch is unstable and restabilises at a Hopf bifurcation (HB₁). The branch becomes unstable at a second Hopf bifurcation (HB₂). At the Hopf bifurcation points, we computed the corresponding eigenvalues for the Heaviside case (7.10)–(7.13) (red panels) and the first 10 eigenvalues obtained via (7.32) (light blue panels). Parameters: $\alpha = 1.25$, and other parameters as in Fig. 7.5

PDE formulation reproduces correctly the Hopf eigenvalues: other PDE eigenvalues (8, in the computations of Fig. 7.6) are clustered around the eigenvalues found using the interface approach (Heaviside firing rate) in Sect. 7.2. For steep sigmoids, the position of HB_{1,2} changes slightly with respect to the Heaviside case.

Close to the HB₂, we expect to find a branch of stable breathers, which can be found by direct numerical simulation of (7.31) with radially-symmetric operators (7.33). The time stepping is done without providing an explicit inverse of the operator \mathcal{L}_1 , but recasting (7.31) as

$$\begin{bmatrix} \alpha^{-1} \mathcal{L}_1 & 0 \\ 0 & 1 \end{bmatrix} \partial_t \begin{bmatrix} u \\ a \end{bmatrix} = \begin{bmatrix} -\mathcal{L}_1 & -g\mathcal{L}_1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} u \\ a \end{bmatrix} + \begin{bmatrix} \mathcal{L}_2 f(u) \\ 0 \end{bmatrix}, \quad (7.37)$$

discretising the left-hand side with a block-diagonal, sparse, time-independent mass matrix and then employing MATLAB's `ode23s` solver with `RelTol=1e-3` and `AbsTol=1e-6`. In the simulations of Fig. 7.7, we started from a stationary spot on the stable branch, decreased the value of h quasi-statically every 2,000 time units (corresponding to at least 100 oscillation cycles of the breathers) and found stable solutions with various amplitudes, spatial extensions and oscillation periods.

We notice that the breathers found with steep sigmoids have smaller amplitudes with respect to the ones found for the ODE model (7.10)–(7.13) (Fig. 7.3). Furthermore, the latter disappear at a limit point, while the former persist for much smaller values of the threshold h . This is to be expected as the ODE model is only valid under the assumption of a slowly varying R . In general we find that

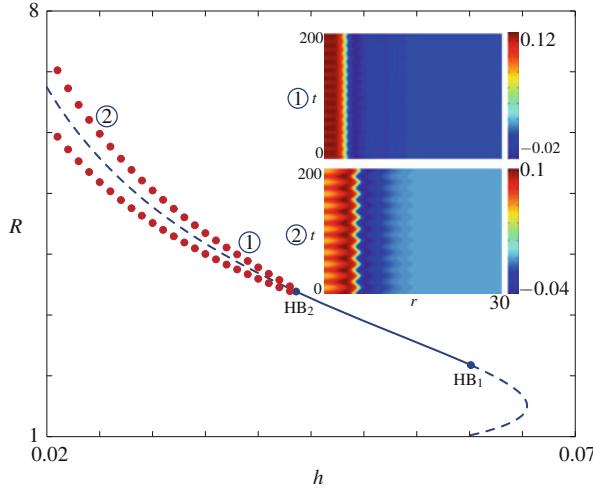


Fig. 7.7 Breathers originating from HB₂ for $\beta = 300$. Direct numerical simulations of (7.31) with radially-symmetric operators (7.33) show the existence of stable breathers, which are identified with the minimum and maximum of $R(t)$ during an oscillation period (red dots). As the threshold value h decreases, we find breathers with larger amplitudes, wider spatial extensions and longer periods. Numerical parameters as in Fig. 7.6

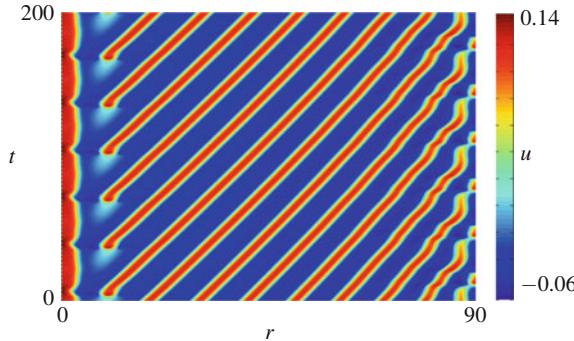


Fig. 7.8 Radial defect found by direct numerical simulation of system (7.31) with linear operators $\mathcal{L}_1, \mathcal{L}_2$ given by (7.33). This solution is found on a large domain, $L = 90$, for $h = 0.034$ and shallow sigmoidal firing rate, $\beta = 50$. Other parameters as in Fig. 7.6

the existence and stability of breathers depends sensitively on the steepness of the sigmoid: for instance, setting $\beta = 30$ and repeating the experiment of Fig. 7.7 leads to a trivial spatially-homogeneous equilibrium; on the other hand, increasing the sigmoid steepness to $\beta = 50$ leads to the formation of radial defects, that is, solutions in which a stable breather core emits periodically a travelling wave (see Fig. 7.8). Since our patterns are radially symmetric, these solutions correspond to a radial source emitting periodically travelling rings. Similar coherent structures were previously analysed in [25], continued in parameter space in [2] and their existence

was also reported in a nonlocal model with linear adaptation and localised stationary input by Folias and Bressloff [15].

7.4 Amplitude Equations for Breathing

Here we focus on a sigmoidal firing rate, such as given by (7.3), and show how to develop a description of the amplitude of a breathing solution just beyond the point of a breathing instability. We closely follow the ideas in [18], which were originally developed for the study of a three component reaction diffusion equation. A related analysis for one dimensional neural field models (and Heaviside firing rate function) can be found in [14] (and see also Chap. 4).

It is convenient to introduce the vector $X = (u, a)$ and the vector field \mathcal{F} , and write the model (7.1) and (7.2) in the form

$$X_t = \mathcal{F}(X; \sigma), \quad (7.38)$$

where σ represents a vector of system parameters, e.g. $\sigma = (\alpha, g, \beta, h)$, and

$$\mathcal{F}(X; \sigma) = MX + \alpha w \otimes F(X), \quad (7.39)$$

with

$$M = \begin{bmatrix} -\alpha & -\alpha g \\ 1 & -1 \end{bmatrix}, \quad F(X) = \begin{bmatrix} f(u) \\ 0 \end{bmatrix}. \quad (7.40)$$

Here the symbol \otimes denotes a two-dimensional spatial convolution.

Linearising about (7.28) gives

$$\partial_t V = \mathcal{L}V, \quad (7.41)$$

where

$$\mathcal{L} = M + \alpha w \otimes f'(q) \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}. \quad (7.42)$$

For separable solutions of the form $V(\mathbf{r}, t) = \phi(\mathbf{r})e^{\lambda t}$, we generate an eigenvalue problem:

$$\mathcal{L}\phi = \lambda\phi. \quad (7.43)$$

Since the operator \mathcal{L} is in general not self-adjoint, then the eigenvalues, λ , and eigenfunctions, $\phi(\mathbf{r})$, may be complex. Because of translation and rotation invariance we expect the existence of an eigenvalue with $\lambda = 0$. The stationary spot is stable if all other eigenvalues have negative real part. For a breathing instability we are interested in the case that a pair of complex conjugate eigenvalues pass through

the imaginary axis under variation of some parameter, namely $\mathcal{L}\phi = \pm i\omega\phi$. Moreover, we shall focus on the case of radially symmetric breathing motions so that $\phi(\mathbf{r}) = \phi(r)$. We shall assume that a stationary spot S exists for a set of parameters $\sigma = \sigma_c$ for which $\mathcal{F}(S; \sigma_c) = 0$. We now introduce a small parameter $\eta \in \mathbb{R}$ and write $\sigma = \sigma_c + \eta(0, 0, 1, 0, \dots, 0)$, where the non-zero entry is associated to the system parameter that we wish to vary (and we only consider co-dimension 1 bifurcations here). In this case

$$X_t = \mathcal{F}(X; \sigma_c) + \eta\gamma(X), \quad (7.44)$$

where $\eta\gamma(X) = \mathcal{F}(X; \sigma_c + \eta) - \mathcal{F}(X; \sigma_c)$. For small η we expect to find a solution of the form

$$X(\mathbf{r}, t) = S(r) + A(t)e^{i\omega t}(\phi(r) + \eta\delta\phi(r)) + \text{cc} + \chi(\mathbf{r}, t), \quad (7.45)$$

where cc denotes the complex conjugate of the previous term. Here $A(t)$ is a slowly evolving amplitude ($A_t \sim \eta A$), $\eta\delta\phi$ is an unknown perturbation, and χ represents a decaying function in an orthogonal space to $\text{span}\{\phi\}$. Substituting into (7.44) and equating terms in $e^{i\omega t}$ gives an equation that relates A and $\delta\phi$:

$$\eta A (\mathcal{L} - i\omega) \delta\phi = \dot{A}\phi - \eta A\gamma'\phi - \frac{1}{2} A^2 \bar{A} \mathcal{F}''' \phi^2 \bar{\phi}. \quad (7.46)$$

Here the multiplication of vectors $\phi^2 \bar{\phi}$ is interpreted component wise, $\bar{\phi}$ denotes the complex conjugate of ϕ , γ' is the Fréchet derivative of γ with respect to X evaluated at $X = S$ and $\mathcal{F}^{(n)}$ is the n th Fréchet derivative of \mathcal{F} with respect to X evaluated at $X = S$:

$$\mathcal{F}^{(n)} = \alpha w \otimes \frac{d^n}{dq^n} f(q) \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}. \quad (7.47)$$

According to the Fredholm alternative, Eq. (7.46) is solvable as long as the right hand side is orthogonal to the kernel of the operator $\mathcal{L} - i\omega$. It is now convenient to introduce the operator $\mathcal{L}^\dagger + i\omega$, where \mathcal{L}^\dagger is adjoint to, and has the same symmetry properties as \mathcal{L} . We shall denote the corresponding zero eigenfunction of $\mathcal{L}^\dagger + i\omega$ by ϕ^\dagger . We define the inner product of two vector functions a and b as

$$\langle a | b \rangle = \int_{\mathbb{R}^2} d\mathbf{r} \bar{a}(\mathbf{r}) \cdot b(\mathbf{r}), \quad (7.48)$$

where the dot \cdot denotes a vector dot product. Projecting (7.46) onto ϕ^\dagger and using the fact that $\langle \phi^\dagger | (\mathcal{L} - i\omega)\delta\phi \rangle = 0$ we obtain an equation for the evolution of the complex amplitude $A(t)$:

$$\dot{A} = A(M_1|A|^2 + M_2\eta), \quad (7.49)$$

where

$$M_1 = \frac{\langle \phi^\dagger | \mathcal{F}''' \phi^2 \bar{\phi} \rangle}{2 \langle \phi^\dagger | \phi \rangle}, \quad M_2 = \frac{\langle \phi^\dagger | \gamma' \phi \rangle}{\langle \phi^\dagger | \phi \rangle}. \quad (7.50)$$

This analysis provides the basis for understanding the bifurcation diagrams obtained numerically in Sect. 7.3 and their criticality. Equation (7.49) has a trivial solution $A = 0$ describing a stationary spot, which for $\text{Re } M_2 < 0$ is a stable focus. If $\text{Re } M_2$ increase through zero the spot becomes unstable. For $\text{Re } M_2 > 0$ there is a non-trivial periodic solution $A(t) = R e^{i\Omega t}$, where

$$R = \sqrt{-\eta \frac{\text{Re } M_2}{\text{Re } M_1}}, \quad \Omega = \eta \text{Im } M_2 + R^2 \text{Im } M_1. \quad (7.51)$$

This non-trivial solution, describing a breather, is stable for $\text{Re } M_1 < 0$ (supercritical bifurcation) and unstable for $\text{Re } M_1 > 0$ (subcritical bifurcation). Considering a variation in the parameter g around some critical value g_c then

$$\gamma' = \left. \frac{\partial M}{\partial g} \right|_{g=g_c} = \begin{bmatrix} 0 & -\alpha \\ 0 & 0 \end{bmatrix}. \quad (7.52)$$

In the Appendix we show that ϕ^\dagger can be written in closed form as a linear transformation of ϕ :

$$\phi^\dagger = f'(q) \begin{bmatrix} 1/(\alpha g) & 0 \\ 0 & -1 \end{bmatrix} \bar{\phi}. \quad (7.53)$$

This means that we only have to numerically solve $\mathcal{L}\phi = i\omega\phi$ to compute the coefficients M_1 and M_2 in (7.49).

As well as a breathing bifurcation it is also possible for a stationary spot to undergo an instability to a travelling spot, via a drift instability. This has been recently studied for the case of a Heaviside firing rate [8]. Next we show how to analyse the case of a smooth firing rate.

7.5 Drifting

Here we adapt an argument in [23] to show how a spot can undergo an instability to a drifting pulse as g is increased through $1/\alpha$. From invariance of the full system (under rotation and translation) there exist Goldstone modes $\phi_i = \partial S(r)/\partial x_i$, $i = 1, 2$, such that

$$\mathcal{L}\phi_i = 0. \quad (7.54)$$

One of the possible destabilisations of the spot S occurs when one of the other modes exactly coincides with ϕ_i under parameter variation. Because of this parameter degeneracy a generalised eigenfunction ψ_i of \mathcal{L} appears:

$$\mathcal{L}\psi_i = -\phi_i. \quad (7.55)$$

The solvability condition for (7.55) leads to an equation defining the bifurcation point in the form

$$\langle \phi_i^\dagger | \phi_i \rangle = 0, \quad (7.56)$$

where ϕ_i^\dagger is the eigenfunction of the operator \mathcal{L}^\dagger with zero eigenvalue. More clearly we write $\mathcal{L}^\dagger \phi_i^\dagger = 0$ and $\mathcal{L}^\dagger \psi_i^\dagger = -\phi_i^\dagger$, normalised by $\langle \psi_i | \phi_j \rangle = \langle \psi_i | \psi_j^\dagger \rangle = 0$, and

$$\langle \phi_i | \psi_j^\dagger \rangle = \langle \psi_i | \phi_j^\dagger \rangle = \pi \delta_{i,j}. \quad (7.57)$$

Using (7.53) the inner product in Eq. (7.56) can be easily calculated, giving

$$0 = \left(\frac{1}{\alpha g} - 1 \right) \int_{\mathbb{R}^2} d\mathbf{r} f'(q(r)) \left(\frac{\partial q(r)}{\partial x_i} \right)^2. \quad (7.58)$$

Hence, a spot will lose stability as g increases through $1/\alpha$ and begin to drift (translate). Note that a model of synaptic depression can also destabilise a spot in favour of a travelling pulse [4].

7.6 Center Manifold Reduction: Particle Description

Here we adapt the technique in [11], originally developed to describe spot dynamics in multi-component reaction-diffusion equations, to derive a reduced description of a slowly moving spot. Beyond a drift instability, and for small η , we expect to find a solution of (7.44) that is a translating spot. In terms of a translation operator $\tau(\mathbf{p})u(\mathbf{r}, t) = u(\mathbf{r} - \mathbf{p}, t)$, $\mathbf{p} \in \mathbb{R}^2$ we may write this solution as

$$X(\mathbf{r}, t) = \tau(\mathbf{p}) \left[S(r) + \sum_{j=1}^2 a_j(t) \psi_j(\mathbf{r}) + \chi(\mathbf{r}, t) \right], \quad (7.59)$$

where \mathbf{p} denotes the location of the spot, $a_{1,2}$ are time dependent amplitudes, and χ is a function in an orthogonal space to $\text{span}\{\phi_i, \psi_i\}$. Differentiating X with respect to t gives

$$X_t = \tau(\mathbf{p}) \left[-\dot{\mathbf{p}} \cdot \nabla_{\mathbf{r}} X + \sum_{j=1}^2 \dot{a}_j \psi_j + \chi_t \right]. \quad (7.60)$$

Using (7.59) we may calculate the first term on the right hand side of (7.60) using

$$\dot{\mathbf{p}} \cdot \nabla_{\mathbf{r}} X = \dot{\mathbf{p}} \cdot \nabla_{\mathbf{r}} S(r) + a_1 \dot{\mathbf{p}} \cdot \nabla_{\mathbf{r}} \psi_1(\mathbf{r}) + a_2 \dot{\mathbf{p}} \cdot \nabla_{\mathbf{r}} \psi_2(\mathbf{r}) + \dot{\mathbf{p}} \cdot \nabla_{\mathbf{r}} \chi(\mathbf{r}, t). \quad (7.61)$$

The corresponding right hand side of (7.44) for the solution (7.59) can be expanded as

$$\tau(\mathbf{p}) \left\{ - \sum_{j=1}^2 a_j \phi_j + \mathcal{L} \chi + \frac{1}{2} \mathcal{F}'' W^2 + \frac{1}{6} \mathcal{F}''' W^3 + \eta \gamma(S(r)) + \eta \gamma'(S(r)) W + \dots \right\}, \quad (7.62)$$

where the N th power of the vector $W = \sum_j a_j \psi_j + \chi$ is interpreted component wise. Taking the inner product of both sides of (7.44) with ψ_i^\dagger gives

$$\begin{aligned} \langle X_t | \psi_1^\dagger \rangle &= -\dot{p}_1 \langle \phi_1 | \psi_1^\dagger \rangle, \\ \langle X_t | \psi_2^\dagger \rangle &= -\dot{p}_2 \langle \phi_2 | \psi_2^\dagger \rangle, \\ \langle \mathcal{F}(X; \sigma_c) + \eta \gamma(X) | \psi_1^\dagger \rangle &= -a_1 \langle \phi_1 | \psi_1^\dagger \rangle, \\ \langle \mathcal{F}(X; \sigma_c) + \eta \gamma(X) | \psi_2^\dagger \rangle &= -a_2 \langle \phi_2 | \psi_2^\dagger \rangle. \end{aligned} \quad (7.63)$$

Using (7.44) and equating expressions in (7.63) gives an equation for the evolution of the spot position as

$$\dot{\mathbf{p}} = \mathbf{a}. \quad (7.64)$$

Hence we may interpret \mathbf{a} as the spot velocity. To determine the dynamics for \mathbf{a} we write χ as a function that is quadratic in the amplitudes a_i and linear in η :

$$\chi = a_1^2 V_1(\mathbf{r}) + a_2^2 V_2(\mathbf{r}) + a_1 a_2 V_3(\mathbf{r}) + \eta V_4(\mathbf{r}). \quad (7.65)$$

Demanding that terms at this order balance requires

$$a_1 \mathbf{a} \cdot \nabla_{\mathbf{r}} \psi_1 + a_2 \mathbf{a} \cdot \nabla_{\mathbf{r}} \psi_2 = \mathcal{L} \chi + \frac{1}{2} \mathcal{F}'' (\mathbf{a} \cdot \psi)^2 + \eta \gamma(S), \quad \psi = (\psi_1, \psi_2). \quad (7.66)$$

Equating terms in $a_i a_j$ and η shows that

$$-\mathcal{L} V_1(\mathbf{r}) = \frac{1}{2} \mathcal{F}'' (\psi_1)^2 + \nabla_{x_1} \psi_1, \quad (7.67)$$

$$-\mathcal{L} V_2(\mathbf{r}) = \frac{1}{2} \mathcal{F}'' (\psi_2)^2 + \nabla_{x_2} \psi_2, \quad (7.68)$$

$$-\mathcal{L} V_3(\mathbf{r}) = \mathcal{F}'' \psi_1 \psi_2 + \nabla_{x_2} \psi_1 + \nabla_{x_1} \psi_2, \quad (7.69)$$

$$-\mathcal{L} V_4(\mathbf{r}) = \gamma(S(r)). \quad (7.70)$$

Here multiplication of vectors $\psi_i \psi_j$ is interpreted component wise. We now take the inner product of both sides of (7.44) with ϕ_i^\dagger :

$$\langle X_t | \phi_1^\dagger \rangle = -\dot{p}_1 \langle \chi_{x_1} | \phi_1^\dagger \rangle + \dot{a}_1 \langle \psi_1 | \phi_1^\dagger \rangle, \quad (7.71)$$

$$\langle X_t | \phi_2^\dagger \rangle = -\dot{p}_2 \langle \chi_{x_2} | \phi_2^\dagger \rangle + \dot{a}_2 \langle \psi_2 | \phi_2^\dagger \rangle, \quad (7.72)$$

$$\begin{aligned} \langle \mathcal{F}(X; \sigma_c) + \eta \gamma(X) | \phi_i^\dagger \rangle &= \langle \mathcal{L} \chi + \frac{1}{2} \mathcal{F}'' W^2 + \frac{1}{6} \mathcal{F}''' W^3 + \eta \gamma(S(r)) \\ &\quad + \eta \gamma'(S(r)) W | \phi_i^\dagger \rangle. \end{aligned} \quad (7.73)$$

Making use of (7.66) and working to only cubic order in the amplitudes so that $W^2 \approx (\mathbf{a} \cdot \psi)^2 + 2(\mathbf{a} \cdot \psi)\chi$ and $W^3 \approx (\mathbf{a} \cdot \psi)^3$ we find that

$$\begin{aligned} \pi \dot{a}_j &= \langle \mathcal{F}''(\mathbf{a} \cdot \psi)\chi | \phi_j^\dagger \rangle + \frac{1}{6} \langle \mathcal{F}'''(\mathbf{a} \cdot \psi)^3 | \phi_j^\dagger \rangle + \eta \langle \gamma'(S)(\mathbf{a} \cdot \psi) | \phi_j^\dagger \rangle \\ &\quad + \langle \mathbf{a} \cdot \nabla_r \chi | \phi_j^\dagger \rangle. \end{aligned} \quad (7.74)$$

Introducing the complex amplitude $a = a_1 + i a_2$ gives us the Stuart-Landau equation

$$\dot{a} = a(M_1|a|^2 + M_2\eta), \quad (7.75)$$

where

$$\pi M_1 = \frac{1}{6} \langle \mathcal{F}''' \psi_1^3 | \phi_1^\dagger \rangle + \langle \mathcal{F}'' \psi_1 V_1^2 | \phi_1^\dagger \rangle + \langle \partial_{x_1} V_1 | \phi_1^\dagger \rangle, \quad (7.76)$$

$$\pi M_2 = \langle \mathcal{F}'' \psi_1 V_4 | \phi_1^\dagger \rangle + \langle \gamma'(S) \psi_1 | \phi_1^\dagger \rangle + \langle \partial_{x_1} V_4 | \phi_1^\dagger \rangle. \quad (7.77)$$

Hence, there is a pitchfork equation with branches of solutions $(a_1, a_2) = (0, 0)$ (a standing spot) and $|a|^2 = -M_2\eta/M_1$ for $M_2\eta/M_1 < 0$ and $M_{1,2} \neq 0$ (a travelling spot). Beyond a drift instability the speed of the travelling spot will scale as $\sqrt{|\eta M_2/M_1|}$ for small η . Treating g as the bifurcation parameter we see that the speed scales as $\sqrt{g - 1/\alpha}$, which compares well with direct numerical simulations near the bifurcation point (not shown).

7.7 Scattering

For anatomical interactions which decay exponentially quickly, such as Mexican or wizard hat functions, then we would expect a neural field model with a single spot solution to also support multiple spots, at least for some large separation between spots. This then begs the question of how spots interact when they come close together. Interestingly it has already been found numerically that spots in a model with spike frequency adaptation can scatter like dissipative solitons [8]. Here we

adapt techniques originally developed by Ei et al. [10, 11], for multi-component reaction diffusion systems, to show that two slowly moving spots will reflect from each other in a head-on collision.

We introduce the sum of two spots with centers offset by a vector \mathbf{h} as

$$S(\mathbf{r}; \mathbf{h}) = S(\mathbf{r}) - S(\mathbf{r} - \mathbf{h}), \quad S(\mathbf{r}) = S(|\mathbf{r}|). \quad (7.78)$$

We then consider solutions of the form

$$X(\mathbf{r}, t) = \tau(\mathbf{p}) [S(\mathbf{r}; \mathbf{h}) + \mathbf{a} \cdot \psi(\mathbf{r}) + \mathbf{b} \cdot \psi(\mathbf{r} - \mathbf{h}) + \chi(\mathbf{r}, t)]. \quad (7.79)$$

We may then adapt the technique of Sect. 7.6, closely following [11], to derive the equations of motion for $(\mathbf{p}, \mathbf{h}, \mathbf{a}, \mathbf{b})$ as

$$\dot{\mathbf{p}} = \mathbf{a} + \mathbf{G}_1(\mathbf{h}), \quad (7.80)$$

$$\dot{\mathbf{h}} = \mathbf{b} - \mathbf{a} + \mathbf{G}_1(\mathbf{h}) - \mathbf{G}_2(\mathbf{h}), \quad (7.81)$$

$$\dot{\mathbf{a}} = \nabla W(\mathbf{a}) + \mathbf{H}_1(\mathbf{h}), \quad (7.82)$$

$$\dot{\mathbf{b}} = \nabla W(\mathbf{b}) + \mathbf{H}_2(\mathbf{h}), \quad (7.83)$$

where

$$\mathbf{G}_1(\mathbf{h}) = \begin{bmatrix} \langle \mathcal{F}(S(\cdot; \mathbf{h})) | \psi_1^\dagger \rangle \\ \langle \mathcal{F}(S(\cdot; \mathbf{h})) | \psi_2^\dagger \rangle \end{bmatrix} \quad \mathbf{G}_2(\mathbf{h}) = \begin{bmatrix} \langle \mathcal{F}(S(\cdot + \mathbf{h}; \mathbf{h})) | \psi_1^\dagger \rangle \\ \langle \mathcal{F}(S(\cdot + \mathbf{h}; \mathbf{h})) | \psi_2^\dagger \rangle \end{bmatrix}, \quad (7.84)$$

$$\mathbf{H}_1(\mathbf{h}) = \begin{bmatrix} \langle \mathcal{F}(S(\cdot; \mathbf{h})) | \phi_1^\dagger \rangle \\ \langle \mathcal{F}(S(\cdot; \mathbf{h})) | \phi_2^\dagger \rangle \end{bmatrix} \quad \mathbf{H}_2(\mathbf{h}) = \begin{bmatrix} \langle \mathcal{F}(S(\cdot + \mathbf{h}; \mathbf{h})) | \phi_1^\dagger \rangle \\ \langle \mathcal{F}(S(\cdot + \mathbf{h}; \mathbf{h})) | \phi_2^\dagger \rangle \end{bmatrix}, \quad (7.85)$$

and

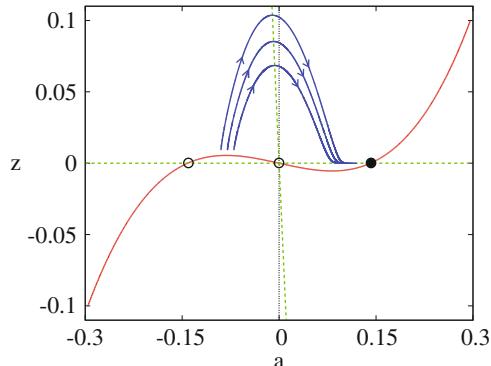
$$W(\mathbf{x}) = \frac{1}{4} M_1 |\mathbf{x}|^4 + \frac{1}{2} \eta M_2 |\mathbf{x}|^2. \quad (7.86)$$

Now for a spot shape like that of (7.23) we may use the asymptotic properties of $K_0(r)$ to see that $q(r) \sim \exp(-r)/\sqrt{r}$ for large r . We expect similar decay properties of spots in the case of steep sigmoidal firing rates. In this case we can use results from [11], valid as $h = |\mathbf{h}| \rightarrow \infty$, to represent the interaction functions in the form

$$G_i(\mathbf{h}) = (-1)^{i-1} G_0 \frac{1}{\sqrt{h}} e^{-h} \mathbf{e}, \quad H_i(\mathbf{h}) = (-1)^{i-1} H_0 \frac{1}{\sqrt{h}} e^{-h} \mathbf{e}, \quad \mathbf{e} = \frac{\mathbf{h}}{h}, \quad (7.87)$$

for constants G_0 and H_0 , which we shall assume to be positive.

Fig. 7.9 Phase plane for the dynamical system (7.89), showing nullclines and some typical trajectories. The consequence for spot-spot interactions is that for small η then slowly moving spots will scatter from each other by reflection. Parameters are $\eta = M_2 = G_0 = H_0 = 0.1$ and $M_1 = -0.5$



Now consider the interaction of two travelling pulses, with one centred at \mathbf{p} and another at $-\mathbf{p}$ moving on a line joining their centres so that the separation between them is $\mathbf{h} = 2\mathbf{p}$. For simplicity we shall take \mathbf{p} to be along the x_1 axis and write $\mathbf{p} = p(1, 0)$ and set $\mathbf{a} = a(1, 0)$. In this case we have that

$$\dot{p} = a + G_0 f(p), \quad \dot{a} = M_1 a^3 + M_2 a \eta + H_0 f(p), \quad (7.88)$$

where $f(p) = e^{-2p}/\sqrt{2p}$. Introducing $z = f(p)$ we may rewrite this dynamical system as

$$\dot{z} = -2zQ(z)[a + G_0 z], \quad \dot{a} = H_0 z - f(a), \quad (7.89)$$

where $Q(z) = 1 + (4f^{-1}(z))^{-1} > 0$ and $f(a) = -M_1 a^3 - M_2 a \eta$. There are stationary solutions of (7.89) in the (a, z) plane at $(v_-, 0)$, $(0, 0)$, and $(v_+, 0)$, where $v_{\pm} = \pm\sqrt{-M_2 \eta / M_1}$ (with $M_1 < 0$ and $M_2 \eta > 0$). Linear stability analysis around a stationary solution gives a pair of eigenvalues that determine stability in the form $\lambda_1 = -2a$ and $\lambda_2 = -f'(a)$. Hence the only stable solution is $(v_+, 0)$. An analysis of the phase plane, see Fig. 7.9, shows that trajectories that start with $a < 0$ and small z (namely spots moving with negative velocity toward each other from far apart) ultimately move to $a = v_+ > 0$ with small z so that the spots reverse their motion and travel with positive velocity away from each other. Hence we expect two travelling spots to reflect from each other, as shown in a simulation of the full model in Fig. 7.10.

7.8 Discussion

In this chapter we have shown how to analyse the properties of spots in planar neural field models with spike frequency adaptation, using a mixture of techniques ranging from direct numerical simulations, through amplitude equations to explicit results

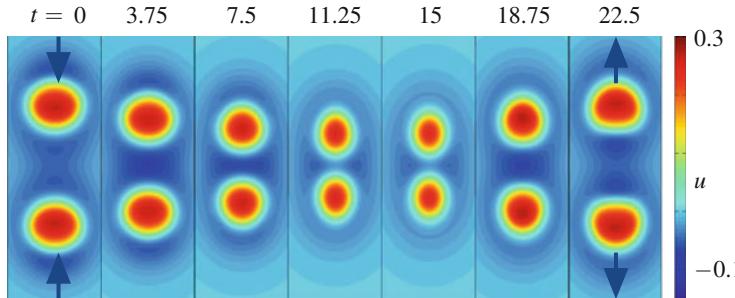


Fig. 7.10 Results of a direct numerical simulation of two scattering spots in the neural field model with Heaviside firing rate and parameters $\gamma = 4$, $\mu = 0.5$, $g = 0.5$ and $\alpha = 5$. The overall spatial domain has a size of 51.2×51.2 of which we see a zoomed set of seven snapshots at indicated times

for the special case of a Heaviside firing rate function using an interface approach. There are a number of natural extensions of this work that may be developed. The scattering theory that we developed is valid only for slowly moving spots, which is expected to be the case when model parameters are near to that defining the onset of a drift instability for a single spot. However, numerical simulations show that fast moving spots may scatter differently to slow moving ones, with the possibility of *fusion* and *annihilation*, as well as repulsion. In this case it is likely that these more complicated scenarios can be understood using the scattor theory of Nishiura et al. [22]. In this framework the scattering process is understood in terms of the stable and unstable manifolds of a certain unstable pattern that has the form of a two-lobed *peanut* shape. These solutions are expected to arise via a symmetry breaking bifurcation of spots to solutions with D_2 symmetry (generated by rotations of π , and reflection across a central axis). The computation of the scattor requires the numerical calculation of non-rotationally symmetric solutions, and the numerical schemes for PDEs that we have described here are easily generalised. Moreover, for the Heaviside firing rate we can use a formulation in [8] to express such solutions purely as line integrals, with a substantial reduction in dimensionality that will facilitate an exhaustive numerical bifurcation analysis. The interaction of drift and peanut modes of instability is known to generate a rotational motion of travelling spots, at least in three component reaction diffusion equations [26]. It should also be possible to extend the center manifold reduction developed here to describe the behaviour of localised solutions in the neighbourhood of such a co-dimension two bifurcation point. Finally, we flag up the utility of the interface approach in two spatial dimensions for analysing not only localised states, but extended solutions such as spirals [20]. All of the above are topics of ongoing research and will be reported upon elsewhere.

Acknowledgements The authors would like to acknowledge useful discussions with Paul Bressloff, Gr  gory Faye, Carlo Laing and David Lloyd that have helped to improve the presentation of the ideas in this chapter.

Appendix

To establish that ϕ^\dagger can be written as a linear transformation of ϕ we proceed by writing $M = P \operatorname{diag}(\lambda_+, \lambda_-) P^{-1}$, where $P = [v_+ \ v_-]$, and v_\pm are the right eigenvectors of M :

$$v_\pm = \begin{bmatrix} \alpha g \\ -(\alpha + \lambda_\pm) \end{bmatrix}, \quad (7.90)$$

with

$$\lambda_\pm = \frac{-(1+\alpha) \pm \sqrt{(1-\alpha)^2 - 4\alpha g}}{2}. \quad (7.91)$$

In this case we may recast the eigenvalue problem $\mathcal{L}\phi = \lambda\phi$ as

$$\begin{bmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{bmatrix} P^{-1}\phi + \alpha w \otimes f'(q) P^{-1} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \phi = \lambda P^{-1}\phi. \quad (7.92)$$

Similarly we may write $M^\dagger = R \operatorname{diag}(\lambda_+, \lambda_-) R^{-1}$, where $R = [w_+ \ w_-]$, and w_\pm are the right eigenvectors of M^\dagger :

$$w_\pm = \begin{bmatrix} 1 \\ \alpha + \lambda_\pm \end{bmatrix}. \quad (7.93)$$

The adjoint operator \mathcal{L}^\dagger can be found as

$$\mathcal{L}^\dagger = M^\dagger + \alpha f'(q) w \otimes \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}. \quad (7.94)$$

By inspection it can be seen that the adjoint equation $\mathcal{L}^\dagger\phi^\dagger = \bar{\lambda}\phi^\dagger$ has a solution

$$\phi^\dagger = f'(q)RP^{-1}\bar{\phi}, \quad (7.95)$$

which can be evaluated to give (7.53). Here we make use of the result that

$$PR^{-1} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} RP^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}. \quad (7.96)$$

References

1. Amari, S.: Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.* **27**, 77–87 (1977)
2. Avitabile, D.: Computation of planar patterns and their stability. PhD thesis, Department of Mathematics, University of Surrey (2008)
3. Blomquist, P., Wyller, J., Einevoll, G.T.: Localized activity patterns in two-population neuronal networks. *Physica D* **206**, 180–212 (2005)
4. Bressloff, P.C., Kilpatrick, Z.P.: Two-dimensional bumps in piecewise smooth neural fields with synaptic depression. *SIAM J. Appl. Math.* **71**, 379–408 (2011)
5. Coombes, S.: Waves, bumps, and patterns in neural field theories. *Biol. Cybern.* **93**, 91–108 (2005)
6. Coombes, S., Owen, M.R.: Exotic dynamics in a firing rate model of neural tissue with threshold accommodation. *AMS Contemp. Math. Fluids Waves: Recent Trends Appl. Anal.* **440**, 123–144 (2007)
7. Coombes, S., Venkov, N.A., Shiu, L., Bojak, I., Liley, D.T.J., Laing, C.R.: Modeling electrocortical activity through improved local approximations of integral neural field equations. *Phys. Rev. E* **76**, 051901 (2007)
8. Coombes, S., Schmidt, H., Bojak, I.: Interface dynamics in planar neural field models. *J. Math. Neurosci.* **2**(1), 9 (2012)
9. Doedel, E., Oldeman, B.: Auto-07p: continuation and bifurcation software for ordinary differential equations. Technical report, Concordia University, Montreal (2009)
10. Ei, S.-I., Mimura, M., Nagayama, M.: Pulse–pulse interaction in reaction–diffusion systems. *Physica D* **165**, 176–198 (2002)
11. Ei, S.-I., Mimura, M., Nagayama, M.: Interacting spots in reaction diffusion systems. *Discret. Contin. Dyn. Syst. Ser. A* **14**, 31–62 (2006)
12. Ermentrout, G.B.: Neural nets as spatio-temporal pattern forming systems. *Rep. Prog. Phys.* **61**, 353–430 (1998)
13. Faye, G., Rankin, J., Chossat, P.: Localized states in an unbounded neural field equation with smooth firing rate function: a multi-parameter analysis. *J. Math. Biol.* **66**(6), 1303–1338. doi:10.1007/s00285-012-0532-y
14. Folias, S.E.: Nonlinear analysis of breathing pulses in a synaptically coupled neural network. *SIAM J. Appl. Dyn. Syst.* **10**, 744–787 (2011)
15. Folias, S.E., Bressloff, P.C.: Breathing pulses in an excitatory neural network. *SIAM J. Appl. Dyn. Syst.* **3**(3), 378–407 (2004)
16. Folias, S.E., Bressloff, P.C.: Breathers in two-dimensional neural media. *Phys. Rev. Lett.* **95**, 208107(1–4) (2005)
17. Goldman-Rakic, P.S.: Cellular basis of working memory. *Neuron* **14**, 477–485, (1995)
18. Gurevich, S.V., Amirashvili, Sh., Purwins, H.-G.: Breathing dissipative solitons in three-component reaction-diffusion systems. *Phys. Rev. E* **74**, 066201(1–7) (2006)
19. Kilpatrick, Z.P., Bressloff, P.C.: Effects of synaptic depression and adaptation on spatiotemporal dynamics of an excitatory neuronal network. *Physica D* **239**, 547–560 (2010)
20. Laing, C.R.: Spiral waves in nonlocal equations. *SIAM J. Appl. Dyn. Syst.* **4**, 588–606 (2004)
21. Laing, C.R., Troy, W.C.: PDE methods for nonlocal models. *SIAM J. Appl. Dyn. Syst.* **2**, 487–516 (2003)
22. Nishiura, Y., Teramoto, T., Ueda, K.-I.: Scattering of traveling spots in dissipative systems. *Chaos* **15**, 047509(1–10) (2005)
23. Owen, M.R., Laing, C.R., Coombes, S.: Bumps and rings in a two-dimensional neural field: splitting and rotational instabilities. *New J. Phys.* **9**, 378 (2007)
24. Pinto, D.J., Ermentrout, G.B.: Spatially structured activity in synaptically coupled neuronal networks: II. Lateral inhibition and standing pulses. *SIAM J. Appl. Math.* **62**, 226–243 (2001)

25. Sandstede, B., Scheel, A.: Defects in oscillatory media: toward a classification. *Dyn. Syst.* **3**, 1–68 (2004)
26. Teramoto, T., Suzuki, K., Nishiura, Y.: Rotational motion of traveling spots in dissipative systems. *Phys. Rev. E* **80**, 046208(1–10) (2009)

Chapter 8

Heterogeneous Connectivity in Neural Fields: A Stochastic Approach

Chris A. Brackley and Matthew S. Turner

Abstract One of the traditional approximations applied in Amari type neural field models is that of a homogeneous isotropic connection function. Incorporation of heterogeneous connectivity into this type of model has taken many forms, from the addition of a periodic component to a crystal-like inhomogeneous structure. In contrast, here we consider stochastic inhomogeneous connections, a scheme which necessitates a numerical approach. We consider both local inhomogeneity, a local stochastic variation of the strength of the input to different positions in the media, and long range inhomogeneity, the addition of connections between distant points. This leads to changes in the well known solutions such as travelling fronts and pulses, which (where these solutions still exist) now move with fluctuating speed and shape, and also gives rise to a new type of behaviour: persistent fluctuations in activity. We show that persistent activity can arise from different mechanisms depending on the connection model, and show that there is an increase in coherence between fluctuations at distant regions as long-range connections are introduced.

8.1 Introduction

Continuum neural field models of the type proposed by Amari [1, 2] (see also Chap. 3) have been used as a model for cortical tissue, describing phenomena such as travelling fronts and pulses of activity [12, 27], stationary and breathing bumps [15, 16, 18, 28], and instabilities leading to pattern formation such as might be responsible for visual hallucinations [6, 22] (see also Chaps. 1, 4, and 7). Much of

C.A. Brackley (✉)

School of Physics and Astronomy, University of Edinburgh, Mayfield Road,
Edinburgh EH9 3JZ, UK
e-mail: cbrackle@ph.ed.ac.uk

M.S. Turner

Department of Physics and Complexity Centre, University of Warwick, Coventry CV4 7AL, UK

this work uses the assumption that neural connectivity in the tissue is homogeneous and isotropic. This simplification may give an adequate first order approximation of the behaviour of the tissue in some situations (e.g. fronts of activity in cortical slices [11]), but it is mainly motivated by the fact that it leads to much more tractable equations. There have been several attempts to include more realistic connections in this type of model: for example travelling fronts in a periodically varying connection function have been studied in Refs. [7, 9, 13, 25], whilst Ref. [8] considers a crystal like structure for connectivity such as might be expected in the visual cortex, and long range point connections have been studied in Refs. [5, 24, 26, 29]. Spatial inhomogeneity can also be introduced via external input to the fields [10, 17, 18].

In this chapter we take a different approach, and consider quenched stochastic inhomogeneous connections which are introduced in addition to a homogeneous component. To achieve this we numerically construct spatially continuous, stochastic connection functions. With such a scheme it is not possible to solve the field equations analytically, so instead we use approximation and numerical methods.

We consider a two population continuum model, given in two dimensions by

$$\begin{aligned} \tau_u \frac{\partial u}{\partial t} + u(\mathbf{x}, t) &= \int_{\Gamma} w(\mathbf{x}, \mathbf{x}') f(u(\mathbf{x}', t)) d\mathbf{x}' - g v(\mathbf{x}, t), \\ \tau_v \frac{\partial v}{\partial t} + v(\mathbf{x}, t) &= f(u(\mathbf{x}, t)), \end{aligned} \quad (8.1)$$

where Γ denotes the extent of the system. The field $u(\mathbf{x}, t)$ describes an excitatory neural population and $v(\mathbf{x}, t)$ local inhibition (which can be interpreted either as an inhibitory neural population, or as nonlinear local feedback). We set the time units of the system by choosing $\tau_u = 1$; τ_v and g give the relative response time and strength of the inhibition. We employ the usual approximation for the firing rate function, taking it to be a step at threshold θ , i.e. $f(u) = \Theta(u - \theta)$ where $0 < \theta < 1$.

For the inhomogeneous connection function $w(\mathbf{x}, \mathbf{x}')$ we consider two different forms. First in Sect. 8.2 we consider a local inhomogeneity where the connection weight varies with position, but there are no long range connections; then in Sect. 8.3 we consider a connection function in which long range connections can be introduced by varying a single parameter. In both cases we examine the well known solutions of propagating fronts and pulses of activity in one dimension, before describing a new type of behaviour—namely persistent fluctuations of activity—in both one and two dimensional systems.

We perform numerical simulations of Eq. (8.1) by discretizing space, and then using a 4th order Runge-Kutta algorithm to solve a set of first order ordinary differential equations. If the connection function $w(\mathbf{x}, \mathbf{x}')$ is chosen carefully, the integral on the right hand side of the equation for $u(\mathbf{x}, t)$ can be written as a convolution. Although this leads to some loss of generality, the convolution theorem can then be exploited—using a fast Fourier transform algorithm [19] the integral can be very efficiently solved.

8.2 Local Inhomogeneity in Connection Density

There is some experimental evidence [21] that the number of synapses between two neurons is a Gaussian function of their separation. Thus we adopt a Gaussian connection function, and add to it a small *inhomogeneous* perturbation. In this section we assume that there are no long distance connections and write our connection function in 1D as

$$w(x, x') = w_H(x - x')[1 + A(w_1(x) + w_2(x'))], \quad (8.2)$$

where w_H is a normalised Gaussian function

$$w_H(y) = \frac{1}{\sqrt{\pi}} e^{-y^2}, \quad (8.3)$$

and the constant A gives the magnitude of the inhomogeneous connections. The unit width of the homogeneous function defines the spatial length scale of the connections (and all lengths given in the rest of the chapter are quoted in these units). For w_1 and w_2 we numerically generate functions which vary stochastically (but continuously) in space with Gaussian statistics, zero mean, and unit mean squared, and which are auto-correlated on a length λ (there is no correlation between the functions). We include two different functions (of x and x') in order to remove any bi-directionality in the connections; for simplicity each is a different stochastic realisation of the function with the same statistics. This can loosely be thought to represent additional (to the homogeneous) connections *into* point x and *out of* point x' . It is this separability of the connection function which allows the integral to be written as a convolution. Note that although the functions are constructed stochastically, they do not vary with time, so the dynamics of the system are entirely deterministic.

The inhomogeneity is therefore characterised by its magnitude A , and correlation length λ . We identify two different regimes, $\lambda < 1$ and $\lambda > 1$. The first represents local heterogeneity in the connections, i.e., on length scales less than the width of w_H ; the second can be interpreted as variation of connection strength on lengths larger than the width of w_H , i.e., locally the connections appear homogeneous, but the overall connection density varies on longer length scales. For the latter case it is important to note that whilst connection *density* varies on long length scales in this regime, we have not included any *long range* connections. We also note that since $u(x, t)$ represents a population of excitatory neurons, only a positive connection function makes physical sense. Thus in the present work we only consider values of A that are small enough so that $w(x, x')$ remains positive for all x, x' .

8.2.1 Travelling Fronts and Pulses

We first examine the effect of local stochastic connectivity on the well known solutions to Eq.(8.1) in 1D, i.e. travelling fronts and pulses [27]. We treat the

inhomogeneity as a small perturbation to the system; initially we add a constant additional connection weight γ to a homogeneous function $w_H(y)$, later replacing this with the inhomogeneous connections $\gamma \rightarrow \gamma(x, x') = A[w_1(x) + w_2(x')]$, and expanding $\gamma(x, x') = \bar{\gamma} + \delta\gamma(x, x')$ about $\bar{\gamma} = 0$. We consider the equations

$$\begin{aligned}\tau_u \frac{\partial u}{\partial t} + u(x, t) &= \int_{\Gamma} w_H(x - x')(1 + \gamma) f(u(x', t)) dx' - g v(x, t), \\ \tau_v \frac{\partial v}{\partial t} + v(x, t) &= f(u(x, t)),\end{aligned}\quad (8.4)$$

where $w_H(y)$ is given in Eq. (8.3). The uniform steady states are found by setting $\partial_t u = \partial_t v = 0$ and assuming no x dependence. These are then given by the pairs of values of u and v which simultaneously solve

$$\begin{aligned}u &= (1 + \gamma) f(u) - g v, \\ v &= f(u).\end{aligned}\quad (8.5)$$

The stability of the points can be shown in the standard way by expanding $u(x, t) = \bar{u} + \delta u e^{\omega t}$ and $v(x, t) = \bar{v} + \delta v e^{\omega t}$ in Eq. (8.4), and then finding the eigenvalues ω . There are stable fixed points at $(\bar{u}_1, \bar{v}_1) = (0, 0)$ and $(\bar{u}_3, \bar{v}_3) = (1 + \gamma - g, 1)$, and an unstable saddle point at $(\bar{u}_2, \bar{v}_2) = (\theta, 1)$; note that the fixed point (\bar{u}_3, \bar{v}_3) only exists if $1 + \gamma - g > \theta$. Travelling wave fronts are possible when there are two stable steady states—the front connects a region in the (\bar{u}_1, \bar{v}_1) state with a region in the (\bar{u}_3, \bar{v}_3) state.

The speed of the front can be found by following Ref. [12]. The equations for $u(x, t)$ and $v(x, t)$ can be solved using the Green's functions

$$\eta_u(s) = \begin{cases} \frac{1}{\tau_u} e^{-s/\tau_u} & s \geq 0, \\ 0 & s < 0, \end{cases} \quad \text{and} \quad \eta_v(s) = \begin{cases} \frac{1}{\tau_v} e^{-s/\tau_v} & s > 0, \\ 0 & s \leq 0, \end{cases} \quad (8.6)$$

and a change of variables $\xi = x - ct$ can be used to transform to a co-moving frame where the front is stationary with shape given by

$$q(\xi) = \int_0^\infty \eta_u(s) [\psi_u(\xi + cs) - g \psi_v(\xi + cs)] ds, \quad (8.7)$$

where

$$\psi_u(\xi) = \int_{-\infty}^{\infty} (1 + \gamma) w_H(y) f(q(\xi - y)) dy, \quad (8.8)$$

$$\psi_v(\xi) = \int_0^\infty \eta_v(s) f(q(\xi + cs)) ds. \quad (8.9)$$

The boundary conditions are $q(0) = \theta$, with $q(\xi) < \theta$ for $\xi > 0$, and $q(\xi) > \theta$ for $\xi < 0$. Since the firing rate is taken to be a step function $f(u) = \Theta(u - \theta)$, these integrals can be solved, and then the front speed c is given by

$$\theta = \begin{cases} \frac{1+\gamma}{2} \left[1 - e^{\frac{1}{4c^2}} \left(1 - \text{erf}\left(\frac{1}{2c}\right) \right) \right] c > 0, \\ \frac{1+\gamma}{2} \left[1 + e^{\frac{1}{4c^2}} \left(1 + \text{erf}\left(\frac{1}{2c}\right) \right) \right] c < 0. \end{cases} \quad (8.10)$$

which can be solved numerically (a different choice of $w_H(y)$ can lead to analytic solutions). Figure 8.1a, b show how c depends on θ and γ ; there is a minimum value of γ below which there is no front solution. There is no dependence on τ_v or g . It can be shown using the Evans' function method [14] that when both stable fixed points (\bar{u}_1, \bar{v}_1) and (\bar{u}_3, \bar{v}_3) exist, the moving front solution is also stable.

So far we have proceeded as in Ref. [12], except for the inclusion of the constant γ which represents the additional connections that are added to the system. The extension to spatially varying connections is simply a matter of the replacement $\gamma \rightarrow \gamma(x, x') = A[w_1(x) + w_2(x')]$. Equation (8.8) becomes

$$\psi_u(\xi) = \int_{-\infty}^{\infty} (1 + \gamma(x, x')) w_H(y) f(q(\xi - y)) dy. \quad (8.11)$$

We assume for the moment that this replacement does not change the fact that there are two stable steady states; we shall discuss what happens if this is not the case at the end of this section. Considering first the regime where λ is shorter than the unit width of the homogeneous connection function and the length $c\tau_v$, we find that the inhomogeneity will be averaged out, and effectively $\gamma(x, x') = 0$. That is to say, the integral in Eq. (8.11) is effectively an average of $\gamma(x, x')$ over the width of the function $w_H(y)$; since in the $\lambda < 1$ regime $\gamma(x, x')$ varies on a length scale much shorter than this and has zero mean, we expect behaviour to be the same as the homogeneous case.

In the limit of large λ (which we argued above is biologically relevant), $w_2(x')$ will vary on length scales much longer than the width of the homogeneous connections $w_H(y)$, and so in the integral we can approximate $w_2(x') \approx w_2(x)$ and then $\gamma(x, x') \approx \gamma(x) = A[w_1(x) + w_2(x)]$. We then proceed to expand $\gamma(x) = \bar{\gamma} + \delta\gamma(x)$ about $\bar{\gamma} = 0$. We argue that the speed c is given by a function $c = h(\gamma)$ (the solution to Eq. (8.10)), and expand to second order, giving

$$c = h(\bar{\gamma}) + \delta\gamma h'(\bar{\gamma}) + \frac{\delta\gamma^2}{2} h''(\bar{\gamma}) + \mathcal{O}(\delta\gamma^3). \quad (8.12)$$

Since the connection function is composed of Gaussian distributed functions with zero mean, we know that $\bar{\gamma} = 0$ and $\langle \delta\gamma^2 \rangle = 2A^2$, $\langle \delta\gamma^4 \rangle = 6A^4$ etc. and the odd moments are zero, with A the amplitude of the inhomogeneity. We can therefore write the mean and variance of the speed

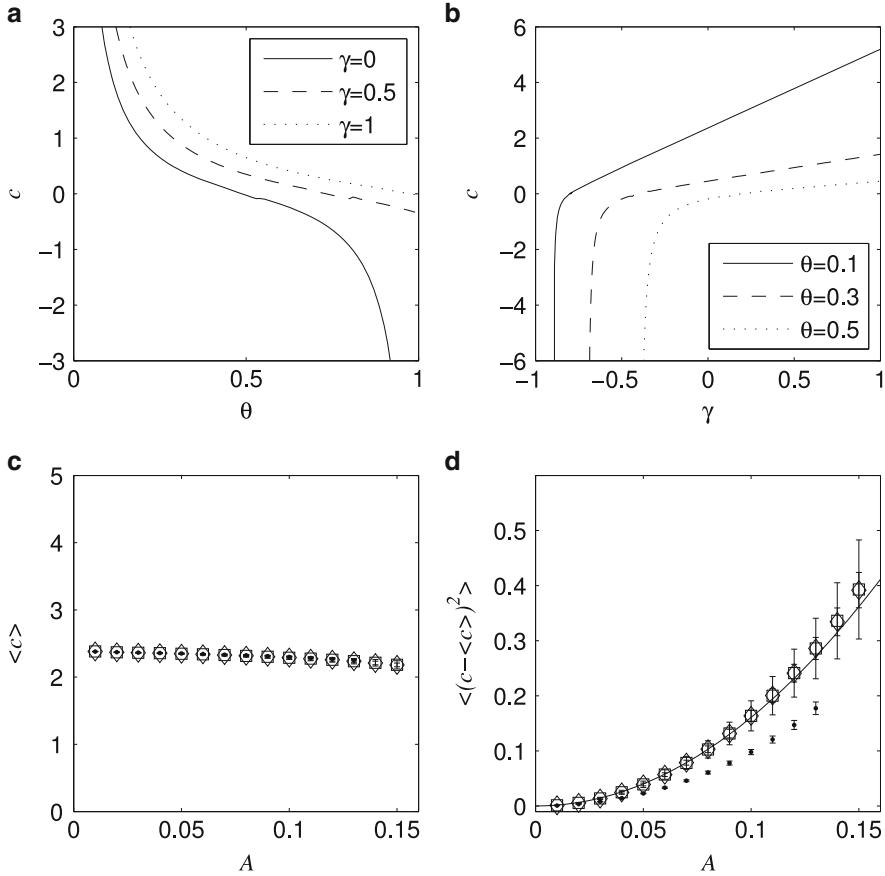


Fig. 8.1 Plots (a) and (b) show, for the homogeneous system given by Eq. (8.4), how the front propagation speed c varies with θ (at constant γ) and γ (at constant θ) respectively. Note that there is a minimum value of γ below which the front solution does not exist. There is no dependence on τ_v or g , and where moving front solutions exist they are stable. Plots (c) and (d) show how the mean and variance of the speed of a front travelling through an inhomogeneous system with two stable steady state solutions depends on the magnitude of the inhomogeneous connections A . Points, diamonds and squares show results for $\lambda = 0.5, 5$ and 10 respectively, averaged over 10 realisations of the stochastic connections. The solid line gives the analytic result for the variance of c derived in the large λ limit (Eq. (8.13)). The other parameters used are $\theta = 0.1$ and $g = 0.2$

$$\begin{aligned} \langle c \rangle &= h(0) + A^2 h''(0) + \mathcal{O}(\delta\gamma^4), \\ \langle (c - \langle c \rangle)^2 \rangle &= 2A^2 h'(0)^2 + \frac{1}{2}A^4 h''(0)^2 + \mathcal{O}(\delta\gamma^4). \end{aligned} \quad (8.13)$$

From inspection of the plot of c against γ (Fig. 8.1b) there is an approximately linear relationship at $\gamma = 0$; thus we approximate the first derivative $h'(0)$ as a constant (found numerically from Eq. (8.10) to be $h'(0) = 2.835$) and neglect higher

derivatives. We therefore expect the mean speed to be independent of A at linear order, and the variance to vary as A^2 .

We test these predictions by using numerical simulations of an inhomogeneous system where there are two stable steady states throughout, and where initial conditions have been chosen so as to lead to travelling fronts (for example an initial condition with a discontinuity in the u variable, $u(x < 0, t = 0) = 1$, $u(x > 0, t = 0) = 0$, $v(x, t = 0) = 0$). Figure 8.1c, d show how the mean and variance of the speed of the front vary with A for several values of λ . The averages are over not only the speed of a single front as it moves through the system, but also over many systems with different realisations of the stochastic connections. We find that for small A the mean speed is approximately constant, but for larger A our approximation fails, as the mean speed starts to decrease. The solid line shows the equation $2h'(0)A^2$ with $h'(0) = 2.835$, and we find that this becomes a better approximation as λ increases.

As well as travelling front solutions, homogeneous models of this type can also support travelling *pulses*. Again we investigate the effect of inhomogeneous connections by first considering a homogeneous additional connection weight γ , before expanding about this. As shown in Ref. [12] the derivation of the speed and shape of a travelling pulse follows in a similar manner to that of the front, but with boundary conditions $q(0) = q(\Delta) = \theta$, $q(\xi) > \theta$ for $0 \leq \xi < \Delta$ and $q(\xi) < \theta$ otherwise; i.e. the pulse has width Δ . Since this system is isotropic, solutions with $c > 0$ and $c < 0$ will be identical under the transformation $x \rightarrow -x$, so we only consider $c > 0$. Analysing Eq. (8.4) with these boundary conditions, we find equations relating c and Δ to g , θ , τ_u and γ (similar to Eq. (8.10) for the fronts). Figure 8.2 shows the dependence of pulse width and speed on γ and g for $\theta = 0.1$ and $\tau_v = 1$. Note that for some parameters there are two branches of solutions. The stability can again be found by constructing an Evans function, which allows identification of a stable and an unstable branch.

To examine whether the system still supports travelling pulses in the presence of inhomogeneous connections we proceed as before, and replace $\gamma \rightarrow \gamma(x) = \bar{\gamma} + \delta\gamma(x)$, and consider the $\lambda > 1$ case. From Fig. 8.2 we find that stable solutions only exist at $\gamma = 0$ for large g , and that they only exist over a narrow range of γ . This means that in a typical system we require large g and small $\langle \delta\gamma^2 \rangle$ for there to be a significantly large regions where pulses are stable. We observe that in such a system there are regions in which we can initiate a travelling pulse which will move with fluctuating speed and width; the pulse cannot propagate into regions in which locally a stable pulse solution does not exist, i.e. a pulse will die if it encounters such a region.

In summary front solutions exist and are stable in a homogeneous system provided there are two stable steady states and γ is greater than some minimum value which depends on the firing threshold θ (Fig. 8.1b). Pulse solutions are only stable for a small range of γ and g . With inhomogeneous connections fronts can still propagate provided there remains two stable steady states, and pulses can be initiated and propagate only in the regions in which they are locally stable. Fluctuations in the speed of fronts and pulses which arise due to stochasticity in firing, rather than in connectivity are studied in Chap. 9 of this book.

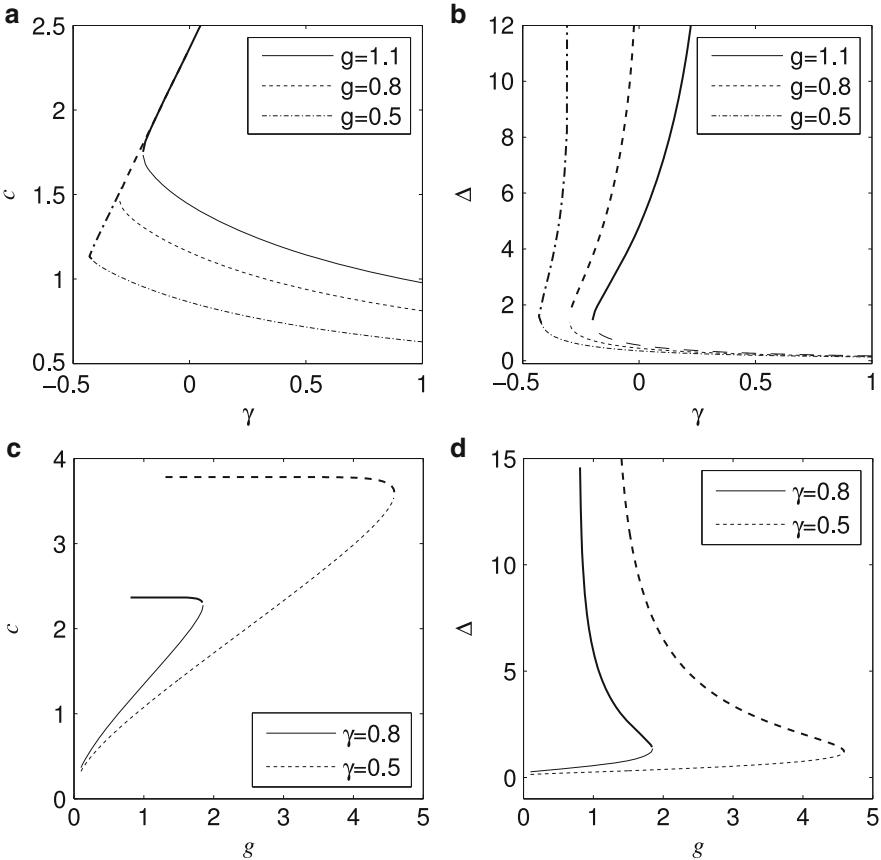


Fig. 8.2 Plots showing how the speed and width of a travelling pulse in the homogeneous system given by Eq. (8.4) vary with the constant γ , for different values of the inhibitory population strength g . The upper branches are found to be stable (heavy lines) and the lower branches are unstable (light lines). Other parameters are $\theta = 0.1$ and $\tau_v = 1$

We have so far examined inhomogeneous systems in which there are either one or two stable steady states everywhere in the system; we now turn our attention to what happens if this is not the case.

If we define the function

$$W(x) = \int_0^L w_H(x - x')[1 + A(w_1(x) + w_2(x'))]dx' - g, \quad (8.14)$$

then from Eq. (8.1) one notes that if $W(x) > \theta$ for all x , then by setting the derivatives to zero and assuming $u(x, t) = u(x)$ the steady states are given by

$$\bar{u}(x) = W(x)f(\bar{u}(x)), \quad (8.15)$$

i.e. $\bar{u}_1(x) = 0$ and $\bar{u}_3(x) = W(x)$. If the condition $W(x) > \theta$ is not true for all x , then in the regions where $W(x) < \theta$ there is only one steady state ($\bar{u}_1 = 0$); for values of x where there are two steady states, the system can exist in the upper state only *locally*. In such a system a new kind of behaviour can be observed. Broadly speaking, we see patches of high activity, patches of low activity and regions which fluctuate between the two. Such persistent fluctuations have not previously been observed in neural field models without the addition of external input.

8.2.2 Persistent Fluctuations

In this section we examine in more detail this new type of fluctuating behaviour. In order to qualitatively understand how the fluctuations arise in systems with stochastic inhomogeneous connections we also examine some deterministic inhomogeneous connection functions. We then quantitatively study the patterns of activity in the stochastic system in the fluctuating state, seeking to understand what this behaviour might mean for information transfer across neural tissue. This involves measurements of the mean and mean squared activity, and spatial and temporal correlations in the activity patterns, and consider how these depend on the properties of the underlying connections. The same fluctuating behaviour is seen in both 1D and 2D systems.

Figure 8.3 shows an example of a 1D system where we observe persistent fluctuations in $u(x, t)$ at some values of x . Also shown is the time evolution of the activity, both at some arbitrarily chosen points in the system, and on average. We find that at some points u appears to oscillate periodically, whilst at others it appears more chaotic. The spatially averaged activity¹ $\langle u(x, t) \rangle_x$ appears to fluctuate chaotically. (In fact a numerical measurement of the Lyapunov exponents of the discretized approximation of the system shows this to be a limit cycle with an extremely long period.)

So, what parameter values are required in order to see the fluctuating behaviour? Firstly, the inhibition strength g must be large enough in order that there are some regions of the system where there are two stable steady states, and some regions where there is only one. We also find that both the amplitude, A , and length scale, λ , of the connections must not be too small, otherwise, for some realisations of the connections, the activity drops to the lower steady state across the whole system. To initiate fluctuations there must be some initial excitation in the regions of the system with two stable steady states, either via the initial conditions or a transient external input. The fluctuations are then seen to persist indefinitely once any external input is removed.

¹Angled brackets and subscripts denote averages. For example $\langle \dots \rangle_{x,t}$ denotes average over both space and time.

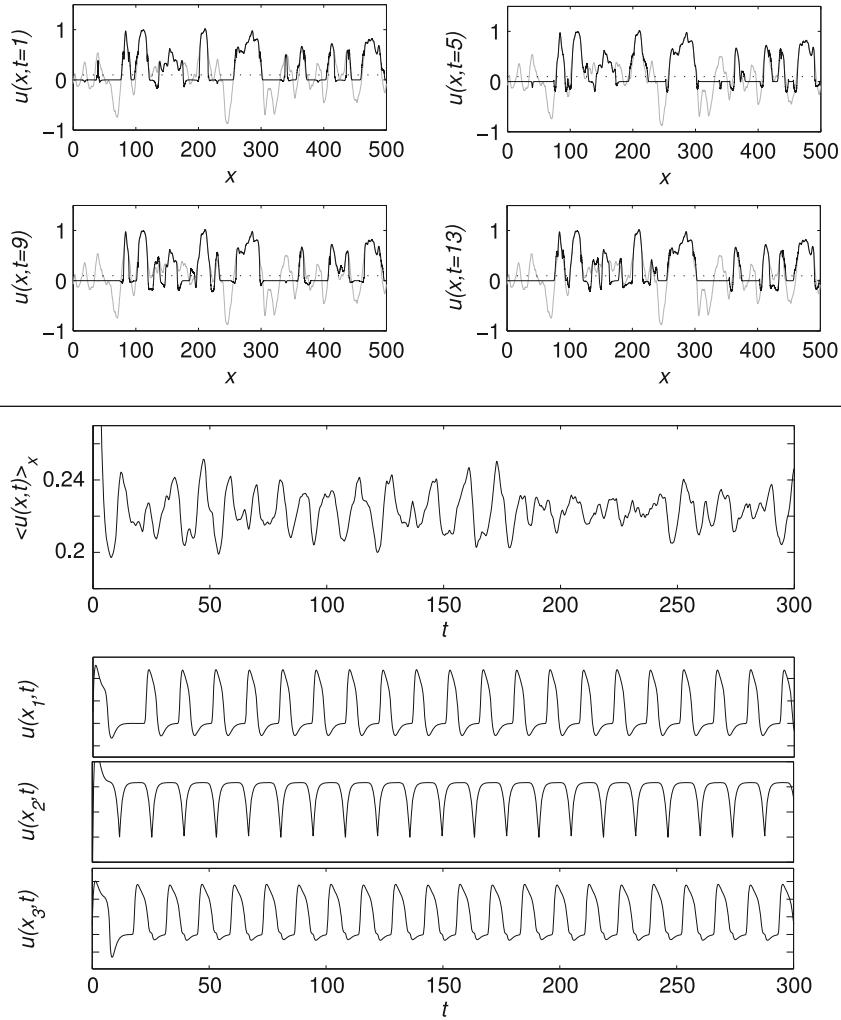


Fig. 8.3 Top: Plots showing snapshots of the activity u as a function of x at different times t , whilst a 1D system is undergoing persistent fluctuations. Black lines show $u(x, t)$, grey lines show $W(x)$ (Eq. (8.14)), and the dotted line shows $\theta = 0.1$. Bottom: Plots showing how u varies in time during persistent fluctuations, at different randomly chosen points x_i and on average. For all plots the other parameters are $g = 0.8$, $\tau_v = 1$, $A = 0.3$ and $\lambda = 5$

The existence of the fluctuating state depends on the particular realisation of the connections, as well as their statistical properties. In order to understand why we *do* or *do not* observe persistent fluctuations for a particular realisation of the inhomogeneity, we examine some carefully chosen *deterministic* inhomogeneous connection functions. Observing the behaviour of this simpler system allows us to more easily understand how the fluctuations arise.

8.2.2.1 Simple Deterministic Connection Function

In 1D, in the large λ limit the field equations can be written

$$\begin{aligned} \tau_u \frac{\partial u}{\partial t} + u(x, t) &= [1 + A(w_1(x) + w_2(x))] \int_0^L w_H(x - x') f(u(x', t)) dx' \\ &\quad - g v(x, t), \\ \tau_v \frac{\partial v}{\partial t} + v(x, t) &= f(u(x, t)), \end{aligned} \tag{8.16}$$

where we have approximated $w_2(x') \approx w_2(x)$, since the width of the function w_H is much less than λ . If instead of using stochastic functions we choose $w_2(x) = 0$, and $w_1(x)$ a piecewise linear function, we can construct a system with a region with two stable steady states ($\bar{u}_1 = 0$, and $\bar{u}_3 = 1$), a region with a single steady state ($\bar{u}_1 = 0$), and a boundary region where $w_1(x)$ has constant gradient. The resulting $W(x)$ as defined by Eq. (8.14) is shown in Fig. 8.4a. If the system is set up with the initial condition $u(x, t) \geq \theta$ for all x , then after a short transient time the result is a region where $u(x, t) = 1$ for all t , a region where $u(x, t) = 0$ for all t , and a boundary region where u fluctuates. That is to say, a “bump” in $w_1(x)$ gives rise to a bump of activity with fluctuating boundaries.

In Fig. 8.4b–e we show how periodic fluctuations of the edges of the activity bump occur. A retracting front of activity forms in the “boundary region”; as this front retracts a small growing “side bump” forms. When the peak of this side bump reaches threshold, u grows rapidly and the direction of movement of the front changes. The front moves out into the region which can only support the $\bar{u}_1 = 0$ steady state before retracting again, and the process repeats. The fluctuations are therefore a consequence of the non-monotonicity of the front at the edges of the highly connected bump, which originates from the presence of the local inhibitory field. A higher density of connections allows the tissue to reside in an active state and this activity spreads into the region with a lower connection density. In this region there are insufficient excitatory connections to sustain the high activity, and as the inhibitory field v increases the front moves back into the highly connected region. The time scale of the fluctuations is determined by the relaxation time of the inhibition τ_v . We also find that the frequency of the fluctuations increases with the gradient of $W(x)$ in the boundary region (as well as depending on θ , τ_v , and g). Also the width of the fluctuating region increases as τ_v increases.

A slightly different choice of $W(x)$ can produce a bump of activity with fluctuating boundaries which emits travelling pulses. As we saw in the previous section there is a narrow range of values of γ for which stable travelling pulse solutions exist. If the region adjacent to the bump has $w_1(x)$ consistent with this, pulses can propagate into it. The rate at which they are emitted depends on the gradient of the edges of the bump in $W(x)$.

Breathing bumps and pulse emitting bumps have also been observed in homogeneous connection models. In Ref. [10] it was shown that stationary fronts and pulses

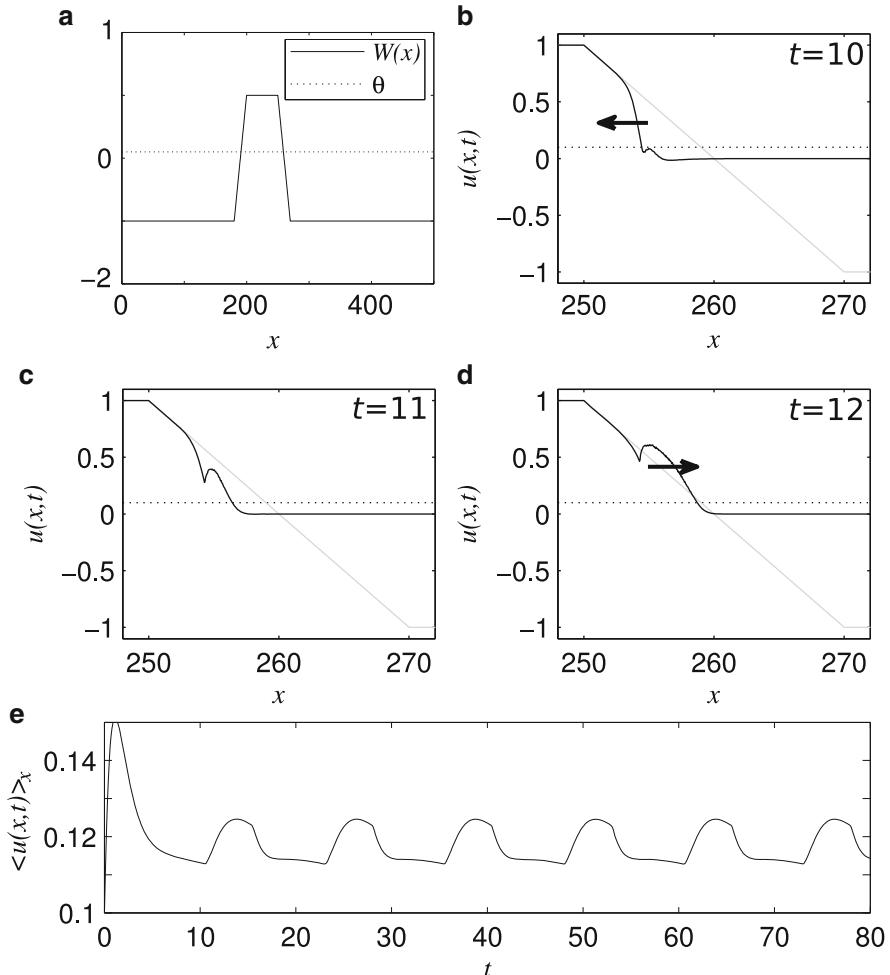


Fig. 8.4 Plot (a) shows an example deterministic $W(x)$ (solid line) with different regions where there are either two or one stable steady states. At the boundary between these regions, $W(x)$ varies continuously and in this example has a constant gradient of 0.1. The dotted line shows the threshold $\theta = 0.1$. Plots (b)–(d) show how the activity fluctuates at the edge of the “bump” in $W(x)$, via the generation of a “side bump”. Solid black lines show $u(x,t)$, grey lines show $W(x)$ and dotted lines show the threshold θ . Arrows show the direction of motion of the front. Plot (e) shows how the mean activity varies with time. After an initial transient there is periodic oscillation

can be generated in a homogeneous connection model with linear feedback via the introduction of a spatially varying external input; if the magnitude of the input is increased, the system goes through a Hopf bifurcation to an oscillating state. In Ref. [15] spike frequency adaption (where the dynamics of the threshold depend on activity) was introduced, giving rise to stationary bumps which go unstable in favour

of breathing solutions. In both of these cases the mechanism for the generation of the oscillations is different to the one described here. Here the oscillations arise as a result of inhomogeneity in the connections.

By examining this simple deterministic inhomogeneity in the connections we obtain a better qualitative understanding of what is taking place in the case of stochastic connection functions. In that case, the observed behaviour is a result of many “breathing bumps” (with characteristic width λ), which fluctuate with a frequency which depends on the local gradient of $W(x)$. These bumps may interact with each other, some emitting pulses. For small A and small λ (i.e. in a regime where regions with $W(x) > \theta$ do not have large spatial extent) the retracting fronts of activity may meet before the “side bumps” have grown to reach threshold. The activity collapses into the lower steady state ($\bar{u}_1 = 0$), and this is typically irreversible. This explains why for some realisations of the stochastic connections (and systems of finite size) we find $u(x, t) \rightarrow 0$ for all x after a short transient time. In general, requirements for fluctuations are that $W(x)$ has non-zero gradient (a requirement for the growth of “side bumps”), that θ must be close enough to $\bar{u}_1 = 0$, and that the bumps in $W(x)$ are wide enough so that the peaks of the side bumps reach threshold before the retracting fronts meet.

8.2.3 Activity Patterns and Correlations

We now return to the case of stochastic connections in 1D, and examine how the properties of $w_1(x)$ and $w_2(x)$ effect the fluctuations. The dependence of the magnitude A and length scale λ of the inhomogeneous connections, on the mean and mean squared activity during persistent fluctuations is shown in Fig. 8.5. As one would expect, increasing A leads to an increase in the mean activity $\langle u(x, t) \rangle_{x,t}$. There is also an increase in the amplitude of the fluctuations, i.e. $\langle u(x, t)^2 \rangle_{x,t}$ increases. In contrast, variation of λ has little impact on the mean and mean squared activity, except at small λ . In that case, the regions where $W(x) > \theta$ are small, and we expect fewer regions where the activity fluctuates; i.e. the majority of the system sits in the lower steady state, leading to a reduced activity on average. We note that for $A \gtrsim 1/\sqrt{2}$ the connection function becomes unphysical as it is likely that $w(x, x')$ will be negative for some x, x' .

The length λ can, however, determine the length and time scales of the fluctuations. We define respectively the spatial and temporal correlation functions

$$C_x(X) = \frac{\langle \delta u_x(x, t) \delta u_x(x + X, t) \rangle_{x,t}}{\langle \delta u_x(x, t)^2 \rangle_{x,t}}, \quad (8.17)$$

$$C_t(\tau) = \frac{\langle \delta u_t(x, t) \delta u_t(x, t + \tau) \rangle_{x,t}}{\langle \delta u_t(x, t)^2 \rangle_{x,t}}, \quad (8.18)$$

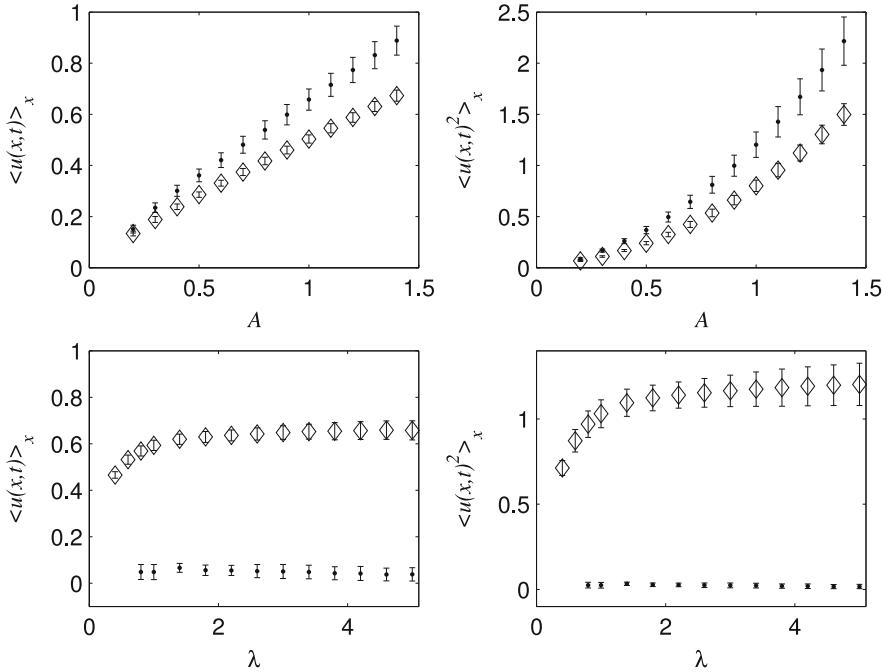


Fig. 8.5 Plots showing how the amplitude A and correlation length λ of the connections effect the spatial mean and mean squared of the activity. Top left and right show the effect of varying A at fixed correlation lengths $\lambda = 5$ (points) and $\lambda = 0.5$ (diamonds). Bottom left and right show the effect of varying λ at fixed $A = 0.1$ (points) and $A = 1$ (diamonds). All results are averages from 10 realisations of the connections, with errors given by the standard deviation over realisations

where $\delta u_x(x, t) = u(x, t) - \langle u(x, t) \rangle_x$ and $\delta u_t(x, t) = u(x, t) - \langle u(x, t) \rangle_t$. As well as averaging over x and t for a single system, we also average over many different simulations with different realisations of the stochastic connections. From the spatial correlation function we can measure a correlation length l , defined to be the length over which $C_x(X)$ drops by a factor e^{-1} . As is shown in Fig. 8.6a–c, for large A the correlation length of the fluctuating activity is close to that of the underlying connection functions.² That is to say, as the strength of the inhomogeneity increases, the activity patterns become more entrained to the structure of the connections.

We turn now to temporal correlations. In the previous subsection we found that for a single “bump” in connection density, the frequency of the fluctuations depends on the gradient of the function $W(x)$. To see whether this is also true in the stochastic case we define a “characteristic gradient” of the inhomogeneous component of the connections

²We note that due to the finite size of the system each realisation of the connections has a correlation length not quite equal to λ , so we also show the measured mean value for the correlation length of the connections.

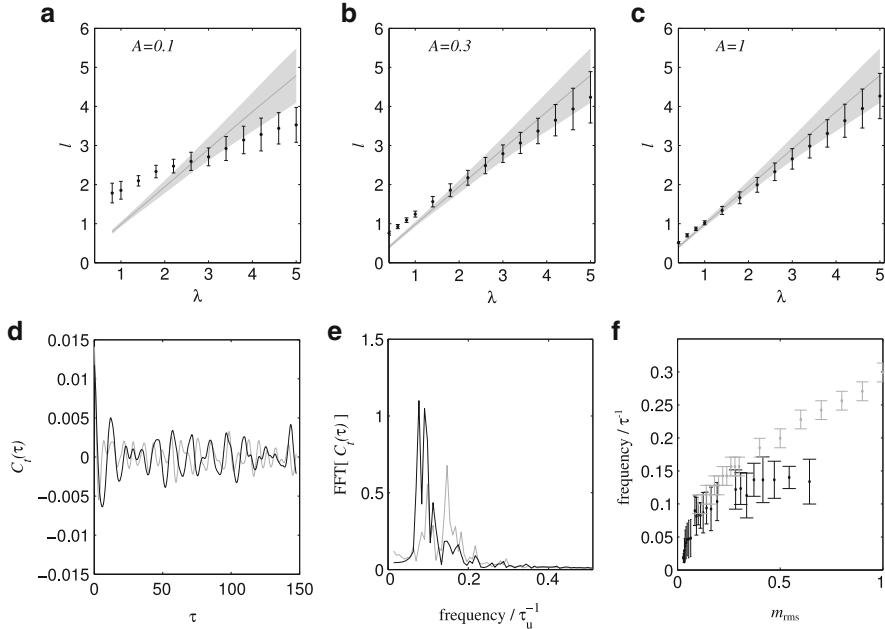


Fig. 8.6 Plots (a)–(c) show how the correlation length of the connections effects the correlation length in $u(x, t)$ for different values of A . Results are averages over 10 realisations of the inhomogeneous connections, with error given by the standard deviation. Also shown (grey line) is the average measured correlation length of the functions $w_1(x)$ and $w_2(x)$; the shaded region shows the standard deviation. Plot (d) shows the time correlation function $C_t(\tau)$ as defined by Eq. (8.18). The grey line is from a system with $\lambda = 0.4$ and the black from $\lambda = 5$, both with $A = 0.3$. Plot (e) shows the fast Fourier transform of the time correlation functions for the same systems. Finally, (f) shows how the characteristic gradient m_{rms} of the connection function effects the principal frequency of the fluctuating activity (black points). We include results from systems with $A = 0.1$, $A = 0.3$ and $A = 1$, with λ varying between 2 and 5 (since our definition of characteristic gradient only makes sense for $\lambda > 1$). For comparison (grey points) we also show the frequency of fluctuations from a system with a deterministic connection function containing a single gradient, as shown in Fig. 8.4a

$$m_{\text{rms}} = A \sqrt{\left\langle \left(\frac{dw_1}{dx} + \frac{dw_2}{dx} \right)^2 \right\rangle}. \quad (8.19)$$

Expanding the square gives

$$\left\langle \left(\frac{dw_1}{dx} + \frac{dw_2}{dx} \right)^2 \right\rangle = \left\langle \left(\frac{dw_1}{dx} \right)^2 + \left(\frac{dw_2}{dx} \right)^2 + 2 \frac{dw_1}{dx} \frac{dw_2}{dx} \right\rangle = 2 \left\langle \left(\frac{dw_1}{dx} \right)^2 \right\rangle,$$

where we have used the fact the $\langle d_x w_1 \rangle = \langle d_x w_2 \rangle$ and $\langle d_x w_1 d_x w_2 \rangle = 0$, because there is no correlation between w_1 and w_2 (d_x denotes derivative with respect to x). From the construction of the functions w_1 and w_2 we have

$$\left\langle \left(\frac{d w_1}{dx} \right)^2 \right\rangle = \frac{1}{\lambda^2}, \quad (8.20)$$

which gives a characteristic gradient of connections in the system of

$$m_{\text{rms}} = \frac{A\sqrt{2}}{\lambda}. \quad (8.21)$$

We note that this quantity is only relevant in the large λ limit, since we have used the approximation $w_2(x') \approx w_2(x)$.

To examine how variation of m_{rms} effects the fluctuations, we consider the Fourier transform of the function $C_t(\tau)$ (Fig. 8.6e), i.e. we look at the frequency of the fluctuations. We take the frequency component with the largest amplitude to be the “principal frequency” of the fluctuations. Figure 8.6f shows how the principal frequency depends on m_{rms} ; also shown is the same measurement for the deterministic inhomogeneity examined in the previous section (Fig. 8.4a). In general as the characteristic gradient of the connections increases, the frequency of the fluctuations increases; for the stochastic connections the frequency reaches a maximum between 0.1 and 0.15 Hz.

8.2.4 Persistent Fluctuations in 2D

In this section we examine results from 2D simulations. Due to the high computational overhead we consider smaller systems and fewer realisations of the connections than in the 1D case. Although this means that the results for the 2D case may be less reliable, reassuringly we see qualitatively the same behaviour as in 1D.

We focus on the regime in which we observe persistent fluctuations; Fig. 8.7 shows snapshots of $u(x, t)$ at different times. We also show the dependence of the correlation length of the stochastic connection functions on the correlations in the activity. The grey solid line shows the measured correlation lengths of the underlying connection functions $w_1(\mathbf{x})$ and $w_2(\mathbf{x}')$. The points show the measured correlation length of the activity, and we see behaviour the same as in the 1D case; as A increases the patterns in activity become slave to the correlations in the underlying connections.

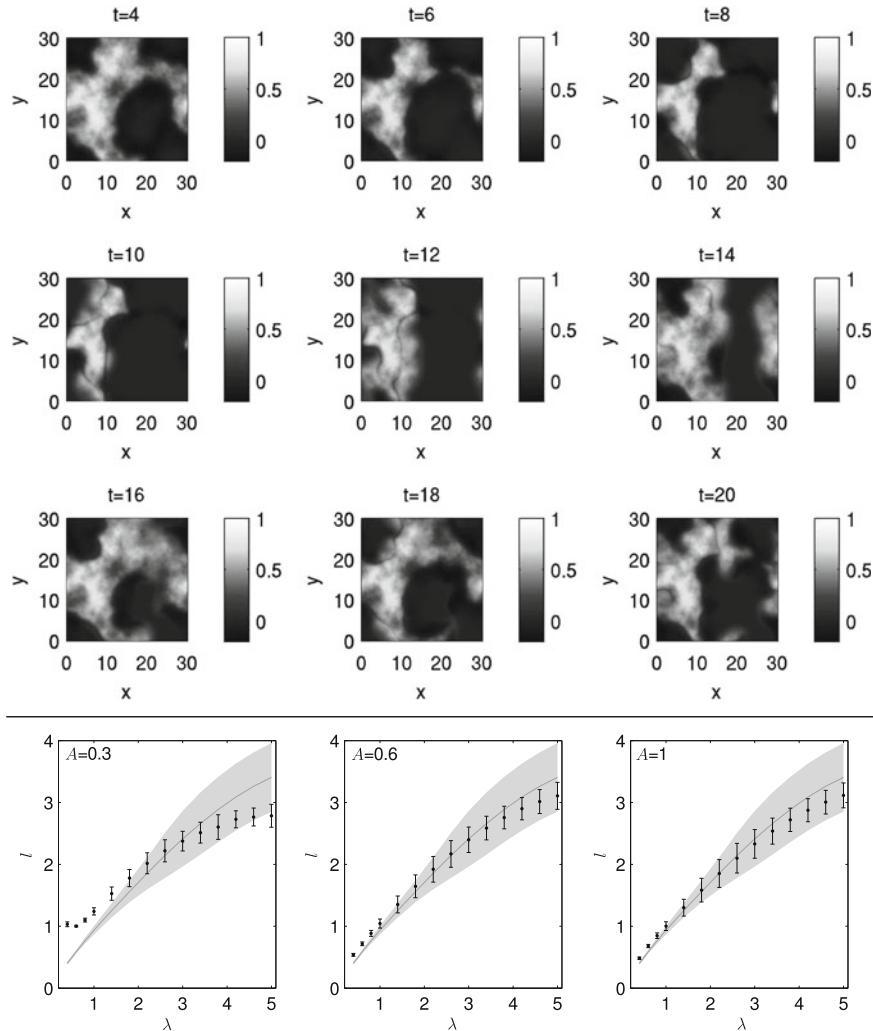


Fig. 8.7 *Top:* Colour plots of $u(x,t)$ for a 2D system undergoing persistent fluctuations. Parameters are $\lambda = 5$, $A = 0.3$, $g = 0.8$, and $\theta = 0.1$. The system is a square of side $L = 30$. *Bottom:* Plot showing how the input correlation length λ of the underlying 2D connections effects the measured correlation length of the fluctuating activity for various A . The grey lines show the measured correlation length of functions $w_1(\mathbf{x})$ and $w_2(\mathbf{x}')$ with the shaded region showing the error in this. All results are averaged over five realisations of the stochastic connections

8.3 Long Range Connections

In the previous section we described a model of inhomogeneous stochastic connections, however these have all been local in nature. That is to say, there are no connections between distant points, only spatial variation in the local density

of connections. As detailed in Ref. [4], stochastic *long range* connections can be introduced into Eq. (8.1) via the following connection function

$$w(\mathbf{x}, \mathbf{x}') = w_H(\mathbf{x} - \mathbf{x}') + Aw_I(|\mathbf{x} - \mathbf{x}'|) [w_1(\mathbf{x}) + w_2(\mathbf{x}')]. \quad (8.22)$$

As before w_H is a homogeneous Gaussian connection function (Eq. (8.3)), and $w_1(\mathbf{x})$ and $w_2(\mathbf{x}')$ are numerically generated stochastic functions representing additional connections into and out of the tissue at point \mathbf{x} and \mathbf{x}' respectively. This time however we choose these functions such that they are always positive. The function w_I can be thought of as an envelope for the inhomogeneous connections; we choose a power law function

$$w_I(y) = \mathcal{N} \frac{|y|^{-\alpha}}{1 + |y|^{-\alpha}}, \quad (8.23)$$

where the exponent α determines the range of the connections, and \mathcal{N} is chosen such that $\int_{\Gamma} w_I(y) dy = 1$. If $\alpha > 1$, w_I is a narrow function, and we recover a local connection model; for $\alpha < 1$, w_I is a wide function and can extend throughout the system, i.e. there are a small number of long range connections between distant points. Figure 8.8a–c show typical realisations of $w(x, x')$ in 1D for different values of α .

In the large α (local connections) regime, wave front solutions exist as in Sect. 8.2, provided two stable steady states solutions exist across the whole system. These are given by

$$\bar{u}(\mathbf{x}) = W(\mathbf{x})f(\bar{u}), \quad (8.24)$$

where we now define

$$W(\mathbf{x}) = 1 + A \int_{\Gamma} w_I(\mathbf{x} - \mathbf{x}') [w_1(\mathbf{x}) + w_2(\mathbf{x}')] d\mathbf{x}' - g. \quad (8.25)$$

For the small α (long range connections) regime activity no longer propagates throughout the system at a finite speed: due the long range connections a local external stimulus can lead to the system entering the upper steady state at all points in the system.

As with the model described in the previous section, for some choice of parameters A and g , the system can support persistent fluctuations of activity. For the case of local connections (large α) fluctuations arise when there are some regions of the system which have two stable steady states, and other regions in which there is only one. There are regions in which $u \rightarrow \bar{u}_1 = 0$, regions in which $u \rightarrow \bar{u}_3$, and connecting regions where u fluctuates.

In the regime of long range connections (small α) there is slightly different behaviour. There are fluctuations of activity at every point throughout the system. There are large large regions (where $W(\mathbf{x}) > \theta$) in which u fluctuates coherently

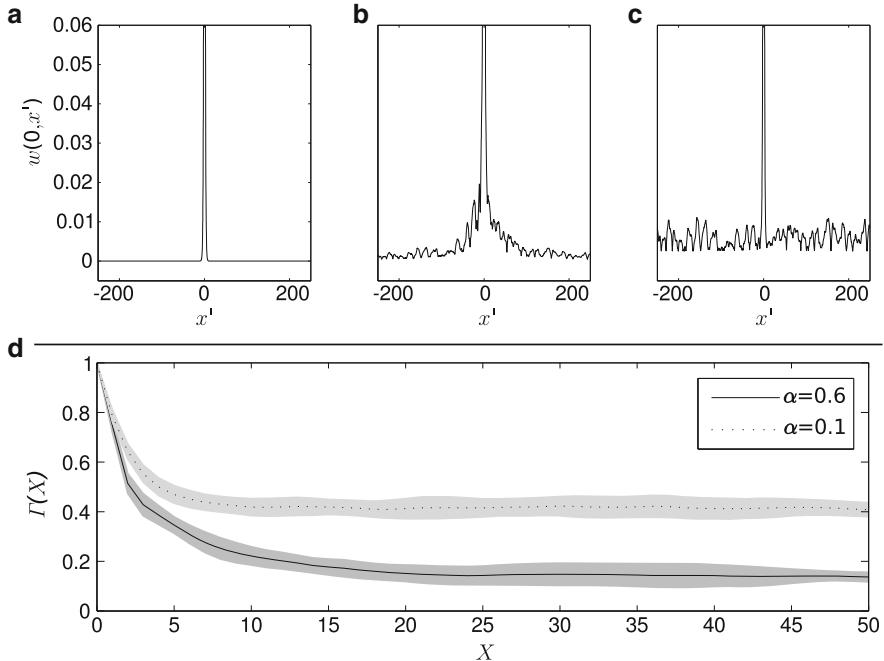


Fig. 8.8 Top: Plots showing the connection function in Eq. (8.22) for a 1D system with (a) $\alpha = 6$, (b) $\alpha = 1$, and (c) $\alpha = 0.1$. Note that the homogeneous peak of w_H is present in each case. In (c) the long range “tails” of the inhomogeneous connections have a small amplitude due to the normalisation \mathcal{N} . Bottom: Average coherence (see Eq. (8.26)) of a 1D system with $A = 1.9$ and $g = 2.9$, averaged over 10 realisations of the connections for $\alpha = 0.6$ (solid line) and $\alpha = 0.1$ (dotted line). The standard deviation is shown by the shaded regions (Figure adapted from Ref. [4])

remaining above θ , and smaller regions where u fluctuates more quickly about θ . As detailed in Ref. [4] the fluctuations originate in regions where $W(\mathbf{x})$ cuts through the threshold θ , much as for the model discussed in the previous section. Here though, due to long range connections spanning the length of the system, this also leads to fluctuations at all other points.

In order to quantify the “coherently” fluctuating regions we use the following quantity to characterize fluctuations at points x and x' :

$$\Gamma(X) = \left\langle \frac{\langle u(x, t)u(x - X, t) \rangle_t^2}{\langle u(x, t)^2 \rangle_t \langle u(x - X, t)^2 \rangle_t} \right\rangle_x, \quad (8.26)$$

where here $X = x - x'$, and as before angled brackets with subscripts denote average over space or time. A value of $\Gamma(X) = 1$ means that points separated by a distance X have activity which is fluctuating synchronously; a small value of $\Gamma(X)$ means that fluctuations in these regions are fluctuating interdependently. Figure 8.8 shows that in a system with $\alpha < 1$ the long range connections give rise to fluctuations

which are coherent over large regions of the system. Similar results are obtained for 2D systems.

8.4 Conclusions

We have demonstrated that travelling front solutions in a continuum neural field model are robust to the addition of small amplitude, inhomogeneous, stochastic (local) connections, provided there is only weak inhibitory feedback. The wavefront connects a region in a quiescent $\bar{u}_1 = 0$ steady state with a region in a spatially varying $\bar{u}_3 = \bar{u}_3(x)$ steady state, and travels with a time varying velocity. This provides a mechanism for the fluctuation in speed of travelling fronts of activity such as observed in “1D” slices of neural tissue (like those studied in Ref. [11]). Via a simple expansion $\gamma(x) = \bar{\gamma} + \delta\gamma(x)$ predictions can be made about the dependence of the resulting distribution of front speeds on the magnitude of the inhomogeneity. The mean speed of a front of activity remains largely unaffected by the inhomogeneity. The variance of the speed is independent of A when the inhomogeneity is correlated on lengths shorter than 1, and grows with A^2 if the connections are correlated on lengths longer than 1 (in units of the width of the homogeneous connections). If the magnitude of the inhomogeneity A or the strength of the inhibition g are large enough to destabilise the upper steady state (i.e. $W(x) < \theta$ for some x), then there will be regions of the system through which the front cannot propagate.

Persistent fluctuations of activity are a new type of behaviour for continuum neural field models. By studying a simple deterministic connection function we have identified the origin of the fluctuations, and explained why such behaviour is observed for some realisations of the connections, but not others. All of the results presented in this chapter arise from initial conditions where $u > \theta$ across the entire system, meaning that the system will show persistent fluctuations if it is able to support them. A more natural scenario would be an initially quiescent system ($u = 0$), which is then excited by a transient, possibly spatially heterogeneous, external input. This could be likened to an idea initially suggested by Hebb in Ref. [20] known as cortical reverberation; Hebb hypothesised that a particular weak input pattern might elicit a large persistent response from a system, whereas a different, stronger, input pattern may have little effect. In terms of the present model, a short lived spatially varying input pattern may excite some regions of the system into the persistently fluctuating state; a different input (for example to a region where there is no upper steady state) may only excite the system for the duration of the input. This can also be linked to models of working memory [30], where a localised stimulus excites a region into an active state, and this high activity is maintained after the stimulation has ceased.

By examining correlations in the fluctuating activity, we find that, except for very small values of λ and A , the patterns follow the underlying connection functions w_1 and w_2 . We expect this in the large λ regime since we can make the

approximation that $w_2(x')$ does not vary within the width of the inhomogeneous connections, and $u(x, t)$ will closely follow $W(x)$. A natural question to ask is “Can we say anything about connectivity from measuring spatial correlations in activity patterns?” This could lead to testable predictions in experimental work, for example using fluorescent dyes [3].

In two dimensional systems with local inhomogeneous connections we find persistent fluctuations which are qualitatively the same as in 1D. Aside from the simple extensions to planar fronts and pulses, one could study, for example, the propagation of high activity from a locally excited region. In a homogeneous system the 2D analogue of planar fronts is an expanding circular region of high activity [23]; in an inhomogeneous model the high activity region would likely be irregularly shaped, and there could be, for example, channels of higher connectivity down which activity could propagate more rapidly, or conversely “barrier” regions with lower connectivity.

Finally we have shown that persistent fluctuations can also be found in systems with long range inhomogeneous connections. By allowing additional connections over long distances, fluctuations can become synchronised in different regions.

Acknowledgements The authors would like to thank George Rowlands, Magnus Richardson and Pierre Sens for helpful discussions.

References

1. Amari, S.-I.: Homogeneous nets of neuron-like elements. *Biol. Cybern.* **17**, 211–220 (1975)
2. Amari, S.-I.: Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.* **27**, 77–87 (1977)
3. Bao, W., Wu, J.-Y.: Propagating wave and irregular dynamics: spatiotemporal patterns of cholinergic theta oscillations in neocortex in vitro. *J. Neurophysiol.* **90**, 333–341 (2003)
4. Brackley, C.A., Turner, M.S.: Persistant fluctuations of activity in un-driven continuum neural field models with power law connections. *Phys. Rev. E* **79**, 011918 (2009)
5. Brackley, C.A., Turner, M.S.: Two-point heterogeneous connections in a continuum neural field model. *Biol. Cybern.* **100**, 371 (2009)
6. Bressloff, P.C.: New mechanism for neural pattern formation. *Phys. Rev. Lett.* **76**, 0031–9007 (1996)
7. Bressloff, P.C.: Traveling fronts and wave propagation failure in an inhomogeneous neural network. *Phys. D Nonlinear Phenom.* **155**(1–2), 83–100 (2001)
8. Bressloff, P.C.: Spatially periodic modulation of cortical patterns by long-range horizontal connections. *Phys. D Nonlinear Phenom.* **185**(3–4), 131–157 (2003)
9. Bressloff, P.C.: From invasion to extinction in heterogeneous neural fields. *J. Math. Neurosci.* **2**(1), 6 (2012)
10. Bressloff, P.C., Folias, S.E., Prat, A., Li, Y.-X.: Oscillatory waves in inhomogeneous neural media. *Phys. Rev. Lett.* **91**(17), 178101 (2003)
11. Cherkin, R.D., Pierce, P.A., Connors, B.W.: Periodicity and directionality in the propagation of epileptiform discharges across neocortex. *J. Neurophysiol.* **60**(5), 1695–1713 (1988)
12. Coombes, S.: Waves, bumps, and patterns in neural field theories. *Biol. Cybern.* **93**, 91–108 (2005)

13. Coombes, S., Laing, C.R.: Pulsating fronts in periodically modulated neural field models. *Phys. Rev. E* **83**, 011912 (2011)
14. Coombes, S., Owen, M.R.: Evans functions for integral neural field equations with heaviside firing rate function. *SIAM J. Appl. Dyn. Syst.* **34**, 574–600 (2004)
15. Coombes, S., Owen, M.R.: Bumps, breathers, and waves in a neural network with spike frequency adaption. *Phys. Rev. Lett.* **94**, 148102 (2005)
16. Coombes, S., Lord, G.J., Owen, M.R.: Waves and bumps in neuronal networks with axo-dendritic synaptic interactions. *Phys. D Nonlinear Phenom.* **178**, 219–241 (2003)
17. Folias, S.E., Bressloff, P.C.: Breathing pulses in an excitatory neural network. *SIAM J. Appl. Dyn. Syst.* **3**, 378–407 (2004)
18. Folias, S.E., Bressloff, P.C.: Breathers in two-dimensional neural media. *Phys. Rev. Lett.* **95**, 208107 (2005)
19. Frigo, M., Johnson, S.G.: The design and implementation of FFTW3. *Proc. IEEE* **93**(2), 216–231 (2005). Special issue on “Program Generation, Optimization, and Platform Adaptation”
20. Hebb, D.O.: *The Organization of Behavior*, 1st edn. Wiley, New York (1949)
21. Hellwig, B.: A quantitative analysis of the local connectivity between pyramidal neurons in layers 2/3 of the rat visual cortex. *Biol. Cybern.* **82**, 111–121 (2000)
22. Hutt, A., Bestehorn, M., Wennekers, T.: Pattern formation in intracortical neuronal fields. *Netw. Comput. Neural Syst.* **14**, 351–368 (2003)
23. Idiart, M.A.P., Abbott, L.F.: Propagation of excitation in neural newtork models. *Network* **4**, 285–294 (1993)
24. Jirsa, V.K., Kelso, J.A.S.: Spatiotemporal pattern formation in neural systems with heterogeneous connection topologies. *Phys. Rev. E* **62**(6), 8462–8465 (2000)
25. Kilpatrick, Z., Folias, S., Bressloff, P.: Traveling pulses and wave propagation failure in inhomogeneous neural media. *SIAM J. Appl. Dyn. Syst.* **7**(1), 161–185 (2008)
26. Pinotsis, D.A., Hansen, E., Friston, K.J., Jirsa, V.: Anatomical connectivity and the resting state activity of large cortical networks. *Neuroimage* **65**(4), 127–138 (2013)
27. Pinto, D.J., Ermentrout, G.B.: Spatially structured activity in synaptically coupled neuronal networks: I. Travelling fronts and pulses. *SIAM J. Appl. Math.* **62**, 206–225 (2001)
28. Pinto, D.J., Ermentrout, G.B.: Spatially structured activity in synaptically coupled neuronal networks: II. Lateral inhibition and standing pulses. *SIAM J. Appl. Math.* **62**, 226–243 (2001)
29. Qubbaj, M.R., Jirsa, V.K.: Neural field dynamics under variation of local and global connectivity and finite transmission speed. *Physica D* **238**, 2331–2346 (2009)
30. Wang, X.-J.: Synaptic reverberation underlying mnemonic persistent activity. *Trends Neurosci.* **24**, 455–463 (2001)

Chapter 9

Stochastic Neural Field Theory

Paul C. Bressloff

Abstract We survey recent work on extending neural field theory to take into account synaptic noise. We begin by showing how mean field theory can be used to represent the macroscopic dynamics of a local population of N spiking neurons, which are driven by Poisson inputs, as a rate equation in the thermodynamic limit $N \rightarrow \infty$. Finite-size effects are then used to motivate the construction of stochastic rate-based models that in the continuum limit reduce to stochastic neural fields. The remainder of the chapter illustrates how methods from the analysis of stochastic partial differential equations can be adapted in order to analyze the dynamics of stochastic neural fields. First, we consider the effects of extrinsic noise on front propagation in an excitatory neural field. Using a separation of time scales, it is shown how the fluctuating front can be described in terms of a diffusive-like displacement (wandering) of the front from its uniformly translating position at long time scales, and fluctuations in the front profile around its instantaneous position at short time scales. Second, we investigate rare noise-driven transitions in a neural field with an absorbing state, which signals the extinction of all activity. In this case, the most probable path to extinction can be obtained by solving the classical equations of motion that dominate a path integral representation of the stochastic neural field in the weak noise limit. These equations take the form of nonlocal Hamilton equations in an infinite-dimensional phase space.

P.C. Bressloff (✉)

Department of Mathematics, University of Utah, 1400 East 155 South, Salt Lake City,
UT 84112, USA

e-mail: bressloff@math.utah.edu

9.1 Introduction

The spike trains of individual cortical neurons *in vivo* tend to be very noisy, having interspike interval (ISI) distributions that are close to Poisson [58]. The main source of intrinsic fluctuations at the single cell level is channel noise, which arises from the variability in the opening and closing of a finite number of ion channels [57]. On the other hand, extrinsic fluctuations in membrane voltage are predominantly due to synaptic noise. That is, cortical neurons are bombarded by thousands of synaptic inputs, many of which are not correlated with a meaningful input and can thus be treated as background synaptic noise [31]. It is not straightforward to determine how noise at the single cell level translates into noise at the population or network level. One approach is to formulate the dynamics of a population of spiking neurons in terms of the evolution of the probability density of membrane potentials – the so-called population density method [1, 16, 17, 21, 36, 39, 46, 48, 52, 53]. Usually, a simple model of a spiking neuron is used, such as the integrate-and-fire model [35], and the network topology is assumed to be either fully connected or sparsely connected. It can then be shown that under certain conditions, even though individual neurons exhibit Poisson-like statistics, the neurons fire asynchronously so that the total population activity evolves according to a mean-field rate equation with a characteristic activation or gain function. This gain firing rate function can then be used to construct rate-based models. Formally speaking, the asynchronous state only exists in the thermodynamic limit $N \rightarrow \infty$, where N determines the size of the population. This then suggests a possible source of intrinsic noise at the network level arises from fluctuations about the asynchronous state due to finite size effects [7, 37, 41, 44, 59].

In this chapter we review recent work on incorporating synaptic noise into continuum neural fields, and methods for analyzing the resulting stochastic dynamics. Neural field equations can be derived under two basic assumptions: (i) the spike trains of individual neurons are decorrelated (asynchronous) so that the total synaptic input to a neuron is slowly varying and deterministic, and (ii) there exists a well-defined continuum limit of the resulting network rate equations. So far there has been no rigorous proof that either of these assumptions hold in large-scale spiking network models of cortex. In particular, there has been no systematic scaling up of population density methods to derive continuum neural field models that take proper account of noise-induced fluctuations and statistical correlations between neurons at multiple spatial and temporal scales. Consequently, current formulations of stochastic neural field theory tend to be phenomenologically based. One approach is to consider a Langevin version of the deterministic neural field equations involving some form of extrinsic spatiotemporal white noise [15, 32, 38], whereas another is to treat the neural field equations as the thermodynamic limit of an underlying master equation [9, 10, 19, 20]. In the latter case, a diffusion approximation leads to an effective Langevin equation with multiplicative noise.

In order to motivate the construction of stochastic neural field equations, we begin by reviewing the population density method for analyzing the stochastic

dynamics of a local population of leaky integrate-and-fire neurons (Sect. 9.2), following along similar lines to Gerstner and Kistler [35]. We then describe how finite-size effects at the level of local populations can be incorporated into stochastic versions of rate-based models (Sect. 9.3). In Sect. 9.4 we show how stochastic PDE methods can be used to analyze traveling waves in stochastic neural fields. Finally, we derive a path integral formulation of a stochastic neural field and show how this can be used to analyze rare events (Sect. 9.5). Note that a related survey of stochastic neural fields can be found in Ref. [12].

9.2 Population Density Method and Mean Field Theory

Integrate-and-fire (IF) neuron models neglect details regarding the spike generation process by reducing the latter to an all-or-nothing threshold event. That is, whenever the membrane potential crosses a firing threshold, the neuron fires a spike, typically modeled as a Dirac delta function, and the membrane potential is reset to some subthreshold value. Although they are less realistic than conductance-based models, they have provided a very useful platform for exploring probabilistic models of spiking neurons [35]. Consider a homogeneous population of N identical uncoupled leaky IF neurons. The membrane potential (internal state) of the i th neuron evolves according to the equation

$$C \frac{dV_i}{dt} = -\frac{V_i(t)}{R} + I_i(t) \quad (9.1)$$

where C, R are the capacitance and resistance of the cell membrane and $I_i(t)$ represents the sum of synaptic and external currents. The form of the action potential is not described explicitly. Spikes are formal events characterized by the ordered sequence of firing times $\{T_i^m, m \in \mathbb{Z}\}$ determined by the threshold crossing conditions

$$T_i^m = \inf\{t, t > T^{m-1} | V_i(t) = \kappa, \dot{V}_i > 0\}, \quad (9.2)$$

where κ is the firing threshold. Immediately after firing, the potential is reset to a value $V_r < \kappa$,

$$\lim_{t \rightarrow [T_i^m]_+} V_i(t) = V_r. \quad (9.3)$$

For simplicity we set $V_r = 0$, $CR = \tau$ and absorb C into $I_i(t)$.

Suppose that all neurons in the population receives the same driving current I^{ext} . Each neuron also receives a stochastic background input consisting of a set of spike trains stimulating different synapses labeled $k = 1, \dots, K$. An input spike of the k th synapse causes a jump of the membrane potential by an amount w_k , and the spikes are generated by a time-dependent Poisson process at a rate $v_k(t)$. This means that in each small time interval $[t, t + \Delta t]$ the probability that a spike

arrives on the k th synapse is $v_k(t)\Delta t$, and each spike is uncorrelated with any other. Although the spike rates $v_k(t)$ are the same for all neurons, the actual spike trains innervating different synapses of different neurons are statistically independent. Let $p(v, t)$ denote the membrane potential density, that is, $p(v, t)dv$ is the fraction of neurons i that have a membrane potential $v \leq V_i(t) \leq v + dv$ in the limit $N \rightarrow \infty$. Whenever a neuron fires its membrane is reset. This has two important implications. First, conservation of probability implies that

$$\int_{-\infty}^{\kappa} p(v, t)dv = 1. \quad (9.4)$$

Second, the fraction of neurons (probability flux $J(\kappa, t)$) that flow across the threshold κ per unit time determines the population activity $A(t)$:

$$J(\kappa, t) = A(t) \equiv \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N \sum_m \delta(t - T_j^m). \quad (9.5)$$

The reset condition means that neurons that “disappear” across threshold “re-enter” the population at the reset potential $v = 0$. This implies that there is an absorbing boundary condition at threshold

$$p(\kappa, t) = 0 \quad (9.6)$$

and a discontinuity in the probability fluxes at reset:

$$J(0^+, t) - J(0^-, t) = A(t). \quad (9.7)$$

The probability density $p(v, t)$ evolves according to the Chapman-Kolmogorov equation [35]¹

$$\frac{\partial p}{\partial t} = \frac{1}{\tau} \frac{\partial}{\partial v} [vp(v, t) + RI_{\text{ext}}p(v, t)] + \sum_k v_k(t)[p(v - w_k, t) - p(v, t)]. \quad (9.8)$$

¹Equation (9.8) and various generalizations have been used to develop numerical schemes for tracking the probability density of a population of synaptically coupled spiking neurons [46, 48], which in the case of simple neuron models, can be considerably more efficient than classical Monte Carlo simulations that follow the states of each neuron in the network. On the other hand, as the complexity of the individual neuron model increases, the gain in efficiency of the population density method decreases, and this has motivated the development of a moment closure scheme that leads to a Boltzmann-like kinetic theory of IF networks [21, 52]. However, as shown in Ref. [39], considerable care must be taken when carrying out the dimension reduction, since it can lead to an ill-posed problem over a wide range of physiological parameters. That is, the truncated moment equations may not support a steady-state solution even though a steady-state probability density exists for the full system. An alternative approach is to develop a mean field approximation as highlighted here.

In the limit of small jump amplitudes w_k , Eq. (9.8) can be approximated by a diffusion equation. That is, Taylor expanding to second order in w_k gives

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial v} \left[\frac{-v}{\tau} + \mu(t) + I_{\text{ext}}(t) \right] p(v, t) + \frac{\sigma^2(t)}{2} \frac{\partial^2}{\partial v^2} p(v, t), \quad (9.9)$$

where $\mu(t)$ is the mean background synaptic input

$$\mu(t) = \sum_k v_k(t) w_k, \quad (9.10)$$

and $\sigma(t)$ determines the size of the membrane fluctuations,

$$\sigma^2(t) = \sum_k v_k(t) w_k^2. \quad (9.11)$$

The Fokker-Planck equation determines the time evolution of the probability density of a membrane potential evolving according to the equivalent stochastic differential equation (Langevin equation)

$$dV = -\frac{V(t)}{\tau} dt + \mu(t) dt + \sigma(t) dW(t), \quad (9.12)$$

where $W(t)$ is a Wiener process,

$$\langle dW(t) \rangle = 0, \quad \langle dW(t) dW(t) \rangle = dt \quad (9.13)$$

In the case of constant rates v_k , the resulting Langevin equation describes the well known Ornstein-Uhlenbeck process.

The Fokker-Planck equation (9.9) can be rewritten as a continuity equation reflecting conservation of probability:

$$\frac{\partial}{\partial t} p(v, t) = -\frac{\partial}{\partial v} J(v, t), \quad \text{for } v \neq \kappa, 0, \quad (9.14)$$

where

$$J(v, t) = \frac{1}{\tau} [-v + \mu(t) + I_{\text{ext}}(t)] p(v, t) - \frac{\sigma^2(t)}{2} \frac{\partial}{\partial v} p(v, t). \quad (9.15)$$

Equations (9.5) and (9.15) together with the absorbing boundary condition (9.6) implies that

$$\frac{\partial}{\partial v} p(\kappa, t) = -\frac{2A(t)}{\sigma^2(t)}. \quad (9.16)$$

Continuity of p at reset, $p(0^+, t) = p(0^-, t)$, together with Eqs. (9.7) and (9.15) show that there is a discontinuity in the first derivative of $p(v, t)$ at $v = 0$:

$$\frac{\partial}{\partial v} p(0^+, t) - \frac{\partial}{\partial v} p(0^-, t) = -\frac{2A(t)}{\sigma^2(t)}. \quad (9.17)$$

Hence, one has to solve the Fokker–Planck equation (9.9) together with the boundary conditions (9.6), (9.16), and (9.17).

Now suppose that the background rates v_k and external input I_{ext} are time-independent so that the total mean input

$$I_0 = I_{\text{ext}} + \sum_k v_k w_k \quad (9.18)$$

is a constant. The steady-state Fokker–Planck equation implies that the flux

$$J(v) = (-v/\tau + I_0)p(v) - \frac{\sigma_0^2}{2} \frac{\partial}{\partial v} p(v), \quad (9.19)$$

with $\sigma_0 = \sum_k v_k w_k^2$ for time-independent rates, is constant except at $v = 0$ where it jumps by an amount A_0 , which is the steady-state population activity. Taking $J(v) = 0$ for $v < 0$, one can solve Eq. (9.19) to obtain the Gaussian distribution

$$p_0(v) = \frac{c_1}{\sigma_0} \exp\left[-\frac{(v/\tau - I_0)^2}{\sigma_0^2}\right], \quad \text{for } v \leq 0 \quad (9.20)$$

for some constant c_1 . However, such a solution cannot be valid for $v > 0$, since it does not satisfy the absorbing boundary condition $p_0(\kappa) = 0$. It turns out that in this domain the solution is of the form [17, 35]

$$p_0(v) = \frac{c_2}{\sigma_0^2} \exp\left[-\frac{(v/\tau - I_0)^2}{\sigma_0^2}\right] \int_v^\kappa \exp\left[\frac{(x/\tau - I_0)^2}{\sigma_0^2}\right] dx, \quad \text{for } 0 < v \leq \kappa \quad (9.21)$$

for some constant c_2 . Equation (9.19) shows that $c_2 = 2J(v)$ for $0 < v \leq \kappa$ with $J(v) = A_0$. Continuity of the solution at $v = 0$ implies that

$$c_1 = \frac{c_2}{\sigma_0} \int_v^\kappa \exp\left[\frac{(x/\tau - I_0)^2}{\sigma_0^2}\right] dx. \quad (9.22)$$

Finally, the constant c_2 is determined by the normalization condition for p . On setting $A_0 = c_2/2\kappa$, one finds a firing rate

$$A_0 = \left[\tau \sqrt{\pi} \int_{-\sqrt{\tau} I_0 / \sigma_0}^{(\kappa - \tau I_0) / \sqrt{\tau} \sigma_0} e^{v^2} (1 + \operatorname{erf}(v)) dv \right]^{-1} \equiv F(I_0), \quad (9.23)$$

where F is the so-called gain function for the population [17, 35, 46].

The above analysis assumed that the neurons were independent of each other so that the only synaptic inputs were from some stochastic background. Now suppose that we have a fully connected network such that there is an additional contribution to the synaptic input into each neuron of the form

$$I_{\text{rec}}(t) = \frac{\Gamma_0}{N} \sum_{j=1}^N \sum_m \delta(t - T_j^m) = \Gamma_0 A(t), \quad (9.24)$$

where Γ_0/N is the strength of connection between any pair of neurons within the population, and we have used the definition (9.5) of the population activity $A(t)$. Suppose that the neuronal population is in a macroscopic state with constant activity $A(t) = A_0$, which is referred to as a state of asynchronous firing. (Formally speaking, such an asynchronous state only makes sense in the thermodynamic limit $N \rightarrow \infty$.) The steady-state activity can then be determined self-consistently from Eq. (9.23) by setting

$$I_0 = I_{\text{ext}} + \left[\sum_k v_k w_k + \Gamma_0 A_0 \right], \quad (9.25)$$

and solving the implicit equation for A_0 in terms of $I_{\text{ext}} + \sum_k v_k w_k$, which leads to an effective gain function $A_0 = F_{\text{eff}}(I_{\text{ext}} + \sum_k v_k w_k)$. One can also determine the stability of the asynchronous state by considering small perturbations of the steady-state probability distribution. One finds that in the limit of low noise, the asynchronous state is unstable and the neurons tend to split up into several synchronized groups that fire alternately. The overall activity then oscillates several times faster than the individual neurons [17, 36, 63]. One of the interesting properties of the asynchronous state from a computational perspective is that the population activity can respond rapidly to a step input [35]. The basic intuition behind this is that in the asynchronous state there will always be a significant fraction of neurons that are sitting close to the firing threshold so that as soon as a step increase in input current occurs they can respond immediately. However, the size of the step has to be at least as large as the noise amplitude σ , since the threshold acts as an absorbing boundary, that is, the density of neurons vanishes as $v \rightarrow \kappa$.

So far noise has been added explicitly in the form of stochastic background activity as described in the paragraph below equation (3). It is also possible for a network of deterministic neurons with fixed random connections to generate its own noise [3, 16, 17, 64]. In particular, suppose that each neuron in the population of N neurons receives input from C randomly selected neurons in the population with $C \ll N$. The assumption of sparse connectivity means that two neurons

share only a small number of common inputs. Hence, if the presynaptic neurons fire stochastically then the input spike trains that arrive at distinct postsynaptic neurons can be treated as statistically independent. Since the presynaptic neurons belong to the same population, it follows that each neuron's output should itself be stochastic in the sense that it should have a sufficiently broad distribution of interspike intervals. This will tend to occur if the neurons operate in a subthreshold regime, that is, the mean total input is below threshold so that threshold crossings are fluctuation driven.

9.3 Stochastic Rate-Based Models

Now suppose that a network of synaptically coupled spiking neurons is partitioned into a set of P homogeneous populations with $N_\alpha = \delta_\alpha N$ neurons in each population, $\alpha = 1, \dots, P$. Let p denote the population function that maps the single neuron index $i = 1, \dots, N$ to the population index α to which neuron i belongs: $p(i) = \alpha$. Furthermore, suppose the synaptic interactions between populations are the same for all neuron pairs. (Relaxing this assumption can lead to additional sources of stochasticity as explored in Refs. [32, 61].) Denote the sequence of firing times of the j th neuron by $\{T_j^m, m \in \mathbb{Z}\}$. The net synaptic current into postsynaptic neuron i due to stimulation by the spike train from presynaptic neuron j , with $p(i) = \alpha, p(j) = \beta$, is taken to have the general form $N_\beta^{-1} \sum_m \Phi_{\alpha\beta}(t - T_j^m)$, where $N_\beta^{-1} \Phi_{\alpha\beta}(t)$ represents the temporal filtering effects of synaptic and dendritic processing of inputs from any neuron of population β to any neuron of population α . (A specific form for $\Phi_{\alpha\beta}(t)$ will be given in Sect. 9.3.1; a more general discussion of different choices of $\Phi_{\alpha\beta}(t)$ can be found in the review of Ref. [12].) Assuming that all synaptic inputs sum linearly, the total synaptic input to the soma of the i th neuron, which we denote by $u_i(t)$, is

$$u_i(t) = \sum_\beta \frac{1}{N_\beta} \sum_{j:p(j)=\beta} \Phi_{\alpha\beta}(t - T_j^m) = \int_{-\infty}^t \sum_\beta \Phi_{\alpha\beta}(t - t') \frac{1}{N_\beta} \sum_{j:p(j)=\beta} a_j(t') dt' \quad (9.26)$$

for all $p(i) = \alpha$, where $a_j(t) = \sum_{m \in \mathbb{Z}} \delta(t - T_j^m)$. That is, $a_j(t)$ represents the output spike train of the j th neuron in terms of a sum of Dirac delta functions. In order to obtain a closed set of equations, we have to determine the firing times T_i^m given by Eq. (9.2), where $V_i(t)$ evolves according to the LIF model (9.1) with $I_i(t) \rightarrow u_i(t)$, or the more general conductance-based model

$$C \frac{dV_i}{dt} = -I_{\text{con},i}(V_i, \dots) + u_i, \quad (9.27)$$

supplemented by additional equations for various ionic gating variables [30]. It follows that, after transients have decayed away, $u_i(t) = u_\alpha(t)$ for all $p(i) = \alpha$ with

$$u_\alpha(t) = \sum_{\beta=1}^P \int_{-\infty}^t \Phi_{\alpha\beta}(t-t') a_\beta(t') dt', \quad (9.28)$$

and $a_\alpha(t)$ is the output activity of the α th population:

$$a_\alpha(t) = \frac{1}{N_\alpha} \sum_{j; p(j)=\beta} a_j(t) \quad (9.29)$$

In general, Eqs. (9.26) and (9.27) are very difficult to analyze. However, considerable simplification can be obtained if the total synaptic current $u_i(t)$ is slowly varying compared to the membrane potential dynamics given by Eq. (9.27). This would occur, for example, if each of the homogeneous subnetworks fired asynchronously as described in Sect. 9.2. One is then essentially reinterpreting the population activity variables $u_\alpha(t)$ and $a_\alpha(t)$ as mean fields of local populations. (Alternatively, a slowly varying synaptic current would occur if the synapses are themselves sufficiently slow [13,28].) Under these simplifying assumptions, one can carry out a short-term temporal averaging of Eq. (9.28) in which the output population activity is approximated by the instantaneous firing rate $a_\alpha(t) = F_\alpha(u_\alpha(t))$ with F_α identified with the population gain function calculated in Sect. 9.2. (In practice, the firing rate function is usually approximated by a sigmoid.) Equation (9.28) then forms the closed system of integral equations

$$u_\alpha(t) = \int_{-\infty}^t \sum_\beta \Phi_{\alpha\beta}(t-t') F_\alpha(u_\beta(t')) dt'. \quad (9.30)$$

The observation that finite-size effects provide a source of noise within a local population then suggests one way to incorporate noise into rate-based models, namely, to take the relationship between population output activity $a_\alpha(t)$ and effective synaptic current $u_\alpha(t)$ to be governed by a stochastic process.

9.3.1 Neural Langevin Equation

The simplest approach is to assume that population activity is a stochastic variable $A_\alpha(t)$ evolving according to a Langevin equation (stochastic differential equation) of the form

$$\tau_\alpha dA_\alpha(t) = [-A_\alpha(t) + F(U_\alpha(t))] dt + \sigma_\alpha dW_\alpha(t) \quad (9.31)$$

with the stochastic current $U_\alpha(t)$ satisfying the integral equation (9.28). Here $W_\alpha(t)$, $\alpha = 1, \dots, P$ denotes a set of P independent Wiener processes with

$$\langle dW_\alpha(t) \rangle = 0, \quad \langle dW_\alpha(t) dW_\beta(t) \rangle = \delta_{\alpha,\beta} dt, \quad (9.32)$$

and σ_α is the strength of noise in the α th population. In general, the resulting stochastic model is non-Markovian. However, if we take $\Phi_{\alpha\beta}(t) = w_{\alpha\beta}\Phi(t)$ with $\Phi(t) = \tau^{-1}e^{-t/\tau}H(t)$ and $H(t)$ the Heaviside function, then we can convert the latter equation to the form

$$\tau dU_\alpha(t) = \left[-U_\alpha(t) + \sum_{\beta=1}^P w_{\alpha\beta} A_\beta(t) \right] dt. \quad (9.33)$$

It is important to note that the time constant τ_α cannot be identified directly with membrane or synaptic time constants. Instead, it determines the relaxation rate of a local population to the mean-field firing rate. In the limit $\tau_\alpha \rightarrow 0$, Eqs. (9.31) and (9.33) reduce to a “voltage-based” rate model perturbed by additive noise:

$$\tau dU_\alpha(t) = \left[-U_\alpha(t) + \sum_{\beta=1}^P w_{\alpha\beta} F(U_\beta(t)) \right] dt + d\tilde{W}_\alpha(t). \quad (9.34)$$

Here $\tilde{W}_\alpha(t) = \sum_{\beta=1}^P w_{\alpha\beta} \sigma_\beta W_\beta(t)$ so that

$$\langle d\tilde{W}_\alpha(t) \rangle = 0, \quad \langle d\tilde{W}_\alpha(t) d\tilde{W}_\beta(t) \rangle = \left[\sum_\gamma w_{\alpha\gamma} w_{\beta\gamma} \sigma_\gamma^2 \right] dt. \quad (9.35)$$

Thus eliminating the dynamics of the firing rate leads to spatially correlated noise for the dynamics of U_α . On the other hand, in the limit $\tau \rightarrow 0$, we obtain a stochastic “activity-based” model

$$\tau_\alpha dA_\alpha(t) = \left[-A_\alpha(t) + F(\sum_\beta w_{\alpha\beta} A_\alpha(t)) \right] dt + \sigma_\alpha dW_\alpha(t). \quad (9.36)$$

Here the dynamical variable A_α represents the firing rate of a local population. For a detailed discussion of the differences between activity-based and voltage-based neural rate equations, see Refs. [12, 30].

9.3.2 Neural Master Equation

An alternative approach to incorporating noise into the population firing rate has been developed in terms of a jump Markov process [9, 10, 19, 20, 47]. Such a

description is motivated by the idea that each local population consists of a discrete number of spiking neurons, and that finite-size effects are a source of intrinsic rather than extrinsic noise [7, 59]. The stochastic output activity of a local population of N neurons is now expressed as $A_\alpha(t) = N_\alpha(t)/(N\Delta t)$ where $N_\alpha(t)$ is the number of neurons in the α th population that fired in the time interval $[t - \Delta t, t]$, and Δt is the width of a sliding window that counts spikes. Suppose that the discrete stochastic variables $N_\alpha(t)$ evolve according to a one-step jump Markov process:

$$N_\alpha(t) \rightarrow N_\alpha(t) \pm 1 : \quad \text{transition rate } \Omega_\alpha^\pm(t), \quad (9.37)$$

in which $\Omega_\alpha^\pm(t)$ are functions of $N_\alpha(t)$ and $U_\alpha(t)$ with $U_\alpha(t)$ evolving according to the integral equation (9.28) or its differential version (9.33). Thus, synaptic coupling between populations occurs via the transition rates. The transition rates are chosen in order to yield a deterministic rate-based model in the thermodynamic limit $N \rightarrow \infty$. One such choice is

$$\Omega_\alpha^+(t) = \frac{N\Delta t}{\tau_\alpha} F(U_\alpha(t)), \quad \Omega_\alpha^-(t) = \frac{N_\alpha(t)}{\tau_\alpha}. \quad (9.38)$$

The resulting stochastic process defined by Eqs. (9.37), (9.38) and (9.33) is an example of a stochastic hybrid system based on a piecewise deterministic process [14]. That is, the transition rates $\Omega_\alpha^\pm(t)$ depend on $U_\alpha(t)$, with the latter itself coupled to the associated jump Markov according to Eq. (9.33), which is only defined between jumps, during which $U_\alpha(t)$ evolves deterministically. (Stochastic hybrid systems also arise in applications to genetic networks [65] and to excitable neuronal membranes [18, 50].) A further simplification is obtained in the limit $\tau \rightarrow 0$, since the continuous variables $U_\alpha(t)$ can be eliminated to give a pure birth-death process for the discrete variables $N_\alpha(t)$. Let $P(\mathbf{n}, t) = \text{Prob}[\mathbf{N}(t) = \mathbf{n}]$ denote the probability that the network of interacting populations has configuration $\mathbf{n} = (n_1, n_2, \dots, n_P)$ at time $t, t > 0$, given some initial distribution $P(\mathbf{n}, 0)$. The probability distribution then evolves according to the birth-death master equation [9, 19, 20]

$$\frac{dP(\mathbf{n}, t)}{dt} = \sum_\alpha \left[(\mathbb{T}_\alpha - 1) (\Omega_\alpha^-(\mathbf{n}) P(\mathbf{n}, t)) + (\mathbb{T}_\alpha^{-1} - 1) (\Omega_\alpha^+(\mathbf{n}) P(\mathbf{n}, t)) \right], \quad (9.39)$$

where $\Omega_\alpha^\pm(\mathbf{n}) = \Omega_\alpha^\pm(t)$ with $N_\alpha(t) = n_\alpha$ and $U_\alpha(t) = \sum_\beta w_{\alpha\beta} n_\beta / (N\Delta t)$, and \mathbb{E}_α is a translation operator: $\mathbb{E}_\alpha^{\pm 1} F(\mathbf{n}) = F(\mathbf{n}_{\alpha\pm})$ for any function F with $\mathbf{n}_{\alpha\pm}$ denoting the configuration with n_α replaced by $n_\alpha \pm 1$. Equation (9.39) is supplemented by the boundary conditions $P(\mathbf{n}, t) \equiv 0$ if $n_\alpha = N_\alpha + 1$ or $n_\alpha = -1$ for some α . The birth-death master equation (9.39) has been the starting point for a number of recent studies of the effects of intrinsic noise on neural fields, which adapt various methods from the analysis of chemical master equations including system size expansions and path integral representations [9, 19, 20]. However, there are a number of potential problems with the master equation formulation. First, there is no unique prescription

for choosing transition rates that yield a given rate-based model in the mean-field limit. Moreover, depending on the choice of how the transition rates scale with the system size N , the statistical nature of the dynamics can be Gaussian-like [9] or Poisson-like [19, 20]. Second, the interpretation of $N_\alpha(t)$ as the number of spikes in a sliding window of width Δt implies that $\tau \gg \Delta t$ so the physical justification for taking the limit $\tau \rightarrow 0$ is not clear. Finally, for large N the master equation can be approximated by a Langevin equation with multiplicative noise (in the sense of Ito), and thus reduces to the previous class of stochastic neural field model [10].

9.3.3 Continuum Limit

So far we have indicated how to incorporate noise into a discrete network of neuronal populations. In order to obtain a corresponding stochastic neural field equation it is now necessary to take an appropriate continuum limit. For simplicity, we will focus on the simplest stochastic rate model given by Eqs.(9.31) and (9.33). The continuum limit of Eq.(9.33) proceeds heuristically as follows. First, set $U_\alpha(t) = U(\alpha\Delta d, t)$, $A_\alpha(t) = A(\alpha\Delta d, t)$ and $w_{\alpha\beta} = \rho\Delta d w(\alpha\Delta d, \beta\Delta d)$ where ρ is a synaptic density and Δd is an infinitesimal length scale. Taking the limit $\Delta d \rightarrow 0$ and absorbing ρ into w gives

$$\tau dU(x, t) = [-U(x, t) + \int_{-\infty}^{\infty} w(x - y)A(y)dy]dt. \quad (9.40)$$

We also assume that the noise strength $\sigma_\alpha = \sigma/\sqrt{\Delta d}$ and define $W_\alpha(t)/\sqrt{\Delta d} = W(\alpha\Delta d, t)$. Taking the limit $\Delta d \rightarrow 0$ in Eq.(9.31) with $\tau_\alpha = \hat{\tau}$ for all α gives

$$\hat{\tau}dA(x, t) = [-A(x, t) + F(U(x, t))]dt + \sigma dW(x, t) \quad (9.41)$$

with

$$\langle dW(x, t) \rangle = 0, \quad \langle dW(x, t)dW(y, t) \rangle = \delta(x - y)dt. \quad (9.42)$$

In the limit $\hat{\tau} \rightarrow 0$ we obtain a stochastic version of a voltage-based neural field equation, namely,

$$\tau dU(x, t) = [-U(x, t) + \int_{-\infty}^{\infty} w(x - y)F(U(y, t))dy]dt + \sigma d\tilde{W}(x, t) \quad (9.43)$$

with

$$\langle d\tilde{W}(x, t) \rangle = 0, \quad \langle d\tilde{W}(x, t)d\tilde{W}(y, t) \rangle = dt \int_{-\infty}^{\infty} w(x - z)w(y - z)dz. \quad (9.44)$$

Similarly, in the limit $\tau \rightarrow 0$ we have a stochastic version of an activity-based neural field equation

$$\hat{\tau} dA(x, t) = \left[-A(x, t) + F \left(\int_{-\infty}^{\infty} w(x-y) A(y, t) dy \right) dt + dW(x, t) \right]. \quad (9.45)$$

From a numerical perspective, any computer simulation would involve rediscretizing space and then solving a time-discretized version of the resulting stochastic differential equation. On the other hand, in order to investigate analytically the effects of noise on spatiotemporal dynamics, it is more useful to work directly with stochastic neural fields. One can then adapt various PDE methods for studying noise in spatially extended systems [55], as illustrated in the next section.

9.4 Traveling Waves in Stochastic Neural Fields

In this section we review some recent work on analyzing traveling waves in stochastic neural fields [15].

9.4.1 Traveling Fronts in a Deterministic Neural Field

Let us begin by briefly reviewing front propagation in a scalar neural field equation of the voltage-based form

$$\tau \frac{\partial u(x, t)}{\partial t} = -u(x, t) + \int_{-\infty}^{\infty} w(x-x') F(u(x', t)) dx'. \quad (9.46)$$

For concreteness F is taken to be a sigmoid function

$$F(u) = \frac{1}{1 + e^{-\gamma(u-\kappa)}} \quad (9.47)$$

with gain γ and threshold κ . In the high-gain limit $\gamma \rightarrow \infty$, this reduces to the Heaviside function

$$F(u) \rightarrow H(u - \kappa) = \begin{cases} 1 & \text{if } u > \kappa \\ 0 & \text{if } u \leq \kappa \end{cases} \quad (9.48)$$

The weight distribution is taken to be a positive (excitatory), even function of x , $w(x) \geq 0$ and $w(-x) = w(x)$, with $w(x)$ a monotonically decreasing function of x for $x \geq 0$. The weight distribution is typically taken to be an exponential weight distribution

$$w(x) = \frac{1}{2\sigma} e^{-|x|/\sigma}, \quad (9.49)$$

where σ determines the range of synaptic connections. The latter tends to range from 100 μm to 1 mm. We fix the units of time and space by setting $\tau = 1, \sigma = 2$.

A homogeneous fixed point solution U^* of Eq. (9.46) satisfies

$$U^* = W_0 F(U^*), \quad W_0 = \int_{-\infty}^{\infty} w(y) dy. \quad (9.50)$$

In the case of a sigmoid function with appropriately chosen gain and threshold, it is straightforward to show that there exists a pair of stable fixed points U^*_{\pm} separated by an unstable fixed point U_0^* , see also Fig. 9.3a. In the high gain limit $F(U) \rightarrow H(U - \kappa)$ with $0 < \kappa < W_0$, the unstable fixed point disappears and $U_+^* = W_0, U_-^* = 0$. As originally shown by Amari [2], an explicit traveling front solution of Eq. (9.46) that links U_+^* and U_-^* can be constructed in the case of a Heaviside nonlinearity. In order to construct such a solution, we introduce the traveling wave coordinate $\xi = x - ct$, where c denotes the wavespeed, and set $u(x, t) = \mathcal{U}(\xi)$ with $\lim_{\xi \rightarrow -\infty} \mathcal{U}(\xi) = U_+^* > 0$ and $\lim_{\xi \rightarrow \infty} \mathcal{U}(\xi) = 0$ such that $\mathcal{U}(\xi)$ only crosses the threshold κ once. Since Eq. (9.46) is equivariant with respect to uniform translations, we are free to take the threshold crossing point to be at the origin, $\mathcal{U}(0) = \kappa$, so that $\mathcal{U}(\xi) < \kappa$ for $\xi > 0$ and $\mathcal{U}(\xi) > \kappa$ for $\xi < 0$. Substituting this traveling front solution into Eq. (9.46) with $F(u) = H(u - \kappa)$ then gives

$$-c\mathcal{U}'(\xi) + \mathcal{U}(\xi) = \int_{-\infty}^0 w(\xi - \xi') d\xi' = \int_{\xi}^{\infty} w(x) dx \equiv \hat{w}(\xi), \quad (9.51)$$

where $\mathcal{U}'(\xi) = d\mathcal{U}/d\xi$. Multiplying both sides of the above equation by $e^{-\xi/c}$ and integrating with respect to ξ leads to the solution

$$\mathcal{U}(\xi) = e^{\xi/c} \left[\kappa - \frac{1}{c} \int_0^{\xi} e^{-y/c} \hat{w}(y) dy \right]. \quad (9.52)$$

Finally, requiring the solution to remain bounded as $\xi \rightarrow \infty$ ($\xi \rightarrow -\infty$) for $c > 0$ implies that κ must satisfy the condition

$$\kappa = \frac{1}{c} \int_0^{\infty} e^{-y/c} \hat{w}(y) dy, \quad (9.53)$$

and thus

$$\mathcal{U}(\xi) = \frac{1}{c} \int_0^{\infty} e^{-y/c} \hat{w}(y + \xi) dy. \quad (9.54)$$

In the case of the exponential weight distribution (9.49), the relationship between wavespeed c and threshold κ for right-moving fronts is

$$c = c_+(\kappa) \equiv \frac{1}{2\kappa} [1 - 2\sigma\kappa] \quad \text{for } \kappa < 0.5. \quad (9.55)$$

A similar expression holds for left-moving fronts ($c < 0$) for which $1 > \kappa > 0.5$. Using Evans function techniques, it can also be shown that the traveling front is stable [24, 66]. Finally, given the existence of a traveling front solution for a Heaviside rate function, it is possible to prove the existence of a unique front in the case of a smooth sigmoid nonlinearity using a continuation method [29].

9.4.2 Stochastic Neural Field with Extrinsic Noise

Let us now consider a stochastic version of the scalar neural field (9.46) given by the neural Langevin equation

$$dU = \left[-U(x, t) + \int_{-\infty}^{\infty} w(x - y) F(U(y, t)) dy \right] dt + \varepsilon^{1/2} g(U(x, t)) dW(x, t). \quad (9.56)$$

We assume that $dW(x, t)$ represents an independent Wiener process such that

$$\langle dW(x, t) \rangle = 0, \quad \langle dW(x, t) dW(x', t') \rangle = 2C([x - x']/\lambda) \delta(t - t') dt dt', \quad (9.57)$$

where $\langle \cdot \rangle$ denotes averaging with respect to the Wiener process. Here λ is the spatial correlation length of the noise such that $C(x/\lambda) \rightarrow \delta(x)$ in the limit $\lambda \rightarrow 0$, and ε determines the strength of the noise, which is assumed to be weak. For the sake of generality, we take the extrinsic noise to be multiplicative rather than additive. Following standard formulations of Langevin equations [34], the extrinsic noise term is assumed to be of Stratonovich form. Note, however, that a Kramers-Moyal expansion of the neural master equation (9.39) yields a Langevin neural field equation with multiplicative noise of the Ito form [9, 10]. The main results highlighted below do not depend on the precise form of the noise.

The effects of multiplicative noise on front propagation can be analyzed using methods previously developed for reaction–diffusion equations [4, 25, 55, 56], as recently shown in Ref. [15]. The starting point of such methods is the observation that multiplicative noise in the Stratonovich sense leads to a systematic shift in the speed of the front (assuming a front of speed c exists when $\varepsilon = 0$). This is a consequence of the fact that $\langle g(U) dW \rangle \neq 0$ even though $\langle dW \rangle = 0$. The former average can be calculated using Novikov’s theorem [45]:

$$\varepsilon^{1/2} \langle g(U(x, t)) \eta(x, t) \rangle = \varepsilon C(0) \langle g'(U) g(U) \rangle. \quad (9.58)$$

The above result also follows from Fourier transforming Eq. (9.56) and evaluating averages using the Fokker–Planck equation in Fourier space [15, 55]. Note that in the limit $\lambda \rightarrow 0$, $C(0) \rightarrow 1/\Delta x$ where Δx is a lattice cut-off, which is typically identified with the step size of the spatial discretization scheme used in numerical simulations. The method developed in Ref. [4] for stochastic PDEs is to construct an approximation scheme that separates out the diffusive effects of noise from the mean drift. Applying a similar method to the neural field equation (9.56) [15], we first rewrite the equation as

$$dU(x, t) = [h(U(x, t)) + \int_{-\infty}^{\infty} w(x - y)F(U(y, t))dy]dt + \varepsilon^{1/2}R(U, x, t), \quad (9.59)$$

where

$$h(U) = -U + \varepsilon C(0)g'(U)g(U) \quad (9.60)$$

and

$$R(U, x, t) = g(U)\eta(x, t) - \varepsilon^{1/2}C(0)g'(U)g(U). \quad (9.61)$$

The stochastic process R has zero mean (so does not contribute to the effective drift) and correlation

$$\langle R(U, x, t)R(U, x', t') \rangle = \langle g(U(x, t))\eta(x, t)g(U(x', t')\eta(x', t') \rangle + \mathcal{O}(\varepsilon^{1/2}). \quad (9.62)$$

The next step in the analysis is to assume that the fluctuating term in Eq. (9.59) generates two distinct phenomena that occur on different time-scales: a diffusive-like displacement of the front from its uniformly translating position at long time scales, and fluctuations in the front profile around its instantaneous position at short time scales [4, 55]. In particular, following Ref. [15], we express the solution U of Eq. (9.59) as a combination of a fixed wave profile U_0 that is displaced by an amount $\Delta(t)$ from its uniformly translating mean position $\xi = x - c_\varepsilon t$, and a time-dependent fluctuation Φ in the front shape about the instantaneous position of the front:

$$U(x, t) = U_0(\xi - \Delta(t)) + \varepsilon^{1/2}\Phi(\xi - \Delta(t), t). \quad (9.63)$$

Here c_ε denotes the mean speed of the front. The wave profile U_0 and associated wave speed c_ε are obtained by solving the modified deterministic equation

$$-c_\varepsilon \frac{dU_0}{d\xi} - h(U_0(\xi)) = \int_{-\infty}^{\infty} w(\xi - \xi')F(U_0(\xi'))d\xi'. \quad (9.64)$$

Both c_ε and U_0 depend non-trivially on the noise strength ε due to the ε -dependence of the function h . Thus, $c_\varepsilon \neq c$ for $\varepsilon > 0$ and $c_0 = c$, where c is the speed of the front in the absence of multiplicative noise. It also follows that the expansion (9.63) is not equivalent to a standard small-noise expansion in ε . Equation (9.64) is chosen so that to leading order, the stochastic variable $\Delta(t)$ undergoes unbiased Brownian motion with a diffusion coefficient $D(\varepsilon) = \mathcal{O}(\varepsilon)$ (see below). Thus $\Delta(t)$ represents the effects of slow fluctuations, whereas Φ represents the effects of fast fluctuations.

The next step is to substitute the decomposition (9.63) into Eq. (9.59) and expand to first order in $\mathcal{O}(\varepsilon^{1/2})$. Imposing Eq. (9.64), after shifting $\xi \rightarrow \xi - \Delta(t)$, and dividing through by $\varepsilon^{1/2}$ then gives

$$d\Phi(\xi, t) = \hat{L} \circ \Phi(\xi, t) dt + \varepsilon^{-1/2} U'_0(\xi) d\Delta(t) + dR(U_0, \xi, t) \quad (9.65)$$

where \hat{L} is the non-self-adjoint linear operator

$$\hat{L} \circ A(\xi) = c_\varepsilon \frac{dA(\xi)}{d\xi} + h'(U_0(\xi)) A(\xi) + \int_{-\infty}^{\infty} w(\xi - \xi') F'(U_0(\xi')) A(\xi') d\xi' \quad (9.66)$$

for any function $A(\xi) \in L_2(\mathbb{R})$. The non-self-adjoint linear operator \hat{L} has a 1D null space spanned by $U'_0(\xi)$, which follows from differentiating equation (9.64) with respect to ξ . We then have the solvability condition for the existence of a nontrivial solution of Eq. (9.66), namely, that the inhomogeneous part is orthogonal to the null space of the adjoint operator. The latter is defined with respect to the inner product

$$\int_{-\infty}^{\infty} B(\xi) \hat{L} A(\xi) d\xi = \int_{-\infty}^{\infty} [\hat{L}^* B(\xi)] A(\xi) d\xi \quad (9.67)$$

where $A(\xi)$ and $B(\xi)$ are arbitrary integrable functions. Hence,

$$\hat{L}^* B(\xi) = -c_\varepsilon \frac{dB(\xi)}{d\xi} + h'(U_0(\xi)) B(\xi) + F'(U_0(\xi)) \int_{-\infty}^{\infty} w(\xi - \xi') B(\xi') d\xi'. \quad (9.68)$$

Taking the null space of \hat{L} to be spanned by the function $\mathcal{V}(\xi)$, we have

$$\int_{-\infty}^{\infty} \mathcal{V}(\xi) [U'_0(\xi) d\Delta(t) + \varepsilon^{1/2} dR(U_0, \xi, t)] d\xi = 0. \quad (9.69)$$

Thus $\Delta(t)$ satisfies the stochastic ODE

$$d\Delta(t) = -\varepsilon^{1/2} \frac{\int_{-\infty}^{\infty} \mathcal{V}(\xi) dR(U_0, \xi, t) d\xi}{\int_{-\infty}^{\infty} \mathcal{V}(\xi) U'_0(\xi) d\xi}. \quad (9.70)$$

Using the lowest order approximation $dR(U_0, \xi, t) = g(U_0(\xi))dW(\xi, t)$, we deduce that (for $\Delta(0) = 0$)

$$\langle \Delta(t) \rangle = 0, \quad \langle \Delta(t)^2 \rangle = 2D(\varepsilon)t \quad (9.71)$$

where $D(\varepsilon)$ is the effective diffusivity

$$D(\varepsilon) = \varepsilon \langle d\Delta(t)d\Delta(t') \rangle = \varepsilon \frac{\int_{-\infty}^{\infty} \mathcal{V}(\xi)^2 g^2(U_0(\xi)) d\xi}{\left[\int_{-\infty}^{\infty} \mathcal{V}(\xi) U'_0(\xi) d\xi \right]^2}. \quad (9.72)$$

9.4.3 Explicit Results for a Heaviside Rate Function

We now illustrate the above analysis by considering a particular example where the effective speed c_ε and diffusion coefficient $D(\varepsilon)$ can be calculated explicitly [15]. That is, take $g(U) = g_0 U$ for the multiplicative noise term and set $F(U) = H(U - \kappa)$. (The constant g_0 has units of $\sqrt{\text{length/time}}$.) The deterministic equation (9.64) for the fixed profile U_0 then reduces to

$$-c_\varepsilon \frac{dU_0}{d\xi} + U_0(\xi)\gamma(\varepsilon) = \int_{-\infty}^{\infty} w(\xi - \xi') H(U_0(\xi') - \kappa) d\xi', \quad (9.73)$$

with

$$\gamma(\varepsilon) = (1 - \varepsilon g_0^2 C(0)), \quad (9.74)$$

The analysis of the wave speeds proceeds along identical lines to the deterministic model. Thus, multiplying both sides of Eq. (9.73) by $e^{-\xi\gamma(\varepsilon)/c_\varepsilon}$ and integrating with respect to ξ gives

$$U_0(\xi) = e^{\xi\gamma(\varepsilon)/c_\varepsilon} \left[\kappa - \frac{1}{c_\varepsilon} \int_0^\xi e^{-y\gamma(\varepsilon)/c_\varepsilon} \hat{w}(y) dy \right]. \quad (9.75)$$

Finally, requiring the solution to remain bounded as $\xi \rightarrow \infty$ ($\xi \rightarrow -\infty$) for $c_\varepsilon > 0$ implies that κ must satisfy the condition

$$\kappa = \frac{1}{c_\varepsilon} \int_0^\infty e^{-y\gamma(\varepsilon)/c_\varepsilon} \hat{w}(y) dy. \quad (9.76)$$

Hence, in the case of the exponential weight distribution (9.49), we have

$$c_\varepsilon = \gamma(\varepsilon)c_+ (\gamma(\varepsilon)\kappa) = \frac{\sigma}{2\kappa} [1 - 2\kappa\gamma(\varepsilon)] \quad (9.77)$$

for $c_\varepsilon > 0$ with $c_+(\kappa)$ defined in Eq.(9.55). Assuming that $0 \leq \gamma(\varepsilon) \leq 1$, we see that multiplicative noise shifts the effective velocity of front propagation in the positive ξ direction.

In order to calculate the diffusion coefficient, it is first necessary to determine the null vector $\mathcal{V}(\xi)$ of the adjoint linear operator \hat{L}^* defined by Eq.(9.68). Setting $F(U) = H(U - \kappa)$ and $g(U) = g_0 U$, the null vector \mathcal{V} satisfies the equation

$$c_\varepsilon \mathcal{V}'(\xi) + \gamma(\varepsilon) \mathcal{V}(\xi) = -\frac{\delta(\xi)}{U'_0(0)} \int_{-\infty}^{\infty} w(\xi') \mathcal{V}(\xi') d\xi'. \quad (9.78)$$

This can be solved explicitly to give [8]

$$\mathcal{V}(\xi) = -H(\xi) \exp(-\Gamma(\varepsilon)\xi), \quad \Gamma(\varepsilon) = \frac{\gamma(\varepsilon)}{c_\varepsilon}. \quad (9.79)$$

We have used the fact that the solution to Eq.(9.73) is of the form

$$U_0(\xi) = \frac{1}{c_\varepsilon} \int_0^\infty e^{-\Gamma(\varepsilon)y} \hat{w}(y + \xi) dy, \quad (9.80)$$

with $\hat{w}(\xi)$ defined in Eq.(9.51) and, hence,

$$U'_0(\xi) = -\frac{1}{c_\varepsilon} \int_0^\infty e^{-\Gamma(\varepsilon)y} w(y + \xi) dy. \quad (9.81)$$

Using Eq.(9.79), Eq.(9.72) reduces to the form

$$D(\varepsilon) = \varepsilon \frac{\int_0^\infty e^{-2\Gamma(\varepsilon)\xi} U_0(\xi)^2 d\xi}{\left[\int_0^\infty e^{-\Gamma(\varepsilon)\xi} U'_0(\xi) d\xi \right]^2}, \quad (9.82)$$

which can be evaluated explicitly for an exponential weight distribution to give

$$D(\varepsilon) = \frac{1}{2} \varepsilon \sigma g_0^2 (1 + \sigma \Gamma(\varepsilon)) \quad (9.83)$$

In Fig.9.1 we show the temporal evolution of a single stochastic wave front, which is obtained by numerically solving the Langevin equation (9.56) for $F(U) = H(U - \kappa)$, $g(U) = U$ and an exponential weight distribution w . In order to numerically calculate the mean location of the front as a function of time, we carry out a large number of level set position measurements. That is, the positions $X_a(t)$ are determined such that $U(X_a(t), t) = a$, for various level set values $a \in (0.5\kappa, 1.3\kappa)$ and the mean location is defined to be $\bar{X}(t) = \mathbb{E}[X_a(t)]$, where the expectation is first taken with respect to the sampled values a and then averaged

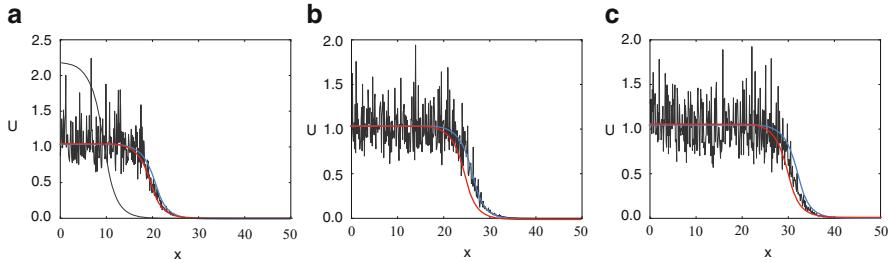


Fig. 9.1 Numerical simulation showing the propagation of a front solution of the stochastic neural field equation (9.56) for Heaviside weight function $F(U) = H(U - \kappa)$ with $\kappa = 0.35$, exponential weight function (9.49) with $\sigma = 2$, and multiplicative noise $g(U) = U$. Noise strength $\epsilon = 0.005$ and $C(0) = 10$. The wave profile is shown at successive times (a) $t = 12$ (b) $t = 18$ and (c) $t = 24$, with the initial profile at $t = 0$ (smooth curve in (a)) given by Eq. (9.80). In numerical simulations we take the discrete space and time steps $\Delta x = 0.1$, $\Delta t = 0.01$. The deterministic part U_0 of the stochastic wave is shown by the more advanced curves and the corresponding solution in the absence of noise ($\epsilon = 0$) is shown by the less advanced curves

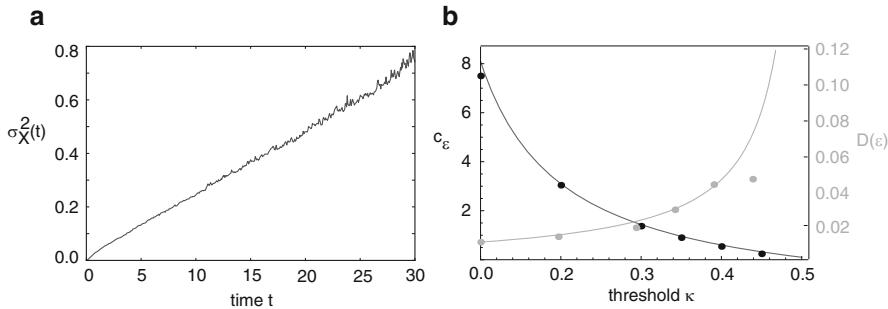


Fig. 9.2 (a) Variance $\sigma_X^2(t)$ of front position as a function of time, averaged over $N = 4,096$ trials. Same parameter values as Fig. 9.1. (b) Plot of wave speed c_ε (black curve) and diffusion coefficient $D(\varepsilon)$ (gray curve) as a function of threshold κ . Numerical results (solid dots) are obtained by averaging over $N = 4,096$ trials starting from the initial condition given by Eq. (9.80). Corresponding theoretical predictions (solid curves) for c_ε and $D(\varepsilon)$ are based on Eqs. (9.77) and (9.82), respectively. Other parameters are as in Fig. 9.1

over N trials. The corresponding variance is given by $\sigma_X^2(t) = \mathbb{E}[(X_a(t) - \bar{X}(t))^2]$. In Fig. 9.2a, $\sigma_X^2(t)$ is plotted as a function of t . It can be seen that it varies linearly with t , consistent with the assumption that there is a diffusive-like displacement of the front from its uniformly translating position at long time scales. The slope of the curve then determines the effective diffusion coefficient according to $\sigma_X^2(t) \sim 2D(\varepsilon)t$. In Fig. 9.2b, the numerically estimated speed and diffusion coefficient are plotted for various values of the threshold κ and are compared to the corresponding theoretical curves obtained using the above analysis. It can be seen that there is excellent agreement with the theoretical predictions provided that κ is not too large. Note that as $\kappa \rightarrow 0.5$, the wave speed decreases towards zero so that the assumption

of relatively slow diffusion breaks down. Note that the analysis of freely propagating fronts can be extended to the case of fronts locked to an externally moving stimulus [15]. One finds that the latter are much more robust to noise, since the stochastic wandering of the mean front profile is now described by an Ornstein–Uhlenbeck process rather than a Wiener process, so that the variance in front position saturates in the long time limit rather than increasing linearly with time.

9.5 Path Integral Representation of a Stochastic Neural Field

Recently, Buice and Cowan [19] have used path integral methods and renormalization group theory to establish that a stochastic neural field with an absorbing state, which evolves according to a birth–death master equation of the form (9.39), belongs to the universality class of directed percolation, and consequently exhibits power law behavior suggestive of many measurements of spontaneous cortical activity in vitro and in vivo [6, 51]. (If a network enters an absorbing state all activity is extinguished.) Although the existence of power law behavior is still controversial [5], the application of path integral methods provides another example of how analytical techniques familiar in the study of PDEs and chemical master equations are being adapted to studies of continuum neural fields. (For reviews on path integral methods for stochastic differential equations see Refs. [22, 60, 67].) In this section, we show how a stochastic neural field with extrinsic noise and an absorbing state can be reformulated as a path integral, and use this to estimate the time to extinction of network activity. A more detailed discussion can be found in Ref. [11].

9.5.1 Pulled Fronts, Absorbing States and Extinction Events

In order to construct a neural field with an absorbing state, it is convenient to consider an activity-based rather than a voltage-based neural field of the form

$$\tau \frac{\partial a(x, t)}{\partial t} = -a(x, t) + F \left(\int_{-\infty}^{\infty} w(x - x') a(x', t) dx' \right). \quad (9.84)$$

For the moment, we consider an unbounded domain with $x \in \mathbb{R}$. We also have the additional constraint that $a(x, t) \geq 0$ for all (x, t) , since the field $a(x, t)$ represents the instantaneous firing rate of a local population of neurons at position x and time t . Suppose that $F(a)$ in Eq. (9.84) is a positive, bounded, monotonically increasing function of a with $F(0) = 0$, $\lim_{a \rightarrow 0^+} F'(a) = 1$ and $\lim_{a \rightarrow \infty} F(a) = \kappa$ for some positive constant κ . For concreteness, we take

$$F(a) = \begin{cases} 0, & a \leq 0 \\ a, & 0 < a \leq \kappa \\ \kappa, & a > \kappa. \end{cases} \quad (9.85)$$

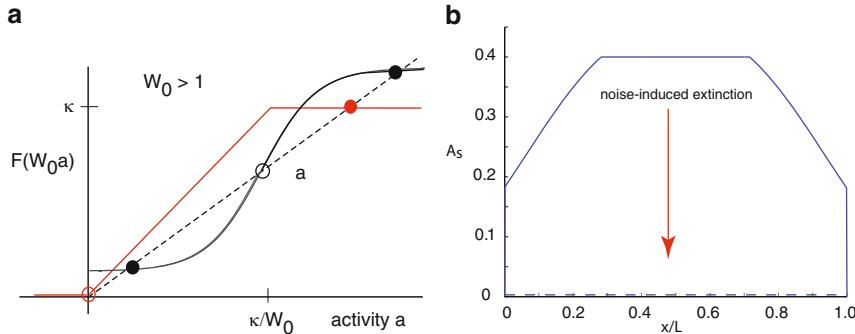


Fig. 9.3 (a) Plots of piecewise linear and sigmoidal firing rate functions. Intercepts of $y = F(W_0 a)$ with straight line $y = a$ determine homogeneous fixed points. Stable (unstable) fixed points indicated by filled (unfilled) circles. (b) Stable steady state solution $a(x, t) = A_s(x)$ of neural field equation (9.84) on a finite spatial domain of length L with boundary conditions $a(0, t) = a(L, t) = 0$. Here $W_0 = 1.2$, $\sigma = 1$, $\kappa = 0.4$ and $L = 5$ in the presence of multiplicative noise, fluctuations can drive the network to the zero absorbing state resulting in the extinction of activity

A homogeneous fixed point solution a^* of Eq. (9.84) satisfies $a^* = F(W_0 a^*)$ with $W_0 = \int_{-\infty}^{\infty} w(y) dy$. In the case of the given piecewise linear firing rate function, we find that if $W_0 > 1$, then there exists an unstable fixed point at $a^* = 0$ (absorbing state) and a stable fixed point at $a^* = \kappa$, see Fig. 9.3a. The construction of a front solution linking the stable and unstable fixed points differs considerably from that considered in neural fields with sigmoidal or Heaviside nonlinearities as considered in Sect. 9.4, where the front propagates into a metastable state. Following the PDE theory of fronts propagating into unstable states [62], we expect there to be a continuum of front velocities and associated traveling wave solutions; the particular velocity selected depends on the initial conditions. Fronts propagating into unstable states can be further partitioned into two broad categories: the so-called *pulled* and *pushed* fronts [62] emerging from sufficiently localized initial conditions. Pulled fronts propagate into an unstable state such that the asymptotic velocity is given by the linear spreading speed v^* , which is determined by linearizing about the unstable state within the leading edge of the front. That is, perturbations around the unstable state within the leading edge grow and spread with speed v^* , thus “pulling along” the rest of the front. On the other hand, pushed fronts propagate into an unstable state with a speed greater than v^* , and it is the nonlinear growth within the region behind the leading edge that pushes the front speeds to higher values. One of the characteristic features of pulled fronts is their sensitivity to perturbations in the leading edge of the wave. This means that standard perturbation methods for studying the effects of spatial heterogeneities [43] or external noise fluctuations [54] break down. The effects of spatial heterogeneities on neural fields that support pulled fronts is explored elsewhere [11, 23].

Consider a traveling wave solution $\mathcal{A}(x - ct)$ of Eq. (9.84) with $\mathcal{A}(\xi) \rightarrow \kappa$ as $\xi \rightarrow -\infty$ and $\mathcal{A}(\xi) \rightarrow 0$ as $\xi \rightarrow \infty$. One can determine the range of velocities c for which such a solution exists by assuming that $\mathcal{A}(\xi) \approx e^{-\lambda\xi}$ for sufficiently large ξ . The exponential decay of the front suggests that we linearize equation (9.84), which in traveling wave coordinates (with $\tau = 1$) takes the form

$$-c \frac{d\mathcal{A}(\xi)}{d\xi} = -\mathcal{A}(\xi) + \int_{-\infty}^{\infty} w(\xi - \xi') \mathcal{A}(\xi') d\xi'. \quad (9.86)$$

However, in order to make the substitution $\mathcal{A}(\xi) \approx e^{-\lambda\xi}$ we need to restrict the integration domain of ξ' to the leading edge of the front. Suppose, for example that $w(x)$ is given by the Gaussian distribution

$$w(x) = \frac{W_0}{\sqrt{2\pi\sigma^2}} e^{-x^2/2\sigma^2}. \quad (9.87)$$

Given the fact that the front solution $\mathcal{A}(\xi)$ is bounded, we introduce a cut-off X with $\sigma \ll X \ll \xi$, and approximate Eq. (9.86) by

$$-c \frac{d\mathcal{A}(\xi)}{d\xi} = -\mathcal{A}(\xi) + \int_{\xi-X}^{\xi+X} w(\xi - \xi') \mathcal{A}(\xi') d\xi'. \quad (9.88)$$

Substituting the exponential solution in (9.86) then yields the dispersion relation $c = c(\lambda)$ with

$$c(\lambda) = \frac{1}{\lambda} \left[\int_{-X}^X w(y) e^{-\lambda y} dy - 1 \right]. \quad (9.89)$$

Finally, we now take the limit $X \rightarrow \infty$ under the assumption that $w(y)$ is an even function to yield

$$c(\lambda) = \frac{1}{\lambda} [\mathcal{W}(\lambda) - 1], \quad (9.90)$$

where $\mathcal{W}(\lambda) = \hat{w}(\lambda) + \hat{w}(-\lambda)$ and $\hat{w}(\lambda) = \int_0^\infty w(y) e^{-\lambda y} dy$ is the Laplace transform of $w(x)$. We are assuming that $w(y)$ decays sufficiently fast as $|y| \rightarrow \infty$ so that the Laplace transform $\hat{w}(\lambda)$ exists for bounded, negative values of λ . This holds in the case of the Gaussian distribution (9.87) for which $\mathcal{W}(\lambda) = W_0 e^{\lambda^2 \sigma^2 / 2}$. Hence,

$$c(\lambda) = \frac{W_0 e^{\lambda^2 \sigma^2 / 2} - 1}{\lambda}. \quad (9.91)$$

If $W_0 > 1$ (necessary for the zero activity state to be unstable) then $c(\lambda)$ is a positive unimodal function with $c(\lambda) \rightarrow \infty$ as $\lambda \rightarrow 0$ or $\lambda \rightarrow \infty$ and a unique minimum at $\lambda = \lambda_0$ with λ_0 . Assuming that the full nonlinear system supports a pulled front then a sufficiently localized initial perturbation (one that decays faster than $e^{-\lambda_0 x}$)

will asymptotically approach the traveling front solution with the minimum wave speed $c_0 = c(\lambda_0)$. Note that $c_0 \sim \sigma$ and $\lambda_0 \sim \sigma^{-1}$.

In the above analysis, the effects of boundary conditions on front propagation were ignored, which is a reasonable approximation if the size L of the spatial domain satisfies $L \gg \sigma$, where σ is the range of synaptic weights. In the case of a finite domain, following passage of an invasive activity front, the network settles into a non-zero stable steady state, whose spatial structure will depend on the boundary conditions. The steady-state equation takes the form

$$a(x) = F \left(\int_0^L w(x - x') a(x') dx' \right). \quad (9.92)$$

In the case of Dirichlet boundary conditions, $a(0, t) = a(L, t) = 0$ with $L \gg \sigma$, the steady-state will be uniform in the bulk of the domain with $a(x) \approx a_0$ except for boundary layers at both ends. Here a_0 is the nonzero solution to the equation $a_0 = F(W_0 a_0)$. An example of a steady-state solution is plotted in Fig. 9.3b. (Note that the sudden drop to zero right on the boundaries reflects the non-local nature of the neural field equation.)

Now suppose some source of extrinsic noise is added to the neural field equation (9.84):

$$dA = \left[-A + F \left(\int_{\Sigma} w(x - y) A(y, t) dy \right) \right] dt + \varepsilon g(A) dW(x, t), \quad (9.93)$$

for $0 \leq t \leq T$ and initial condition $A(x, 0) = \Phi(x)$. Here ε determines the noise strength and $\Sigma = [0, L]$ denotes the spatial domain of the neural field. We will assume that $g(0) = 0$ so that the uniform zero activity state $A \equiv 0$ is an absorbing state of the system; any noise-induced transition to this state would then result in extinction of all activity. An example of multiplicative noise that vanishes at $A = 0$ is obtained by carrying out a diffusion approximation of the neural master equation previously introduced by Buice et al. [19, 20], see Bressloff [9, 10]. Based on the analysis of stochastic traveling waves in Sect. 9.4, we would expect the noise to induce a stochastic wandering of a pulled front solution of the corresponding deterministic equation. However, in the case of stochastic PDEs, it has previously been shown that the stochastic wandering of a pulled front about its mean position is subdiffusive with $\text{var}\Delta(t) \sim t^{1/2}$, in contrast to the diffusive wandering of a front propagating into a metastable state for which $\text{var}\Delta(t) \sim t$ [54]. Such scaling is a consequence of the asymptotic relaxation of the leading edge of the deterministic pulled front. Since pulled front solutions of the neural field equation (9.84) exhibit similar dynamics, it suggests that there will also be subdiffusive wandering of these fronts in the presence of multiplicative noise. This is indeed found to be the case [15]. Another consequence of the noise is that it can induce a transition from the quasi-uniform steady state to the zero absorbing state. In the case of weak noise, the time to extinction is exponentially large and can be estimated using path integral methods as outlined below, see also Ref. [11].

9.5.2 Derivation of Path Integral Representation

In order to develop a framework to study rare extinction events in the weak noise limit, we construct a path integral representation of the stochastic Langevin equation (9.93) along the lines of Ref. [11]. We will assume that the multiplicative noise is of Ito form [34]. For reviews on path integral methods for stochastic differential equations, see Refs. [22, 60, 67]. Discretizing both space and time with $A_{i,m} = A(m\Delta d, i\Delta t)$, $W_{i,m}\sqrt{\Delta t/\Delta d} = dW(m\Delta d, i\Delta t)$, $w_{mn}\Delta d = w(m\Delta d, n\Delta d)$ gives

$$A_{i+1,m} - A_{i,m} = \left[-A_{i,m} + F(\Delta d \sum_n w_{mn} A_{i,n}) \right] \Delta t + \frac{\sqrt{\Delta t}}{\sqrt{\Delta d}} g(A_{i,m}) W_{i,m} + \Phi_m \delta_{i,0},$$

where $i = 0, 1, \dots, N$ for $T = N\Delta t$, $n = 0, \dots, \hat{N}$ for $L = \hat{N}\Delta d$, and

$$\langle W_{i,m} \rangle = 0, \quad \langle W_{i,m} W_{i',m'} \rangle = \delta_{i,i'} \delta_{m,m'}. \quad (9.94)$$

Let \mathbf{A} and \mathbf{W} denote the vectors with components $A_{i,m}$ and $W_{i,m}$ respectively. Formally, the conditional probability density function for \mathbf{A} given a particular realization of the stochastic process \mathbf{W} (and initial condition Φ) is

$$P[\mathbf{A}|\mathbf{W}] = \prod_n \prod_{i=0}^N \delta \left(A_{i+1,m} - A_{i,m} + \left[A_{i,m} - F \left(\Delta d \sum_n w_{mn} A_{i,n} \right) \right] \Delta t - \frac{\sqrt{\Delta t}}{\sqrt{\Delta d}} g(A_{i,m}) W_{i,m} - \Phi_m \delta_{i,0} \right). \quad (9.95)$$

Inserting the Fourier representation of the Dirac delta function,

$$\delta(A_{i,m}) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i \tilde{U}_{i,m} A_{i,m}} d\tilde{U}_{i,m}, \quad (9.96)$$

gives

$$\begin{aligned} P[\mathbf{A}|\mathbf{W}] &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp \left\{ -i \sum_{i,m} \tilde{U}_{i,m} (A_{i+1,m} - A_{i,m}) \right\} \\ &\quad \times \exp \left\{ -i \sum_{i,m} \tilde{U}_{i,m} \left[A_{i,m} - F \left(\Delta d \sum_n w_{mn} A_{i,n} \right) \right] \Delta t \right\} \\ &\quad \times \exp \left\{ i \sum_{i,m} \tilde{U}_{i,m} \left(\frac{\sqrt{\Delta t}}{\sqrt{\Delta d}} g(A_{i,m}) W_{i,m} + \Phi_m \delta_{i,0} \right) \right\} \prod_{n=0}^{\hat{N}} \prod_{j=0}^N \frac{d\tilde{U}_{j,n}}{2\pi}. \end{aligned} \quad (9.97)$$

Each $W_{i,m}$ is independently drawn from a Gaussian probability density function $P(W_{i,m}) = (2\pi)^{-1/2} e^{-W_{i,m}^2/2}$. Hence, setting

$$P[\mathbf{A}] = \int_{-\infty}^{\infty} P[\mathbf{A}|\mathbf{W}] \prod_{j,n} P(W_{j,n}) dW_{j,n}$$

and performing the integration with respect to $W_{j,n}$ by completing the square, gives

$$\begin{aligned} P[\mathbf{A}] &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp \left\{ -i \sum_{i,m} \tilde{U}_{i,m} (A_{i+1,m} - A_{i,m}) \right\} \\ &\quad \times \exp \left\{ -i \sum_{i,m} \tilde{U}_{i,m} \left[A_{i,m} - F \left(\Delta d \sum_n w_{mn} A_{i,n} \right) \right] \Delta t \right\} \\ &\quad \times \exp \left\{ \sum_{i,m} \left([i \tilde{U}_{i,m}]^2 g^2(A_{i,m}) \frac{\Delta t}{2 \Delta d} + i \tilde{U}_{i,m} \Phi_m \delta_{i,0} \right) \right\} \prod_{n=0}^N \prod_{j=0}^{\hat{N}} \frac{d \tilde{U}_{j,n}}{2\pi}. \end{aligned} \tag{9.98}$$

Finally, taking the continuum limits $\Delta d \rightarrow 0$, and $\Delta t \rightarrow 0$, $N, \hat{N} \rightarrow \infty$ for fixed T, L with $A_{i,m} \rightarrow A(x, t)$ and $i \tilde{U}_{i,m} / \Delta d \rightarrow \tilde{U}(x, t)$ results in the following path integral representation of a stochastic neural field:

$$P[A] = \int e^{-S[A, \tilde{U}]} \mathcal{D}\tilde{U} \tag{9.99}$$

with

$$\begin{aligned} S[A, \tilde{U}] &= \int_{\Sigma} \int_0^T \tilde{U}(x, t) \left[A_t(x, t) + A(x, t) - F \left(\int_{\Sigma} w(x-y) A(y, t) dy \right) \right. \\ &\quad \left. - \Phi(x) \delta(t) - \frac{1}{2} \tilde{U}(x, t) g^2(A(x, t)) \right] dt dx. \end{aligned} \tag{9.100}$$

Here $\mathcal{D}\tilde{U}$ denotes the path-integral measure

$$\mathcal{D}\tilde{U} = \lim_{N, \hat{N} \rightarrow \infty} \prod_{n=0}^{\hat{N}} \prod_{j=0}^N \frac{d \tilde{U}_{j,n} \Delta d}{2\pi i}.$$

Given the probability functional $P[A]$, a path integral representation of various moments of the stochastic field A can be constructed [22]. For example, the mean field is

$$\langle\langle A(x, t) \rangle\rangle = \int A(x, t) e^{-S[A, \tilde{U}]} \mathcal{D}A \mathcal{D}\tilde{U}, \tag{9.101}$$

whereas the two-point correlation is

$$\langle\langle A(x, t)A(x', t') \rangle\rangle = \int A(x, t)A(x', t')e^{-S[A, \tilde{U}]} \mathcal{D}A \mathcal{D}\tilde{U}. \quad (9.102)$$

Another important characterization of the stochastic neural field is how the mean activity (and other moments) respond to small external inputs (linear response theory). First, suppose that a small external source term $h(x, t)$ is added to the right-hand side of the deterministic version ($g \equiv 0$) of the field equation (9.93). Linearizing about the time-dependent solution $A(x, t)$ of the unperturbed equation ($h \equiv 0$) leads to an inhomogeneous linear equation for the perturbed solution $\varphi(x, t) = A^h(x, t) - A(x, t)$:

$$\frac{\partial \varphi}{\partial t} = -\varphi(x, t) + F' \left(\int_{\Sigma} w(x - y)A(y, t)dy \right) \int_{\Sigma} w(x - y)\varphi(y, t)dy + h(x, t). \quad (9.103)$$

Introducing the deterministic Green's function or propagator $\mathcal{G}_0(x, t; x', t')$ according to the adjoint equation

$$\begin{aligned} -\frac{\partial \mathcal{G}_0}{\partial t'} &= \delta(x - x')\delta(t - t') - \mathcal{G}_0(x, t; x', t') \\ &\quad + \int_{\Sigma} F' \left(\int_{\Sigma} w(y - z)A(z, t')dz \right) w(x' - y)\mathcal{G}_0(x, t; y, t')dy \end{aligned} \quad (9.104)$$

with $\mathcal{G}_0(x, t; x', t') = 0$ for $t \leq t'$ (causality), the linear response is given by

$$\varphi(x, t) = \int_{-\Sigma} \int_{-\Sigma}^t \mathcal{G}_0(x, t; x', t')h(x', t')dt'dx'. \quad (9.105)$$

In other words, in terms of functional derivatives

$$\frac{\delta A(x, t)}{\delta h(x', t')} = \mathcal{G}_0(x, t; x', t'). \quad (9.106)$$

Now suppose that a corresponding source term $\int dx \int dt h(x, t)\tilde{U}(x, t)$ is added to the action (9.100). The associated Green's function for the full stochastic model is defined according to

$$\mathcal{G}(x, t; x', t') \equiv \frac{\delta \langle\langle A(x, t) \rangle\rangle}{\delta h(x', t')} = \langle\langle A(x, t)\tilde{U}(x', t') \rangle\rangle \quad (9.107)$$

with

$$\lim_{t \rightarrow t'_+} \mathcal{G}(x, t; x', t') = \delta(x - x')$$

and $\mathcal{G}(x, t; x', t') = 0$ for $t \leq t'$. The above analysis motivates the introduction of the generating functional

$$Z[J, \tilde{J}] = \int e^{-S[A, \tilde{U}] + \int_{\Sigma} \int_0^T [\tilde{A}(x, t) J(x, t) + \tilde{J}(x, t) A(x, t)] dx dt} \mathcal{D}A \mathcal{D}\tilde{U}. \quad (9.108)$$

Various moments of physical interest can then be obtained by taking functional derivatives with respect to the “current sources” J, \tilde{J} . For example,

$$\langle\langle A(x, t) \rangle\rangle = \frac{\delta}{\delta \tilde{J}(x, t)} Z[J, \tilde{J}] \Big|_{J=\tilde{J}=0} \quad (9.109)$$

$$\langle\langle A(x, t) A(x', t') \rangle\rangle = \frac{\delta}{\delta \tilde{J}(x, t)} \frac{\delta}{\delta \tilde{J}(x', t')} Z[J, \tilde{J}] \Big|_{J=\tilde{J}=0} \quad (9.110)$$

$$\langle\langle A(x, t) \tilde{A}(x', t') \rangle\rangle = \frac{\delta}{\delta \tilde{J}(x, t)} \frac{\delta}{\delta J(x', t')} Z[J, \tilde{J}] \Big|_{J=\tilde{J}=0}. \quad (9.111)$$

9.5.3 Hamiltonian–Jacobi Dynamics and Population Extinction in the Weak-Noise Limit

In Ref. [11], the path-integral representation of the generating functional (9.108) is used to estimate the time to extinction of a metastable non-trivial state. That is, following along analogous lines to previous studies of reaction–diffusion equations [27, 42], the effective Hamiltonian dynamical system obtained by extremizing the associated path integral action is used to determine the most probable or optimal path to the zero absorbing state. (Alternatively, one could consider a WKB approximation of solutions to the corresponding functional Fokker–Planck equation or master equation [26, 33, 40].) In the case of the neural field equation, this results in extinction of all neural activity. For a corresponding analysis of a neural master equation with x -independent steady states, see Refs. [9, 10].

The first step is to perform the rescalings $\tilde{A} \rightarrow \tilde{A}/\sigma^2$ and $\tilde{J} \rightarrow \tilde{J}/\sigma^2$, so that the generating functional (9.108) becomes

$$Z[J, \tilde{J}] = \int e^{-\sigma^{-2} S[A, \tilde{U}] + \sigma^{-2} \int_{\Sigma} \int_0^T [\tilde{A}(x, t) J(x, t) + \tilde{J}(x, t) A(x, t)] dt dx} \mathcal{D}A \mathcal{D}\tilde{U}. \quad (9.112)$$

In the limit $\sigma \rightarrow 0$, the path integral is dominated by the “classical” solutions $q(x, t), p(x, t)$, which extremize the exponent or action of the generating functional:

$$\left. \frac{\delta S[A, \tilde{A}]}{\delta A(x, t)} \right|_{\tilde{A}=p, A=q} = -\tilde{J}(x, t), \quad \left. \frac{\delta S[A, \tilde{A}]}{\delta \tilde{A}(x, t)} \right|_{\tilde{A}=p, A=q} = -J(x, t). \quad (9.113)$$

In the case of zero currents $J = \tilde{J} = 0$, these equations reduce to

$$\frac{\partial q(x, t)}{\partial t} = \frac{\delta \mathcal{H}[q, p]}{\delta p(x, t)}, \quad \frac{\partial p(x, t)}{\partial t} = -\frac{\delta \mathcal{H}[q, p]}{\delta q(x, t)}, \quad (9.114)$$

where

$$S[q, p] = \int_0^T \left[\int_{\Sigma} p(x, t) \dot{q}(x, t) dx - \mathcal{H}[q, p] \right] dt - \int_{\Sigma} p(x, 0) \Phi(x) dx,$$

such that

$$\begin{aligned} \mathcal{H}[q, p] = \int_{\Sigma} p(x, t) & \left[-q(x, t) + F \left(\int_{\Sigma} w(x-y) q(y, t) dy \right) \right. \\ & \left. + \frac{1}{2} p(x, t) g^2(q(x, t)) \right] dx \end{aligned} \quad (9.115)$$

Equations (9.114) take the form of a Hamiltonian dynamical system in which q is a “coordinate” variable, p is its “conjugate momentum” and \mathcal{H} is the Hamiltonian functional. Substituting for \mathcal{H} leads to the explicit Hamilton equations

$$\frac{\partial q(x, t)}{\partial t} = -q(x, t) + F \left(\int_{\Sigma} w(x-y) q(y, t) dy \right) + p(x, t) g^2(q(x, t)) \quad (9.116)$$

$$\begin{aligned} \frac{\partial p(x, t)}{\partial t} = -p(x, t) & + \int_{\Sigma} F' \left(\int_{\Sigma} w(y-z) q(z, t) dz \right) w(y-x) p(y, t) dy \\ & + p^2(x, t) g(q(x, t)) g'(q(x, t)) \end{aligned} \quad (9.117)$$

It can be shown that $q(x, t), p(x, t)$ satisfy the same boundary conditions as the physical neural field $A(x, t)$ [49]. Thus, in the case of periodic boundary conditions, $q(x + L, t) = q(x, t)$ and $p(x + L, t) = p(x, t)$. It also follows from the Hamiltonian structure of Eqs. (9.116) and (9.117) that there is an integral of motion given by the conserved “energy” $E = \mathcal{H}[q, p]$.

The particular form of \mathcal{H} implies that one type of solution is the zero energy solution $p(x, t) \equiv 0$, which implies that $q(x, t)$ satisfies the deterministic scalar neural field equation (9.84). In the $t \rightarrow \infty$ limit, the resulting trajectory in the infinite dimensional phase space converges to the steady state solution $\mathcal{O}_+ = [q_s(x), 0]$, where $q_s(x)$ satisfies Eq. (9.92). The Hamiltonian formulation of extinction events then implies that the most probable path from $[q_s(x), 0]$ to the absorbing state is the unique zero energy trajectory that starts at \mathcal{O}_+ at time $t = -\infty$ and approaches another fixed point $\mathcal{P} = [0, p_e(x)]$ at $t = +\infty$ [27, 49]. In other words, this so-called activation trajectory is a heteroclinic connection $\mathcal{O}_+ \mathcal{P}$ (or instanton solution) in the functional phase space $[q(x), p(x)]$. It can be seen from Eq. (9.117) that the activation trajectory is given by the curve

$$p(x) = \mathcal{F}_x[q] \equiv -\frac{-q(x) + F(\int_{\Sigma} w(x-y)q(y)dy)}{g(q(x))^2} \quad (9.118)$$

so that

$$p_e(x) = \lim_{q \rightarrow 0} \mathcal{F}_x[q] = -(W_0 - 1) \lim_{q \rightarrow 0} \frac{q(x)}{g(q(x))^2}, \quad (9.119)$$

assuming $F(q) \sim q$ for $0 < q \ll 1$. Note that the condition that $p_e(x)$ exists and is finite is equivalent to the condition that there exists a stationary solution to the underlying functional Fokker–Planck equation – this puts restrictions on the allowed form for g . For the zero energy trajectory emanating from \mathcal{O}_+ at $t = -\infty$, the corresponding action is given by

$$S_0 = \int_{-\infty}^{\infty} \int_{\Sigma} p(x, t) \dot{q}(x, t) dx dt, \quad (9.120)$$

and up to pre-exponential factors, the estimated time τ_e to extinction from the steady-state solution $q_s(x)$ is given by [27, 49]

$$\ln \tau_e \approx \sigma^{-2} S_0. \quad (9.121)$$

For x -dependent steady-state solutions $q_s(x)$, which occur for Dirichlet boundary conditions and finite L , one has to solve Eqs. (9.116) and (9.117) numerically. Here we will consider the simpler case of x -independent solutions, which occur for periodic boundary conditions or Dirichlet boundary conditions in the large L limit (where boundary layers can be neglected). Restricting to x -independent state transitions, the optimal path is determined by the Hamilton equations (9.116) and (9.117):

$$\dot{q} = -q + F(W_0 q) + pg^2(q) \quad (9.122)$$

$$-\dot{p} = -p + W_0 F'(W_0 q) p + p^2 g(q) g'(q) \quad (9.123)$$

In Fig. 9.4 we plot the various constant energy solutions of the Hamilton equations (9.122) and (9.123) for the differentiable rate function $F(q) = \tanh(q)$ and multiplicative noise factor $g(q) = q^s$. In the first case $p_e = 0.2$ and in the second $p_e = 0$. The zero-energy trajectories are highlighted as thicker curves. Let us first consider the case $s = 1/2$ for which $p_e = 0.2$, see Fig. 9.4a. As expected, one zero-energy curve is the line $p = 0$ along which Eq. (9.122) reduces to the x -independent version of Eq. (9.84). If the dynamics were restricted to the one-dimensional manifold $p = 0$ then the non-zero fixed point $\mathcal{O}_+ = (q_0, 0)$ with $q_0 = F(W_0 q_0)$ would be stable. However, it becomes a saddle point of the full dynamics in the (q, p) plane, reflecting the fact that it is metastable when fluctuations are taken into account. A second zero-energy curve is the absorbing

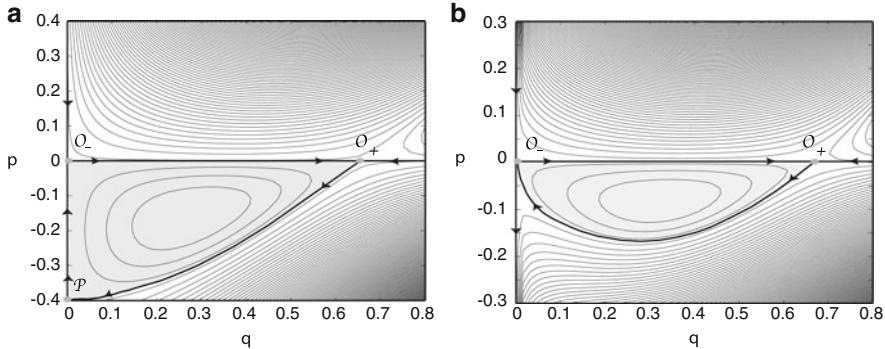


Fig. 9.4 Phase portrait of constant energy trajectories for the Hamiltonian system given by Eqs. (9.122) and (9.123) with $F(q) = \tanh(q)$ and $g(q) = q^s$ for $q > 0$. Zero-energy trajectories are indicated by thick curves. The stable and unstable fixed points of the mean-field dynamics are denoted by \mathcal{O}_+ and \mathcal{O}_- . (a) $s = 1/2$: There exists a non-zero fluctuational fixed point \mathcal{P} that is connected to \mathcal{O}_+ via a zero-energy heteroclinic connection. The curve $\mathcal{O}_+ P$ is the optimal path from the metastable state to the absorbing state. (b) $s = 1/4$: There is no longer a fluctuational fixed point \mathcal{P} so the optimal path is a direct heteroclinic connection between \mathcal{O}_+ and \mathcal{O}_-

line $q = 0$ which includes two additional hyperbolic fixed points denoted by $\mathcal{O}_- = (0, 0)$ and $\mathcal{P} = (0, p_e)$ in Fig. 9.4. \mathcal{O}_- occurs at the intersection with the line $p = 0$ and corresponds to the unstable zero activity state of the deterministic dynamics, whereas \mathcal{P} is associated with the effects of fluctuations. Moreover, there exists a third zero-energy curve, which includes a heteroclinic trajectory joining \mathcal{O}_- at $t = -\infty$ to the fluctuational fixed point \mathcal{P} at $t = +\infty$. This heteroclinic trajectory represents the optimal (most probable) path linking the metastable fixed point to the absorbing boundary. For $s < 1/2$, $p_e = 0$ and the optimal path is a heteroclinic connection from \mathcal{O}_+ to \mathcal{O}_- . In both cases, the extinction time τ_e is given by Eq. (9.121) with

$$S_0 = \int_{\mathcal{O}_+}^{\mathcal{P}} pdq, \quad (9.124)$$

where the integral evaluated along the heteroclinic trajectory from \mathcal{O}_+ to \mathcal{P} , which is equal to the area in the shaded regions of Fig. 9.4.

Note that since the extinction time is exponentially large in the weak noise limit, it is very sensitive to the precise form of the action S_0 and thus the Hamiltonian \mathcal{H} . This implies that when approximating the neural master equation of Buice et al. [19, 20] by a Langevin equation of the form (9.93) with $\sigma \sim 1/\sqrt{N}$, where N is the system size, the resulting Hamiltonian differs from that obtained directly from the master equation and can thus generate a poor estimate of the extinction time. This can be shown either by comparing the path integral representations of the generating functional for both stochastic processes or by comparing the WKB approximation of the master equation and corresponding Fokker–Planck equation. This particular issue is discussed elsewhere for neural field equations [9, 10].

Acknowledgements This work was supported by the National Science Foundation (DMS-1120327).

References

1. Abbott, L.F., van Vreeswijk, C.: Asynchronous states in networks of pulse-coupled oscillators. *Phys. Rev. E* **48**(2), 1483–1490 (1993)
2. Amari, S.: Dynamics of pattern formation in lateral inhibition type neural fields. *Biol. Cybern.* **27**, 77–87 (1977)
3. Amit, D.J., Brunel, N.: Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cereb Cortex* **7**, 237–252 (1997)
4. Armero, J., Casademunt, J., Ramirez-Piscina, L., Sancho, J.M.: Ballistic and diffusive corrections to front propagation in the presence of multiplicative noise. *Phys. Rev. E* **58**, 5494–5500 (1998)
5. Bedard, C., Destexhe, A.: Macroscopic models of local field potentials the apparent 1/f noise in brain activity. *Biophys. J.* **96**, 2589–2603 (2009)
6. Beggs, J.M., Plenz, D.: Neuronal avalanches are diverse and precise activity patterns that are stable for many hours in cortical slice cultures. *J. Neurosci.* **24**, 5216–5229 (2004)
7. Boustani, S.E., Destexhe, A.: A master equation formalism for macroscopic modeling of asynchronous irregular activity states. *Neural Comput.* **21**, 46–100 (2009)
8. Bressloff, P.C.: Traveling fronts and wave propagation failure in an inhomogeneous neural network. *Physica D* **155**, 83–100 (2001)
9. Bressloff, P.C.: Stochastic neural field theory and the system-size expansion. *SIAM J. Appl. Math.* **70**, 1488–1521 (2009)
10. Bressloff, P.C.: Metastable states and quasicycles in a stochastic Wilson-Cowan model of neuronal population dynamics. *Phys. Rev. E* **85**, 051,903 (2010)
11. Bressloff, P.C.: From invasion to extinction in heterogeneous neural fields. *J. Math. Neurosci.* **2**, 6 (2012)
12. Bressloff, P.C.: Spatiotemporal dynamics of continuum neural fields. *J. Phys. A* **45**, 033,001 (109pp.) (2012)
13. Bressloff, P.C., Coombes, S.: Dynamics of strongly coupled spiking neurons. *Neural Comput.* **12**, 91–129 (2000)
14. Bressloff, P.C., Newby, J.M.: Metastability in a stochastic neural network modeled as a velocity jump markov process. *SIAM J. Appl. Dyn. Syst.* **12**, 1394–1435 (2013)
15. Bressloff, P.C., Webber, M.A.: Front propagation in stochastic neural fields. *SIAM J. Appl. Dyn. Syst.* **11**, 708–740 (2012)
16. Brunel, N.: Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *J. Comput. Neurosci.* **8**, 183–208 (2000)
17. Brunel, N., Hakim, V.: Fast global oscillations in networks of integrate-and-fire neurons with low firing rates. *Neural Comput.* **11**, 1621–1671 (1999)
18. Buckwar, E., Riedler, M.G.: An exact stochastic hybrid model of excitable membranes including spatio-temporal evolution. *J. Math. Biol.* **63**, 1051–1093 (2011)
19. Buice, M., Cowan, J.D.: Field-theoretic approach to fluctuation effects in neural networks. *Phys. Rev. E* **75**, 051,919 (2007)
20. Buice, M., Cowan, J.D., Chow, C.C.: Systematic fluctuation expansion for neural network activity equations. *Neural Comput.* **22**, 377–426 (2010)
21. Cai, D., Tao, L., Shelley, M., McLaughlin, D.W.: An effective kinetic representation of fluctuation-driven neuronal networks with application to simple and complex cells in visual cortex. *Proc. Natl. Acad. Sci. USA* **101**, 7757–7562 (2004)
22. Chow, C.C., Buice, M.: Path integral methods for stochastic differential equations (2011). arXiv nlin/105966v1

23. Coombes, S., Laing, C.R.: Pulsating fronts in periodically modulated neural field models. *Phys. Rev. E* **83**, 011,912 (2011)
24. Coombes, S., Owen, M.R.: Evans functions for integral neural field equations with Heaviside firing rate function. *SIAM J. Appl. Dyn. Syst.* **4**, 574–600 (2004)
25. de Pasquale, F., Gorecki, J., Pojelański, J.: On the stochastic correlations in a randomly perturbed chemical front. *J. Phys. A* **25**, 433 (1992)
26. Dykman, M.I., Mori, E., Ross, J., Hunt, P.M.: Large fluctuations and optimal paths in chemical kinetics. *J. Chem. Phys. A* **100**, 5735–5750 (1994)
27. Elgart, V., Kamenev, A.: Rare event statistics in reaction–diffusion systems. *Phys. Rev. E* **70**, 041,106 (2004)
28. Ermentrout, G.B.: Reduction of conductance-based models with slow synapses to neural nets. *Neural Comput.* **6**, 679–695 (1994)
29. Ermentrout, G.B., McLeod, J.B.: Existence and uniqueness of travelling waves for a neural network. *Proc. R. Soc. Edinb. A* **123**, 461–478 (1993)
30. Ermentrout, G.B., Terman, D.: *Mathematical Foundations of Neuroscience*. Springer, Berlin (2010)
31. Faisal, A.A., Selen, L.P.J., Wolpert, D.M.: Noise in the nervous system. *Nat. Rev. Neurosci.* **9**, 292 (2008)
32. Faugeras, O., Touboul, J., Cessac, B.: A constructive mean–field analysis of multi–population neural networks with random synaptic weights and stochastic inputs. *Front. Comput. Neurosci.* **3**, 1–28 (2009)
33. Freidlin, M.I., Wentzell, A.D.: *Random perturbations of dynamical systems*. Springer, New York (1984)
34. Gardiner, C.W.: *Handbook of Stochastic Methods*, 4th edn. Springer, Berlin (2009)
35. Gerstner, W., Kistler, W.: *Spiking Neuron Models*. Cambridge University Press, Cambridge (2002)
36. Gerstner, W., Van Hemmen, J.L.: Coherence and incoherence in a globally coupled ensemble of pulse–emitting units. *Phys. Rev. Lett.* **71**(3), 312–315 (1993)
37. Ginzburg, I., Sompolinsky, H.: Theory of correlations in stochastic neural networks. *Phys. Rev. E* **50**, 3171–3191 (1994)
38. Hutt, A., Longtin, A., Schimansky-Geier, L.: Additive noise induces turing transitions in spatial systems with application to neural fields and the Swift-Hohenberg equation. *Physica D* **237**, 755–773 (2008)
39. Ly, C., Tranchina, D.: Critical analysis of a dimension reduction by a moment closure method in a population density approach to neural network modeling. *Neural Comput.* **19**, 2032–2092 (2007)
40. Maier, R.S., Stein, D.L.: Limiting exit location distribution in the stochastic exit problem. *SIAM J. Appl. Math.* **57**, 752–790 (1997)
41. Mattia, M., Guidice, P.D.: Population dynamics of interacting spiking neurons. *Phys. Rev. E* **66**, 051,917 (2002)
42. Meerson, B., Sasorov, P.V.: Extinction rates of established spatial populations. *Phys. Rev. E* **83**, 011,129 (2011)
43. Mendez, V., Fort, J., Rotstein, H.G., Fedotov, S.: Speed of reaction-diffusion fronts in spatially heterogeneous media. *Phys. Rev. E* **68**, 041,105 (2003)
44. Meyer, C., van Vreeswijk, C.: Temporal correlations in stochastic networks of spiking neurons. *Neural Comput.* **14**, 369–404 (2002)
45. Novikov, E.A.: Functionals and the random-force method in turbulence theory. *Sov. Phys. JETP* **20**, 1290 (1965)
46. Nykamp, D., Tranchina, D.: A population density method that facilitates large–scale modeling of neural networks: analysis and application to orientation tuning. *J. Comput. Neurosci.* **8**, 19–50 (2000)
47. Ohira, T., Cowan, J.D.: Stochastic neurodynamics and the system size expansion. In: Ellacott, S., Anderson, I.J. (eds.) *Proceedings of the First International Conference on Mathematics of Neural Networks*, pp. 290–294. Academic (1997)

48. Omurtag, A., Knight, B.W., Sirovich, L.: On the simulation of large populations of neurons. *J. Comput. Neurosci.* **8**, 51–63 (2000)
49. Ovaskainen, O., Meerson, B.: Stochastic models of population extinction. *Trends Ecol. Evol.* **25**, 643–652 (2010)
50. Pakdaman, K., Thieullen, M., Wainrib, G.: Fluid limit theorems for stochastic hybrid systems with application to neuron models. *J. Appl. Probab.* **24**, 1 (2010)
51. Plenz, D., Thiagarajan, T.C.: The organizing principles of neuronal avalanches: cell assemblies in the cortex? *Trends Neurosci.* **30**, 101–110 (2007)
52. Rangan, A.V., Kovacic, G., Cai, D.: Kinetic theory for neuronal networks with fast and slow excitatory conductances driven by the same spike train. *Phys. Rev. E* **77**, 041,915 (2008)
53. Renart, A., Brunel, N., Wang, X.J.: Mean-field theory of irregularly spiking neuronal populations and working memory in recurrent cortical networks. In: Feng, J. (ed.) *Computational Neuroscience: A Comprehensive Approach*, pp. 431–490. CRC, Boca Raton (2004)
54. Rocco, A., Ebert, U., van Saarloos, W.: Subdiffusive fluctuations of “pulled” fronts with multiplicative noise. *Phys. Rev. E* **65**, R13–R16 (2000)
55. Sagues, F., Sancho, J.M., Garcia-Ojalvo, J.: Spatiotemporal order out of noise. *Rev. Mod. Phys.* **79**, 829–882 (2007)
56. Schimansky-Geier, L., Mikhailov, A.S., Ebeling, W.: Effects of fluctuations on plane front propagation in bistable nonequilibrium systems. *Ann. Phys.* **40**, 277 (1983)
57. Smith, G.D.: Modeling the stochastic gating of ion channels. In: Fall, C., Marland, E.S., Wagner, J.M., Tyson, J.J. (eds.) *Computational Cell Biology*, chap. 11. Springer, New York (2002)
58. Softky, W.R., Koch, C.: The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPS. *J. Neurosci.* **13**, 334–350 (1993)
59. Soula, H., Chow, C.C.: Stochastic dynamics of a finite-size spiking neural network. *Neural Comput.* **19**, 3262–3292 (2007)
60. Tauber, U.C.: Field-theory approaches to nonequilibrium dynamics. *Lect. Notes Phys.* **716**, 295–348 (2007)
61. Touboul, J., Hermann, G., Faugeras, O.: Noise-induced behaviors in neural mean field dynamics. *SIAM J. Appl. Dyn. Syst.* **11**(1), 49–81 (2012)
62. van Saarloos, W.: Front propagation into unstable states. *Phys. Rep.* **386**, 29–222 (2003)
63. van Vreeswijk, C., Abbott, L.F.: Self-sustained firing in populations of integrate-and-fire neurons. *SIAM J. Appl. Math.* **53**(1), 253–264 (1993)
64. van Vreeswijk, C., Sompolinsky, H.: Chaotic balanced state in a model of cortical circuits. *Neural Comput.* **10**, 1321–1371 (1998)
65. Zeisler, S., Franz, U., Wittich, O., Liebscher, V.: Simulation of genetic networks modelled by piecewise deterministic markov processes. *IET Syst. Biol.* **2**, 113–135 (2008)
66. Zhang, L.: On the stability of traveling wave solutions in synaptically coupled neuronal networks. *Differ. Integral Equ.* **16**, 513–536 (2003)
67. Zinn-Justin, J.: *Quantum Field Theory and Critical Phenomena*, 4th edn. Oxford University Press, Oxford (2002)

Chapter 10

On the Electrodynamics of Neural Networks

Peter beim Graben and Serafim Rodrigues

Abstract We present a microscopic approach for the coupling of cortical activity, as resulting from proper dipole currents of pyramidal neurons, to the electromagnetic field in extracellular fluid in presence of diffusion and Ohmic conduction. Starting from a full-fledged three-compartment model of a single pyramidal neuron, including shunting and dendritic propagation, we derive an observation model for dendritic dipole currents in extracellular space and thereby for the dendritic field potential that contributes to the local field potential of a neural population. Under reasonable simplifications, we then derive a leaky integrate-and-fire model for the dynamics of a neural network, which facilitates comparison with existing neural network and observation models. In particular, we compare our results with a related model by means of numerical simulations. Performing a continuum limit, neural activity becomes represented by a neural field equation, while an observation model for electric field potentials is obtained from the interaction of cortical dipole currents with charge density in non-resistive extracellular space as described by the Nernst-Planck equation. Our work consistently satisfies the widespread dipole assumption discussed in the neuroscientific literature.

10.1 Introduction

Hans Berger's 1924 discovery of the human *electroencephalogram* (EEG) [11] lead to a tremendous research enterprise in clinical, cognitive and computational neurosciences [58]. However, one of the yet unresolved problems in the biophysics

P. beim Graben (✉)

Bernstein Center for Computational Neuroscience, Göttingen, Germany Department of German Studies and Linguistics, Humboldt-Universität zu Berlin, Berlin, Germany

S. Rodrigues

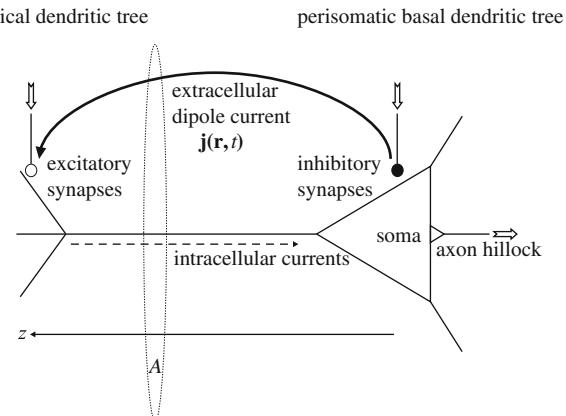
School of Computing and Mathematics, Centre for Robotics and Neural Systems, University of Plymouth, Plymouth, UK

of neural systems is understanding the proper coupling of complex neural network dynamics to the electromagnetic field, that is macroscopically measurable by means of neural mass potentials, such as *local field potential* (LFP) or electroencephalogram. One requirement for this understanding is a *forward model* that links the ‘hidden’ activities of billions of neurons in mammalian brains and their propagation through neural networks to experimentally accessible quantities such as LFP and EEG. Utilizing terminology from theoretical physics, we call the operationally accessible quantities *observables* and an integrative forward model an *observation model*. Yet, there is an ongoing debate in the literature whether field effects, i.e. the *feedback* from mass potentials to neural activity, plays a functional role in the self-organization of cortical activity (e.g. [40]). Such field effects have recently been demonstrated via experiments on ephaptic interaction [31]. Thus a theoretical framework for observation forward and feedback models is mandatory in order to describe the coupling between neural network activity and the propagation of extracellular electromagnetic fields in clinical, computational and cognitive neurosciences, e.g. for treatment of epilepsy [55] or for modeling cognition-related brain potentials [8, 46].

Currently, there is ample evidence that the generators of neural field potentials, such as cortical LFP and EEG are the cortical pyramidal cells (sketched in Fig. 10.1). They exhibit a long dendritic trunk separating mainly excitatory synapses at the apical dendritic tree from mainly inhibitory synapses at the perisomatic basal dendritic tree [23, 60]. When both kinds of synapses are simultaneously active, inhibitory synapses generate current sources and excitatory synapses current sinks in extracellular space, causing the pyramidal cell to behave as a microscopic dipole surrounded by its characteristic electrical field. This dendritic dipole field is conveniently described by its associated electrodynamic potential, the *dendritic field potential (DFP)*. Dendritic fields superimpose to the field of a cortical dipole layer, which is measurable as cortical LFP, due to the geometric arrangement of pyramidal cells in a cortical column. There, pyramidal cells exhibit an axial symmetry and they are aligned in parallel to each other, perpendicular to the cortex’ surface, thus forming a palisade of cell bodies and dendritic trunks. Eventually, cortical LFP gives rise to the EEG measurable at the human’s scalp [27, 53, 58].

Weaving the above phenomena into a mathematical and biophysical plausible observation model that represents correctly the multi-spatiotemporal characteristics of LFP is a non-trivial task. The difficulty results from the complexity of brain processes that operate at several spatial and temporal scales. On one hand the organization of the brain, from single neuron scales to that of whole brain regions, changes its connectivity from almost probabilistic to highly structured as discussed above in the case of the cortical columns. On the other hand, temporal dynamics in time scales ranges from milliseconds for discrete events like spikes to hours and even longer for synaptic plasticity and learning. Hence, there is strong interaction between the different spatiotemporal scales [12, 45], which directly contribute to complex oscillatory dynamics, e.g., to *mixed-mode oscillations* [25, 29]. Thus it is not clear how and when to break down complex brain processes into simpler ‘building blocks’ where analysis can be made. Despite these peculiarities, various

Fig. 10.1 Sketch of a cortical pyramidal neuron with extracellular current dipole between spatially separated excitatory (open bullet) and inhibitory synapses (filled bullet). Neural in- and outputs are indicated by the jagged arrows. The z -axis points toward the scull. Current density j is given by dendritic current I_1 through cross section area A as described in the text



mathematical and computational approaches have been proposed in order to establish coarse-graining techniques and how to move from one scale to another.

Most studies for realistically simulating LFP, typically for the extracellular fluid in the vicinity of a neuron, have been attempted by means of compartmental models [2, 48, 54, 57] where every compartment contributes a portion of extracellular current to the DFP that is given by Coulomb's equation in conductive media [5, 7, 53]. However, because compartmental models are computationally expensive, large-scale neural network simulations preferentially employ point models, based either on conductance [36, 50] or population models [39, 56, 64, 65] where neural mass potentials are estimated either through sums (or rather differences) of postsynaptic potentials [24] or of postsynaptic currents [50]. In particular, the model of Mazzoni et al. [50] led to a series of recent follow-up studies [51, 52] that address the correlations between numerically simulated or experimentally measured LFP/EEG and spike rates by means of statistical modeling and information theoretic measures.

To adequately explain field potentials measured around the dendritic tree of an individual cortical pyramidal cell (DFP), in extracellular space of a cortical module (LFP), or at a human's scalp (EEG), Maxwell's electromagnetic field equations, specifically the continuity equation describing conservation of charge have to be taken into account. However, coupling the activity of discrete neural networks to the continuous electromagnetic field is difficult since neural network topology is not embedded into physical space as an underlying metric manifold. This can be circumvented by employing continuous neural networks as investigated in *neural field theory (NFT)* [1, 15, 18, 38, 41, 66]. In fact previous studies [42, 47] gave the first reasonable accounts for such couplings in NFT population models that are motivated by the corresponding assumptions for neural mass models (cf. Chap. 17 in this volume). Jirsa et al. [42] relate the impressed current density in extracellular space to neural field activity. On the other hand, Liley et al. [47] consider LFP as average somatic membrane potential being proportional to the neural field. Their model found a number of successful applications [13, 14, 22] (see also Chap. 14). However, both approaches [42, 47] are not concerned with the microscopic geometry

around the field generators, the cortical pyramidal cells. Therefore, they do not take pyramidal dipole currents into account.

Another problem with the aforementioned neural field approaches is that the extracellular space has either been completely neglected, or only implicitly been taken into account by assuming that cortical LFP is proportional to either membrane potentials or synaptic currents as resulting from a purely resistive medium. That means, dipole currents in the extracellular space have been completely abandoned. However, recent studies indicate that at least the resistive property of the extracellular space is crucial [49], but more interestingly, it has been revealed that diffusion currents, represented by their corresponding Warburg impedances [59], cannot be neglected in extracellular space as they may substantially contribute to the characteristic power spectra of neural mass potentials [3, 4, 6, 27].

In this chapter, we outline a theoretical framework for the microscopic coupling of continuous neural networks, i.e. neural fields, to the electromagnetic field, properly described by dipole currents of cortical pyramidal neurons and diffusion effects in extracellular space. As a starting point we use a three-compartment model for a single pyramidal cell [7, 26, 63] and derive the evolution law for the activity of a neural network. These derivations additionally include observation equations for the extracellular dipole currents, which explicitly incorporate extracellular resistivity and diffusion. Subsequently, we demonstrate that our approach can be related to previous modeling strategies, by considering reasonable simplifications. Herein, we intentionally and specifically simplify our approach to a leaky integrate-and-fire (LIF) model for the dynamics of a neural network, which then shows the missing links that previous modeling approaches failed to incorporate to account for a proper dipole LFP observation model. In particular, we compare our results with the related model by Mazzoni et al. [50] by means of numerical simulations. Moreover, performing the continuum limit (yet *à la physique*) for the network yields an Amari-type neural field equation [1] coupled to the Maxwell equations in extracellular fluid, while an observation model for electric field potentials is obtained from the interaction of cortical dipole currents with charge density in non-resistive extracellular space as described by the Nernst-Planck equation. Thereby, our work provides for the first time a biophysically plausible observation model for the Amari-type neural field equations and crucially, it gives estimates for the local field potentials that satisfy the widespread dipole assumption discussed in the neuroscientific literature.

10.2 Pyramidal Neuron Model

Inspired by earlier attempts to approximate the complex shape of cortical pyramidal neurons by essentially three passively coupled compartments [7, 26, 63], we reproduce in Fig. 10.2 the electronic equivalent circuit of beim Graben [7] for the i th pyramidal cell (Fig. 10.1) in a population of P pyramidal neurons here. This parsimonious circuit allows to derive our observation model. It comprises one compartment for the apical dendritic tree where p_i excitatory synapses are situated

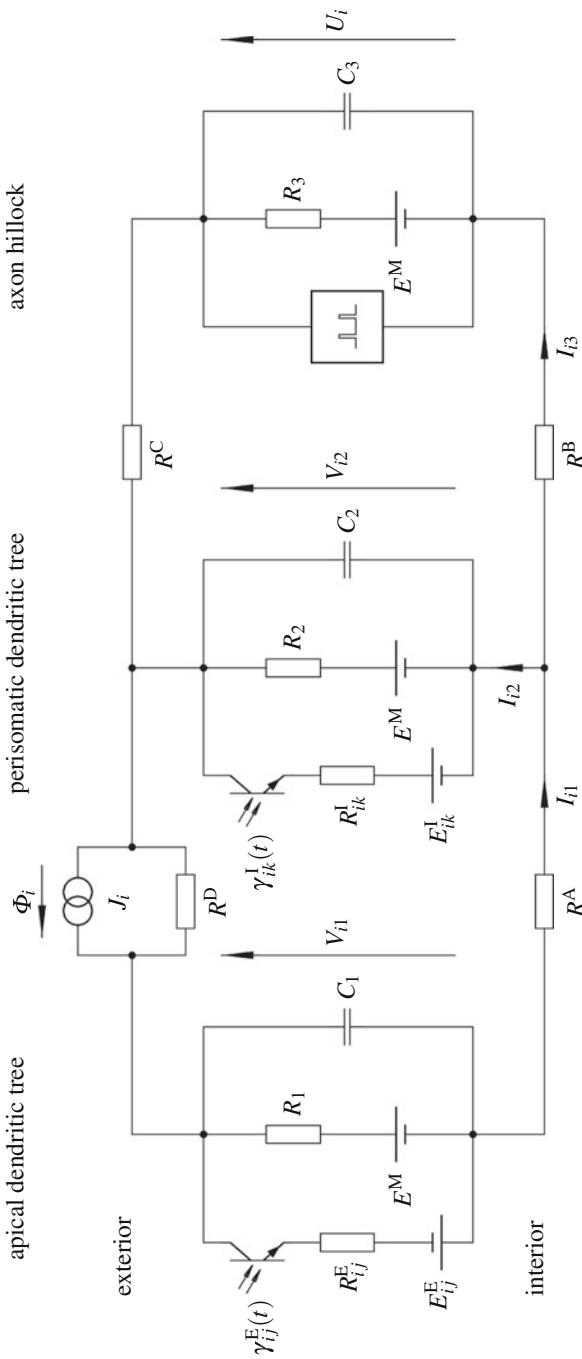


Fig. 10.2 Equivalent circuit for a three-compartment pyramidal neuron model

(for the sake of simplicity, we only show one synapse here), another one for the soma and perisomatic basal dendritic tree, populated with q_i mainly inhibitory synapses (again, only one synapse is shown here), and a third one for the axon hillock where membrane potential is converted into spike trains by means of an integrate-and-fire mechanism. Note that nonlinear fire mechanisms of Hodgkin-Huxley type can be incorporated as well. In total we consider N populations of neurons, arranged in two-dimensional layers $\Gamma_n \subset \mathbb{R}^2$ ($i = n, \dots, N$). Neurons in layers 1 to M should be excitatory, neurons in layers $M + 1$ to N should be inhibitory and layer 1 exclusively contains the P cortical pyramidal cells in our simplified treatment. The total number of neurons should be K .

Excitatory synapses are schematically represented by the left-most branch of Fig. 10.2 as ‘phototransistors’ [7] in order to indicate that they comprise quantated resistors, namely ion channels whose resistance depends on the concentration of ligand molecules which are either extracellular neurotransmitters or intracellular metabolites [44]. There, the excitatory postsynaptic current (EPSC) at a synapse between a neuron j from layers 1 to M and neuron i is given as

$$I_{ij}^E(t) = \frac{\gamma_{ij}^E(t)}{R_{ij}^E}(V_{i1}(t) - E_{ij}^E). \quad (10.1)$$

Here, the time-dependent function $\gamma_{ij}^E(t)$ reflects the neurotransmitter-gated opening of postsynaptic ion channels. Usually, this function is given as a sum of characteristic excitatory impulse response functions $\eta^E(t)$ that is elicited by one presynaptic spike, i.e.

$$\gamma(t) = \sum_v \eta(t - t_v) \quad (10.2)$$

where t_v denote the ordered spike arrival times. Moreover, R_{ij}^E comprises the maximum synaptic conductivity as well as the electrotonic distance between the synapse between neuron j and i and i ’s trigger zone, both expressed as resistance. $V_{i1}(t)$ is the membrane potential of neuron i ’s compartment 1, i.e. the apical dendritic tree and E_{ij}^E is the excitatory reversal potential of the synapse $j \rightarrow i$. We can conveniently express $\gamma(t)$ through the spike rate [10, 18]

$$a(t) = \sum_v \delta(t - t_v) \quad (10.3)$$

by means of a temporal convolution (‘ $*$ ’ denotes convolution product)

$$\gamma(t) = \int_{-\infty}^t \eta(t - t') a(t') dt' = (\eta * a)(t). \quad (10.4)$$

Furthermore, the apical dendritic compartment 1 is characterized by a particular leakage resistance R_1 and by a capacity C_1 , reflecting the temporary charge storage

capacity of the membrane. Both, R_1 and C_1 are correlated with the compartment's membrane area [26]. The battery E^M denotes the Nernst resting potential [43, 62].

The middle branch of Fig. 10.2 describes the inhibitory synapses (also displayed as ‘phototransistors’ [7]) between a neuron k from layers $M + 1$ to N and neuron i . Here, inhibitory postsynaptic currents (IPSC)

$$I_{ik}^I(t) = \frac{\gamma_{ik}^I(t)}{R_{ik}^I}(V_{i2}(t) - E_{ik}^I), \quad (10.5)$$

described by a similar channel opening function $\gamma^I(t)$, shunt the excitatory branch with the trigger zone when compartment's 2 membrane potential $V_{i2}(t)$ is large due to previous excitation. Also Eqs. (10.2) and (10.3) hold for another postsynaptic impulse response function $\eta^I(t)$, characteristic for inhibitory synapses. The resistance of the current paths along the cell plasma is given by R_{ik}^I . Finally, E_{ik}^I denotes the inhibitory reversal potential of the synapse $k \rightarrow i$. Also the somatic and perisomatic dendritic compartment 2 possesses its specific leakage resistance R_2 and capacity C_2 .

The cell membrane at the axon hillock [36] itself is represented by the branch at the right hand side described by another RC-element consisting of R_3 and C_3 . Action potentials, $\delta(t - t_v)$, are generated by a leaky integrate-and-fire mechanism [50] as indicated by a ‘black box’ when the membrane potential $U_i(t)$ crosses a certain threshold $\theta_i > 0$ from below at time t_v , i.e.

$$U_i(t_v) \geq \theta_i. \quad (10.6)$$

Afterwards, membrane potential is reset to some steady-state potential [50]

$$U_i(t_{v+1}) \leftarrow E. \quad (10.7)$$

and the integration of the differential equations can be restarted at time $t = t_{v+1} + \tau_{rp}$ after interrupting the dynamics for a refractory period τ_{rp} .

The three compartments are coupled through longitudinal resistors, R^A , R^B , R^C , R^D where R^A , R^B denote the resistivity of the cell plasma and R^C , R^D that of extracellular space [37]. Yet, in extracellular space not only Ohmic but also diffusion currents are present [3, 4, 6, 32–34, 61]. These are taken into account by the current source J_i connected in parallel to R^D . However, for convenience, diffusion currents in the extracellular space between compartments 2 and 3 are disregarded following an adiabatic approximation.

Finally, the membrane potentials at compartments 1, V_{i1} , 2, V_{i2} , and 3, U_i , as the dynamical state variable as well as the DFP Φ_i are shown in Fig. 10.2. The latter drops along the extracellular resistor R^D . For the aim of calculation, the mesh currents I_{i1} (current in the apical compartment 1 of neuron i), I_{i2} (current in somatic and perisomatic compartment 2 of neuron i) and I_{i3} (the leaky integrate-and-fire (LIF) current in compartment 3 of neuron i) are indicated.

The circuit in Fig. 10.2 obeys Kirchhoff's laws; first for currents:

$$I_{i1} = C_1 \frac{dV_{i1}}{dt} + \frac{V_{i1} - E^M}{R_1} + \sum_{j=1}^{p_i} I_{ij}^E \quad (10.8)$$

$$I_{i2} = C_2 \frac{dV_{i2}}{dt} + \frac{V_{i2} - E^M}{R_2} + \sum_{k=1}^{q_i} I_{ik}^I \quad (10.9)$$

$$I_{i3} = C_3 \frac{dU_i}{dt} + \frac{U_i - E^M}{R_3} \quad (10.10)$$

$$I_{i3} = I_{i1} - I_{i2}, \quad (10.11)$$

and second, for voltages:

$$V_{i1} = (R^A + R^D)I_{i1} + (R^B + R^C)I_{i3} + U_i - R^D J_i \quad (10.12)$$

$$V_{i2} = (R^B + R^C)I_{i3} + U_i \quad (10.13)$$

$$\Phi_i = R^D(I_{i1} - J_i), \quad (10.14)$$

where the postsynaptic currents I_{ij}^E and I_{ik}^I are given through (10.1) and (10.5). Here, p_i is the number of excitatory and q_i the number of inhibitory synapses connected to neuron i , respectively.

Subtracting (10.13) from (10.12) yields the current along the pyramidal cell's dendritic trunk

$$I_{i1} = \frac{V_{i1} - V_{i2} + R^D J_i}{R^A + R^D}. \quad (10.15)$$

The circuit described by Eqs. (10.8–10.14) shows that the neuron i is likely to fire when the excitatory synapses are activated. Then, the LIF current I_{i3} equals the dendritic current I_{i1} . If, by contrast, also the inhibitory synapses are active, the dendritic current I_{i1} follows the shortcut between the apical and perisomatic dendritic trees and only a portion could evoke spikes at the trigger zone (Eq. (10.10)). On the other hand, the large dendritic current I_{i1} , diminished by some diffusion current J_i , flowing through the extracellular space of resistance R_i^D , gives rise to a large DFP Φ_i .

In order to simplify the following derivations, we first gauge the resting potential to $E^M = 0$. Then, excitatory synapses are characterized by $E_{ij}^E > 0$, while inhibitory synapses obey $E_{ik}^I < 0$. Combining Eqs. (10.8–10.13) entails

$$C_1 \frac{dV_{i1}}{dt} + \frac{V_{i1}}{R_1} + \sum_{j=1}^{p_i} I_{ij}^E = \frac{V_{i1} - V_{i2} + R^D J_i}{R^A + R^D} \quad (10.16)$$

$$C_2 \frac{dV_{i2}}{dt} + \frac{V_{i2}}{R_2} + \sum_{k=1}^{q_i} I_{ik}^I = \frac{V_{i1} - V_{i2} + R^D J_i}{R^A + R^D} - \frac{V_{i2} - U_i}{R^B + R^C} \quad (10.17)$$

$$C_3 \frac{dU_i}{dt} + \frac{U_i}{R_3} = \frac{V_{i2} - U_i}{R^B + R^C}. \quad (10.18)$$

10.2.1 General Solution of the Circuit Equations

Next we follow Bressloff's [16] argumentation and regard the compartmental voltages as auxiliary variables that are merged into a two-dimensional vector $\mathbf{V}_i = (V_{i1}, V_{i2})^T$ which is subject to elimination. We only keep Eq. (10.18) as the evolution law of the entire state vector $\mathbf{U} = (U_i)_{i=1,\dots,K}$ of the neural network. Inserting the postsynaptic currents from (10.1) and (10.5) into Eqs. (10.16, 10.17) and temporarily assuming an arbitrary time-dependence for the functions $\gamma(t)$ from Eq. (10.2) (in fact, the $\gamma(t)$ are given through the presynaptic spike rates and are thus nonlinear functions of the entire state \mathbf{U}), we obtain a system of two inhomogeneous linear differential equations that can be compactly written in matrix form as

$$\frac{d}{dt} \mathbf{V}_i(t) = \mathbf{H}_i(t) \cdot \mathbf{V}_i(t) + \mathbf{G}_i(t), \quad (10.19)$$

with

$$\mathbf{H}_i(t) = \begin{pmatrix} \frac{1}{C_1} \left(-\frac{1}{R_1} + \frac{1}{R^A+R^D} - \sum_j \frac{\gamma_{ij}^E(t)}{R_{ij}^E} \right) & -\frac{1}{C_1(R^A+R^D)} \\ \frac{1}{C_2(R^A+R^D)} & \frac{1}{C_2} \left(-\frac{1}{R_2} - \frac{1}{R^A+R^D} - \frac{1}{R^B+R^C} - \sum_k \frac{\gamma_{ik}^I(t)}{R_{ik}^I} \right) \end{pmatrix} \quad (10.20)$$

and

$$\mathbf{G}_i(t) = \begin{pmatrix} \sum_j \frac{\gamma_{ij}^E(t) E_{ij}^E}{C_1 R_{ij}^E} + \frac{R^D}{C_1(R^A+R^D)} J_i(t) \\ \sum_k \frac{\gamma_{ik}^I(t) E_{ik}^I}{C_2 R_{ik}^I} + \frac{R^D}{C_2(R^A+R^D)} J_i(t) + \frac{1}{C_2(R^B+R^C)} U_i(t) \end{pmatrix}. \quad (10.21)$$

Initial conditions start with $\mathbf{V} = 0$ in the infinite past $t = -\infty$ for the sake of convenience.

Obviously, the time-dependence of the transition matrix $\mathbf{H}(t)$ is due to the shunting terms $\gamma(t)$. In order to solve (10.19) one first considers the associated homogeneous differential equation

$$\frac{d}{dt} \mathbf{W}_i(t) = \mathbf{H}_i(t) \cdot \mathbf{W}_i(t) \quad (10.22)$$

whose general solutions are given as the columns of

$$\mathbf{W}_i(t) = \mathbb{T} e^{t \mathbf{H}_i(t)}, \quad (10.23)$$

where \mathbb{T} denotes the *time-ordering operator* [20, 21]. Using the integral (10.23), a particular solution of the inhomogeneous equation (10.19) is then obtained by the variation of parameter method as

$$\mathbf{V}_i(t) = \int_{-\infty}^t \mathbf{X}_i(t, t') \cdot \mathbf{G}_i(t') dt' \quad (10.24)$$

with matrix-valued Green's function

$$\mathbf{X}_i(t, t') = \mathbf{W}_i(t) \cdot \mathbf{W}_i(t')^{-1}. \quad (10.25)$$

Therefore, the compartmental voltages are obtained as

$$V_{i\alpha}(t) = \sum_{\beta=1}^2 \int_{-\infty}^t \chi_{i\alpha\beta}(t, t') g_{i\beta}(t') dt' = \sum_{\beta=1}^2 \chi_{i\alpha\beta} * g_{i\beta} \quad (10.26)$$

with components $\mathbf{X}_i(t, t') = (\chi_{i\alpha\beta}(t, t'))_{\alpha\beta}$ and $\mathbf{G}_i(t') = (g_{i\beta}(t'))_{\beta}$, $\alpha, \beta = 1, 2$.

10.2.2 Observation Model

In order to derive the general observation equations for the DFP of the three-compartment model, we insert the formal solutions (10.26) and the inhomogeneity (10.21) into Eq. (10.15) and obtain

$$\begin{aligned} I_{i1}(t) &= \frac{1}{R^A + R^D} \int_{-\infty}^t (\chi_{i11}(t, t') - \chi_{i21}(t, t')) \left[\sum_j \frac{E_{ij}^E}{C_1 R_{ij}^E} \gamma_{ij}^E(t) + \frac{R^D}{C_1(R^A + R^D)} J_i(t) \right] + \\ &(\chi_{i12}(t, t') - \chi_{i22}(t, t')) \left[\sum_k \frac{E_{ik}^I}{C_2 R_{ik}^I} \gamma_{ik}^I(t) + \frac{R^D}{C_2(R^A + R^D)} J_i(t) + \frac{1}{C_2(R^B + R^C)} U_i(t) \right] dt' + \\ &\frac{R^D}{R^A + R^D} J_i(t), \end{aligned} \quad (10.27)$$

which can be reshaped by virtue of the convolutions (10.4) to

$$\begin{aligned}
I_{i1} = & \frac{1}{R^A + R^D} \left[\sum_j \frac{E_{ij}^E}{C_1 R_{ij}^E} (\chi_{i11} - \chi_{i21}) * \eta^E * a_j + \frac{R^D}{C_1(R^A + R^D)} (\chi_{i11} - \chi_{i21}) * J_i \right] + \\
& \frac{1}{R^A + R^D} \left[\sum_k \frac{E_{ik}^I}{C_2 R_{ik}^I} (\chi_{i12} - \chi_{i22}) * \eta^I * a_k + \frac{R^D}{C_2(R^A + R^D)} (\chi_{i12} - \chi_{i22}) * J_i + \right. \\
& \left. \frac{1}{C_2(R^B + R^C)} (\chi_{i12} - \chi_{i22}) * U_i \right] + \frac{R^D}{R^A + R^D} J_i . \quad (10.28)
\end{aligned}$$

Introducing new impulse response functions that simultaneously account for synaptic transmission (η) and dendritic propagation (χ) by

$$\psi_{i\alpha 1} = \chi_{i\alpha 1} * \eta^E \quad (10.29)$$

$$\psi_{i\alpha 2} = \chi_{i\alpha 2} * \eta^I \quad (10.30)$$

yields

$$\begin{aligned}
I_{i1} = & \frac{1}{R^A + R^D} \left[\sum_j \frac{E_{ij}^E}{C_1 R_{ij}^E} (\psi_{i11} - \psi_{i21}) * a_j + \frac{R^D}{C_1(R^A + R^D)} (\chi_{i11} - \chi_{i21}) * J_i \right] + \\
& \frac{1}{R^A + R^D} \left[\sum_k \frac{E_{ik}^I}{C_2 R_{ik}^I} (\psi_{i12} - \psi_{i22}) * a_k + \frac{R^D}{C_2(R^A + R^D)} (\chi_{i12} - \chi_{i22}) * J_i + \right. \\
& \left. \frac{1}{C_2(R^B + R^C)} (\chi_{i12} - \chi_{i22}) * U_i \right] + \frac{R^D}{R^A + R^D} J_i . \quad (10.31)
\end{aligned}$$

Eventually we obtain the DFP of neuron i as the potential dropping along the resistor R^D caused by the current through it (Eq. (10.14)), i.e.

$$\begin{aligned}
\Phi_i = & \frac{R^D}{R^A + R^D} \left\{ \sum_j \frac{E_{ij}^E}{C_1 R_{ij}^E} (\psi_{i11} - \psi_{i21}) * a_j + \sum_k \frac{E_{ik}^I}{C_2 R_{ik}^I} (\psi_{i12} - \psi_{i22}) * a_k + \right. \\
& \left[\frac{R^D}{C_1(R^A + R^D)} (\chi_{i11} - \chi_{i21}) + \frac{R^D}{C_2(R^A + R^D)} (\chi_{i12} - \chi_{i22}) - R^A \delta \right] * J_i + \\
& \left. \frac{1}{C_2(R^B + R^C)} (\chi_{i12} - \chi_{i22}) * U_i \right\} \quad (10.32)
\end{aligned}$$

10.2.3 Neurodynamics

Equation (10.32) reveals that the DFP is driven by the neuron's state variable U_i , by the entirety of postsynaptic potentials caused by spike trains a_i and by the diffusion currents J_i . The state variables and the spike trains are given by the network's evolution equation that is straightforwardly derived along the lines of Bressloff [16] again. To this end, we insert $V_{i2}(t)$ as the solution of (10.26) into the remaining Eq. (10.18) to get

$$C_3(R^B + R^C) \frac{dU_i}{dt} + \left(1 + \frac{R^B + R^C}{R_3}\right) U_i = \chi_{i21} * g_{i1} + \chi_{i22} * g_{i2}. \quad (10.33)$$

Next, we insert the inhomogeneity (10.21) again and obtain

$$\begin{aligned} C_3(R^B + R^C) \frac{dU_i}{dt} + \left(1 + \frac{R^B + R^C}{R_3}\right) U_i &= \sum_j \frac{E_{ij}^E}{C_1 R_{ij}^E} \chi_{i21} * \gamma_{ij}^E + \\ \frac{R^D}{C_1(R^A + R^D)} \chi_{i21} * J_i + \sum_k \frac{E_{ik}^I}{C_2 R_{ik}^I} \chi_{i22} * \gamma_{ik}^I + \frac{R^D}{C_2(R^A + R^D)} \chi_{i22} * J_i + \\ \frac{1}{C_2(R^B + R^C)} \chi_{i22} * U_i. \end{aligned} \quad (10.34)$$

Utilizing the convolutions (10.4) once more, yields

$$\begin{aligned} C_3(R^B + R^C) \frac{dU_i}{dt} + \left(1 + \frac{R^B + R^C}{R_3}\right) U_i &= \sum_j \frac{E_{ij}^E}{C_1 R_{ij}^E} \chi_{i21} * \eta^E * a_j + \\ \sum_k \frac{E_{ik}^I}{C_2 R_{ik}^I} \chi_{i22} * \eta^I * a_k + \frac{R^D}{C_1(R^A + R^D)} \chi_{i21} * J_i + \frac{R^D}{C_2(R^A + R^D)} \chi_{i22} * J_i + \\ \frac{1}{C_2(R^B + R^C)} \chi_{i22} * U_i. \end{aligned} \quad (10.35)$$

which becomes

$$\begin{aligned} C_3(R^B + R^C) \frac{dU_i}{dt} + \left(1 + \frac{R^B + R^C}{R_3}\right) U_i - \frac{1}{C_2(R^B + R^C)} \chi_{i22} * U_i = \\ \sum_j \frac{E_{ij}^E}{C_1 R_{ij}^E} \psi_{i21} * a_j + \sum_k \frac{E_{ik}^I}{C_2 R_{ik}^I} \psi_{i22} * a_k + \frac{R^D}{R^A + R^D} \left[\frac{1}{C_1} \chi_{i21} + \frac{1}{C_2} \chi_{i22} \right] * J_i \end{aligned} \quad (10.36)$$

after inserting the Green's functions (10.29) and (10.30) again. Equation (10.36) together with (10.3), (10.6) and (10.7) determine the dynamics of a network with three-compartment pyramidal neurons.

10.3 Leaky Integrate-and-Fire Model

The most serious difficulty for dealing with the neurodynamical evolution equations (10.36, 10.3, 10.6, 10.7) and the DFP observation equation (10.32) is the inhomogeneity of the matrix Green's function $\mathbf{X}_i(t, t')$ involved through the time-ordering operator and the time-dependence of $\mathbf{H}_i(t)$.

10.3.1 Simplification

In a first approximation \mathbf{H}_i becomes time-independent by neglecting the shunting terms [20, 21]. Then, the matrix Green's function $\mathbf{X}_i(t, t')$ becomes

$$\mathbf{X}(t, t') = \mathbf{X}(t - t') = e^{(t-t')\mathbf{H}} = \mathbf{Q}^{(t-t')} \quad (10.37)$$

with

$$\mathbf{Q} = e^{\mathbf{H}} \quad (10.38)$$

and

$$\mathbf{H} = \begin{pmatrix} \frac{1}{C_1} \left(-\frac{1}{R_1} + \frac{1}{R^A + R^D} \right) & -\frac{1}{C_1(R^A + R^D)} \\ \frac{1}{C_2(R^A + R^D)} & \frac{1}{C_2} \left(-\frac{1}{R_2} - \frac{1}{R^A + R^D} - \frac{1}{R^B + R^C} \right) \end{pmatrix}, \quad (10.39)$$

i.e. the transition matrix \mathbf{H} , and consequently also the Green's function, do not depend on the actual neuron index i any more. In this case, analytical methods can be employed [16].

However, for the present purpose, we employ another crucial simplification by choosing the electrotonic parameters in such a way that $\chi_{22}(t) \approx \delta(t)$. By virtue of this choice the dendritic filtering of compartment 2 is completely neglected. Then, the neural evolution equation (10.36) turns into

$$\begin{aligned} C_3(R^B + R^C) \frac{dU_i}{dt} + \left(1 + \frac{R^B + R^C}{R_3} - \frac{\hat{\tau}}{C_2(R^B + R^C)} \right) U_i = \\ \sum_j \frac{E_{ij}^E}{C_1 R_{ij}^E} \psi_{21} * a_j + \sum_k \frac{E_{ik}^I}{C_2 R_{ik}^I} \psi_{22} * a_k + \frac{R^D}{R^A + R^D} \left[\frac{1}{C_1} \chi_{21} + \frac{1}{C_2} \delta \right] * J_i \end{aligned}$$

$$C_3(R^B + R^C) \frac{dU_i}{dt} + \frac{C_2 R_3 (R^B + R^C) + C_2 (R^B + R^C)^2 - \hat{\tau} R_3}{C_2 R_3 (R^B + R^C)} U_i = \\ \sum_j \frac{E_{ij}^E}{C_1 R_{ij}^E} \psi_{21} * a_j + \sum_k \frac{E_{ik}^I}{C_2 R_{ik}^I} \psi_{22} * a_k + \frac{R^D}{R^A + R^D} \left[\frac{1}{C_1} \chi_{21} + \frac{1}{C_2} \delta \right] * J_i \quad (10.40)$$

where all kernels lost their neuron index i . Additionally, some time constant $\hat{\tau}$ results from the temporal convolution. Multiplying next with

$$r = \frac{C_2 R_3 (R^B + R^C)}{C_2 R_3 (R^B + R^C) + C_2 (R^B + R^C)^2 - \hat{\tau} R_3} \quad (10.41)$$

yields a leaky integrate-and-fire (LIF) model

$$\tau \frac{dU_i}{dt} + U_i = \sum_j w_{ij}^E \psi_{21} * a_j + \sum_k w_{ik}^I \psi_{22} * a_k + \kappa \left[\frac{1}{C_1} \chi_{21} + \frac{1}{C_2} \delta \right] * J_i \quad (10.42)$$

where we have introduced the following parameters:

- *Time constant*

$$\tau = r C_3 (R^B + R^C) \quad (10.43)$$

- *Excitatory synaptic weights*

$$w_{ij}^E = r \frac{E_{ij}^E}{C_1 R_{ij}^E} > 0 \quad (10.44)$$

- *Inhibitory synaptic weights*

$$w_{ik}^I = r \frac{E_{ik}^I}{C_2 R_{ik}^I} < 0 \quad (10.45)$$

- *Diffusion coefficient*

$$\kappa = r \frac{R^D}{R^A + R^D} \cdot \quad (10.46)$$

Moreover, we make the same approximation for the DFP (10.32) and obtain

$$\Phi_i = \frac{R^D}{R^A + R^D} \left\{ \sum_j \frac{E_{ij}^E}{C_1 R_{ij}^E} (\psi_{11} - \psi_{21}) * a_j + \sum_k \frac{E_{ik}^I}{C_2 R_{ik}^I} (\psi_{12} - \delta) * a_k + \left[\frac{R^D}{C_1(R^A + R^D)} (\chi_{11} - \chi_{21}) + \frac{R^D}{C_2(R^A + R^D)} (\chi_{12} - \delta) - R^A \delta \right] * J_i + \frac{1}{C_2(R^B + R^C)} (\chi_{12} - \delta) * U_i \right\} \quad (10.47)$$

as the observation equation of the LIF model.

Equations (10.42, 10.47) still exhibit some redundancy, seeing that the kernel ψ_{21} always relates to excitatory synapses while the kernel ψ_{22} refers to inhibitory synapses. We could thus absorb these kernel indices into the presynaptic neuron indices by introducing new kernels

$$\psi_j = \begin{cases} \psi_{21} & : j \text{ excitatory} \\ \psi_{22} & : j \text{ inhibitory} \end{cases} \quad (10.48)$$

$$\zeta_j = \begin{cases} \psi_{11} - \psi_{21} & : j \text{ excitatory} \\ \psi_{12} - \delta & : j \text{ inhibitory} \end{cases} \quad (10.49)$$

These kernels entail

$$\tau \frac{dU_i}{dt} + U_i = \sum_j w_{ij} \psi_j * a_j + \kappa \left[\frac{1}{C_1} \chi_{21} + \frac{1}{C_2} \delta \right] * J_i \quad (10.50)$$

$$\Phi_i = \frac{R^D}{r(R^A + R^D)} \left\{ \sum_j w_{ij} \zeta_j * a_j + \left[\frac{R^D}{C_1(R^A + R^D)} (\chi_{11} - \chi_{21}) + \frac{R^D}{C_2(R^A + R^D)} (\chi_{12} - \delta) - R^A \delta \right] * J_i + \frac{1}{C_2(R^B + R^C)} (\chi_{12} - \delta) * U_i \right\} \quad (10.51)$$

after also dropping the redundant excitatory/inhibitory superscripts. Thus, the indices i, j now extend over the entire network of K units.

Because the relevance of diffusion currents is controversially discussed in the literature [3, 6, 32–34], we could provisionally neglect these for further simplification: $J_i = 0$ which leads to

$$\tau \frac{dU_i}{dt} + U_i = \sum_j w_{ij} \psi_j * a_j \quad (10.52)$$

$$\Phi_i = \frac{R^D}{r(R^A + R^D)} \left\{ \sum_j w_{ij} \zeta_j * a_j + \frac{1}{C_2(R^B + R^C)} (\chi_{12} - \delta) * U_i \right\}. \quad (10.53)$$

10.3.2 Simulation

We have extensively discussed the system (10.52, 10.53) under the further assumption $\chi_{12} = \delta$ in a recent paper [9], and herein we present further numerical simulations under different external stimulation input. In particular, we simulate a cortical tissue as a LIF network (10.52), comprising of 1,000 interneurons and 4,000 pyramidal neurons interconnected randomly via an Erdős-Rényi graph with connection probability of 0.2. We refer the reader to [9] on how the somatic, dendritic and extracellular electrotonic parameters are estimated and how these are related to the phenomenological parameters of Mazzoni et al. [50]. All other parameters such as steady state voltages, refractory period, synaptic latencies, thresholds and others can also be found therein. Thalamic inputs are the only source of noise, which attempt to account for both cortical heterogeneity and spontaneous activity. This is achieved by modeling a two level noise, where the first level is an Ornstein-Uhlenbeck process superimposed with a constant or periodic signal and the second level is a time varying inhomogeneous Poisson process. Thus, we have the following time varying rate, $\lambda(t)$, that feeds into an inhomogeneous Poisson process:

$$\tau_n \frac{dn(t)}{dt} = -n(t) + \sigma_n \sqrt{\frac{2}{\tau_n}} \eta(t) \quad (10.54)$$

$$\lambda(t) = [c_0 + n(t)]_+ \quad (10.55)$$

where $\eta(t)$ represents Gaussian white noise, c_0 represents a constant signal (but equally could be periodic or other), and the operation $[\cdot]_+$ is the threshold-linear function, $[x]_+ = x$ if $x > 0$, $[x]_+ = 0$ otherwise, which circumvents negative rates. The constant signal c_0 can range between 1.2 and 2.6 spikes/ms. Note also that periodic or more complex ‘naturalistic’ signals can be applied, but we have herein kept it simple just for illustrative purposes. The parameters of the Ornstein-Uhlenbeck process are $\tau_n = 16$ ms and the standard deviation $\sigma_n = 0.4$ spikes/ms; also refer to [50] for these parameters.

The network simulations were run under the *Brian Simulator*, which is a Python based environment [35]. We focus on resistive extracellular case and compare between our DFP Φ_i measure (10.53) and the Mazzoni LFP measure (MPLB) defined herein as the sum of the moduli of excitatory and inhibitory synaptic currents:

$$V_i^{\text{MPLB}} = \sum_j |w_{ij}^E \psi_{21} * a_j| + \sum_k |w_{ik}^I \psi_{22} * a_k| \quad (10.56)$$

In addition, Mazzoni et al. [50] take the sum of V_i^{MPLB} across all pyramidal neurons. To provide a comparison we will also consider the sum of our proposed DFP measure (10.53), but also contrast it with its average. Thus we compare the following models of local field potentials:

$$L_1 = \sum_i V_i^{\text{MPLB}} \quad (10.57)$$

$$L_2 = \sum_i \Phi_i \quad (10.58)$$

$$L_3 = \frac{1}{P} \sum_i \Phi_i, \quad (10.59)$$

where P is the number of pyramidal neurons. Subsequently, we run the network for 2 s with three different noise levels, specifically, receiving constant signals with rates 1.2, 1.6 and 2.4 spikes/ms as depicted in Fig. 10.3. We report two main striking differences between LFP measures Eqs. (10.57), (10.58) and (10.59), namely in frequency and in amplitude. The L_1 responds instantaneously to the spiking network activity by means of high frequency oscillations. Moreover, L_1 exhibits a large amplitude overestimating experimental LFP/EEG measurements that typically vary between 0.5 and 2 mV [45, 58]. In contrast, L_2 and L_3 respond more smoothly to population activity and it is noticeable that our LFP estimates represent the low-pass filtered thalamic input. Clearly, both L_2 and L_3 have same time profile, however, the L_3 measures comparably with experimental LFP, that is, in the order of millivolts (although its not contained within the experimental range 0.5–2 mV). Thus we do concede that further work is required to improve our estimates. A minor improvement could be attained by applying a weighted average to mimic the distance of an electrode to a particular neuron. However, more biophysical modeling work is in demand as other neural effects, such as neuron-glia interaction and other effects, could be required to bring down these estimates to experimental results.

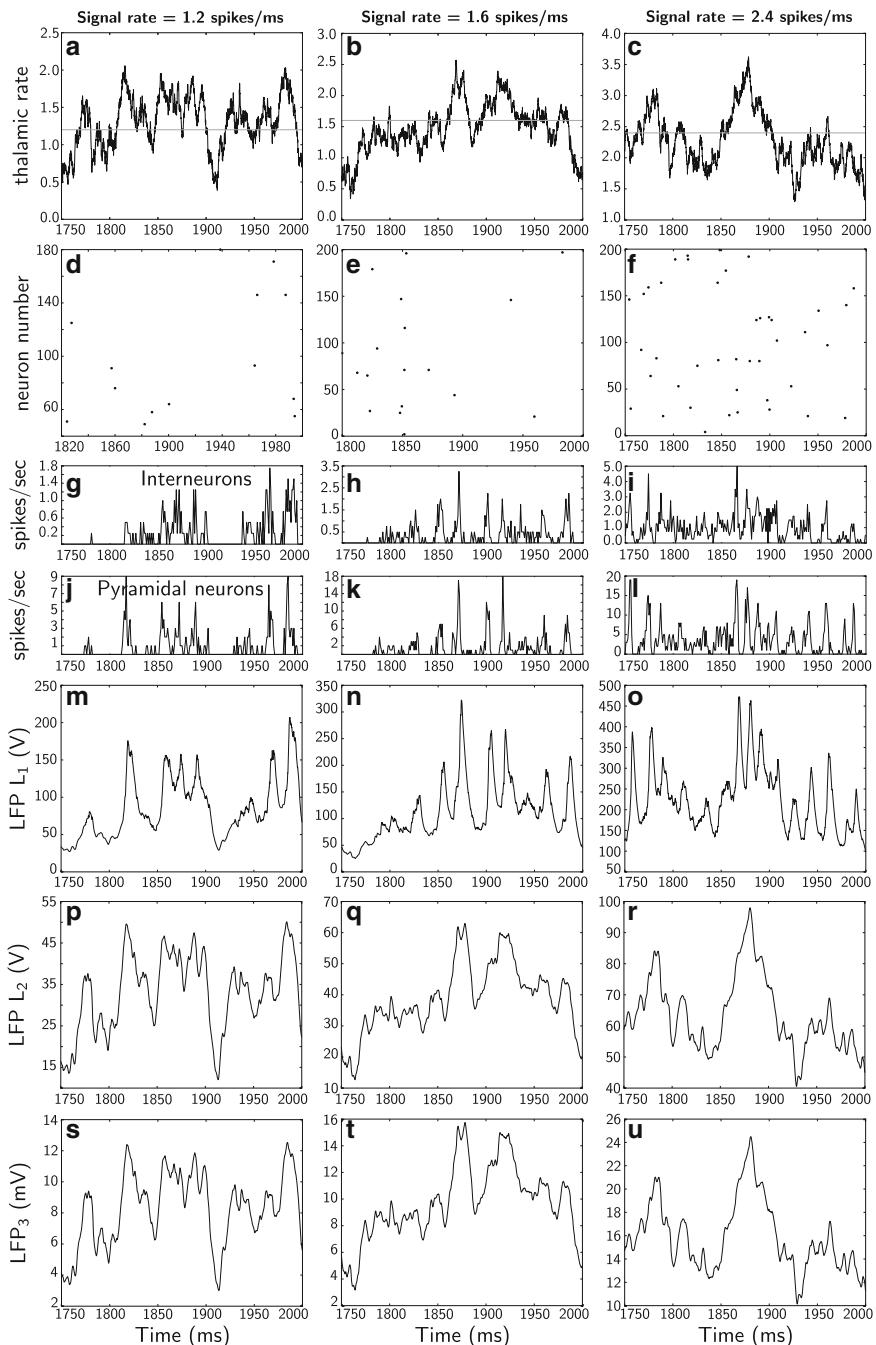


Fig. 10.3 Dynamics of the network and LFP comparisons: the three columns represent different runs of the network for three different rates, 1.2, 1.6 and 2.4 spikes/ms. In each column, all panels

10.4 Continuum Neural Field Model

So far we have considered the electrical properties of neural networks containing cortical pyramidal cells by means of equivalent circuits of a three-compartment model. In order to link these properties to the electromagnetic field in extracellular space, we need an embedding of the network topology into physical metric space \mathbb{R}^3 . This is most easily achieved in the continuum limit of neural field theory.

10.4.1 Rate Model

Starting with the approximation from Sect. 10.3, we first transform our LIF approach into a rate model. According to Eq. (10.3), a spike train is represented by a sum over delta functions. In order to obtain the number of spikes in a time interval $[0, t]$, one has to integrate Eq. (10.3), yielding

$$n(t) = \int_0^t a(t') dt' .$$

Then, the instantaneous spike rate per unit time is formally regained as the original signal Eq. (10.3), through

$$\frac{d}{dt} n(t) = a(t) . \quad (10.60)$$

A spike train $a(t)$ arriving at the presynaptic terminal of an axon causes changes in the conductivity of voltage-gated calcium channels. Therefore, calcium current flows into the synaptic button evoking the release of neurotransmitter into the synaptic cleft which is essentially a stochastic Bernoulli process [7, 44], where the probability $P(k)$ for releasing k transmitter vesicles obeys a binomial distribution

$$P(k) = \binom{Y}{k} p^k (1-p)^{Y-k} , \quad (10.61)$$

with Y the number of allocated vesicles in the button and p the elementary probability that an arriving action potential releases one vesicle.

Fig. 10.3 (continued) show the same 250 ms (extracted from 2 s simulations). The top panels (**a–c**) represent thalamic inputs with the different rates. The second top panels (**d–f**) correspond to a raster plot of the activity of 200 pyramidal neurons. Panels (**g–i**) depict average instantaneous firing rate (computed on a 1 ms bin) of interneurons and panels (**j–l**) correspond to average instantaneous firing rate of pyramidal neurons. Panels (**m–o**) show the LFP L_1 (Eq. (10.57)) which is the Mazzoni et al. measure [50]. Panels (**p–r**) and (**s–u**) depict our proposed LFP measures L_2 (Eq. (10.58)) and L_3 (Eq. (10.59)), respectively. Note and compare the different order of magnitudes between the three LFP measures

In the limit of large numbers, the binomial distribution can be replaced by a normal distribution

$$P(k) = \frac{1}{\sqrt{2\pi y(1-p)}} \exp\left[-\frac{(k-y)^2}{2y(1-p)}\right], \quad (10.62)$$

where $y = Yp$ is the average number of allocated transmitter vesicles. Due to this stochasticity of synaptic transmission, even the dynamics of a single neuron should be treated in terms of statistic ensembles in probability theory. Hence, we describe the state variables $U_i(t)$ by a normal distribution density $\rho(u, t)$ with mean $\bar{U}(t)$ and variance σ^2 , and determine the firing probability as

$$r(t) = \Pr(U(t) \geq \theta) = \int_{\theta}^{\infty} \rho(u, t) du = \frac{1}{2} \operatorname{erfc}\left(\frac{\theta - \bar{U}}{\sqrt{2}\sigma}\right), \quad (10.63)$$

with ‘erfc’ as the complementary error function accounting for the cumulative probability distribution. Thereby, the stochastic threshold dynamics is characterized by the typical sigmoidal activation functions. In computational neuroscience, the complementary error function is often approximated by the logistic function

$$f(u) = \frac{1}{1 + e^{-\gamma(u-\theta)}}$$

with parameters gain γ and threshold θ . Using f as nonlinear activation function, the firing probability $r(t) = f(U(t))$ for mean membrane potential U is closely related to the instantaneous spike rate $a(t)$ (Eq. (10.60)) via

$$a(t) = a_{\max} r(t) = a_{\max} f(U) \quad (10.64)$$

with maximal spike rate a_{\max} which can be absorbed by f :

$$f(u) = \frac{a_{\max}}{1 + e^{-\gamma(u-\theta)}}. \quad (10.65)$$

Inserting (10.64) and (10.65) into our LIF model (10.50) yields a leaky integrator rate (LIR) model [10, 19]

$$\tau \frac{dU_i}{dt} + U_i = \sum_j w_{ij} \psi_j * f(U_j) + \kappa \left[\frac{1}{C_1} \chi_{21} + \frac{1}{C_2} \delta \right] * J_i. \quad (10.66)$$

10.4.2 Neuroelectrodynamics

Next we perform the continuum limit $U_i(t) \rightarrow u_n(\mathbf{x}, t)$ à la physique under the assumption of identical neural properties within each population. The sum in Eq. (10.66) may converge under suitable smoothness assumptions upon the synaptic weight matrix w_{ij} and the convolution kernels.¹ Then a continuous two-dimensional vector $\mathbf{x} \in \Gamma_n$ replaces the neuron index i , while n becomes a population index. The population layers Γ_n become two-dimensional manifolds embedded in three-dimensional physical space such that $\mathbf{x} \in \Gamma_n$ is a two-dimensional projection of a vector $\mathbf{r} \in C \subset \mathbb{R}^3$ (C denoting cortex). Or, likewise, $\mathbf{r} = (\mathbf{x}, z)$, as indicated in Fig. 10.1.

As a result, Eq. (10.66) passes into the Amari equation [1]

$$\tau \frac{\partial}{\partial t} u_i(\mathbf{x}, t) + u_i(\mathbf{x}, t) = \sum_k \int_{\Gamma_k} w_{ik}(\mathbf{x}, \mathbf{x}') \psi(\mathbf{x}', t) * f(u_k(\mathbf{x}', t)) d\mathbf{x}' + h_i(\mathbf{x}, t) \quad (10.67)$$

with input current $h_i(\mathbf{x}, t)$ delivered to neuron layer i at site $\mathbf{x} \in \Gamma_i$. The synaptic weight kernels $w_{ik}(x, x')$ and the synaptic-dendritic impulse response $\psi(x', t)$ are obtained from the synaptic weight matrix, and from the Green's functions $\psi_j(t)$, respectively.

This neural field equation is complemented by the continuum limit of the extracellular dendritic dipole current density through cross section area A with normal vector \mathbf{n}_A , shown in Fig. 10.1, which is obtained from (10.31), i.e.

$$\begin{aligned} \mathbf{j}(\mathbf{r}, t) &= \lim_{i \rightarrow \mathbf{x}} \frac{\mathbf{n}_A}{A} I_{i1} = \\ &= \sum_k \int_{\Gamma_k} \tilde{w}_{1k}(\mathbf{r}, \mathbf{r}') \xi(\mathbf{r}', t) * f(u_k(\mathbf{r}', t)) + \xi_1(t) * \mathbf{j}^D(\mathbf{r}, t) + \xi_2(t) * u_1(\mathbf{r}, t), \end{aligned} \quad (10.68)$$

where we have introduced a modified synaptic weight kernel \tilde{w} and two new convolution kernels ξ_j that are related to the electrotonic parameters of the discrete model (10.31). The proper diffusion current $\mathbf{j}^D(\mathbf{r}, t)$ must be related to the gradient of the charge density $\rho(\mathbf{r}, t)$ according to Fick's law

$$\mathbf{j}^D(\mathbf{r}, t) = -\mathbf{D}^D(\mathbf{r}, t) \nabla \rho(\mathbf{r}, t), \quad (10.69)$$

¹A rigorous treatment of the continuum limit for neural networks requires techniques from stochastic analysis such as mean-field approaches or system-size expansions as carried out by Faugeras et al. [30] and Bressloff [17] (see also Chap. 9).

where the diffusion tensor $\mathbf{D}^D(\mathbf{r}, t)$ accounts for inhomogeneities and unisotropies of extracellular fluid, as being reflected by *diffusion tensor imaging* (DTI) [4,5]. For layer 1 of pyramidal neurons the input is then given by the diffusion current $\mathbf{j}^D(\mathbf{r}, t)$. Therefore, the input to the Amari equation (10.67) becomes

$$h_i(\mathbf{x}, t) = -\delta_{ik}\kappa A \mathbf{D}^D(\mathbf{r}, t) \cdot \nabla \rho(\mathbf{r}, t) \quad (10.70)$$

with Kronecker's δ_{ik} and appropriately redefined κ .

For further treatment of the electrodynamics of neural fields in linear but inhomogeneous and unisotropic media, we need Ohm's law

$$\mathbf{j}^Q(\mathbf{r}, t) = \boldsymbol{\sigma}(\mathbf{r}, t) \cdot \mathbf{E}(\mathbf{r}, t), \quad (10.71)$$

where $\boldsymbol{\sigma}(\mathbf{r}, t)$ is the conductivity tensor and \mathbf{E} the electric field strength. In case of negligible magnetic fields, we can introduce the dendritic field potential φ via

$$\mathbf{E} = -\nabla \varphi. \quad (10.72)$$

The diffusion current (10.69) and Ohmic current (10.71) together obey the Nernst-Planck equation [43,62]

$$\mathbf{j} = -\mathbf{D}^D \cdot \nabla \rho + \boldsymbol{\sigma} \cdot \mathbf{E}. \quad (10.73)$$

In the diffusive and conductive extracellular fluid, we additionally have

- *Einstein's relation* [28]

$$\mathbf{D}^D = k_B T q \boldsymbol{\mu} \quad (10.74)$$

- *Conductivity*

$$\boldsymbol{\sigma} = \boldsymbol{\mu} \rho, \quad (10.75)$$

where k_B denotes the thermodynamic Boltzmann constant, T temperature, q the ion charge, and $\boldsymbol{\mu}$ the ion's mobility tensor related to the fluid's viscosity [43,62]. Inserting (10.74, 10.75) into (10.73) yields

$$\mathbf{j} = -k_B T q \boldsymbol{\mu} \cdot \nabla \rho + \boldsymbol{\mu} \cdot \mathbf{E} \rho. \quad (10.76)$$

This form of the Nernst-Planck equation has to be augmented by a continuity equation

$$\nabla \cdot \mathbf{j} + \frac{\partial \rho}{\partial t} = 0 \quad (10.77)$$

reflecting the conservation of charge as a consequence of Maxwell's equations, and by the first Maxwell equation

$$\nabla \cdot \mathbf{D} = \rho \quad (10.78)$$

where

$$\mathbf{D} = \epsilon \cdot \mathbf{E} \quad (10.79)$$

introduces the electrical permittivity tensor ϵ .

Inserting (10.79) into (10.78) first gives

$$\epsilon \cdot (\nabla \cdot \mathbf{E}) = \rho - (\nabla \cdot \epsilon) \cdot \mathbf{E} . \quad (10.80)$$

Next, we take the divergence of the Nernst-Planck equation (10.76), which yields after consideration of the continuity equation (10.77)

$$\begin{aligned} \nabla \cdot \mathbf{j} &= -k_B T q \nabla \cdot (\boldsymbol{\mu} \cdot \nabla \rho) + \nabla \cdot (\boldsymbol{\mu} \cdot \mathbf{E} \rho) \\ -\epsilon \frac{\partial \rho}{\partial t} &= -k_B T q \epsilon \cdot [(\nabla \cdot \boldsymbol{\mu}) \cdot (\nabla \rho) + \boldsymbol{\mu} \Delta \rho] + \\ \epsilon \cdot (\nabla \cdot \boldsymbol{\mu}) \cdot \mathbf{E} \rho &+ \epsilon \cdot \boldsymbol{\mu} \cdot (\nabla \cdot \mathbf{E}) \rho + \epsilon \cdot \boldsymbol{\mu} \cdot \mathbf{E} \cdot \nabla \rho . \end{aligned}$$

Introducing the commutator $[\epsilon, \boldsymbol{\mu}] = \epsilon \cdot \boldsymbol{\mu} - \boldsymbol{\mu} \cdot \epsilon$, we can write

$$\begin{aligned} -\epsilon \frac{\partial \rho}{\partial t} &= -k_B T q \epsilon \cdot [(\nabla \cdot \boldsymbol{\mu}) \cdot (\nabla \rho) + \boldsymbol{\mu} \Delta \rho] + \epsilon \cdot (\nabla \cdot \boldsymbol{\mu}) \cdot \mathbf{E} \rho + \\ [\epsilon, \boldsymbol{\mu}] \cdot (\nabla \cdot \mathbf{E}) \rho &+ \boldsymbol{\mu} \cdot \rho^2 - \boldsymbol{\mu} \cdot (\nabla \cdot \epsilon) \cdot \mathbf{E} \rho + \epsilon \cdot \boldsymbol{\mu} \cdot \mathbf{E} \cdot \nabla \rho , \end{aligned}$$

where we have also utilized (10.80).

Using the Nernst-Planck equation (10.76) once more, we eliminate the electric field

$$\mathbf{E} = \boldsymbol{\mu}^{-1} \cdot \left(\frac{\mathbf{j} + k_B T q \boldsymbol{\mu} \cdot \nabla \rho}{\rho} \right) \quad (10.81)$$

thus

$$\begin{aligned} -\epsilon \frac{\partial \rho}{\partial t} &= -k_B T q \epsilon \cdot [(\nabla \cdot \boldsymbol{\mu}) \cdot (\nabla \rho) + \boldsymbol{\mu} \Delta \rho] + \epsilon \cdot \nabla (\ln \boldsymbol{\mu}) \cdot (\mathbf{j} + k_B T q \boldsymbol{\mu} \cdot \nabla \rho) + \\ [\epsilon, \boldsymbol{\mu}] \cdot \nabla \cdot \left\{ \boldsymbol{\mu}^{-1} \cdot \left(\frac{\mathbf{j} + k_B T q \boldsymbol{\mu} \cdot \nabla \rho}{\rho} \right) \right\} \rho &+ \boldsymbol{\mu} \cdot \rho^2 - \boldsymbol{\mu} \cdot (\nabla \cdot \epsilon) \cdot \boldsymbol{\mu}^{-1} \cdot (\mathbf{j} + k_B T q \boldsymbol{\mu} \cdot \nabla \rho) + \\ \epsilon \cdot \left(\frac{\mathbf{j} + k_B T q \boldsymbol{\mu} \cdot \nabla \rho}{\rho} \right) \cdot \nabla \rho . \quad (10.82) \end{aligned}$$

The solution of (10.82) provides the extracellular charge density $\rho(\mathbf{x}, t)$ dependent on the extracellular driving currents \mathbf{j} that are delivered by the neural field equations (10.67, 10.68, 10.70). Inserting both $\rho(\mathbf{x}, t)$ and \mathbf{j} into (10.81) yields the DFP

$$-\nabla\varphi = \boldsymbol{\mu}^{-1} \cdot \left(\frac{\mathbf{j} + k_B T q \boldsymbol{\mu} \cdot \nabla \rho}{\rho} \right) \quad (10.83)$$

via (10.72).

These equations of neuroelectrodynamics can be considerably simplified by assuming a homogeneous and isotropic medium. In that case (10.82) reduces to

$$-\epsilon \frac{\partial \rho}{\partial t} = -k_B T q \epsilon \cdot \boldsymbol{\mu} \Delta \rho + \boldsymbol{\mu} \cdot \rho^2 + \epsilon \cdot \left(\frac{\mathbf{j} + k_B T q \boldsymbol{\mu} \cdot \nabla \rho}{\rho} \right) \cdot \nabla \rho, \quad (10.84)$$

which is a kind of Fokker-Planck equation for the charge density. Taking only the first term of the r.h.s. into account, we obtain a diffusion equation whose stationary solution gives rise to the Warburg impedance of extracellular space [3, 4, 6, 59].

10.5 Discussion

In this contribution we outlined a biophysical theory for the coupling of microscopic neural activity to the electromagnetic field as described by the Maxwell equations, in order to adequately explain neural field potentials, such as DFP, LFP, and EEG. To that aim we have started from the widely accepted assumption, that cortical LFP/EEG mostly reflect extracellular dipole currents of pyramidal cells [53, 58]. This assumption has lead us to recent work suggesting that both Ohmic and diffusion currents contribute to LFP/EEG generation [3, 6, 32–34]. In addition, the assumption has placed a further challenge in that the geometry of the cortical pyramidal cells should be incorporated. Accounting for the geometry of the cell seemed to imply that one loses the computational efficiency of point models and we had to resort to compartmental models. However, herein we have proposed a framework showing how to circumvent these apparent difficulties to finally derive a biophysically plausible observation model for the Amari neural field equation [1], with additional dipole currents coupled to the Maxwell's equations.

We have first proposed a full-fledged three-compartment model of a single pyramidal cell decomposed into the apical dendritic tree for the main of excitatory synapses, the soma and the perisomatic dendritic tree that harbors mainly the inhibitory synapses, and the axon hillock exhibiting the neural spiking mechanism. In addition, the extracellular space was represented by incorporating both Ohmic and diffusive impedances, thus assuming that the total current through the extracellular fluid is governed by the Nernst-Planck equation. This has enabled us to account for the Warburg impedance. From this starting point and successive

simplifications we have derived the evolution law of the circuit, represented as an integro-differential equation. In the continuum limit this evolution law went into the Amari neural field equation, augmented by an observation equation for dendritic dipole currents, that are coupled to Maxwell's equations for the electromagnetic field in extracellular fluid.

Moreover we have demonstrated how to simplify and derive from our proposed three-compartment model a standard LIF network which then have enabled us to compare our LFP measure with other LFP measures found in the literature. Herein, we specifically have chosen to compare with the Mazzoni et al. work [50], that proposed the LFP to be the sum of the moduli of inhibitory and excitatory currents. Thus, we have proceeded by mapping our biophysical electrotonic parameters to the phenomenological parameters implemented in Mazzoni's LFP model [50]. However, now with the advantage that our LFP measure accounts for the extracellular currents and the geometry of the cell. Subsequently, we have compared different simulation runs between our LIF network model and that of Mazzoni et al. [50]. This comparison indicates that the Mazzoni et al. model systematically overestimates LFP amplitude by almost one order of magnitude and also systematically overestimates LFP frequencies. For more detailed discussion we refer the interested reader to [9].

At the present stage, we note that there is still a long way to fully explain the spatiotemporal characteristics of LFP and EEG. For example, the polarity reversals observed in experimental LFP/EEG as an electrode traverses different cortical layers are not accounted for in our current model. This is explained with the direction of the dipole currents, which is constrained, in the sense that current sources are situated at the perisomatic and current sinks are situated at apical dendritic tree. Taking this polarity as positive also entails positive DFP and LFP that could only change in strength. However, it is straightforward to adapt our model by fully incorporating cortical layers III and VI, for example. Yet another aspect that was not looked in the present work, was that of *ephaptic interactions* [31, 37, 40, 55] between neurons and the LFP which could act via a *mean-field coupling* as an *order parameter* thereby entraining the local populations to synchronized activity. A possible biophysical basis for this phenomena could be polarization of neurons induced by electric fields that are generated by ionic charges. As a consequence, this could alter the voltage dependent conductances, triggering changes of the neuronal dynamics, such as spiking and the activity of glia cells. We have not accounted for this effect in a biophysical sense yet, however, we could phenomenologically describe this mean-field coupling through a modulation of firing thresholds as outlined in [7].

Acknowledgements We thank Axel Hutt, Jamie Sleigh, Viktor Jirsa and Dimitris Pinotsis for helpful comments improving this chapter. This research was supported by a DFG Heisenberg fellowship awarded to PbG (GR 3711/1-2).

References

1. Amari, S.I.: Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.* **27**, 77–87 (1977)
2. Bazhenov, M., Lonjers, P., Skorheim, S., Bedard, C., Destexhe, A.: Non-homogeneous extracellular resistivity affects the current-source density profiles of up–down state oscillations. *Philos. Trans. R. Soc. Lond. A* **369**(1952), 3802–3819 (2011)
3. Bédard, C., Destexhe, A.: Macroscopic models of local field potentials and the apparent $1/f$ noise in brain activity. *Biophys. J.* **96**(7), 2589–2603 (2009)
4. Bédard, C., Destexhe, A.: Modeling local field potentials and their interaction with the extracellular medium. In: Brette, R., Destexhe, A. (eds.) *Handbook of Neural Activity Measurement*, pp. 136–191. Cambridge University Press, Cambridge (2012)
5. Bédard, C., Kröger, H., Destexhe, A.: Modeling extracellular field potentials and the frequency-filtering properties of extracellular space. *Biophys. J.* **86**(3), 1829–1842 (2004)
6. Bédard, C., Rodrigues, S., Contreras, N.R.D., Destexhe, A.: Evidence for frequency-dependent extracellular impedance from the transfer function between extracellular and intracellular potentials. *J. Comput. Neurosci.* **29**, 389–403 (2010)
7. beim Graben, P.: Foundations of neurophysics. In: beim Graben, P., Zhou, C., Thiel, M., Kurths, J. (eds.) *Lectures in Supercomputational Neuroscience: Dynamics in Complex Brain Networks*, pp. 3–48. Springer, Berlin (2008)
8. beim Graben, P., Potthast, R.: A dynamic field account to language-related brain potentials. In: Rabinovich, M., Friston, K., Varona, P. (eds.) *Principles of Brain Dynamics: Global State Interactions*, pp. 93–112. MIT, Cambridge (2012)
9. beim Graben, P., Rodrigues, S.: A biophysical observation model for field potentials of networks of leaky integrate-and-fire neurons. *Front. Comput. Neurosci.* **6**(100) (2013). doi:10.3389/fncom.2012.00100
10. beim Graben, P., Liebscher, T., Kurths, J.: Neural and cognitive modeling with networks of leaky integrator units. In: beim Graben, P., Zhou, C., Thiel, M., Kurths, J. (eds.) *Lectures in Supercomputational Neuroscience: Dynamics in Complex Brain Networks*, pp. 195–223. Springer, Berlin (2008)
11. Berger, H.: Über das Elektroenzephalogramm des Menschen. *Archiv für Psychiatrie* **87**, 527–570 (1929)
12. Berry, M.J., Warland, D.K., Meister, M.: The structure and precision of retinal spike trains. *Proc. Natl. Acad. Sci. USA* **94**(10), 5411–5416 (1997)
13. Bojak, I., Liley, D.T.J.: Modeling the effects of anesthesia on the electroencephalogram. *Phys. Rev. E* **71**(4), 041902 (2005)
14. Bojak, I., Oostendorp, T., Reid, A., Kötter, R.: Connecting mean field models of neural activity to EEG and fMRI data. *Brain Topogr.* **23**, 139–149 (2010)
15. Breakspear, M., Roberts, J.A., Terry, J.R., Rodrigues, S., Mahant, N., Robinson, P.A.: A unifying explanation of primary generalized seizures through nonlinear brain modeling and bifurcation analysis. *Cereb. Cortex* **16**, 1296–1313 (2006)
16. Bressloff, P.C.: Dynamics of compartmental model recurrent neural networks. *Phys. Rev. E* **50**(3), 2308–2319 (1994)
17. Bressloff, P.: Stochastic neural field theory and the system-size expansion. *SIAM J. Appl. Math.* **70**(5), 1488–1521 (2010)
18. Bressloff, P.C.: Spatiotemporal dynamics of continuum neural fields. *J. Phys. A* **45**(3), 033001 (2012)
19. Bressloff, P.C., Taylor, J.G.: Discrete time leaky integrator network with synaptic noise. *Neural Netw.* **4**(6), 789–801 (1991)
20. Bressloff, P., Taylor, J.: Compartmental-model response function for dendritic trees. *Biol. Cybern.* **70**, 199–207 (1993)
21. Bressloff, P.C., Taylor, J.G.: Spatiotemporal pattern processing in a compartmental-model neuron. *Phys. Rev. E* **47**, 2899–2912 (1993)

22. Coombes, S., Venkov, N.A., Shiao, L., Bojak, I., Liley, D.T.J., Laing, C.R.: Modeling electrocortical activity through improved local approximations of integral neural field equations. *Phys. Rev. E* **76**(5), 051901 (2007)
23. Creutzfeldt, O.D., Watanabe, S., Lux, H.D.: Relations between EEG phenomena and potentials of single cortical cells. I. Evoked responses after thalamic and epicortical stimulation. *Electroencephalogr. Clin. Neurophysiol.* **20**(1), 1–18 (1966)
24. David, O., Friston, K.J.: A neural mass model for MEG/EEG: coupling and neuronal dynamics. *NeuroImage* **20**, 1743–1755 (2003)
25. Desroches, M., Guckenheimer, J., Krauskopf, B., Kuehn, C., Osinga, H.M., Wechselberger, M.: Mixed-mode oscillations with multiple time scales. *SIAM Rev.* **54**(2), 211–288 (2012)
26. Destexhe, A.: Simplified models of neocortical pyramidal cells preserving somatodendritic voltage attenuation. *Neurocomputing* **38–40**, 167–173 (2001)
27. Destexhe, A., Bédard, C.: Local field potential. *Scholarpedia* **8**(8), 10713 (2013)
28. Einstein, A.: Eine neue Bestimmung der Moleküldimensionen. *Ann. Phys.* **19**, 289 (1906)
29. Erchova, I., McGonigle, D.J.: Rhythms of the brain: an examination of mixed mode oscillation approaches to the analysis of neurophysiological data. *Chaos* **8**(1), 015115 (2008)
30. Faugeras, O.D., Touboul, J.D., Cessac, B.: A constructive mean-field analysis of multi population neural networks with random synaptic weights and stochastic inputs. *Front. Comput. Neurosci.* **3**(1) (2008). doi:10.3389/neuro.10.001.2009
31. Fröhlich, F., McCormick, D.A.: Endogenous electric fields may guide neocortical network activity. *Neuron* **67**(1), 129–143 (2010)
32. Gabriel, C., Gabriel, S., Corthout, E.: The dielectric properties of biological tissues: I. Literature survey. *Phys. Med. Biol.* **41**(11), 2231 (1996)
33. Gabriel, S., Lau, R.W., Gabriel, C.: The dielectric properties of biological tissues: II. Measurements in the frequency range 10 Hz to 20 GHz. *Phys. Med. Biol.* **41**(11), 2251 (1996)
34. Gabriel, S., Lau, R.W., Gabriel, C.: The dielectric properties of biological tissues: III. Parametric models for the dielectric spectrum of tissues. *Phys. Med. Biol.* **41**(11), 2271 (1996)
35. Goodman, D., Brette, R.: The Brian simulator. *Front. Neurosci.* **3**(2), 192–197 (2009)
36. Hodgkin, A.L., Huxley, A.F.: A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.* **117**, 500–544 (1952)
37. Holt, G.R., Koch, C.: Electrical interactions via the extracellular potential near cell bodies. *J. Comput. Neurosci.* **6**, 169–184 (1999)
38. Hutt, A., Atay, F.M.: Analysis of nonlocal neural fields for both general and gamma-distributed connectivities. *Physica D* **203**, 30–54 (2005)
39. Jansen, B.H., Rit, V.G.: Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biol. Cybern.* **73**, 357–366 (1995)
40. Jefferys, J.G.R.: Nonsynaptic modulation of neuronal activity in the brain: electric currents and extracellular ions. *Physiol. Rev.* **75**(4), 689–723 (1995)
41. Jirsa, V.K., Haken, H.: A derivation of a macroscopic field theory of the brain from the quasi-microscopic neural dynamics. *Physica D* **99**, 503–526 (1997)
42. Jirsa, V.K., Jantzen, K.J., Fuchs, A., Kelso, J.A.S.: Spatiotemporal forward solution of the EEG and MEG using network modeling. *IEEE Trans. Med. Imaging* **21**(5), 493–504 (2002)
43. Johnston, D., Wu, S.M.S.: Foundations of Cellular Neurophysiology. MIT, Cambridge (1997)
44. Kandel, E.R., Schwartz, J.H., Jessel, T.M. (eds.): Principles of Neural Science. Appleton & Lange, East Norwalk (1991)
45. Lakatos, P., Shah, A., Knuth, K., Ulbert, I., Karmos, G., Schroeder, C.: An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *J. Neurophysiol.* **94**(3), 1904–1911 (2005)
46. Laszlo, S., Plaut, D.C.: A neurally plausible parallel distributed processing model of event-related potential word reading data. *Brain Lang.* **120**(3), 271–281 (2012)
47. Liley, D.T.J., Cadusch, P.J., Dafilis, M.P.: A spatially continuous mean field theory of electrocortical activity. *Netw. Comput. Neural Syst.* **13**, 67–113 (2002)
48. Lindén, H., Tetzlaff, T., Potjans, T.C., Pettersen, K.H., Grün, S., Diesmann, M., Einevoll, G.T.: Modeling the spatial reach of the LFP. *Neuron* **72**(5), 859–872 (2011)

49. Logothetis, N.K., Kayser, C., Oeltermann, A.: In vivo measurement of cortical impedance spectrum in monkeys: implications for signal propagation. *Neuron* **55**(5), 809–823 (2007)
50. Mazzoni, A., Panzeri, S., Logothetis, N.K., Brunel, N.: Encoding of naturalistic stimuli by local field potential spectra in networks of excitatory and inhibitory neurons. *PLoS Comput. Biol.* **4**(12), e1000239 (2008)
51. Mazzoni, A., Whittingstall, K., Brunel, N., Logothetis, N.K., Panzeri, S.: Understanding the relationships between spike rate and delta/gamma frequency bands of LFPs and EEGs using a local cortical network model. *NeuroImage* **52**(3), 956–972 (2010)
52. Mazzoni, A., Brunel, N., Cavallari, S., Logothetis, N.K., Panzeri, S.: Cortical dynamics during naturalistic sensory stimulations: experiments and models. *J. Physiol.* **105**(1–3), 2–15 (2011)
53. Nunez, P.L., Srinivasan, R.: *Electric Fields of the Brain: The Neurophysics of EEG*. Oxford University Press, New York (2006)
54. Pettersen, K., Hagen, E., Einevoll, G.: Estimation of population firing rates and current source densities from laminar electrode recordings. *J. Comput. Neurosci.* **24**, 291–313 (2008)
55. Richardson, K.A., Schiff, S.J., Gluckman, B.J.: Control of traveling waves in the mammalian cortex. *Phys. Rev. Lett.* **94**, 028103 (2005)
56. Rodrigues, S., Chizhov, A., Marten, F., Terry, J.: Mappings between a macroscopic neural-mass model and a reduced conductance-based model. *Biol. Cybern.* **102**(5), 361–371 (2010)
57. Sargsyan, A.R., Papatheodoropoulos, C., Kostopoulos, G.K.: Modeling of evoked field potentials in hippocampal CA1 area describes their dependence on NMDA and GABA receptors. *J. Neurosci. Methods* **104**, 143–153 (2001)
58. Schomer, D.L., Lopes da Silva, F.H. (eds.): *Niedermayer's Electroencephalography. Basic Principles, Clinical Applications, and Related Fields*. Lippincott Williams and Wilkins, Philadelphia (2011)
59. Skale, S., Dolecek, V., Slemnik, M.: Substitution of the constant phase element by Warburg impedance for protective coatings. *Corros. Sci.* **49**(3), 1045–1055 (2007)
60. Spruston, N.: Pyramidal neurons: dendritic structure and synaptic integration. *Nat. Rev. Neurosci.* **9**, 206–221 (2008)
61. Syková, E., Nicholson, C.: Diffusion in brain extracellular space. *Physiol. Rev.* **88**(4), 1277–1340 (2008)
62. Tuckwell, H.C.: *Introduction to Theoretical Neurobiology*, vol. 1. Cambridge University Press, Cambridge (1988)
63. Wang, X.J., Tegnér, J., Constantinidis, C., Goldman-Rakic, P.S.: Division of labor among distinct subtypes of inhibitory neurons in a cortical microcircuit of working memory. *Proc. Natl. Acad. Sci. USA* **101**(5), 1368–1373 (2004)
64. Wendling, F., Bellanger, J.J., Bartolomei, F., Chauvel, P.: Relevance of nonlinear lumped-parameter models in the analysis of depth-EEG epileptic signals. *Biol. Cybern.* **83**, 367–378 (2000)
65. Wilson, H.R., Cowan, J.D.: Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.* **12**(1), 1–24 (1972)
66. Wilson, H.R., Cowan, J.D.: A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* **13**, 55–80 (1973)

Part II

Applications of Neural Fields

Chapter 11

Universal Neural Field Computation

Peter beim Graben and Roland Potthast

Abstract Turing machines and Gödel numbers are important pillars of the theory of computation. Thus, any computational architecture needs to show how it could relate to Turing machines and how stable implementations of Turing computation are possible. In this chapter, we implement universal Turing computation in a neural field environment. To this end, we employ the canonical symbogram representation of a Turing machine obtained from a Gödel encoding of its symbolic repertoire and generalized shifts. The resulting nonlinear dynamical automaton (NDA) is a piecewise affine-linear map acting on the unit square that is partitioned into rectangular domains. Instead of looking at point dynamics in phase space, we then consider functional dynamics of probability distribution functions (p.d.f.s) over phase space. This is generally described by a Frobenius-Perron integral transformation that can be regarded as a neural field equation over the unit square as feature space of a Dynamic Field Theory (DFT). Solving the Frobenius-Perron equation yields that uniform p.d.f.s with rectangular support are mapped onto uniform p.d.f.s with rectangular support, again. We call the resulting representation *dynamic field automaton*.

P. beim Graben (✉)

Department of German Studies and Linguistics, Humboldt-Universität zu Berlin, Berlin, Germany

Bernstein Center for Computational Neuroscience Berlin, Humboldt-Universität zu Berlin,
Berlin, Germany

e-mail: peter.beim.graben@hu-berlin.de

R. Potthast

Department of Mathematics and Statistics, University of Reading, Reading, Berkshire, UK

Deutscher Wetterdienst, Offenbach, Germany

11.1 Introduction

Studying the computational capabilities of neurodynamical systems has commenced with the groundbreaking 1943 article of McCulloch and Pitts [27] on networks of idealized two-state neurons that essentially behave as logic gates. Because nowadays computers are nothing else than large-scale networks of logic gates, it is clear that computers can in principle be build up by neural networks of McCulloch-Pitts units. This has also been demonstrated by a number of theoretical studies reviewed in [46]. However, even the most powerful modern workstation is, from a mathematical point of view, only a finite state machine due to its rather huge, though limited memory, while a universal computer, formally codified as a Turing machine [20, 51], possesses an unbounded memory tape.

Using continuous-state units with a sigmoidal activation function, Siegelmann and Sontag [43] were able to prove that a universal Turing machine can be implemented by a recurrent neural network of about 900 units, most of them describing the machine's control states, while the tape is essentially represented by a plane spanned by the activations of just two units. The same construction, employing a Gödel code [17, 19] for the tape symbols, has been previously used by Moore [29, 30] for proving the equivalence of nonlinear dynamical automata and Turing machines. Along a different vain, deploying sequential cascaded networks, Pollack [36] and later Moore [31] and Tabor [48, 49] introduced and further generalized dynamical automata as nonautonomous dynamical systems. An even further generalization of dynamical automata, where the tape space becomes represented by a function space, led Moore and Crutchfield [32] to the concept of a quantum automaton (see [6] for a review and some unified treatment of these different approaches).

Quite remarkably, another paper from McCulloch and Pitts published in 1947 [34] already set up the groundwork for such functional representations in continuous neural systems. Here, those pioneers investigated distributed neural activation over cortical or subcortical maps representing visual or auditory feature spaces. These neural fields are copied onto many layers, each transforming the field according to a particular member of a symmetry group. For these, a number of field functionals is applied to yield a group invariant that serves for subsequent pattern detection. As early as in this publication, we already find all necessary ingredients for a *Dynamic Field Architecture*: a layered system of neural fields defined over appropriate feature spaces [14, 42] (see also Chaps. 12 and 13 in this volume).

We begin this chapter with a general exposition of dynamic field architectures in Sect. 11.2 where we illustrate how variables and structured data types on the one hand and algorithms and sequential processes on the other hand can be implemented in such environments. In Sect. 11.3 we review known facts about nonlinear dynamical automata and introduce dynamic field automata from a different perspective. The chapter is concluded with a short discussion about universal computation in neural fields.

11.2 Principles of Universal Computation

As already suggested by McCulloch and Pitts [34] in 1947, a neural, or likewise, *dynamic field architecture* is a layered system of Dynamic Neural Fields $u_i(x, t) \in \mathbb{R}$ where $1 \leq i \leq n$ ($i, n \in \mathbb{N}$) indicates the layer, $x \in D$ denotes spatial position in a suitable d -dimensional *feature space* $D \subset \mathbb{R}^d$ and $t \in \mathbb{R}_0^+$ time. Usually, the fields obey the Amari neural field equation [2]

$$\tau_i \frac{\partial u_i(x, t)}{\partial t} = -u_i(x, t) + h(x) + \sum_{j=1}^n \int_D w_{ij}(x, y) f(u_j(y, t)) dy + p_i(x, t), \quad (11.1)$$

where τ_i is a characteristic time scale of the i -th layer, $h(x)$ the unique resting activity, $w_{ij}(x, y)$ the synaptic weight kernel for a connection to site x in layer i from site y in layer j ,

$$f(u) = \frac{1}{1 + e^{-\beta(u-\theta)}} \quad (11.2)$$

is a sigmoidal activation function with gain β and threshold θ , and $p_i(x, t)$ external input delivered to site x in layer i at time t . Note, that a two-layered architecture could be conveniently described by a one-layered complex neural field $z(x, t) = u_1(x, t) + iu_2(x, t)$ as used in [6, 7, 11].

Commonly, Eq.(11.1) is often simplified in the literature by assuming one universal time constant τ , by setting $h = 0$ and by replacing p_i through appropriate initial, $u_i(x, 0)$, and boundary conditions, $u_i(\partial D, t)$. With these simplifications, we have to solve the Amari equation

$$\tau \frac{\partial u_i(x, t)}{\partial t} = -u_i(x, t) + \sum_{j=1}^n \int_D w_{ij}(x, y) f(u_j(y, t)) dy \quad (11.3)$$

for initial condition $u_i(x, 0)$, stating a computational task. Solving that task is achieved through a transient dynamics of Eq.(11.3) that eventually settles down either in an attractor state or in a distinguished terminal state $U_i(x, T)$, after elapsed time T . Mapping one state into another, which again leads to a transition to a third state and so on, we will see how the field dynamics can be interpreted as a kind of universal computation, carried out by a program encoded in the particular kernels $w_{ij}(x, y)$, which are in general heterogeneous, i.e. they are not pure convolution kernels: $w_{ij}(x, y) \neq w_{ij}(|x - y|)$ [5, 22].

11.2.1 Variables and Data Types

How can *variables* be realized in a neural field environment? At the hardware-level of conventional digital computers, variables are sequences of bytes stored

in random access memory (RAM). Since a byte is a word of eight bits and since nowadays RAM chips have about 2–8 GB, the computer’s memory appears as an approximately $8 \times 4 \cdot 10^9$ binary matrix, similar to an image of black-white pixels. It seems plausible to regard this RAM image as a discretized neural field, such that the value of $u(x, t)$ at $x \in D$ could be interpreted as a particular instantiation of a variable. However, this is not tenable for at least two reasons. First, such variables would be highly volatile as bits might change after every processing cycle. Second, the required function space would be a ‘mathematical monster’ containing highly discontinuous functions that are not admitted for the dynamical law (11.3). Therefore, variables have to be differently introduced into neural field computers by assuring temporal stability and spatial smoothness.

We first discuss the second point. Possible solutions of the neural field equation (11.3) must belong to appropriately chosen function spaces that allow the storage and retrieval of variables through binding and unbinding operations. A variable is stored in the neural field by binding its value to an address and its value is retrieved by the corresponding unbinding procedure. These operations have been described in the framework of *Vector Symbolic Architectures* [16, 44] and applied to Dynamic Neural Fields by beim Graben and Potthast [6] through a three-tier top-down approach, called *Dynamic Cognitive Modeling*, where variables are regarded as instantiations of data types of arbitrary complexity, ranging from primitive data types such as characters, integers, or floating numbers, over arrays (strings, vectors and matrices) of those primitives, up to structures and objects that allow the representation of lists, frames or trees. These data types are in a first step decomposed into *filler/role bindings* [44] which are sets of ordered pairs of sets of ordered pairs etc., of so-called fillers and roles. Simple fillers are primitives whereas roles address the appearance of a primitive in a complex data type. These addresses could be, e.g., array indices or tree positions. Such filler/role bindings can recursively serve as complex fillers bound to other roles. In a second step, fillers and roles are identified with particular basis functions over suitable feature spaces while the binding is realized through functional tensor products with subsequent compression (e.g. by means of convolution products) [35, 45].

Since the complete variable allocation of a conventional digital computer can be viewed as an instantiation of only one complex data type, namely an array containing every variable at a particular address, it is possible to map a total variable allocation onto a compressed tensor product in function space of a dynamic field architecture. Assuming that the field u encodes such an allocation, a new variable φ in its functional tensor product representation is stored by binding it first to a new address ψ , yielding $\varphi \otimes \psi$ and second by superimposing it with the current allocation, i.e. $u + \varphi \otimes \psi$. Accordingly, the value of φ is retrieved through an unbinding $\langle \psi^+, u \rangle$ where ψ^+ is the adjoint of the address ψ where φ is bound to. These operations require further underlying structure of the employed function spaces that are therefore chosen as Banach or Hilbert spaces where either adjoint or bi-orthogonal basis functions are available (see [4, 6, 7, 11, 38] for examples).

The first problem was the volatility of neural fields. This has been resolved using attractor neural networks [18, 21] where variables are stabilized as asymptotically

stable fixed points (as in Chap. 3). Since a fixed point is defined through $\dot{u}_i(x, t) = 0$, the field obeys the equation

$$u_i(x, t) = \sum_{j=1}^n \int_D w_{ij}(x, y) f(u_j(y, t)) dy . \quad (11.4)$$

This is often achieved by means of a particularly chosen kernel $w_{ii}(\|x - y\|)$ with local excitation and global inhibition, often called lateral inhibition kernels [14, 42].

11.2.2 Algorithms and Sequential Processes

Conventional computers run programs that dynamically change variables. Programs perform algorithms that are sequences of instructions, including operations upon variables, decisions, loops, etc. From a mathematical point of view, an algorithm is an element of an abstract algebra that has a representation as an operator on the space of variable allocations, which is well-known as *denotational semantics* in computer science [50]. The algebra product is the concatenation of instructions being preserved in the representation which is thereby an algebra homomorphism [4, 6]. Concatenating instructions or composing operators takes place step-by-step in discrete time. Neural field dynamics, as governed by Eq. (11.3), however requires continuous time. How can sequential algorithms be incorporated into the continuum of temporal evolution?

Looking first at conventional digital computers again suggests a possible solution: computers are clocked. Variables remain stable during a clock cycle and gating enables instructions to access variable space. A similar approach has recently been introduced to dynamic field architectures by Sandamirskaya and Schöner [40, 41]. Here a sequence of neural field activities is stored in a stack of layers, each stabilized by a lateral inhibition kernel. One state is destabilized by a gating signal provided by a condition-of-satisfaction mechanism playing the role of the ‘clock’ in this account. Afterwards, the decaying pattern in one layer, excites the next relevant field in a subsequent layer.

Another solution, already outlined in our Dynamic Cognitive Modeling framework [6], identifies the intermediate results of a computation with saddle fields that are connected along their respective stable and unstable manifolds to form stable heteroclinic sequences [1, 5, 39]. We have utilized this approach in [7] for a dynamic field model of syntactic language processing. Moreover, the chosen model of winnerless competition among neural populations [15] allowed us to explicitly construct the synaptic weight kernel from the filler/role binding of syntactic phrase structure trees [7].

11.3 Dynamic Field Automata

In this section we elaborate our recent proposal on *dynamic field automata* [8] by crucially restricting function spaces to spaces with Haar bases which are piecewise constant fields $u(x, t)$ for $x \in D$, i.e.

$$u(x, t) = \begin{cases} \alpha(t) & : x \in A(t) \\ 0 & : x \notin A(t) \end{cases} \quad (11.5)$$

with some time-dependent amplitude $\alpha(t)$ and a possibly time-dependent domain $A(t) \subset D$. Note, that we consider only one-layered neural fields in the sequel for the sake of simplicity.

For such a choice, we first observe that the application of the nonlinear activation function f yields another piecewise constant function over D :

$$f(u(x, t)) = \begin{cases} f(\alpha(t)) & : x \in A(t) \\ f(0) & : x \notin A(t), \end{cases} \quad (11.6)$$

which can be significantly simplified by the choice $f(0) = 0$, that holds, e.g., for the linear identity $f = \text{id}$, for the Heaviside step function $f = \Theta$ or for the hyperbolic tangens, $f = \tanh$.

With this simplification, the input integral of the neural field becomes

$$\int_D w(x, y) f(u(y, t)) \, dy = \int_{A(t)} w(x, y) f(\alpha(t)) \, dy = f(\alpha(t)) \int_{A(t)} w(x, y) \, dy. \quad (11.7)$$

When we additionally restrict ourselves to piecewise constant kernels as well, the last integral becomes

$$\int_{A(t)} w(x, y) \, dy = w|A(t)| \quad (11.8)$$

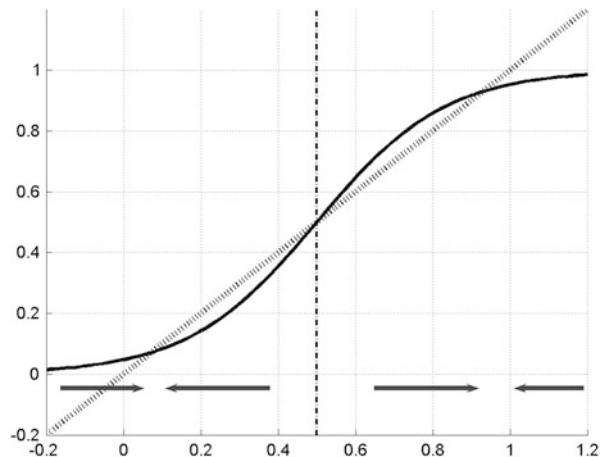
with w as constant kernel value and $|A(t)|$ the measure (i.e. the volume) of the domain $A(t)$. Inserting (11.7) and (11.8) into the fixed point equation (11.4) yields

$$u_0 = |A(t)| \cdot w \cdot f(u_0) \quad (11.9)$$

for the fixed point u_0 . Next, we carry out a linear stability analysis

$$\begin{aligned} \dot{u} &= -u + |A(t)|w f(u) \\ &= -(u_0 + (u - u_0)) + |A(t)|w \left(f(u_0) + f'(u_0) \cdot (u - u_0) \right) + O(|u - u_0|^2) \\ &= \left(-1 + |A(t)|w f'(u_0) \right) \cdot (u - u_0) + O(|u - u_0|^2). \end{aligned} \quad (11.10)$$

Fig. 11.1 Stability of piecewise constant neural field $u_0(x, t)$ over a domain $A \subset D$. Shown are the sigmoidal activation function $f(u)$ (solid) and u (dotted) for comparison. The axes here are given in terms of absolute numbers without unit as employed in Eqs. (11.2) or (11.3)



Thus, we conclude that if $|A(t)|w f'(u_0) < 1$, then $\dot{u} < 0$ for $u > u_0$ and conversely, $\dot{u} > 0$ for $u < u_0$ in a neighborhood of u_0 , such that u_0 is an asymptotically stable fixed point of the neural field equation.

Of course, linear stability analysis is a standard tool to investigate the behavior of dynamic fields around fixed points [5]. For our particular situation it is visualized in Fig. 11.1. When the solid curve displaying $|A(t)|w f(u)$ is above u (the dotted curve), then the dynamics (11.10) leads to an increase of u , indicated by the arrows pointing to the right. In the case where $|A(t)|w f(u) < u$, a decrease of u is obtained from (11.10). This is indicated by the arrows pointing to the left. When we have three points where the curves coincide, Fig. 11.1 shows that the setting leads to two stable fixed-points of the dynamics. When the activity field $u(x)$ reaches any value close to these fixed points, the dynamics leads them to the fixed-point values u_0 .

11.3.1 Turing Machines

For the construction of dynamic field automata through neural fields we consider discrete time that might be supplied by some clock mechanism. This requires the stabilization of the fields (11.5) within one clock cycle which can be achieved by self-excitation with a nonlinear activation function f as described in (11.10), leading to stable excitations as long as we do not include inhibitive elements, where a subsequent state would inhibit those states which were previously excited.

Next we briefly summarize some concepts from theoretical computer science [6, 20, 51]. A *Turing machine* is formally defined as a 7-tuple $M_{TM} = (Q, \mathbf{N}, \mathbf{T}, \delta, q_0, b, F)$, where Q is a finite set of machine control states, \mathbf{N} is another finite set of tape symbols, containing a distinguished ‘blank’ symbol b , $\mathbf{T} \subset \mathbf{N} \setminus \{b\}$ is input alphabet, and

$$\delta : Q \times \mathbf{N} \rightarrow Q \times \mathbf{N} \times \{L, R\} \quad (11.11)$$

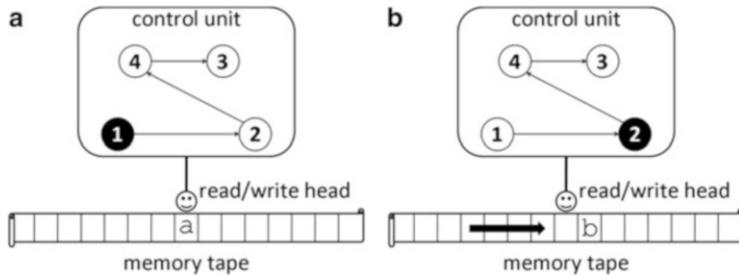


Fig. 11.2 Example state transition from (a) to (b) of a Turing machine with $\delta(1, a) = (2, b, R)$

is a partial state transition function, the so-called ‘machine table’, determining the action of the machine when $q \in Q$ is the current state at time t and $a \in N$ is the current symbol at the memory tape under the read/write head. The machine moves then into another state $q' \in Q$ at time $t + 1$ replacing the symbol a by another symbol $a' \in N$ and shifting the tape either one place to the left (‘L’) or to the right (‘R’). Figure 11.2 illustrates such a state transition. Finally, $q_0 \in Q$ is a distinguished initial state and $F \subset Q$ is a set of ‘halting states’ that are assumed when a computation terminates [20].

A Turing machine becomes a time- and state-discrete dynamical system by introducing *state descriptions*, which are triples

$$s = (\alpha, q, \beta) \quad (11.12)$$

where $\alpha, \beta \in N^*$ are strings of tape symbols to the left and to the right from the head, respectively. N^* contains all strings of tape symbols from N of arbitrary, yet finite, length, delimited by blank symbols b . Then, the transition function can be extended to state descriptions by

$$\delta^* : S \rightarrow S, \quad (11.13)$$

where $S = N^* \times Q \times N^*$ now plays the role of a phase space of a discrete dynamical system. The set of tape symbols and machine states then becomes a larger alphabet $A = N \cup Q$.

Moreover, state descriptions can be conveniently expressed by means of bi-infinite ‘dotted sequences’

$$s = \dots a_{i-3} a_{i-2} a_{i-1} . a_{i_0} a_{i_1} a_{i_2} \dots \quad (11.14)$$

with symbols $a_{i_k} \in A$. In Eq. (11.14) the dot denotes the observation time $t = 0$ such that the symbol left to the dot, a_{i-1} , displays the current state, dissecting the string s into two one-sided infinite strings $s = (\alpha', \beta)$ with $\alpha' = a_{i-1} a_{i-2} a_{i-3} \dots$ as the left-hand part in reversed order and $\beta = a_{i_0} a_{i_1} a_{i_2} \dots$

In symbolic dynamics, a cylinder set [28] is a subset of the space $\mathbf{A}^{\mathbb{Z}}$ of bi-infinite sequences from an alphabet \mathbf{A} that agree in a particular building block of length $n \in \mathbb{N}$ from a particular instance of time $t \in \mathbb{Z}$, i.e.

$$C(n, t) = [a_{i_1}, \dots, a_{i_n}] = \{s \in \mathbf{A}^{\mathbb{Z}} \mid s_{t+k-1} = a_{i_k}, \quad k = 1, \dots, n\} \quad (11.15)$$

is called n -cylinder at time $t \in \mathbb{Z}$. When now $t < 0, n > |t| + 1$ the cylinder contains the dotted word $w = s_{-1}.s_0$ and can therefore be decomposed into a pair of cylinders $(C'(|t|, t), C(|t| + n - 1, 0))$ where C' denotes reversed order of the defining strings again.

A *generalized shift* [29, 30] emulating a Turing machine is a pair $M_{GS} = (\mathbf{A}^{\mathbb{Z}}, \Psi)$ where $\mathbf{A}^{\mathbb{Z}}$ is the space of dotted sequences with $s \in \mathbf{A}^{\mathbb{Z}}$ and $\Psi : \mathbf{A}^{\mathbb{Z}} \rightarrow \mathbf{A}^{\mathbb{Z}}$ is given as

$$\Psi(s) = \sigma^{F(s)}(s \oplus G(s)) \quad (11.16)$$

with

$$F : \mathbf{A}^{\mathbb{Z}} \rightarrow \mathbb{Z} \quad (11.17)$$

$$G : \mathbf{A}^{\mathbb{Z}} \rightarrow \mathbf{A}^e, \quad (11.18)$$

where $\sigma : \mathbf{A}^{\mathbb{Z}} \rightarrow \mathbf{A}^{\mathbb{Z}}$ is the left-shift known from symbolic dynamics [26], $F(s) = l$ dictates a number of shifts to the right ($l < 0$), to the left ($l > 0$) or no shift at all ($l = 0$), $G(s)$ is a word w' of length $e \in \mathbb{N}$ in the domain of effect (DoE) replacing the content $w \in \mathbf{A}^d$, which is a word of length $d \in \mathbb{N}$, in the domain of dependence (DoD) of s , and $s \oplus G(s)$ denotes this replacement function.

A generalized shift becomes a Turing machine by interpreting a_{i-1} as the current control state q and a_{i_0} as the tape symbol currently underneath the head. Then the remainder of α is the tape left to the head and the remainder of β is the tape right to the head. The DoD is the word $w = a_{i-1}.a_{i_0}$ of length $d = 2$.

As an instructive example we consider a toy model of syntactic language processing. In order to process a sentence such as “the dog chased the cat”, linguists often derive a *context-free grammar* (CFG) from a phrase structure tree (see [10] for a more detailed example). In our case such a CFG could consist of *rewriting rules*

$$S \rightarrow NP VP \quad (11.19)$$

$$VP \rightarrow V NP \quad (11.20)$$

$$NP \rightarrow \text{the dog} \quad (11.21)$$

$$V \rightarrow \text{chased} \quad (11.22)$$

$$NP \rightarrow \text{the cat} \quad (11.23)$$

where the left-hand side always presents a nonterminal symbol to be expanded into a string of nonterminal and terminal symbols at the right-hand side. Omitting the

Table 11.1 Sequence of state transitions of the generalized shift processing the well-formed string “the dog chased the cat” ($\text{NP } \text{V } \text{NP}$). The operations are indicated as follows: “predict (X)” means prediction according to rule (X) of the context-free grammar; attach means cancellation of successfully predicted terminals both from stack and input; and “accept” means acceptance of the string as being well-formed

Time	State	Operation
0	$S . \text{NP } \text{V } \text{NP}$	predict (11.19)
1	$\text{VP } \text{NP} . \text{NP } \text{V } \text{NP}$	attach
2	$\text{VP} . \text{V } \text{NP}$	predict (11.20)
3	$\text{NP } \text{V} . \text{V } \text{NP}$	attach
4	$\text{NP} . \text{NP}$	attach
5	$\epsilon . \epsilon$	accept

lexical rules (11.21–11.23), we regard the symbols NP , V , denoting ‘noun phrase’ and ‘verb’, respectively, as terminals and the symbols S (‘sentence’) and VP (‘verbal phrase’) as nonterminals.

A generalized shift processing this grammar is then prescribed by the mappings

$$\begin{aligned} S.a &\mapsto \text{VP } \text{NP}.a \\ \text{VP}.a &\mapsto \text{NP } \text{V}.a \\ a.a &\mapsto \epsilon.\epsilon \end{aligned} \quad (11.24)$$

where the left-hand side of the tape is now called ‘stack’ and the right-hand side ‘input’. In (11.24) $a \in T$ stands for an arbitrary input symbol. The empty word is indicated by ϵ . Note the reversed order for the stack left of the dot. The first two operations in (11.24) are *predictions* according to a rule of the CFG while the last one is an attachment of input material with already predicted material, to be understood as a matching step.

With this machine table, a *parse* of the sentence “the dog chased the cat” ($\text{NP } \text{V } \text{NP}$) is then obtained in Table 11.1.

11.3.2 Nonlinear Dynamical Automata

Applying a Gödel encoding [6, 17, 19]

$$x = \psi(\alpha') := \sum_{k=1}^{\infty} \psi(a_{i-k}) b_L^{-k} \quad (11.25)$$

$$y = \psi(\beta) := \sum_{k=0}^{\infty} \psi(a_{i_k}) b_R^{-k-1}$$

to the pair $s = (\alpha', \beta)$ from the Turing machine state description (11.14) where $\psi(a_j) \in \mathbb{N}_0$ is an integer Gödel number for symbol $a_j \in A$ and $b_L, b_R \in \mathbb{N}$ are the

numbers of symbols that could appear either in α' or in β , respectively, yields the so-called symbol plane or *symbologram representation* $\mathbf{x} = (x, y)^T$ of s in the unit square X [13, 23].

The symbologram representation of a generalized shift is a *nonlinear dynamical automaton* (NDA) [6, 9, 10] which is a triple $M_{NDA} = (X, \mathcal{P}, \Phi)$ where (X, Φ) is a time-discrete dynamical system with phase space $X = [0, 1]^2 \subset \mathbb{R}^2$, the unit square, and flow $\Phi : X \rightarrow X$. $\mathcal{P} = \{D_v | v = (i, j), 1 \leq i \leq m, 1 \leq j \leq n, m, n \in \mathbb{N}\}$ is a rectangular partition of X into pairwise disjoint sets, $D_v \cap D_\mu = \emptyset$ for $v \neq \mu$, covering the whole phase space $X = \bigcup_v D_v$, such that $D_v = I_i \times J_j$ with real intervals $I_i, J_j \subset [0, 1]$ for each bi-index $v = (i, j)$. The cells D_v are the domains of the branches of Φ which is a piecewise affine-linear map

$$\Phi(\mathbf{x}) = \begin{pmatrix} a_x^v \\ a_y^v \end{pmatrix} + \begin{pmatrix} \lambda_x^v & 0 \\ 0 & \lambda_y^v \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}, \quad (11.26)$$

when $\mathbf{x} = (x, y)^T \in D_v$. The vectors $(a_x^v, a_y^v)^T \in \mathbb{R}^2$ characterize parallel translations, while the matrix coefficients $\lambda_x^v, \lambda_y^v \in \mathbb{R}_0^+$ mediate either stretchings ($\lambda > 1$), squeezings ($\lambda < 1$), or identities ($\lambda = 1$) along the x - and y -axes, respectively.

Hence, the NDA's dynamics, obtained by iterating an orbit $\{\mathbf{x}_t \in X | t \in \mathbb{N}_0\}$ from initial condition \mathbf{x}_0 through

$$\mathbf{x}_{t+1} = \Phi(\mathbf{x}_t) \quad (11.27)$$

describes a symbolic computation by means of a generalized shift [29, 30] when subjected to the coarse-graining \mathcal{P} .

The domains of dependence and effect (DoD and DoE) of an NDA, respectively, are obtained as images of cylinder sets under the Gödel encoding (11.25). Each cylinder possesses a lower and an upper bound, given by the Gödel numbers 0 and $b_L - 1$ or $b_R - 1$, respectively. Thus,

$$\begin{aligned} \inf(\psi(C'(|t|, t))) &= \psi(a_{i_{|t|}}, \dots, a_{i_1}) \\ \sup(\psi(C'(|t|, t))) &= \psi(a_{i_{|t|}}, \dots, a_{i_1}) + b_L^{-|t|} \\ \inf(\psi(C(|t| + n - 1, 0))) &= \psi(a_{i_{|t|+1}}, \dots, a_{i_n}) \\ \sup(\psi(C(|t| + n - 1, 0))) &= \psi(a_{i_{|t|+1}}, \dots, a_{i_n}) + b_R^{-|t|-n+1}, \end{aligned}$$

where the suprema have been evaluated by means of geometric series [9]. Thereby, each part cylinder C is mapped onto a real interval $[\inf(C), \sup(C)] \subset [0, 1]$ and the complete cylinder $C(n, t)$ onto the Cartesian product of intervals $R = I \times J \subset [0, 1]^2$, i.e. onto a rectangle in unit square. In particular, the empty cylinder, corresponding to the empty tape $\epsilon.\epsilon$ is represented by the complete phase space $X = [0, 1]^2$.

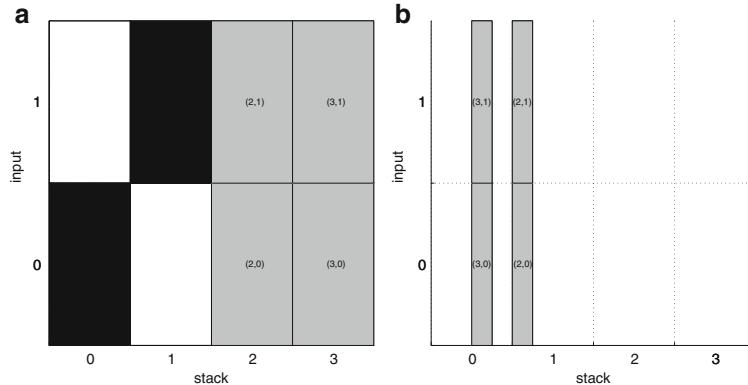


Fig. 11.3 Symbologram of the NDA processing the string “the dog chased the cat” (NP V NP). **(a)** Domains of dependence (DoD) of actions: identity (white), predict (gray), and attach (black). **(b)** Domains of effect (DoE): images of prediction (gray), black rectangles from **(a)** are mapped onto the whole unit square during attachment

Fixing the prefixes of both part cylinders and allowing for random symbolic continuation beyond the defining building blocks, results in a cloud of randomly scattered points across a rectangle R in the symbologram [10]. These rectangles are consistent with the symbol processing dynamics of the NDA, while individual points $\mathbf{x} \in [0, 1]^2$ no longer have an immediate symbolic interpretation. Therefore, we refer to arbitrary rectangles $R \in [0, 1]^2$ as to NDA macrostates, distinguishing them from NDA microstates \mathbf{x} of the underlying dynamical system.

Coming back to our language example, we create an NDA from an arbitrary Gödel encoding. Choosing

$$\psi(\text{NP}) = 0 \quad (11.28)$$

$$\psi(\text{V}) = 1 \quad (11.29)$$

$$\psi(\text{VP}) = 2 \quad (11.30)$$

$$\psi(\text{S}) = 3 \quad (11.31)$$

we have $b_L = 4$ stack symbols and $b_R = 2$ input symbols. Thus, the symbologram is partitioned into eight rectangles. Figure 11.3 displays the resulting (a) DoD and (b) DoE.

11.3.3 Neural Field Computation

Next we replace the NDA point dynamics in phase space by functional dynamics in Banach space. Instead of iterating clouds of randomly prepared initial conditions according to a deterministic dynamics, we consider the deterministic dynamics of

probability measures over phase space. This higher level of description that goes back to Koopman et al. [24, 25] has recently been revitalized for dynamical systems theory [12].

The starting point for this approach is the conservation of probability as expressed by the Frobenius-Perron equation [33]

$$\rho(\mathbf{x}, t) = \int_X \delta(\mathbf{x} - \Phi^{t-t'}(\mathbf{x}')) \rho(\mathbf{x}', t') d\mathbf{x}' , \quad (11.32)$$

where $\rho(\mathbf{x}, t)$ denotes a probability density function over the phase space X at time t of a dynamical system, $\Phi^t : X \rightarrow X$ refers to either a continuous-time ($t \in \mathbb{R}_0^+$) or discrete-time ($t \in \mathbb{N}_0$) flow and the integral over the delta function expresses the probability summation of alternative trajectories all leading into the same state \mathbf{x} at time t .

In the case of an NDA, the flow is discrete and piecewise affine-linear on the domains D_v as given by Eq. (11.26). As initial probability distribution densities $\rho(\mathbf{x}, 0)$ we consider uniform distributions with rectangular support $R_0 \subset X$, corresponding to an initial NDA macrostate,

$$u(\mathbf{x}, 0) = \frac{1}{|R_0|} \chi_{R_0}(\mathbf{x}) , \quad (11.33)$$

where

$$\chi_A(\mathbf{x}) = \begin{cases} 0 & : \mathbf{x} \notin A \\ 1 & : \mathbf{x} \in A \end{cases} \quad (11.34)$$

is the characteristic function for a set $A \subset X$. A crucial requirement for these distributions is that they must be consistent with the partition \mathcal{P} of the NDA, i.e. there must be a bi-index $v = (i, j)$ such that the support $R_0 \subset D_v$.

Inserting (11.33) into the Frobenius-Perron equation (11.32) yields for one iteration

$$u(\mathbf{x}, t+1) = \int_X \delta(\mathbf{x} - \Phi(\mathbf{x}')) u(\mathbf{x}', t) d\mathbf{x}' . \quad (11.35)$$

In order to evaluate (11.35), we first use the product decomposition of the involved functions:

$$u(\mathbf{x}, 0) = u_x(x, 0) u_y(y, 0) \quad (11.36)$$

with

$$u_x(x, 0) = \frac{1}{|I_0|} \chi_{I_0}(x) \quad (11.37)$$

$$u_y(y, 0) = \frac{1}{|J_0|} \chi_{J_0}(y) \quad (11.38)$$

and

$$\delta(\mathbf{x} - \Phi(\mathbf{x}')) = \delta(x - \Phi_x(\mathbf{x}'))\delta(y - \Phi_y(\mathbf{x}')) , \quad (11.39)$$

where the intervals I_0, J_0 are the projections of R_0 onto x - and y -axes, respectively. Correspondingly, Φ_x and Φ_y are the projections of Φ onto x - and y -axes, respectively. These are obtained from (11.26) as

$$\Phi_x(\mathbf{x}') = a_x^v + \lambda_x^v x' \quad (11.40)$$

$$\Phi_y(\mathbf{x}') = a_y^v + \lambda_y^v y' . \quad (11.41)$$

Using this factorization, the Frobenius-Perron equation (11.35) separates into

$$u_x(x, t+1) = \int_{[0,1]} \delta(x - a_x^v - \lambda_x^v x') u_x(x', t) dx' \quad (11.42)$$

$$u_y(y, t+1) = \int_{[0,1]} \delta(y - a_y^v - \lambda_y^v y') u_y(y', t) dy' \quad (11.43)$$

Next, we evaluate the delta functions according to the well-known lemma

$$\delta(f(x)) = \sum_{l:\text{simple zeros}} |f'(x_l)|^{-1} \delta(x - x_l) , \quad (11.44)$$

where $f'(x_l)$ indicates the first derivative of f in x_l . Equation (11.44) yields for the x -axis

$$x_v = \frac{x - a_x^v}{\lambda_x^v} , \quad (11.45)$$

i.e. one zero for each v -branch, and hence

$$|f'(x_v)| = \lambda_x^v . \quad (11.46)$$

Inserting (11.44), (11.45) and (11.46) into (11.42), gives

$$\begin{aligned} u_x(x, t+1) &= \sum_v \int_{[0,1]} \frac{1}{\lambda_x^v} \delta\left(x' - \frac{x - a_x^v}{\lambda_x^v}\right) u_x(x', t) dx' \\ &= \sum_v \frac{1}{\lambda_x^v} u_x\left(\frac{x - a_x^v}{\lambda_x^v}, t\right) \end{aligned}$$

Next, we take into account that the distributions must be consistent with the NDA's partition. Therefore, for given $\mathbf{x} \in D_v$ there is only one branch of Φ contributing a simple zero to the sum above. Hence,

$$u_x(x, t + 1) = \frac{1}{\lambda_x^v} u_x \left(\frac{x - a_x^v}{\lambda_x^v}, t \right). \quad (11.47)$$

Our main finding is now that the evolution of uniform p.d.f.s with rectangular support according to the NDA dynamics Eq. (11.35) is governed by

$$u(\mathbf{x}, t) = \frac{1}{|\Phi^t(R_0)|} \chi_{\Phi^t(R_0)}(\mathbf{x}), \quad (11.48)$$

i.e. uniform distributions with rectangular support are mapped onto uniform distributions with rectangular support [8].

For the proof we first insert the initial uniform density distribution (11.33) for $t = 0$ into Eq. (11.47), to obtain by virtue of (11.37)

$$u_x(x, 1) = \frac{1}{\lambda_x^v} u_x \left(\frac{x - a_x^v}{\lambda_x^v}, 0 \right) = \frac{1}{\lambda_x^v} \frac{1}{|I_0|} \chi_{I_0} \left(\frac{x - a_x^v}{\lambda_x^v} \right).$$

Deploying (11.34) yields

$$\chi_{I_0} \left(\frac{x - a_x^v}{\lambda_x^v} \right) = \begin{cases} 0 & : \frac{x - a_x^v}{\lambda_x^v} \notin I_0 \\ 1 & : \frac{x - a_x^v}{\lambda_x^v} \in I_0. \end{cases}$$

Let now $I_0 = [p_0, q_0] \subset [0, 1]$ we get

$$\begin{aligned} \frac{x - a_x^v}{\lambda_x^v} &\in I_0 \\ \iff p_0 &\leq \frac{x - a_x^v}{\lambda_x^v} \leq q_0 \\ \iff \lambda_x^v p_0 &\leq x - a_x^v \leq \lambda_x^v q_0 \\ \iff a_x^v + \lambda_x^v p_0 &\leq x \leq a_x^v + \lambda_x^v q_0 \\ \iff \Phi_x(p_0) &\leq x \leq \Phi_x(q_0) \\ \iff x &\in \Phi_x(I_0), \end{aligned}$$

where we made use of (11.40). Moreover, we have

$$\lambda_x^v |I_0| = \lambda_x^v (q_0 - p_0) = q_1 - p_1 = |I_1|$$

with $I_1 = [p_1, q_1] = \Phi_x(I_0)$. Therefore,

$$u_x(x, 1) = \frac{1}{|I_1|} \chi_{I_1}(x) .$$

The same argumentation applies to the y -axis, such that we eventually obtain

$$u(\mathbf{x}, 1) = \frac{1}{|R_1|} \chi_{R_1}(\mathbf{x}) , \quad (11.49)$$

with $R_1 = \Phi(R_0)$ the image of the initial rectangle $R_0 \subset X$. Thus, the image of a uniform density function with rectangular support is a uniform density function with rectangular support again.

Next, assume (11.48) is valid for some $t \in \mathbb{N}$. Then it is obvious that (11.48) also holds for $t + 1$ by inserting the x -projection of (11.48) into (11.47) using (11.37), again. Then, the same calculation as above applies when every occurrence of 0 is replaced by t and every occurrence of 1 is replaced by $t + 1$. By means of this inductive proof we have implemented an NDA by a dynamically evolving field. Therefore, we call this representation dynamic field automaton (DFA).

The Frobenius-Perron equation (11.35) can be regarded as a time-discretized Amari dynamic neural field equation (11.3). Discretizing time according to Euler's rule with increment $\Delta t = \tau$ where τ is the time constant of the Amari equation (11.3) yields

$$\begin{aligned} \tau \frac{u(\mathbf{x}, t + \tau) - u(\mathbf{x}, t)}{\tau} + u(\mathbf{x}, t) &= \int_D w(\mathbf{x}, \mathbf{y}) f(u(\mathbf{y}, t)) d\mathbf{y} \\ u(\mathbf{x}, t + \tau) &= \int_D w(\mathbf{x}, \mathbf{y}) f(u(\mathbf{y}, t)) d\mathbf{y} . \end{aligned}$$

For $\tau = 1$ and $f(u) = u$ the Amari equation becomes the Frobenius-Perron equation (11.35) when we set

$$w(\mathbf{x}, \mathbf{y}) = \delta(\mathbf{x} - \Phi(\mathbf{y})) \quad (11.50)$$

where Φ is the NDA mapping from Eq. (11.27). This is the general solution of the kernel construction problem [6, 38]. Note that Φ is not injective, i.e. for fixed \mathbf{x} the kernel is a sum of delta functions encoding the influence from different parts of the space $X = [0, 1]^2$.

Finally we carry out the whole construction for our language example. This yields the field dynamics depicted in Fig. 11.4.

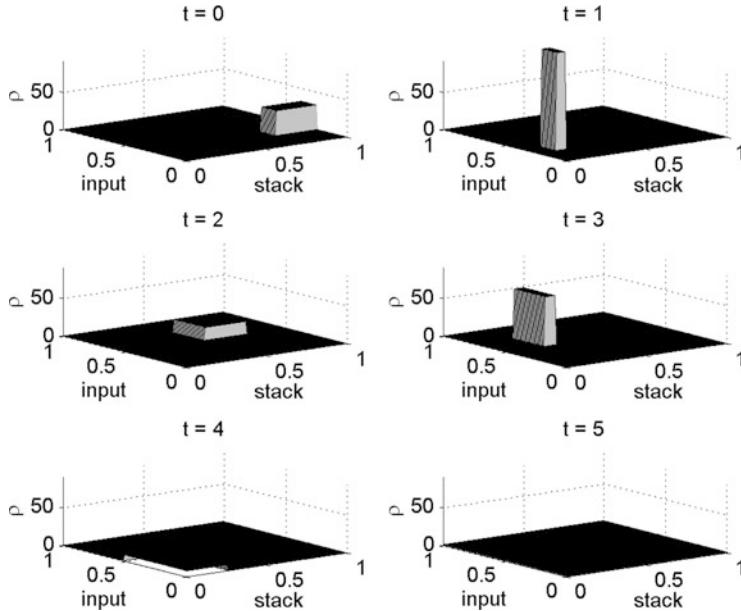


Fig. 11.4 Dynamic field automaton for processing the string “the dog chased the cat” (NP V NP) according to Table 11.1. The NDA states become rectangular supports of uniform distributions which are mapped onto uniform distributions with rectangular supports during discrete temporal evolution

11.4 Discussion

Turing machines and Gödel numbers are important pillars of the theory of computation [20, 47]. Thus, any computational architecture needs to show how it could relate to Turing machines and in what way stable implementations of Turing computation is possible. In this chapter, we addressed the question how universal Turing computation could be implemented in a neural field environment as described by the Amari field equation (11.1). To this end, we employed the canonical symbogram representation [13, 23] of the machine tape as the unit square, resulting from a Gödel encoding of sequences of states.

The action of the Turing machine on a state description is given by a state flow on the unit square which led to a Frobenius-Perron equation (11.32) for the evolution of uniform probability densities. We have implemented this equation in the neural field space by a piecewise affine-linear kernel geometry on the unit square which can be expressed naturally within a neural field framework. We also showed that stability of states and dynamics both in time as well as its encoding for finite programs is achieved by the approach.

However, our construction essentially relied upon discretized time that could be provided by some clock mechanism. The crucial problem of stabilizing states within

every clock cycle could be principally solved by established methods from dynamic field architectures. In such a time-continuous extension, an excited state, represented by a rectangle in one layer, will only excite a subsequent state, represented by another rectangle in another layer when a condition-of-satisfaction is met [40, 41]. Otherwise rectangular states would remain stabilized as described by Eq. (11.10). All these problems provide promising prospects for future research.

Acknowledgements We thank Slawomir Nasuto and Serafim Rodrigues for helpful comments improving this chapter. This research was supported by a DFG Heisenberg fellowship awarded to PbG (GR 3711/1-2).

References

1. Afraimovich, V.S., Zhigulin, V.P., Rabinovich, M.I.: On the origin of reproducible sequential activity in neural circuits. *Chaos* **14**(4), 1123–1129 (2004)
2. Amari, S.I.: Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.* **27**, 77–87 (1977)
3. Arbib, M.A. (ed.): *The Handbook of Brain Theory and Neural Networks*, 1st edn. MIT, Cambridge (1995)
4. beim Graben, P., Gerth, S.: Geometric representations for minimalist grammars. *J. Log. Lang. Inf.* **21**(4), 393–432 (2012)
5. beim Graben, P., Hutt, A.: Attractor and saddle node dynamics in heterogeneous neural fields. *EPJ Nonlinear Biomed. Phys.* **2**, 4 (2014). doi:10.1140/epjnbp17
6. beim Graben, P., Potthast, R.: Inverse problems in dynamic cognitive modeling. *Chaos* **19**(1), 015103 (2009)
7. beim Graben, P., Potthast, R.: A dynamic field account to language-related brain potentials. In: Rabinovich, M., Friston, K., Varona, P. (eds.) *Principles of Brain Dynamics: Global State Interactions*, chap. 5, pp. 93–112. MIT, Cambridge (2012)
8. beim Graben, P., Potthast, R.: Implementing Turing machines in dynamic field architectures. In: Bishop, M., Erden, Y.J. (eds.) *Proceedings of AISB12 World Congress 2012 – Alan Turing 2012*, vol. 5th AISB Symposium on Computing and Philosophy: Computing, Philosophy and the Question of Bio-Machine Hybrids, Birmingham, pp. 36–40 (2012). <http://arxiv.org/abs/1204.5462>
9. beim Graben, P., Jurish, B., Saddy, D., Frisch, S.: Language processing by dynamical systems. *Int. J. Bifurc. Chaos* **14**(2), 599–621 (2004)
10. beim Graben, P., Gerth, S., Vasishth, S.: Towards dynamical system models of language-related brain potentials. *Cogn. Neurodynamics* **2**(3), 229–255 (2008)
11. beim Graben, P., Pinotsis, D., Saddy, D., Potthast, R.: Language processing with dynamic fields. *Cogn. Neurodynamics* **2**(2), 79–88 (2008)
12. Budišić, M., Mohr, R., Mezić, I.: Applied Koopmanism. *Chaos* **22**(4), 047510 (2012)
13. Cvitanović, P., Gunaratne, G.H., Procaccia, I.: Topological and metric properties of Hénon-type strange attractors. *Phys. Rev. A* **38**(3), 1503–1520 (1988)
14. Erlhagen, W., Schöner, G.: Dynamic field theory of movement preparation. *Psychol. Rev.* **109**(3), 545–572 (2002)
15. Fukai, T., Tanaka, S.: A simple neural network exhibiting selective activation of neuronal ensembles: from winner-take-all to winners-share-all. *Neural Comput.* **9**(1), 77–97 (1997)
16. Gayler, R.W.: Vector symbolic architectures are a viable alternative for Jackendoff's challenges. *Behav. Brain Sci.* **29**, 78–79 (2006)

17. Gödel, K.: Über formal unentscheidbare Sätze der *principia mathematica* und verwandter Systeme I. *Monatshefte für Mathematik und Physik* **38**, 173–198 (1931)
18. Hertz, J.: Computing with attractors. In: Arbib, M.A. (ed.): *The Handbook of Brain Theory and Neural Networks*, 1st edn, pp. 230–234. MIT, Cambridge (1995)
19. Hofstadter, D.R.: Gödel, Escher, Bach: An Eternal Golden Braid. Basic Books, New York (1979)
20. Hopcroft, J.E., Ullman, J.D.: *Introduction to Automata Theory, Languages, and Computation*. Addison-Wesley, Menlo Park (1979)
21. Hopfield, J.J.: Neurons with graded response have collective computational properties like those of two-state neurons. *Proc. Natl. Acad. Sci. USA* **81**(10), 3088–3092 (1984)
22. Jirsa, V.K., Kelso, J.A.S.: Spatiotemporal pattern formation in neural systems with heterogeneous connection topologies. *Phys. Rev. E* **62**(6), 8462–8465 (2000)
23. Kennel, M.B., Buhl, M.: Estimating good discrete partitions from observed data: symbolic false nearest neighbors. *Phys. Rev. Lett.* **91**(8), 084,102 (2003)
24. Koopman, B.O.: Hamiltonian systems and transformations in Hilbert space. *Proc. Natl. Acad. Sci. USA* **17**, 315–318 (1931)
25. Koopman, B.O., von Neumann, J.: Dynamical systems of continuous spectra. *Proc. Natl. Acad. Sci. USA* **18**, 255–262 (1932)
26. Lind, D., Marcus, B.: *An Introduction to Symbolic Dynamics and Coding*. Cambridge University Press, Cambridge (1995)
27. McCulloch, W.S., Pitts, W.: A logical calculus of ideas immanent in nervous activity. *Bull. Math. Biophys.* **5**, 115–133 (1943)
28. McMillan, B.: The basic theorems of information theory. *Ann. Math. Stat.* **24**, 196–219 (1953)
29. Moore, C.: Unpredictability and undecidability in dynamical systems. *Phys. Rev. Lett.* **64**(20), 2354–2357 (1990)
30. Moore, C.: Generalized shifts: unpredictability and undecidability in dynamical systems. *Nonlinearity* **4**, 199–230 (1991)
31. Moore, C.: Dynamical recognizers: real-time language recognition by analog computers. *Theor. Comput. Sci.* **201**, 99–136 (1998)
32. Moore, C., Crutchfield, J.P.: Quantum automata and quantum grammars. *Theor. Comput. Sci.* **237**, 275–306 (2000)
33. Ott, E.: *Chaos in Dynamical Systems*. Cambridge University Press, New York (1993)
34. Pitts, W., McCulloch, W.S.: How we know universals: the perception of auditory and visual forms. *Bull. Math. Biophys.* **9**, 127–147 (1947)
35. Plate, T.A.: Holographic reduced representations. *IEEE Trans. Neural Netw.* **6**(3), 623–641 (1995)
36. Pollack, J.B.: The induction of dynamical recognizers. *Mach. Learn.* **7**, 227–252 (1991). Also published in [37], pp. 283–312
37. Port, R.F., van Gelder, T. (eds.): *Mind as Motion: Explorations in the Dynamics of Cognition*. MIT, Cambridge (1995)
38. Pothast, R., beim Graben, P.: Inverse problems in neural field theory. *SIAM J. Appl. Dyn. Syst.* **8**(4), 1405–1433 (2009)
39. Rabinovich, M.I., Huerta, R., Varona, P., Afraimovich, V.S.: Transient cognitive dynamics, metastability, and decision making. *PLoS Comput. Biol.* **4**(5), e1000,072 (2008)
40. Sandamirskaya, Y., Schöner, G.: Dynamic field theory of sequential action: a model and its implementation on an embodied agent. In: *Proceedings of the 7th IEEE International Conference on Development and Learning (ICDL)*, Monterey, pp. 133–138 (2008)
41. Sandamirskaya, Y., Schöner, G.: An embodied account of serial order: how instabilities drive sequence generation. *Neural Netw.* **23**(10), 1164–1179 (2010)
42. Schöner, G.: Neural systems and behavior: dynamical systems approaches. In: Smelser, N.J., Baltes, P.B. (eds.) *International Encyclopedia of the Social & Behavioral Sciences*, pp. 10571–10575. Pergamon, Oxford (2002)
43. Siegelmann, H.T., Sontag, E.D.: On the computational power of neural nets. *J. Comput. Syst. Sci.* **50**(1), 132–150 (1995)

44. Smolensky, P.: Tensor product variable binding and the representation of symbolic structures in connectionist systems. *Artif. Intell.* **46**(1–2), 159–216 (1990)
45. Smolensky, P.: Harmony in linguistic cognition. *Cogn. Sci.* **30**, 779–801 (2006)
46. Sontag, E.D.: Automata and neural networks. In: Arbib, M.A. (ed.): *The Handbook of Brain Theory and Neural Networks*, 1st edn, pp. 119–123. MIT, Cambridge (1995)
47. Spencer, M.C., Tanay, T., Roesch, E.B., Bishop, J.M., Nasuto, S.J.: Abstract platforms of computation. AISB 2013, Exeter (2013)
48. Tabor, W.: Fractal encoding of context-free grammars in connectionist networks. *Expert Syst.: Int. J. Knowl. Eng. Neural Netw.* **17**(1), 41–56 (2000)
49. Tabor, W., Cho, P.W., Szkludlarek, E.: Fractal analysis illuminates the form of connectionist structural gradualness. *Top. Cogn. Sci.* **5**, 634–667 (2013)
50. Tennent, R.D.: The denotational semantics of programming languages. *Commun. ACM* **19**(8), 437–453 (1976)
51. Turing, A.M.: On computable numbers, with an application to the *Entscheidungsproblem*. *Proc. Lond. Math. Soc.* **2**(42), 230–265 (1937)

Chapter 12

A Neural Approach to Cognition Based on Dynamic Field Theory

Jonas Lins and Gregor Schöner

Abstract How may cognitive function emerge from the different dynamic properties, regimes, and solutions of neural field equations? To date, this question has received much less attention than the purely mathematical analysis of neural fields. Dynamic Field Theory (DFT) aims to bridge the ensuing gap, by bringing together neural field dynamics with principles of neural representation and fundamentals of cognition. This chapter provides review of each of these aspects. We show how dynamic fields can be viewed as mathematical descriptions of activation patterns in neural populations that arise due to sensory and motor events; how field dynamics in DFT give rise to a set of stable states and associated instabilities that provide the elementary building blocks for cognitive processes; and how these properties can be brought to bear in the construction of neurally grounded process models of cognition. We conclude that DFT provides a valuable framework for linking mathematical descriptions of neural activity to actual sensory, motor, and cognitive functionality and behavioral signatures thereof.

12.1 Introduction

Much theoretical work has been dedicated to studying neural field equations at an abstract, mathematical level, focusing on the dynamic properties of the solutions (this book provides review of many of the latest efforts in this direction). Much less attention has been directed at understanding how function, in particular, how cognition may emerge from the different dynamic regimes and solutions of neural fields. This has left a gap between the mathematical models of neural fields that capture neurophysiology and theoretical models of neural function. Where is

Jonas Lins (✉) • Gregor Schöner

Institute for Neural Computation, Ruhr-Universität Bochum, 44780 Bochum, Germany

e-mail: Jonas.Lins@ini.rub.de, Gregor.Schoner@ini.rub.de

cognition in all the neural dynamics? May universal principles be identified that capture how the neural substrate rises to the demands of cognition?

Addressing these questions requires a theory that seamlessly integrates schemes of neural representation, fundamentals of cognition, and neural field dynamics. To ground the theory in neurophysiology, we must examine how neural activity captures and represents specific features of the world. To identify which properties of neural mechanisms enable them to support cognitive processes, we need a clear definition of cognition. Finally, a mathematical formalization must be chosen that endows neural fields with appropriate dynamical properties. In addition, such a mathematical theory must account for data from behavioral experiments that are observable indices of underlying cognitive processes.

In this chapter, we first address how patterns of neural activity represent attributes of sensory stimuli and of motoric actions. This provides the foundation for an operational theory of cognition based on dynamic neural fields, that we sketch next. We review core concepts of Dynamic Field Theory (DFT; [26]), a theoretical framework that implements elementary forms of cognition as process models in neural field architectures, explains behavioral data, and generates testable predictions. We then briefly address the critical features of cognition, and discuss how DFT accounts for these properties. We conclude the chapter by describing an exemplary DFT architecture that illustrates how the sketched principles may be applied to model higher-level cognitive function.

12.2 Grounding DFT in Neurophysiology

To get a sense for how neural fields may represent percepts, actions, or cognitive states, we start with single neurons and then move to populations of neurons within a given area of the brain.

Neural tuning is the classical concept that links the activity of a neuron, located somewhere within the neural networks of the brain, with the external conditions to which the organism is exposed, either through sensory stimulation or through an action initiated by the organism. Most neurons in the higher nervous system are active only while a stimulus or motor parameter (e.g., color, shape, or movement direction) is within a restricted range. Within this range, the discharge rate is often a non-monotonic function of the parameter. These functions, called tuning curves, are obtained by plotting discharge rate against the manipulated parameter. In many cases, tuning curves are Gaussian or alike (although more complex schemes exist), centered around a “preferred” value of the parameter. For instance, neurons in visual cortex might respond vigorously to a particular direction of visual motion, while the spike rate falls off gracefully when the direction deviates more and more from that direction.

Tuning is found in the brain for a wealth of parameters. Classical examples include tuning to the position of stimuli on sensory surfaces, such as the location of a visual stimulus on the retina or of a tactile stimulus on the skin (tuning curves

for spatial position are equivalent to receptive field profiles; [19, 28]). In other cases, it is motor space, such as the target position of a saccade [21] or the direction of a hand movement [8], that determines the activation of neurons. Cells may also be tuned to non-spatial feature dimensions like orientation [15] or color [4]. As a more complex example, neurons in visual area V4 are tuned to the curvature of object boundaries at specific angular positions relative to the object center [24]. In general, neurons tend to be sensitive and tuned to more than one dimension at the same time.

It is apparent from these examples that neurons often signal information about specific aspects of the sensed environment or of behavioral events. For brain areas close to the sensory or motor surfaces, it is relatively straightforward to determine tuning curves, by recording from neurons in a number of different sample conditions. Schwartz, Kettner and Georgopoulos [27], for instance, recorded the activity of 568 motor cortical neurons of monkeys while these performed hand movements in different directions. In each trial, the monkeys reached from a central position to one of eight possible targets that were distributed in three-dimensional space. In the great majority of the recorded cells, discharge rate depended on the direction of movement. The tuning curves were well described by cosine functions of the angle formed between the current movement direction and the cell's preferred movement direction.

Tuning in motor cortex is therefore broad, relatively uniform, and the preferred values are broadly distributed, covering the entire dimension. As a result, the tuning curves of different neurons strongly overlap [8]. This is common in most areas of the brain and suggests that sensory or motor parameters are represented by neural populations: For any specific value of a sensory or motor parameter, say, a particular reaching direction, a large ensemble of neurons is active. The activation pattern induced by any individual stimulus or motor condition is best characterized as a distribution of activation within a neural population. But does the entire distribution matter, or do only the neurons contribute whose preferred values are closest to current parameter value, the neurons at the very center of the distribution?

According to the population coding hypothesis, information about currently coded parameter values is indeed represented jointly by all active neurons, with each neuron contributing according to its level of activation. Georgopoulos, Kettner, and Schwartz [10] tested this hypothesis for the coding of movement direction in motor cortex. Using the same experimental data as Schwartz et al. ([27]; see above), they determined a population vector [9] for each reaching movement and compared it to the actual movement direction. The population vector is a weighted vector sum of the preferred direction vectors of all active neurons, each preferred direction vector being weighted with the neuron's current spike rate (this is the theoretical mean of the distribution of population activation in circular statistics). The population vector turned out to be an excellent predictor of movement direction. Importantly, including more neurons in the population vector yielded more precise predictions, suggesting that indeed all neurons contributed to the behavioral outcome of the activation pattern.

Other findings lend more direct support to the population coding hypothesis by demonstrating that weakly activated neurons contribute to the coded estimate. In

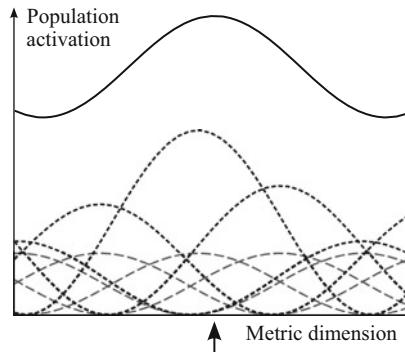


Fig. 12.1 To construct the distribution of population activation (DPA) for a particular parameter value (arrow), the tuning curves of the individual neurons (long dashed lines) are weighted by the respective neurons' firing rates. The weighted curves (short dashed lines) are then summed to obtain the DPA (solid line) over the coded dimension. Since neurons with preferred values close to the specified parameter value have higher firing rates, their curves contribute more strongly to the DPA. This leads to a peak at the position corresponding to the specified value, indicating that this value is currently represented by the population

the superior colliculus, for instance, saccade targets are coded in a topographic map of retinal space. Saccades are rapid eye movements that serve to fixate a target position. Each saccade is accompanied by a blob of activity within the neural map, the position of which specifies the retinal target. Crucially, saccadic endpoints can be influenced by pharmacologically deactivating peripheral regions of the activity blob, even though neurons in these regions are only weakly activated, compared to the cells in the blob center [21]. Similarly, stimulation experiments in the middle temporal visual area (MT) show that the perception of visual motion direction is readily influenced by artificially induced activity peaks in the neural map of movement direction, even when the artificial peak is far from the visually induced peak [11]. Activity seems to be integrated across the whole map.

Apart from corroborating the population coding hypothesis, the population vector method is a first step toward an interpretation of neural population activity. However, the population vector reduces the entire distribution of activity to one single value, discarding potentially meaningful information about its exact shape. Behaviorally relevant information potentially contained in multiple peaks of the distribution or in the shape of activation peaks is lost. Detecting the impact of these properties requires appropriate experimental paradigms and a method for constructing activity distributions from the firing of discrete neurons.

A way to do this is to construct the distribution of population activation (DPA; [5]). Although different variants of this method have been used [2, 3, 16], the basic rationale is to compute distributions from entire tuning curves rather than from the discrete preferred values of the neurons (Fig. 12.1).

First, the tuning curves of the recorded neurons are determined from a set of reference conditions (e.g., a sample of movement directions) and their amplitudes

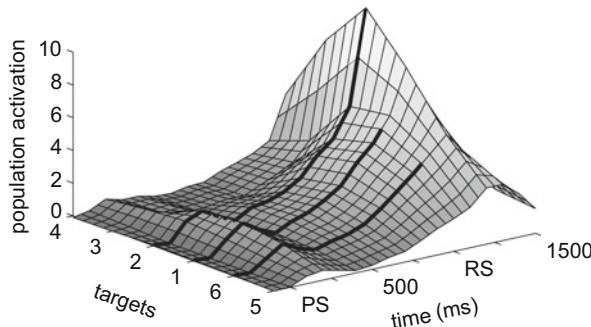


Fig. 12.2 Temporal evolution of a DPA during a reaching task, constructed from the firing of about 100 motor cortical neurons of a monkey [2]. The DPA is defined over the space of possible movement directions (targets). The DPA shown was obtained in a trial where the directional precue consisted of three adjacent lights (i.e., high directional uncertainty). The position of the three lights is indicated in the plot by *thick black lines*. From the occurrence of the preparatory signal (PS) onwards, a broad peak of activation emerges that is centered over the precued directions. It remains above activity baseline throughout the preparatory period. As the response signal (RS) occurs, activity sharpens, increases and shifts position, resulting in a pronounced peak centered around the final movement direction (Adapted from [2])

are normalized. The DPA of any particular test condition is then obtained as a weighted sum of these tuning curves. The weighting factors are the neurons' average spike rates in the test condition. The sum is normalized by the number of neurons and additional normalization steps compensate for uneven sampling from the distribution of preferred values. The DPA obtained this way is defined over the same parameter dimension as the tuning curves.

Bastian, Schöner, and Riehle [2] demonstrated that the shape of a DPA correlated with behavioral constraints. In the behavioral paradigm, monkeys reached from a central button to one of six target lights arranged around it in a hexagonal shape. Each trial started with the monkey pressing and holding the center button. A preparatory period followed, in which varied amounts of information about the upcoming movement direction were provided, by turning on one, two, or three adjacent lights. After one second one of the cued lights was turned to a different color as a definite response signal, prompting the monkey to move its hand to that light.

Immediately after the onset of the preparatory signal, a peak developed in the DPA, centered over the precued directions (Fig. 12.2).

This activation remained above baseline throughout the preparatory period and increased slowly after the response signal was supplied. Thus, information about the potential movement directions was retained throughout this period. Moreover, the shape of the peak reflected the precision of prior information: more informative precues (i.e., fewer cue lights) led to higher and sharper peaks. When the definite response signal occurred, the peak sharpened and shifted toward the position corresponding to the cued movement direction, reaching its maximum height about

100 ms before movement initiation. In this final stage, the shape of the peak was approximately equal for all conditions. In addition, Bastian and colleagues found that the shape of the DPA predicted reaction times (measured as the time from the occurrence of the response signal to movement onset): Broader, less pronounced peaks during the preparatory period corresponded to longer reaction times than sharp, pronounced peaks. Apparently, it took more time for broader peaks to reach sufficient concentration and height to initiate motor action. These findings show that the shape of the DPA impacts on behavior and that DPA shape may reflect different degrees of certainty or precision.

Using very similar techniques, Cisek and Kalaska [3] showed that multimodal DPAs may express different discrete choices of parameter values. They found that DPAs in premotor cortex can simultaneously represent two precued movement directions (only one of which is later realized). Clearly, the population vector is unable to represent such multi-valued information.

To summarize, peaks in DPAs pertain to macroscopically relevant perceptual or behavioral conditions, and the exact shape of the distribution carries information that may observably impact behavior. Thus, DPAs provide an appropriate level of consideration to assess the functional relevance of neural activity patterns.

12.3 Dynamic Field Theory

Dynamic Field Theory (DFT) builds on the finding that the relevant information is carried by distributions of activation among populations of neurons rather than by single cells. Via the DPA method, DFT is tightly linked to the physiology of population coding. DFT describes the evolution in time of activation patterns in neural populations. The activation patterns are modeled as Dynamic Neural Fields (DNFs) that are defined over continuous metric dimensions and evolve continuously in time. The fields may be defined over virtually any perceptual, behavioral or cognitive dimension, such as color, retinal position, tone pitch, movement direction, or allocentric spatial position. Special focus is laid on modeling lateral neural interactions within the fields, endowing them with a particular set of stable attractor states. These stable states correspond to meaningful representational conditions, such as the presence or absence of a particular value along the coded dimension. Instabilities that lead to switches between the different stable states are brought about by sufficient changes in the configuration of the external input a field receives.

The particular mathematical form of field dynamics adopted by DFT has first been analyzed by Amari ([1]; see also [12, 34]):

$$\tau \dot{u}(x, t) = -u(x, t) + s(x, t) + h + \int w(x - x') \sigma(u(x', t)) dx' \quad (12.1)$$

Here, $u(x, t)$ is the field of activation, defined over the metric dimension, x , and time, t . From a neurophysiological viewpoint, the activation, u , can be interpreted

as a correlate to the mean membrane potential of a group of neurons. The time scale of the relaxation process is determined by τ . The field has a constant resting level, h , and may receive localized patterns of external input, $s(x, t)$. The last term describes lateral interactions between different field sites. Here, σ is a sigmoidal function implementing a soft threshold for field output, and w is an interaction kernel that specifies the strength of interactions between different field sites as a function of their metric distance. The kernel typically has a Mexican hat shape, implementing local excitation and surround inhibition, usually with added global inhibition. This means that field sites coding for similar parameter values excite each other, while mutual inhibition predominates between field sites that code for very different values. The sigmoidal threshold function ensures that only sufficiently activated field sites generate output and impact on other sites. The field output can be viewed as corresponding to the mean spike rate of a group of neurons.

In the absence of supra-threshold activation, no output is generated. In this case, the entire field relaxes to the stable attractor that is set by the resting level (which usually resides well below the output threshold). A flat distribution indicates the absence of any specific information about the coded dimension.

When weak, localized input is applied, the attractor at the respective field site is shifted toward the output threshold. As long as the threshold is not reached, though, the field state remains purely input-driven and activation thus simply traces the shape of the input (Fig. 12.3a). Although there is now some structure to the distribution, this state still indicates the absence of conclusive information.

If, in contrast, the localized input is sufficiently strong to push a section of the field above threshold, output is generated and lateral interaction kicks in. Provided the parameters of the interaction kernel are within an appropriate range, lateral interaction promotes the formation of a localized peak of activation (Fig. 12.3b). Local excitation further elevates activation around the input position, whereas more distant field sites are depressed by global inhibition and/or surround inhibition, which prevents the peak from dispersing. Due to these properties we refer to this as a self-stabilized peak.

We call the transition from a sub-threshold solution to a self-stabilized peak the detection instability, because it corresponds to the decision that a coherent, well-defined item is present in the input stream. Peaks are units of representation in this sense, indicating that a particular parameter value is present in the sensory environment, as part of a motor plan, or as the contents of memory. The encoded value itself – what is being perceived, planned, or memorized – is specified by the position of the peak along the metric dimension. Peak height and width, on the other hand, may reflect certainty, intensity and precision (analogous to DPAs).

Conversely, we call it the reverse detection instability when an existing self-stabilized peak vanishes. This happens when the localized input that brought about the peak is sufficiently reduced in strength. For example, when the input is removed entirely, the peak attractor becomes unstable and disappears, while the resting level attractor reappears, to which the system then relaxes. Decreasing the input strength successively will also eventually trigger the reverse detection instability. However, local excitation to a degree shields existing peaks from decaying. The system will

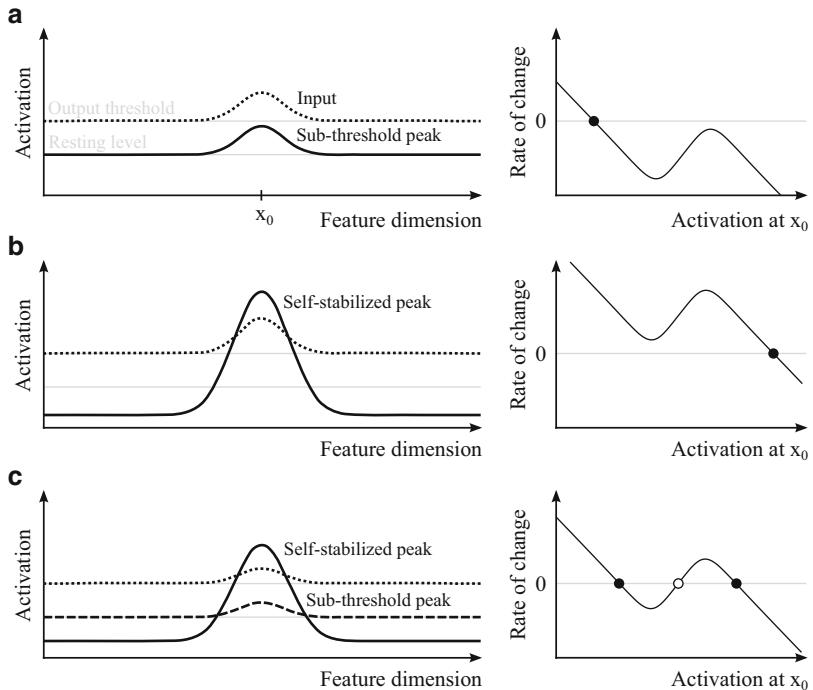


Fig. 12.3 *Left column:* Stable states reached by dynamic neural fields (solid lines and long dashed line) as a result of localized Gaussian inputs of different strengths (dotted lines). *Right column:* Corresponding plots of the rate of change as a function of activation at the peak position, x_0 (note that these plots are only approximate, as they do not take into account the impact of other field sites on the rate of change at x_0 via lateral interactions). Attractors are marked by filled dots, repellors by open dots. **(a)** Weak input results in a purely input-driven sub-threshold peak, which is a monostable attractor state. **(b)** High levels of input that bring activation above threshold result in output generation and lateral interactions, thus leading to a self-stabilized peak. This state as well is monostable. **(c)** For intermediate input strengths the system reaches a bistable state. The current state then depends on the system's prior state. Here, the self-stabilized peak (solid line) corresponds to the attractor on the right side, which is reached from high levels of activation. The sub-threshold peak (long dashed line) corresponds to the left attractor, which is reached from low levels of activation

thus stick to the detection decision across a range of input strengths that would not have triggered the detection instability in the first place. The system is bistable over this range, with the peak attractor and the input-driven attractor coexisting (Fig. 12.3c). The field state then depends on which basin of attraction it resided in prior to the change of input strength. In other words, the state of the system depends on its activation history, it displays hysteresis.

Hysteresis stabilizes decisions against random fluctuations and perturbations. In the nervous system, such fluctuations may arise due to the inherent variability in neural firing or as the result of currently ongoing but unrelated neural processes.

Under such conditions a lack of hysteresis would lead to constantly fluctuating decisions for near-threshold input. It is thus unsurprising that signatures of hysteresis are a common finding in behavioral experiments (for review, see [13]). For instance, perceptual hysteresis has been reported for single-element apparent-motion [14]. In the experiment, participants were shown two squares with differing luminance. While the participants were watching, the squares constantly exchanged their luminance values, which created a percept of either flicker or apparent motion between them. In each trial participants reported at multiple time points which of the two percepts they currently experienced. The decisive variable predicting whether motion or flicker was perceived was background-relative-luminance contrast (BRLC). BRLC is the strength of luminance change between individual frames in relation to how much the spots' average luminance differs from background luminance. High BRLC led to motion percepts more frequently than low BRLC. Hysteresis was observed when BRLC was changed continuously during the trials in a descending or ascending manner. The BRLC value at which the motion percept was lost in descending trials was lower than the value at which the motion percept was established in ascending trials. This suggests that within a certain range of BRLC values motion perception is bistable.

Another fundamental attractor state in DFT is the self-sustained peak state. self-sustained peaks arise in much the same way as self-stabilized ones: When localized, excitatory input brings activation above threshold, output is generated, driving lateral interactions that support peak formation. The difference lies in the balance of excitation and inhibition. A field supports self-sustained peaks if local excitation is sufficiently strong, relative to inhibition, to by itself prevent peaks from decaying after the input is removed. In this regime, peaks decay only when the level of activation is sufficiently decreased, locally or globally, by external inhibitory input or by endogenous inhibitory interactions. Otherwise, self-sustained peaks may persist indefinitely, even in the absence of localized input. The self-sustained regime enables DNFs to support the functionality of the neural process of working memory (see also, [6], and the original [7]), which will be illustrated in Sect. 12.5.

We have so far considered only single localized inputs. But natural environments are usually richly structured – visual scenes are cluttered with objects, auditory signals arrive from multiple directions, and so forth – which in turn implies a variety of potential behavioral goals and movement targets. Several inputs may be equally salient due to, say, equal brightness, contrast, or loudness. Under such conditions multiple stimuli compete for processing and behavioral impact. In terms of DNFs, this amounts to a field receiving multiple localized inputs at the same time. DFT provides dynamic mechanisms that mediate selection in such situations.

The selection of saccade targets is a well-studied example, which has been addressed in detail by DFT modeling efforts. We base the following considerations on those efforts, mainly on a model by Wilimzig, Schneider, and Schöner [33], which in turn complements prior modeling work [20, 31].

Saccades are rapid eye movements that serve to quickly fixate targets in visual space. Saccade trajectories are planned prior to the initiation of the movement and are not adjusted afterwards. The metrics of saccades are specified in the superior

colliculus, a mid-brain structure that integrates cortical and direct visual input in a topographic map of visual space. Activation bumps in this map specify the vertical and horizontal extent of saccades. It is thought that the superior colliculus plays an important role in both target selection and saccade initiation. The DNFs in the model sketched below can be viewed as roughly corresponding to the respective neural populations in the superior colliculus.

The model consists of a selection level and an initiation level. Each level is constituted by a DNF with local excitation and global inhibition, defined over the space of saccadic endpoints (i.e., retinal space). The space is modeled as one-dimensional, which is sufficient to capture most experimental paradigms. We will restrict our considerations to the selection level, where selection between different visual targets occurs.

Items in the visual field are fed into the selection field as localized, Gaussian-shaped input patterns. When a single target item generates sufficient input, the detection instability occurs, resulting in the formation of a self-stabilized peak of activation in the field. Although the model of Wilimzig et al. accounts for several experimental findings with regard to the case of a single target as well, we are here primarily interested in situations in which at least two targets are presented simultaneously. What happens in the double target case depends on the exact configuration of the two inputs and on activation biases that may be caused by stochastic perturbations or by imbalances between the two stimuli.

We first consider the case of two visual items that are equally salient and spatially remote from each other. Also, we will assume that there are no stochastic perturbations, and thus no random activation biases. Due to their equal saliency, both items generate input of the exact same strength. Furthermore, the large distance between the items ensures that there is no interaction between the resulting peaks, except for homogenous global inhibition. As a consequence, two supra-threshold peaks emerge that are somewhat less pronounced than in the single input case, as the sum of global inhibition is larger. The resulting field state is a fixed point of the system, but it is not stable.

This becomes apparent when an activation bias is introduced. One source of such imbalance are stochastic perturbations caused by neuronal variability or by other ongoing neural processes (the saccade model implements stochastic variability as Gaussian white noise). Random fluctuations of activity may provoke selection decisions by strengthening or weakening one of the competing peaks. Another source of imbalance is the relative strength of the inputs themselves. In a visual context, the strength of an input may be associated, for instance, to the brightness or contrast of a stimulus. Regardless of its source, an imbalance in favor of one of the peaks leads to increased activation and excitation around this location, increasing the height of the peak above that of its competitor. The ensuing increase in global inhibition suppresses the weaker peak, eventually reducing it to an input-driven bump. The single-peak state resulting from this selection decision is bistable, with both peak attractors coexisting (Fig. 12.4a).

However, multiple inputs do not always result in selection, but may also lead to fusion. This has been shown empirically for the case of saccades by Ottes, van

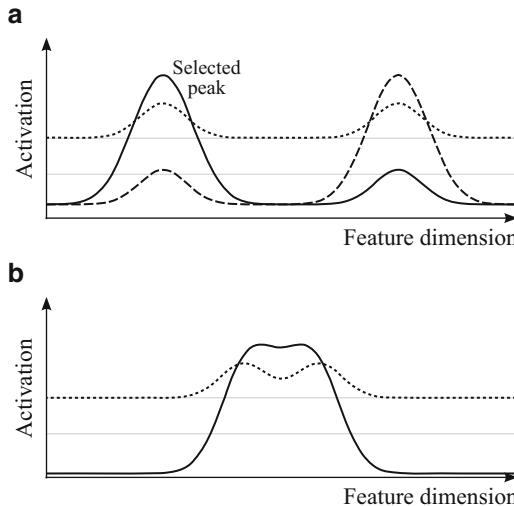


Fig. 12.4 Stable states reached by dynamic neural fields (*solid lines* and *long dashed line*) as a result of different patterns of localized Gaussian input (*dotted lines*). **(a)** Competition between peaks occurs when two inputs are applied at distant positions. Only at one location is a self-stabilized peak formed (*solid line*), while the other is suppressed by inhibition. The state resulting from this selection decision is bistable, with the alternative state (*long dashed line*) continuing to coexist as an attractor. Which state is reached depends on the field's prior activation history, imbalances between the inputs, and noise. **(b)** Two close inputs can result in a monostable fused peak state, with a single peak at an average location between the inputs

Gisbergen, and Eggermont [23]. Their participants made saccades from a fixation point to a green stimulus whenever it appeared. In some trials, a red stimulus appeared alongside the green stimulus. Although the participants were instructed to ignore the red item, the first saccade they made often landed at an average position between the two stimuli. Such averaging saccades occurred much more frequently when the stimuli were spatially close than when they were widely separated.

This phenomenon is as well captured by the saccade model. It is, in fact, a feature of DNFs in general. As observed by Ottes et al., whether selection or fusion occurs depends to a large part on the inputs' metrics. Specifically, two inputs tend to result in a single peak at an average position if they are so close to each other that the regions of input-induced activation are subject to mutual excitation. In that case, the activation propagates from the two input positions towards the center between them, eventually forming a single peak (Fig. 12.4b). The state with fused peak is monostable for very close inputs. When the distance between them is increased after a fused peak has already established, it becomes bistable at some point. That is, although the fused peak persists, applying the same input configuration to a previously inactive field would result in a selection decision. If the distance is increased even more, the attractor of the fused peak eventually becomes unstable

and disappears. The field then relaxes to the now monostable selection state. We call this the fusion/selection instability.

In the model of Wilimzig et al. inhibition is not implemented via the same interaction kernel as excitation, but it is instead mediated by an additional layer of interneurons. This layer receives excitatory input from the main, purely excitatory field and projects back global inhibition to it. On the one hand, this implementation of inhibition was chosen to accord to Dale’s law, which states, roughly put, that each neuron releases the same set of neurotransmitters at all of its synapses. Assuming that DNFs capture homogeneous neural populations, this means that a field’s output can be either excitatory or inhibitory, but not both. On the other hand, using a separate inhibitory field has the effect of delaying the impact of inhibition compared to that of local excitation. Due to the different roles of excitation and inhibition in the fusion and selection of peaks this gives rise to a specific association between the latency and the type of saccades. Saccades with lower latencies are more likely to target an average position between the stimuli, while later saccades tend to select one target. This effect has also been found empirically [23].

Note that the maximum number of peaks that a DNF can support depends on the balance of excitation and inhibition. Fields in DFT are not generally constrained to a single peak. However, the number of peaks is usually quite limited through inhibition. This is particularly relevant when modeling explicit capacity limits in cognition, such as those in working memory or attentional function.

12.4 DFT as an Approach to Cognition

Through the dynamic properties described in the previous section, DNFs acquire capabilities that are at the core of cognition: making decisions and maintaining the outcome of these decisions. Detection decisions make neural representations to some extent independent from the continuous input stream. Selection decisions further decouple the contents of neural representations from the immediate input, by separating items into those that impact processing and those that are ignored. The specific stability properties of DNFs ensure that the outcomes of these decisions are shielded from changes in the input, and retained as long as needed.

The paramount importance of these capabilities for cognition is perhaps best illustrated by considering an alternative approach, one that may seem more straightforward at first glance, but that faces profound issues when it comes to cognition – exactly because it lacks the capabilities described above.

To start with, the nervous system is immersed in a continuously changing environment, facing a continuous stream of sensory input. A simple way to guide behavior based on this type of input is to use some form of continuous closed-loop control system. That is, a mechanism that continuously maps sensory input and feedback to motor action, according to an appropriate function. Such systems may perform intriguingly complex control tasks. Many simple biological organisms work this way, as well as many systems of the human body, such as the regulation

of blood pressure. Could cognition be nothing more than the input-driven evolution of activation in a system of this type? To put it more radically, may cognition work on the same principles as the regulation of blood pressure? Such a view has indeed been advanced by some (e.g., [32]).

The problem with such systems is that the linkage between their different subparts is seamless from input to output. The value of every variable is uniquely specified at any point in time and there are no discontinuities between the input and the output stage. As a consequence, subparts of such systems cannot shield their own state from the continuous input stream or from the impact of other subparts. So the output of the system is basically a transformed version of the input stream that is tightly coupled to the input at all times. There is also no way of how such a system may represent the absence of information. Control variables cannot be “empty”. Conversely, the emergence of new states is always strongly affected by previous states. Due to these properties, the emergence of decisive hallmarks of cognition and behavior cannot be explained by such systems.

One of these hallmarks is the discreteness of behavioral events. How may discrete behavioral events be initiated and terminated on the basis of purely continuous processes? There must be a gap somewhere in between continuous input, intermediate processes, and behavior, that cannot be explained in terms of systems as the one described above.

On a closer look, the problem does not apply to motor action alone, but extends to those capabilities of the brain that are often regarded as “higher” forms of cognition. For example, mental imagery, working memory, sequence generation – all these faculties have in common that their functioning requires a degree of independence from the current sensory or motor environment. Working memory initially requires sensory input to store, but after storage has been achieved, it requires that the stored information be shielded from being overwritten by new input. Imagery is essentially defined by the independence of a perceptual brain state from current sensory input. Actions in a controlled sequence that works toward some distal goal (e.g., making coffee) need to be shielded against distracting input that would trigger unrelated behavior (e.g., taking a cup and cleaning it).

Thus, at the heart of cognition lies the nervous system’s capability to generate, maintain, and act upon inner states that are, to a degree, independent from current sensory input. Mechanisms are needed that decouple the representations upon which cognitive operations are carried out from the immediate sensed world (and from each other). On the other hand, behavior and cognition still need to be closely linked to the sensory surfaces, else we would think and act completely aloof. This is the core assumption of the stance of embodied cognition [25]. Cognition and behavior are still flexible, in that they can be updated online if relevant new input is detected. So what is needed is a balance between decoupling and coupling that allows only certain input to impact on downstream systems, cognition, and action, but which nonetheless allows these decisions to be changed if appropriate.

DFT effectively implements these demands. Note that the decisive capabilities, detection, selection, and appropriate stability, are realized in each individual field. This means that elementary forms of cognition happen at already very low levels

of computation rather than being dependent on complex architectures. We have considered some concrete examples for this in the previous section (e.g., in the context of the saccade model). However, the stability of individual fields also allows for the construction of modular field architectures, which may implement more complex cognitive tasks and employ layers farther removed from the immediate sensory and motor surfaces. We conclude the chapter by considering such an architecture.

12.5 Modeling Visual Working Memory and Change Detection with Dynamic Neural Fields

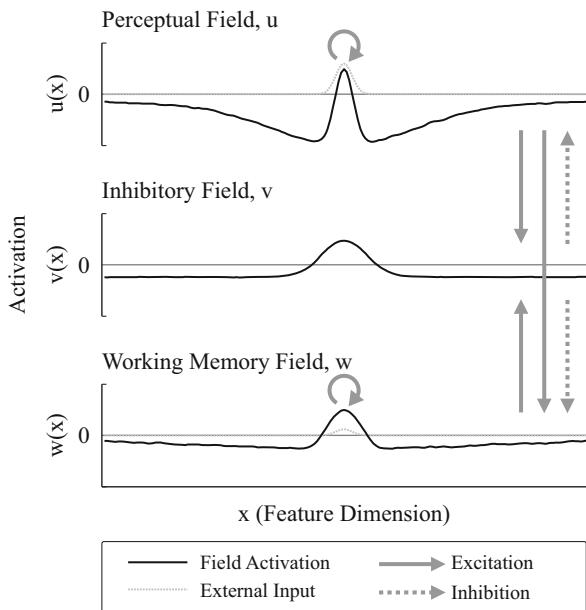
Our considerations are based on a model originally proposed by Johnson and colleagues [17, 18] that is rooted in a general DFT approach to visual and spatial cognition [29, 30]. The model addresses the two closely linked cognitive domains of visual working memory and change detection. Visual working memory stores recent visual input over durations in the order of seconds and makes this information available to other processes. Change detection means comparing the contents of visual working memory to newly incoming visual input. Combining visual working memory and change detection yields a strategy for detecting changes in visual scenes despite the frequent interruptions of the visual input stream by saccades and blinks.

Change detection can be probed experimentally by showing to the participant a display with several simple visual items, such as colored dots, that differ along at least one feature dimension (e.g., [22]). After a short delay of normally less than a second a test display is presented that is either identical to the previous one or in which one of the items has changed with respect to one feature (e.g., an item may have changed color). Participants then indicate whether or not they perceive a change, by responding “different” or “same”.

To perform this task, the items in the first display need to be perceived and encoded into working memory. The resulting representation must then be shielded from new input and maintained over the delay. Finally, the contents of visual working memory must be compared to the test display, which requires integrating working memory and perception.

As a first step in accomplishing this, a system must be capable of retaining information in the absence of input. self-stabilized peaks are well-suited for perception, since they are quite tightly linked to the presence of input, but they decay when the input is removed for a longer period of time. working memory therefore requires a different dynamic regime, namely, the self-sustained one. Because DNFs can operate in only one regime at a time, perception and working memory require separate fields. The model by Johnson and colleagues thus employs a perceptual field and a working memory field (Fig. 12.5).

Fig. 12.5 The three-layer DNF architecture of visual working memory and change detection [18]. See text for details



The different functional roles of these fields arise from their different sources of excitatory input (dashed gray line and solid arrows in Fig. 12.5) and the different dynamic regimes in which they operate. The perceptual field receives direct feature input from the visual scene (taking the form of Gaussians) and operates in a self-stabilized regime. The working memory field receives its main excitatory input from the perceptual field (and weak direct visual input) and operates in a self-sustained regime. The integration of perception and working memory is achieved through a shared layer of inhibitory interneurons. This layer operates in a purely input driven regime and receives excitatory input from both the perceptual and the working memory field. In turn, the inhibitory layer sends back broad (but localized) inhibition to both other fields (dashed arrows in Fig. 12.5). That is, surround inhibition in both the perceptual and the working memory field is mediated by the inhibitory interlayer, as well as mutual inhibition between the perceptual and the working memory field. All fields are defined over a metrically scaled visual feature dimension.

We first consider how a single feature input (corresponding to, say, a single colored item) is encoded perceptually, encoded to working memory, maintained over a delay, and compared to a subsequent input. For this we refer to the simulation results shown in Fig. 12.6.

In model terms, presenting a visual item amounts to providing localized input to the perceptual field (and a much weaker version of the same input to the working memory field). If the input is sufficiently strong, the perceptual field undergoes the detection instability, leading to a self-stabilized peak (Fig. 12.6, $t = 400$). This step corresponds to the perceptual encoding of the input feature value. Once the peak has established, the perceptual field provides localized input to the working

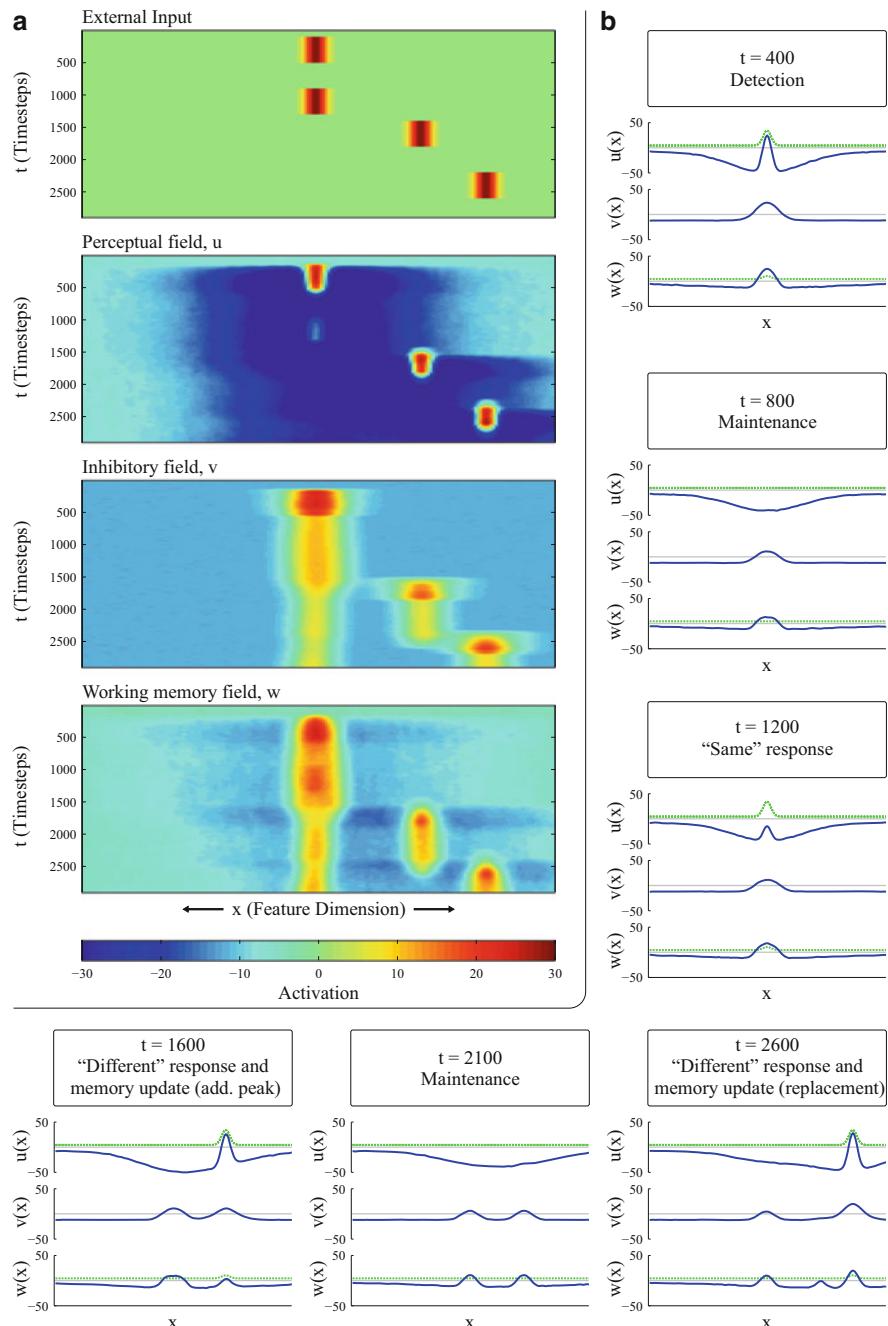


Fig. 12.6 Simulation of change detection in the three-layer DNF architecture. **(a)** Evolution of field activation over simulation time steps, in response to the pattern of feature input shown in the topmost plot. **(b)** Snapshots of the model state at selected time steps. At $t = 400$ the single

memory field, leading to a self-sustained working memory peak. This corresponds to encoding the feature value into working memory. As soon as the external input is removed, the perceptual peak destabilizes and decays, while the working memory peak is maintained over the delay in the absence of input (Fig. 12.6, $t = 800$).

The change detection functionality naturally emerges from this setup. Mediated by the inhibitory layer, the sustained working memory peak leads to inhibition of the perceptual field at the field site corresponding to the feature value held in memory. The resulting activation trough is critical, because it ensures that new input to the perceptual field reaches threshold only if the test item is sufficiently different from the value held in memory. The change detection mechanism thus comes into effect in a completely autonomous manner when new input arrives. If the test item is very similar to the retained one, the visual input to the perceptual field coincides with the center of the trough of inhibition (Fig. 12.6, $t = 1,200$). This makes it unlikely that the output threshold is reached. Accordingly, the absence of a peak in the perceptual field and the concurrent presence of a peak in the working memory field at test time indicate that no change has been detected (“same” response). If, in contrast, the test item is metrically sufficiently different from the first one, the visual input peak is somewhat displaced from the center of the trough (Fig. 12.6, $t = 1,600$). It thus impacts on a field site where inhibition is less pronounced, so that the output threshold is reached more easily. Therefore, supra-threshold activation in the perceptual field at test time means that a change has been detected (“different” response).

Note that the same/different decision can be made explicit by introducing two self-excitatory, mutually inhibitory dynamical nodes, a “same” node and a “different” node. The “different” node receives summed activation from the perceptual field, while the “same” node receives summed activation from the working memory field. To force a decision at test time, a boost of activation is applied to both nodes, leading to the selection of one alternative, depending on the ratio of supra-threshold activation in the perceptual and the working memory field. These nodes were introduced to enable comparisons of the model performance with behavioral data [17]. For simplicity we omit this detail in our considerations.

Fig. 12.6 (continued) localized input (*dashed green line*) has led to a self-stabilized peak in the perceptual field, which projects to both the working memory field and the inhibitory field. Due to the self-sustaining regime of the working memory field, the peak there is maintained in the absence of input ($t = 800$). Through the inhibitory field it creates a trough of inhibition in the perceptual field. If at test the same or a very similar item is shown ($t = 1,200$) the input-driven hump in the perceptual field coincides with a region of the trough where there is strong inhibition. This makes it unlikely that the output threshold is reached. If, in contrast, the difference between the test item and the retained item is sufficiently large ($t = 1,600$), the input impacts on a less strongly inhibited region of the perceptual field so that the threshold is reached more easily. The projection of the new perceptual peak to the working memory field may lead to the updating of working memory with an additional peak that is then as well maintained over periods without input ($t = 2,100$). Alternatively, an existing memory peak may be suppressed and replaced if the new input is relatively close to an existing one ($t = 2,600$)

Apart from signaling change, a peak in the perceptual field may have the effect of updating the contents of working memory. Depending on the metric distance between the already existing peaks and the new input, this can either mean that a completely new peak is established (for remote items; Fig. 12.6, $t = 1,600$), leading to a stable multi-peak solution (Fig. 12.6, $t = 2,100$), or that an existing peak destabilizes and is replaced by the new peak (for close items; Fig. 12.6, $t = 2,600$).

The functionality we have described so far generalizes to multiple items. When several different inputs are applied in the encoding phase, a multi-peak solution arises and persists in the working memory field (Fig. 12.7a, b). Change detection then works in the same way as for single items. However, with a larger number of items, interactions between the retained peaks can lead to effects not observed in the case of a single item. For example, because the level of inhibition increases with the number of peaks in working memory, there is a capacity limit with respect to the total number of items that can be retained (about four in this particular model). This limit can lead to the deletion of existing peaks by new input (“forgetting”), to incomplete encoding of multi-item displays, or to failure to encode new input. The all-or-none property of working memory in the model – a stable peak is either formed or not – is consistent with behavioral data [35].

A counter-intuitive prediction made by the model is that change detection should be enhanced when two metrically close items are retained in working memory and a new, slightly different item is presented. This is because inhibition in the model is not global, but local, and tied to the position of peaks, so that nearby peaks inhibit each other more strongly than other peaks. Close peaks in working memory are thus less pronounced than more isolated ones (Fig. 12.7b).

Because the peaks are smaller, they also lead to slightly less inhibition and a shallower trough in the perceptual field. This makes it more likely that new input within the area of the trough reaches threshold and generates a “different” response. Therefore, the same degree of difference between a test item and a retained item can result in either a “same” or a “different” response, depending on whether the nearest working memory peak is relatively isolated (Fig. 12.7c) or has other peaks in its vicinity (Fig. 12.7d). This prediction has been confirmed empirically [17].

Other behavioral evidence that has been successfully captured by the model includes the selection of inputs for encoding into working memory (e.g., based on saliency) and the mutual repulsion of retained values along the feature dimension.

To summarize, the model provides a neurally plausible process account for visual working memory and change detection, captures diverse behavioral data, and has been a source of new, testable predictions. This example illustrates that, by virtue of its modularity, DFT is well suited to capture not only elementary aspects of cognition, but also more complex (or “higher”) cognitive acts.

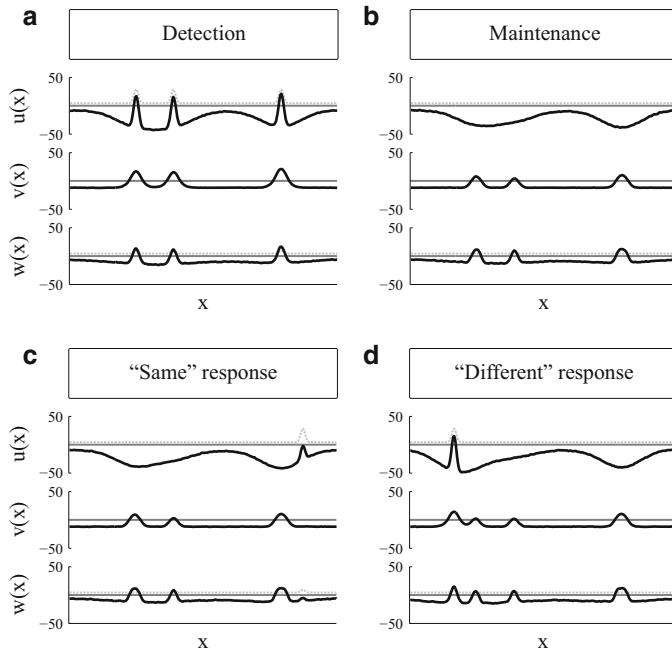


Fig. 12.7 Enhancement of change detection in the three-layer architecture due to similar items being held in working memory. **(a)** Encoding of three items into working memory, two close ones and an isolated one. **(b)** The two close peaks in working memory inhibit each other and are thus less pronounced than the isolated one. In turn, the trough of inhibition in the perceptual field caused by the isolated peak is slightly more pronounced than the trough caused by the joint impact of the two close peaks. **(c)** If the item shown at test is similar to the isolated peak, the input coincides with the deeper trough, making an erroneous “same” response more probable. **(d)** Here the test item is instead similar to one of the close items. Although the degree of similarity is the same as in **(c)**, the threshold is reached more easily since the new input falls into a region of the shallower trough, having to overcome less inhibition

12.6 Conclusions

We have reviewed how neural fields may be viewed as mathematical descriptions of distributions of population activation. Their dynamics, captured in Dynamic Field Theory, leads to a set of stable states, sub-threshold solutions, self-stabilized peaks, self-sustained peaks, and associated instabilities, the detection, selection, and memory instability. From these, cognitive properties of dynamic neural processes emerge. By linking neural field dynamics to behavioral signatures of sensory, motor, and cognitive function, DFT provides an interface between neurally grounded process models and cognition. The instabilities of DFT provide critical properties of cognitive processes, most prominently, the capability to both isolate cognitive states from distractor input or interaction, while at the same time maintaining the capacity

to link cognitive processes to ongoing sensory and motor processes as well as to other concurrent cognitive processes.

Much work remains to be done to ground all cognition in neural processing. DFT has helped make the first steps, emphasizing the embodied nature of cognition. The frontier now is to move such principles toward higher cognition.

Acknowledgements The authors acknowledge support from the German Federal Ministry of Education and Research within the National Network Computational Neuroscience Bernstein Fokus: “Learning behavioral models: From human experiment to technical assistance”, grant FKZ 01GQ0951.

References

1. Amari, S.: Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.* **27**(2), 77–87 (1977)
2. Bastian, A., Schöner, G., Riehle, A.: Preshaping and continuous evolution of motor cortical representations during movement preparation. *Eur. J. Neurosci.* **18**(7), 2047–2058 (2003)
3. Cisek, P., Kalaska, J.F.: Neural correlates of reaching decisions in dorsal premotor cortex: specification of multiple direction choices and final selection of action. *Neuron* **45**(5), 801–814 (2005)
4. Conway, B.R., Tsao, D.Y.: Color-tuned neurons are spatially clustered according to color preference within alert macaque posterior inferior temporal cortex. *Proc. Natl. Acad. Sci.* **106**(42), 18034–18039 (2009)
5. Erlhagen, W., Bastian, A., Jancke, D., Riehle, A., Schöner, G.: The distribution of neuronal population activation (DPA) as a tool to study interaction and integration in cortical representations. *J. Neurosci. Methods* **94**(1), 53–66 (1999)
6. Fuster, J.M.: Cortex and Mind: Unifying Cognition. Oxford University Press, Oxford (2005)
7. Fuster, J.M., Alexander, G.E.: Neuron activity related to short-term memory. *Science* **173**(3997), 652–654 (1971)
8. Georgopoulos, A.P., Kalaska, J.F., Caminiti, R., Massey, J.T.: On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex. *J. Neurosci.* **2**(11), 1527–1537 (1982)
9. Georgopoulos, A.P., Caminiti, R., Kalaska, J.F., Massey, J.T.: Spatial coding of movement: a hypothesis concerning the coding of movement direction by motor cortical populations. *Exp. Brain Res.* **49**, 327–336 (1983)
10. Georgopoulos, A.P., Kettner, R.E., Schwartz, A.B.: Primate motor cortex and free arm movements to visual targets in three-dimensional space. ii. Coding of the direction of movement by a neuronal population. *J. Neurosci.* **8**(8), 2928–2937 (1988)
11. Groh, J.M., Born, R.T., Newsome, W.T.: How is a sensory map read out? Effects of microstimulation in visual area MT on saccades and smooth pursuit eye movements. *J. Neurosci.* **17**(11), 4312–4330 (1997)
12. Grossberg, S.: A theory of human memory: self-organization and performance of sensory-motor codes, maps, and plans. In: Snell, F.M., Rosen, R. (eds.) *Progress in Theoretical Biology*: Vol. 5, pp. 500–639. Academic, New York/London (1978)
13. Hock, H.S., Schöner, G.: A neural basis for perceptual dynamics. In: Huys, R., Jirsa, V.K. (eds.) *Nonlinear Dynamics in Human Behavior*, pp. 151–177. Springer, Berlin/Heidelberg (2010)
14. Hock, H.S., Kogan, K., Espinoza, J.K.: Dynamic, state-dependent thresholds for the perception of single-element apparent motion: bistability from local cooperativity. *Percept. Psychophys.* **59**(7), 1077–1088 (1997)

15. Hubel, D.H., Wiesel, T.N.: Receptive fields and functional architecture of monkey striate cortex. *J. Physiol.* **195**(1), 215–243 (1968)
16. Jancke, D., Erlhagen, W., Dinse, H.R., Akhavan, A.C., Giese, M., Steinhage, A., Schöner, G.: Parametric population representation of retinal location: neuronal interaction dynamics in cat primary visual cortex. *J. Neurosci.* **19**(20), 9016–9028 (1999)
17. Johnson, J.S., Spencer, J.P., Luck, S.J., Schöner, G.: A dynamic neural field model of visual working memory and change detection. *Psychol. Sci.* **20**(5), 568–577 (2009)
18. Johnson, J.S., Spencer, J.P., Schöner, G.: A layered neural architecture for the consolidation, maintenance, and updating of representations in visual working memory. *Brain Res.* **1299**, 17–32 (2009)
19. Jones, J.P., Palmer, L.A.: The two-dimensional spatial structure of simple receptive fields in cat striate cortex. *J. Neurophysiol.* **58**(6), 1187–1211 (1987)
20. Kopecz, K., Schöner, G.: Saccadic motor planning by integrating visual information and pre-information on neural dynamic fields. *Biol. Cybern.* **73**(1), 49–60 (1995)
21. Lee, C., Rohrer, W.H., Sparks, D.L.: Population coding of saccadic eye movements by neurons in the superior colliculus. *Nature* **332**(6162), 357–360 (1988)
22. Luck, S.J., Vogel, E.K.: The capacity of visual working memory for features and conjunctions. *Nature* **390**(6657), 279–281 (1997)
23. Ottes, F.P., van Gisbergen, J.A., Eggermont, J.J.: Latency dependence of colour-based target vs nontarget discrimination by the saccadic system. *Vis. Res.* **25**(6), 849–862 (1985)
24. Pasupathy, A., Connor, C.E.: Shape representation in area v4: position-specific tuning for boundary conformation. *J. Neurophysiol.* **86**(5), 2505–2519 (2001)
25. Riegler, A.: When is a cognitive system embodied? *Cogn. Syst. Res.* **3**(3), 339–348 (2002)
26. Schöner, G.: Dynamical systems approaches to cognition. In: Sun, R. (ed.) *The Cambridge Handbook of Computational Psychology*, pp. 101–126. Cambridge University Press, Cambridge (2008)
27. Schwartz, A.B., Kettner, R.E., Georgopoulos, A.P.: Primate motor cortex and free arm movements to visual targets in three-dimensional space. i. Relations between single cell discharge and direction of movement. *J. Neurosci.* **8**(8), 2913–2927 (1988)
28. Sherrington, C.S.: *The integrative action of the nervous system*. Yale University Press, New Haven (1906)
29. Simmering, V.R., Schutte, A.R., Spencer, J.P.: Generalizing the dynamic field theory of spatial cognition across real and developmental time scales. *Brain Res.* **1202**, 68–86 (2008)
30. Spencer, J.P., Simmering, V.R., Schutte, A.R., Schöner, G.: What does theoretical neuroscience have to offer the study of behavioral development? Insights from a dynamic field theory of spatial cognition. In: Plumert, J.M., Spencer, J.P. (eds.) *The Emerging Spatial Mind*, pp. 320–361. Oxford University Press, Oxford/New York (2007)
31. Trappenberg, T.P., Dorris, M.C., Munoz, D.P., Klein, R.M.: A model of saccade initiation based on the competitive integration of exogenous and endogenous signals in the superior colliculus. *J. Cogn. Neurosci.* **13**(2), 256–271 (2001)
32. van Gelder, T., Port, R.: It's about time: an overview of the dynamical approach to cognition. In: Port, R., van Gelder, T. (eds.) *Mind as Motion*, pp. 1–43. MIT, Cambridge (1995)
33. Wilimzig, C., Schneider, S., Schöner, G.: The time course of saccadic decision making: dynamic field theory. *Neural Netw.* **19**(8), 1059–1074 (2006)
34. Wilson, H.R., Cowan, J.D.: Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.* **12**(1), 1–24 (1972)
35. Zhang, W., Luck, S.J.: Discrete fixed-resolution representations in visual working memory. *Nature* **453**(7192), 233–235 (2008)

Chapter 13

A Dynamic Neural Field Approach to Natural and Efficient Human-Robot Collaboration

Wolfram Erlhagen and Estela Bicho

Abstract A major challenge in modern robotics is the design of autonomous robots that are able to cooperate with people in their daily tasks in a human-like way. We address the challenge of natural human-robot interactions by using the theoretical framework of Dynamic Neural Fields (DNFs) to develop processing architectures that are based on neuro-cognitive mechanisms supporting human joint action. By explaining the emergence of self-stabilized activity in neuronal populations, Dynamic Field Theory provides a systematic way to endow a robot with crucial cognitive functions such as working memory, prediction and decision making. The DNF architecture for joint action is organized as a large scale network of reciprocally connected neuronal populations that encode in their firing patterns specific motor behaviors, action goals, contextual cues and shared task knowledge. Ultimately, it implements a context-dependent mapping from observed actions of the human onto adequate complementary behaviors that takes into account the inferred goal of the co-actor. We present results of flexible and fluent human-robot cooperation in a task in which the team has to assemble a toy object from its components.

13.1 Introduction

Recent advances in robotics technology make the design of socially interactive robots that work closely with ordinary people in their day-to-day work a realistic goal [20]. Research in such human-centered robotics requires to address a wealth

W. Erlhagen (✉)

Department of Mathematics and Applications, Center for Mathematics, University of Minho,
Braga, Portugal

e-mail: wolfram.erlhagen@math.uminho.pt

E. Bicho

Department of Industrial Electronics, Centre Algoritmi, University of Minho, Braga, Portugal
e-mail: estela.bicho@dei.uminho.pt

of new interdisciplinary topics from cognitive psychology, artificial intelligence and neuroscience that go well beyond traditional mathematical issues of robotics research for industrial applications [38]. As fundamentally social beings, we are experts in joint activity in order to realize a common goal. We therefore have high expectancies about an engaging and pleasant interaction with another agent. Humans achieve their remarkable fluent organization of joint activity in routine tasks, such as preparing the dinner table, by continuously monitoring the partner's actions, and predicting them effortlessly in terms of their outcomes [40]. Based on this prediction, an adequate complementary action can be timely selected among all potential behaviors that the task currently affords. To ensure user acceptance, a socially interactive robot that is supposed to substitute a human in a cooperative task should equally contribute to the coordination and synchronization of behaviors among the co-actors. It is thus crucial to endow the robot with high-level cognitive functions such as action understanding, decision making and memory.

Given the large variety of disciplines involved in the emerging field of human-friendly robotics, it is perhaps not surprising that different design approaches toward more natural human-robot interaction have been proposed. Conceptually, they may be broadly classified in top-down, symbolic views on human-like (social) intelligence and more bottom-up, neurodynamics and embodied notions [30]. The predominant top-down approach is inspired by traditional artificial intelligence (AI) models that address the complex problem of selecting an adequate complementary behavior as a sequence of logical operations performed on discrete symbols. The robotics implementations are thus based on formal logic and formal linguistic systems [32]. Good examples are architectures inspired by the theoretical framework of joint intention theory [1, 8, 28]. This framework provides a rigorous logical treatment of how sub-plans of individual agents committed to a common task can be meshed into joint activity. A defining feature of the symbolic approach is that information processing is set up in stages from perception to cognition to action. A perceptual subsystem first converts sensory information about external events into inner symbols to represent the state of the world. Next, this information is used along with representations of current goals, memories of past events and beliefs about the partner's intention to decide about the course of action. On this planning level, actions are formulated as logical operators with preconditions and effects that change the world in a discrete fashion and instantaneously. The abstract plan is then transformed into motor representations of the robotics system that are finally used to generate arm and hand trajectories in order to realize the plan.

The symbolic, disembodied view on how to decide what to do has provided many impressive examples of intelligent behaviors in artificial agents (for review see [46]). However, it is now widely recognized by the robotics and cognitive science communities that the symbolic framework based on serial stages of processing has notoriously problems to cope with real-time interactions in dynamic environments [26, 30, 32]. In human-robot interaction tasks, the robot has to reason about a world that may change at any instance of time due to actions taken by the user. Even if we consider that the processing in the perceptual and decision modules would allow to continuously update the robot's plan in accordance with the user's intention, the

extra processing step needed to embody the abstract action plan in the autonomous robot would challenge the fluent and seemingly effortless coordination of decisions and actions that characterizes human joint action in familiar tasks.

In order to advance toward a more online view of high-level social cognition, our group at the University of Minho has developed and tested over the last couple of years a neurodynamics approach based on the theoretical framework of Dynamic Neural Fields (DNF) [12]. The DNF model for natural human-robot interaction that we present in this chapter implements known neuro-cognitive processing mechanisms supporting dynamic social interactions in humans and other primates [40]. Converging lines of experimental evidence in behavioral and neuro-cognitive studies suggest that the interaction between sensory, cognitive and motor processes in the brain is much more interactive and integrated as previously thought. For instance, neural correlates of decision making seem to be inconsistent with the notion that a central decision maker completes its operation before activating the motor structures to perform the action plan [25]. Instead, the process of action selection may be best understood as a winner-takes-all competition between multiple neuronal population representations of motor behaviors that the environment currently affords [7]. The advantage of such a dynamic competition process for flexible behavior is obvious. Since the flow of sensory information is continuously used to partially specify several potential actions, the system is prepared to quickly adjust to a changing world. Different neural pathways carrying different sources of information demonstrate the tight coupling between visual and motor systems (for review see [35]). For instance, according to the concept of object affordances [24], the perception of a graspable object immediately activates to some extent the neuronal representations of potential motor interactions with that object. The final decision to execute a certain action, represented by a sufficiently activated subpopulation, may depend on additional contextual cues and the current behavioral goal. Very important for social interactions, an impressive body of experimental evidence from behavioral and neurophysiological studies investigating action and perception in a social context shows that when we observe other's actions corresponding motor representations in our motor system become activated (for a recent review see [36]). In a cooperative joint action context like transferring an object to a partner, this automatic action resonance mechanism has been interpreted as evidence that the likelihood of performing a complementary motor program is increased, that is, the 'receiver' immediately prepares a complementary grasping behavior that ensures a safe and robust object transfer [33]. For more complex joint action settings for which the mapping from observed actions onto adequate complementary behaviors is not as clear, the observer has first to predict the partner's ongoing action in terms of the future effects in the environment. The action resonance mechanism is believed to support also the high-level cognitive functionality of action understanding and goal inference [36]. The key idea here is that the observer internally simulates the outcome of perceived actions using his/her own motor representations that have become associated with representations of action goals during learning and practice. The notion that motor representations are crucially involved in a higher-cognitive function like generating expectations

about the future is clearly inconsistent with serial information processing theories of cognitive behavior.

The DNF model of cooperative joint action is organized as a large scale network of reciprocally connected neuronal populations that encode in their firing patterns specific motor behaviors, action goals, contextual cues and shared task knowledge [5, 6]. Although some level of functional modularity exists in the network, it is important to notice that the formation and maintenance of a behavioral decision is not represented in the discharge pattern of “motor” neurons alone, but is distributed among all currently active populations in the network.

The activity in each local population evolves continuously in time under the influence of external input from connected neuronal pools or the sensory system and recurrent excitatory and inhibitory interactions within the population. Central for the design of cognitive agents, the recurrent interactions support the existence of self-sustained bumps of activation. Persistent population activity allows us for instance to implement a working memory function in the robot to cope with temporally missing sensory information, or to simulate future environmental inputs that may inform the current decision process about a goal-directed behavior [12].

As a specific mathematical formulation of a DNF, we adopt Amari’s model for pattern formation in neural populations since it allows analytical treatment [2]. This is an important advantage when trying to design a complex robot control architecture for real-world experiments.

The chapter is organized as follows: first, we give an overview about the neuro-cognitive foundations of the DNF model and describe its mathematical implementation. We then illustrate the coordination of actions and decisions between human user and robot organized by the network dynamics in a joint action task in which the two teammates have to jointly assemble a toy object from its components.

13.2 Dynamic Neural Field Model of Joint Action

As a working definition, joint action can be regarded as any form of social interaction whereby two or more individuals coordinate their actions in space and time to bring about a change in the environment. Crucial building blocks for successful joint action coordination are the capacities to recognize actions performed by others, and to integrate predicted effects of own and others’ behaviors in the action selection process [40]. What are the neural bases of efficient social interactions? The discovery of the so-called mirror neuron system first in monkey and later in human gives strong support for the hypothesis that observing actions performed by another individual elicit a motor activation in the brain of the observer similar to that which occurs when the observer plans his/her own goal-directed action (for a recent review see [36]). This automatic action resonance mechanism has given rise to the hypothesis that covert motor simulations support action understanding in a social context without the costs that are associated with conscious mental processes or explicit communication.

Mirror neurons in premotor cortex of monkeys (area F5) become active both when the monkey performs a specific motor act like grasping an object and when it observes another individual making a similar action. Importantly, for most mirror neurons the congruency between the observed and the executed motor act is relatively broad. This suggests that their discharge is not related to the fine details of the movements but codes the goal of the observed or executed motor act.

Object manipulation tasks typically involve a series of action phases like reaching, grasping, lifting, holding and placing that are bounded by specific sensory events defining subgoals of the task [18]. Distinct populations of mirror neurons are assumed to represent the functional goals of these successive action phases. Mirror neurons have been also described in areas PFG and PF of the inferior parietal lobe (IPL). These areas are anatomically connected with premotor area F5 and with higher visual areas in the superior temporal sulcus (STS). STS neurons discharge during hand-object interactions similar to those encoded by F5 neurons. The difference seems to be that STS neurons do not discharge during overt movements. STS neurons thus might provide mirror neurons with a visual description of goal-directed motor acts.

The hypothesis that the discharge of neuronal populations in the STS-PFG/PF-F5 circuit plays a key role in action understanding and goal inference has obtained strong support from a series of neurophysiological experiments. It has been shown for instance that grasping mirror neurons are activated also when the critical part of the observed action, the hand-object interaction, is hidden behind a screen and can thus only be inferred from additional contextual information (e.g., the presence of a graspable object behind the occluding surface [45]). In a recent study, Fogassi and colleagues [19] reported that IPL mirror neurons, in addition to recognize the goal of an observed motor act, discriminate identical grasping behaviors according to the final goal of the action sequence in which the motor act is embedded (e.g., grasping for eating versus grasping for placing in a container). They further argued that because the discriminated motor act is part of a specific chain of motor primitives associated with a specific goal representation most likely in prefrontal cortex (PFC), the monkey could predict at the time of the grasping the ultimate goal of the observed action and, thus read the motor intention of the acting individual. Of course, the discrimination of the grasping behavior is only possible because of an additional contextual cue (e.g., the presence of a container in the scene). This suggests that the simulation process in IPL mirror neurons is not exclusively shaped by input from STS but also depends on input from goal and object representations.

Figure 13.1 sketches the multi-layered dynamic field model of joint action consisting of various neural populations that are associated through hand-coded synaptic links (not all are shown to avoid crowding). As a central part, it integrates a previous DNF model of action understanding and goal-directed imitation inspired by the mirror system [15]. Ultimately, the distributed network implements a flexible mapping between observed and executed actions that takes into account the inferred goal of the co-actor, contextual cues and shared task knowledge.

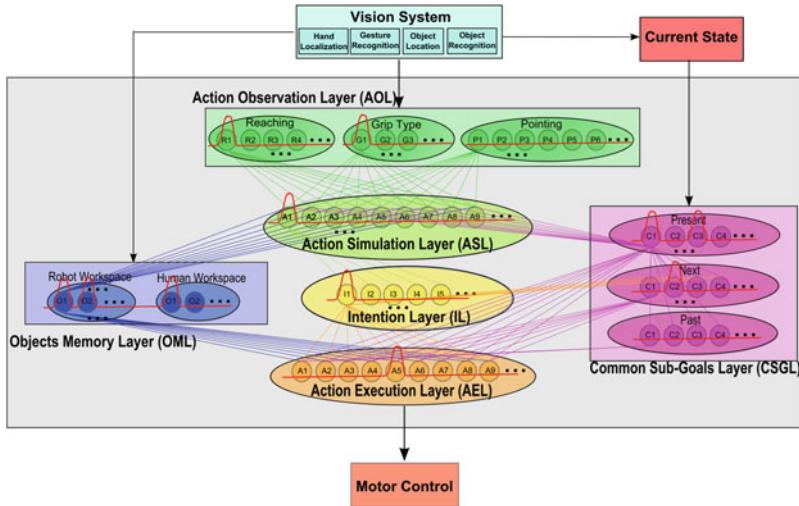


Fig. 13.1 Joint action model consisting of a distributed network of interconnected neural populations. It implements a flexible mapping from observed actions (layer AOL) onto complementary actions (layer AEL) taking into account the inferred action goal of the partner (layer IL), contextual cues (layer OML) and shared task knowledge (layer CSGL). The goal inference capacity is based on motor simulation (layer ASL)

An observed hand movement that is recognized by the vision system as a particular movement primitive (e.g. a whole hand-grasping-from above) is represented by suprathreshold activity of a specific neuronal population in the action observation layer (AOL). Input from AOL to corresponding populations in the action simulation layer (ASL) may activate together with input from the object memory layer (OML) and the common sub-goals layer (CSGL) specific chains of movement primitives that are linked to neuronal representation of the ultimate action goal in the intention layer (IL) [16]. Suprathreshold population activity in IL will drive one or more associated populations in the action execution layer (AEL) that represent possible complementary motor behaviors. Similar to ASL, the motor behaviors are organized in chains of motor primitives like reaching-grasping-placing. There are different ways how to represent the temporal order and the timing of motor sequences in the dynamic field framework [17, 37]. To simplify the present robotics experiments with its emphasis on competitive action selection, we have not modeled these chains as a sequential activation of individual neural populations, but represent the entire motor behavior by a single pool of neurons.

The final decision in AEL depends not only on the input from IL but also on input from OML and CSGL. OML contains neuronal population representations of the various objects in the scene. It is organized in two layers that discriminate whether a specific object is within the user's or within the robot's reachable space. Input from OML automatically pre-activates neural representations of associated motor behaviors in AEL. Specifically for the joint assembly task, possible object-directed

behaviors include the transfer of the object to the co-actor or a direct placement of the object as part of the assembly work. In addition, communicative gestures like for instance pointing to the specific component may be used in joint activity to attract the co-actor's attention [4]. Efficient task performance requires to carry out the steps in the task in the correct order, without repeating or omitting actions in the sequence. This behavioral planning heavily depends on the predicted consequences of intended actions (i.e. a change in the state of the target object [43]). The common subgoals layer CSGL contains neuronal representation of desired end results of individual assembly steps that can be realized by associated motor representations in AEL and that are recognized by the vision system. Neurophysiological evidence suggests that in sequential tasks, distinct subpopulations in PFC represent already achieved subgoals and subgoals that have still to be accomplished [22]. In line with this finding, CSGL contains two connected DNF layers with population representations of past and future events. Input from the vision system about the achievement of a specific subgoal activates the corresponding population in the past layer, which in turn inhibits the corresponding goal representation and simultaneously excites one or more populations in the future layer. They represent in their activity patterns predicted end result of subsequent assembly steps that the current state of the assembly work affords. Important for the fluency of the team behavior, the updating of subgoals in CGSL may not only be triggered by direct input from the vision system but also by input from IL representing the inferred motor intention of the co-actor. This allows the observer to prepare future actions in response to anticipated rather than observed action outcomes [5, 6].

13.3 Model Details

In their seminal work, Wilson and Cowan [48] and Amari [2] introduced Dynamic Neural Fields as rate models of cortical population dynamics that abstract from the biophysical details of neural firing. The architecture of this model class reflects the hypothesis that strong excitatory and inhibitory interactions within local populations that receive synaptic input from multiple connected neuronal pools form a basic mechanism of cortical information processing. As shown in numerous simulation studies, dynamic neural field models are powerful enough to reproduce neural population dynamics observed in neurophysiological experiments (e.g., [14]), and to understand the basic mechanisms underling a large variety of experimental findings on the perceptual and behavioral level (for review see [39]).

For the design of the robot control architecture for natural human-robot interactions, we adopt the model of a single layer of a homogeneous neural network consisting of excitatory and inhibitory neurons proposed by Amari [2]. This model allows for a rigorous analysis of the existence and stability of characteristics solutions such as local excitations or “bumps”. In the following, we give a brief overview about the techniques developed by Amari, and explain the adaptations we have made to cope with the specific needs of the robotics implementations.

The dynamics of each population in the distributed network shown in Fig. 13.1 is governed by the equation:

$$\begin{aligned} \tau \frac{\partial u(x, t)}{\partial t} &= -u(x, t) + S(x, t) \\ &+ \int_{-\infty}^{\infty} w(x - x') f(u(x', t)) dx' - h \end{aligned} \quad (13.1)$$

where $u(x, t)$ is the average activity of neuron x at time t and parameter $\tau > 0$ defines the time scale of the field dynamics. The globally inhibitory input $h > 0$ determines the resting state to which the activity of neuron x relaxes without external input $S(x, t) \geq 0$. The integral term in Eq. 13.1 describes the interactions within the populations which are chosen of lateral-inhibition type:

$$w(x) = A \exp(-x^2/2\sigma^2) - w_{inhib} \quad (13.2)$$

where $A > 0$ and $\sigma > 0$ describe the amplitude and the standard deviation of a Gaussian, respectively. For simplicity, the long-range inhibitory interactions are assumed to be constant, $w_{inhib} > 0$, implementing a competition between subpopulations that are sufficiently separated in space. Note that distinct neural populations encoding entire temporal behaviors like grasping, holding or placing seem to be spatially segregated in the mirror neuron areas [35]. Interpreting the metric of neural interactions in anatomical space like in Amari's original model is thus possible. However, the metric distance might be also defined in an abstract psychological space [41]. In this case, functionally distinct behaviors associated with specific goals would be represented by spatially separate, competing pools of neurons whereas similar motor behaviors associated with the same goal (e.g., grasping with different grip types) would be represented by partially overlapping populations.

Amari assumes for his analysis of pattern formation that the output function $f(u)$, which gives the firing rate of a neuron with input u , is the Heaviside step function, i.e., $f(u) = 0$ for $u \leq 0$ and $f(u) = 1$ otherwise. To model a more gradually increasing impact of the recurrent interactions on the population dynamics we apply a smooth and differentiable output function of sigmoid shape with slope β and threshold u_0 :

$$f(u) = \frac{1}{1 + \exp(-\beta(u - u_0))}. \quad (13.3)$$

It has been shown by Kishimoto and Amari [29] that many of the results concerning the existence and stability of localized activity patterns obtained with a step output function take over to the more general case of the sigmoid.

The model parameters are chosen to guarantee that the population dynamics is bi-stable, that is, the attractor state of a stable "bump" coexists with a stable homogeneous resting state. A sufficiently strong transient input $S(x, t)$ may drive

the neural population beyond threshold, $f(u) > u_0$. The resting state loses stability and a localized activation pattern evolves. In the various layers of the network model, these bumps represent memorized information about object location, the inferred action goal of the co-actor or a decision for a specific complementary behavior. Weaker external input signals from connected populations lead to a subthreshold activation pattern for which the contribution of recurrent interactions is negligible. It is important to note, however, that this preshaping by weak input may nevertheless influence the robot's behavior. Since the level of pre-activation affects the rate at which a suprathreshold activation pattern rises [13], a pre-activated population has a computational advantage over a population at resting level and thus has a higher probability to influence the decision process in AEL.

For the case of a step output function, the conditions for the existence and stability of a single bump of length a in the presence of a stationary external input $S(x)$ can be easily derived following Amari's approach (see Chap. 3 for details). For the robotics experiments we are specifically interested in the existence of localized excitation in response to symmetric, bell-shaped input. Given the definition

$$W(x) = \int_0^x w(x')dx' \quad (13.4)$$

the length a of the bump satisfies in this case

$$S(x_0 + a/2) = h - W(a) \quad (13.5)$$

where x_0 denotes the position of the maximum $S(x)$. If $h > 0$ is chosen such that

$$W_m = \max_{x>0} W(x) > h \quad (13.6)$$

holds, there exist two solutions \hat{a} and a , with $\hat{a} < a$, but only the larger excitation pattern is stable (for details see [2]).

We assume that the time dependent input from a connected population u_j to a target population u_i has a separable form $S_i(x, t) = S(x)g_j(t)$ where $S(x)$ is modeled as a Gaussian function and $g_j(t) = 1$ if $f(u_j) > u_0$ and $g_j(t) = 0$ otherwise. In other words, a stationary input is applied during the period of suprathreshold activity in u_j . Numerical studies show that the evolving localized activation in u_j could have been directly used as input pattern as well. However, assuming a constant input shape allows us to closely follow Amari's analysis. The total input from all connected populations and external sources (e.g., vision system, also modeled as Gaussian signal) to u_i is then given by

$$S_i(x, t) = k \sum_j g_j(t) A_j \exp(-(x - x_i)^2 / 2\sigma^2) \quad (13.7)$$

where $k > 0$ is a scale factor to guarantee that the total external input remains small compared with the recurrent interactions within the local population.

To model different cognitive functions like working memory or decision making in the various layers of the model, we specifically adapt the basic field equation given by Eq. 13.1. To implement in OML, AOL and CSGL a working memory function, it is important that a bump remains after cessation of the transient stimulus that has initially driven its evolution. The condition $W_m > h > 0$ guarantees the existence of a stable bump for $S(x) = 0$ which, however, has a slightly smaller width compared to the bump in the presence of input. We call this solution self-sustained to distinguish it from a suprathreshold activity pattern that becomes self-stabilized only because of the presence of external input. In this case, equation $S(x_0 + a/2) = h - W(a)$ has a solution which represents a stable localized activation but $h > W_m$ holds, that is, the field dynamics is in the mono-stable regime and suprathreshold activity will decay to rest state without external support.

To represent and memorize simultaneously multiple items, a multi-bump solution is required. An interaction kernel with long-range, constant inhibition (Eq. 13.2) may sustain multiple localized activity patterns without external inputs with additional stabilization mechanisms [12, 44]. For simplicity, we have used for the current robotics experiments kernels with limited spatial range to exclude mutual competition between multiple memories. An alternative solution that we are currently exploring for the robotics work is to use coupling functions with multiple zero-crossings, modeling excitatory interactions also at larger distances [17, 31].

The memory is continuously updated in accordance with input from the vision system indicating a change in the external world (e.g., a new location of a specific object). To implement the “forgetting” process, we use a simple first-order dynamics with an appropriate time scale for the (local) adaptation of the inhibitory input h to destabilize an existing bump [3]:

$$\frac{dh}{dt} = -r_{h,min}c_h(h - h_{min}) - r_{h,max}(1 - c_h)(h - h_{max}) \quad (13.8)$$

where $|h_{max}| < W_m$ and $|h_{min}| > W_m$ are the two limit values for h that define the bi-stable and the mono-stable regime, respectively. The rate of change for destabilizing a memory function in case of an existing bump ($c_h = 1$) or restoring in the absence of a bump ($c_h = 0$) is given by the parameters $r_{h,min} > 0$ and $r_{h,max} > 0$.

To meet the real-time constraints of action selection and goal inference in a continuously changing environment, we apply in ASL, AEL and in the CSGL layer representing future subtasks a field dynamics with self-stabilized rather than self-sustained activation patterns. A decision to select a certain motor behavior that takes into account the most likely goal of the co-actor’s current action, is temporally stabilized by sufficient strong support of external and internal evidence, but will automatically lose stability if this evidence changes in favor of a competing behavior.

13.4 Setup of Human-Robot Experiments

To test the dynamic neural field model of joint action in human-robot experiments, we have adopted a joint assembly paradigm in which the team has to construct a toy ‘vehicle’ from components that are initially distributed on a table (Fig. 13.2).

The toy object consists of a round platform with an axle on which two wheels have to be attached and fixed with a nut. Subsequently, four columns that differ in their color have to be plugged into corresponding holes in the platform. The placing of another round object on top of the columns finishes the task. The components were designed to limit the workload for the vision and the motor system of the robot.

It is assumed that each teammate is responsible to assemble one side of the toy. Since the working areas of the human and the robot do not overlap, the spatial distribution of components on the table obliges the team to coordinate and synchronize handing-over sequences. In addition, some assembly steps require that one co-worker helps the other by fixating a part in a certain position. It is further assumed that both teammates know the construction plan and keep track of the subtasks which have been already completed by the team. The prior knowledge about the sequential execution of the assembly work is represented in the connectivity between the two layers of CSGL encoding already achieved and still to be accomplished assembly steps. Since the sequential order of tasks execution is not unique, at each stage of the construction the execution of several subtasks may be simultaneously possible.

The humanoid robot ARoS used in the experiments has been built in our lab. It consists of a stationary torus on which a 7 Degrees of Freedom (DOFs) AMTEC arm (Schunk GmbH) with a 3-fingers dexterous gripper (Barrett Technology Inc.) and a stereo camera head are mounted. A speech synthesizer (Microsoft Speech SDK 5.1) allows the robot to communicate the result of its goal inference and decision making processes to the human user [4].

The information about object class, position and pose is provided by the vision system. The object recognition combines color-based segmentation with template matching derived from earlier learning examples [47]. The same technique is also used for the classification of object-directed, static hand postures such as grasping and communicative gestures such as pointing.

The selection of a specific complementary behavior in AEL has to be translated into a collision-free arm and hand trajectory. As an important constraint for efficient joint action coordination, the robotics motion should be perceived by the user as smooth and goal-directed. To achieve realistic temporal motor behaviors like reaching, grasping and manipulating objects we apply a global planning technique in posture space. It is formalized as a nonlinear optimization problem and allows us to integrate constraints obtained from human reaching and grasping movements such as for instance bell-shaped velocity profiles of the joints (for details see [10]).

Fig. 13.2 Joint action scenario: the human-robot team has to assemble a ‘toy vehicle’ from components that are initially distributed on a table



13.5 Results

In the following we illustrate the coordination of decisions and actions between the human and the robot in the joint assembly task by presenting video snapshots of the interactions and the associated neuronal population representations in the model network. In the examples shown, we focus for simplicity on the initial phase of the construction to explain from the perspective of the robot the impact of action observation on action selection in varying context.¹ As summarized in Table 13.1, there are 9 possible goal-directed sequences and communicative gestures that distinct populations in AEL and ASL represent.

At any time of the human-robot interaction only a few of these action alternatives are simultaneously possible, that is, are supported by input from connected populations. Figure 13.3 illustrates the competition between action alternatives in AEL and the decisions linked to overt behavior of the robot.² It is important to notice, however, that the competition process in ASL and AEL also works for more complex scenarios with a larger set of possible complementary behaviors (e.g., a household scenario [34], full construction of the ‘toy vehicle’ [6]). The number of competing action representations only affects the time it takes to stabilize a suprathreshold activation pattern representing a decision [13].

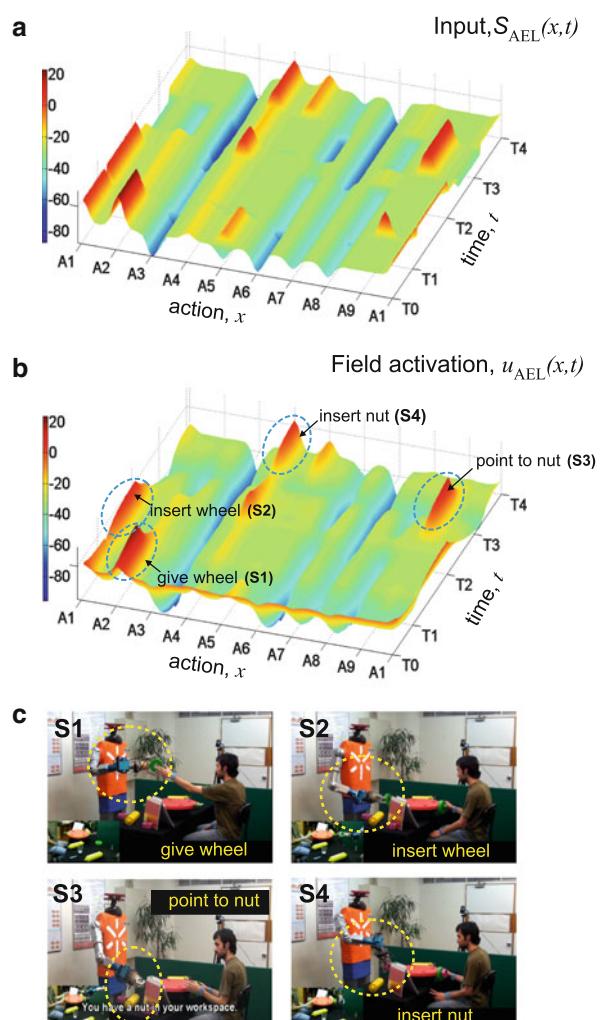
¹But see <http://www.youtube.com/watch?v=A0qemfXnWiE> for a video with the complete construction task.

²Video of the human-robot interactions depicted in Fig. 13.3 can be found in http://dei-s1.dei.uminho.pt/pessoas/estela/Videos/JAST/Video_Fig4_Aros_Human_Toy_Vehicle.mpg.

Table 13.1 Goal-direct sequences and communicative gestures

Action	Sequence of motor primitives	Short description
A_1	reach wheel → grasp → attach	attach wheel
A_2	reach wheel → grasp → handover	give wheel
A_3	reach hand → grasp wheel → attach	receive wheel to attach
A_4	reach nut → grasp → attach	attach nut
A_5	reach nut → grasp → handover	give nut
A_6	reach hand → grasp nut → attach	receive nut to attach
A_7	hold out hand	request piece
A_8	point to wheel	point to wheel
A_9	point to nut	point to nut

Fig. 13.3 Sequence of decisions in AEL and corresponding robot behavior: (a) Temporal evolution of total input to AEL. (b) Temporal evolution of field activity showing the competition process and the sequence of decisions ‘give wheel’, ‘insert wheel’, ‘point to nut’ and ‘insert nut’. (c) The four snapshots illustrate corresponding events of the human-robot interactions



13.5.1 Selection Based on an Anticipatory Model of Action Observation and Shared Task Knowledge

A cornerstone of fluent human social interactions is the ability to predict the outcomes of others' action sequences. It allows individuals to prepare actions in responses to events in the environment that will occur only a considerable time ahead. For a robot that is supposed to assist a human user in a shared task, a goal inference capacity should be used to select an action that best serves the user's future needs. But even if the human co-worker hesitates and does not show any overt behavior, a fluent team performance requires that the robot is able to take initiative and to select an action in accordance with the shared task knowledge.

These cognitive capacities are tested in the experiment depicted in Fig. 13.4 (video snapshots) and Fig. 13.5 (field activities). The experiment starts by placing the platform on the table. The vision input updates the task representation in CSGL and the activity of two populations representing the possible subgoals of attaching the wheels become suprathreshold. Initially, the two wheels are located in the working area of the human while the two nuts are located in the workspace of the robot. As shown in snapshots S1–S2 (Fig. 13.4), the human reaches and grasps a wheel. At the moment of the grasping, ARoS anticipates that the co-actor's motor intention is to mount the wheel on his side. It immediately decides to reach for a nut to hold it out for the human since according to the assembly plan it is the component that he will need next.

The capacity to infer the goal of the user at the time of grasping is possible because how the partner grasps an object conveys information about what he intends do with it. The robot has sequences of motor primitives in its motor repertoire that associate the type of grasping with specific final goals. A grasping from above is used to attach a wheel to the axle whereas using a side grip is the most comfortable and secure way to hand the wheel over to the co-actor. The observation of an above grip (represented in AOL) together with information about the currently active subgoal (attach wheel on the user's side in CSGL) trigger an activation peak in ASL that represents the simulation of the corresponding 'reaching-grasping-inserting' chain (see panel a in Fig. 13.5, time interval T0–T1), which automatically activates the underlying goal, 'insert wheel', in the intention layer (see panel b in Fig. 13.5, time interval T0–T1; see also snapshot S1 in Fig. 13.4). Whenever the activation pattern in IL rises above threshold it initiates a dynamic updating process in the second layer of CSGL, which represents the next possible subgoal(s) for the team (see panel c in Fig. 13.5; see also snapshot S2 in Fig. 13.4, time interval T0–T1). The shared task representation allows the robot to select a complementary action that serves the user's future goal of fixing the wheel with a nut, i.e. the evolving activation pattern in AEL (panel d in Fig. 13.5, time interval T0–T1) reflects the decision to 'give a nut' to the human.

Since the robot has no wheel in its working area, an alternative decision would be to request a wheel from the user to attach it on its side of the platform. The robot's choice to first serve the human is the result of slight differences in input

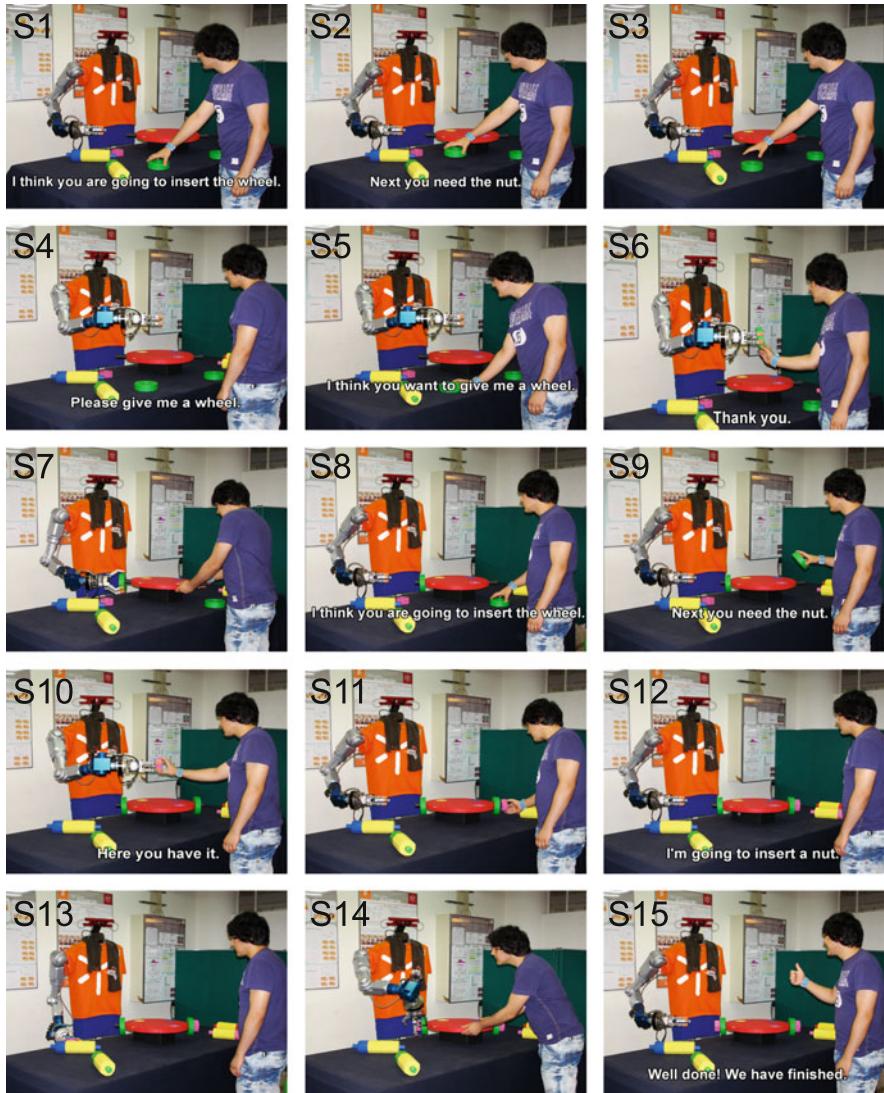


Fig. 13.4 Video snapshots that illustrate the capacity of the robot to infer goals, take initiative and anticipate the user’s future needs

strength from populations in CSGL to associated action representations in AEL. These differences favor the execution of the user’s subtasks over the subtasks that are under the control of the robot.

However, as illustrated in snapshot S3 (Fig. 13.4), in this experiment the human does not attach the wheel. Instead he places the wheel back on the table, then hesitates and does not show any object-directed action. As a consequence, no

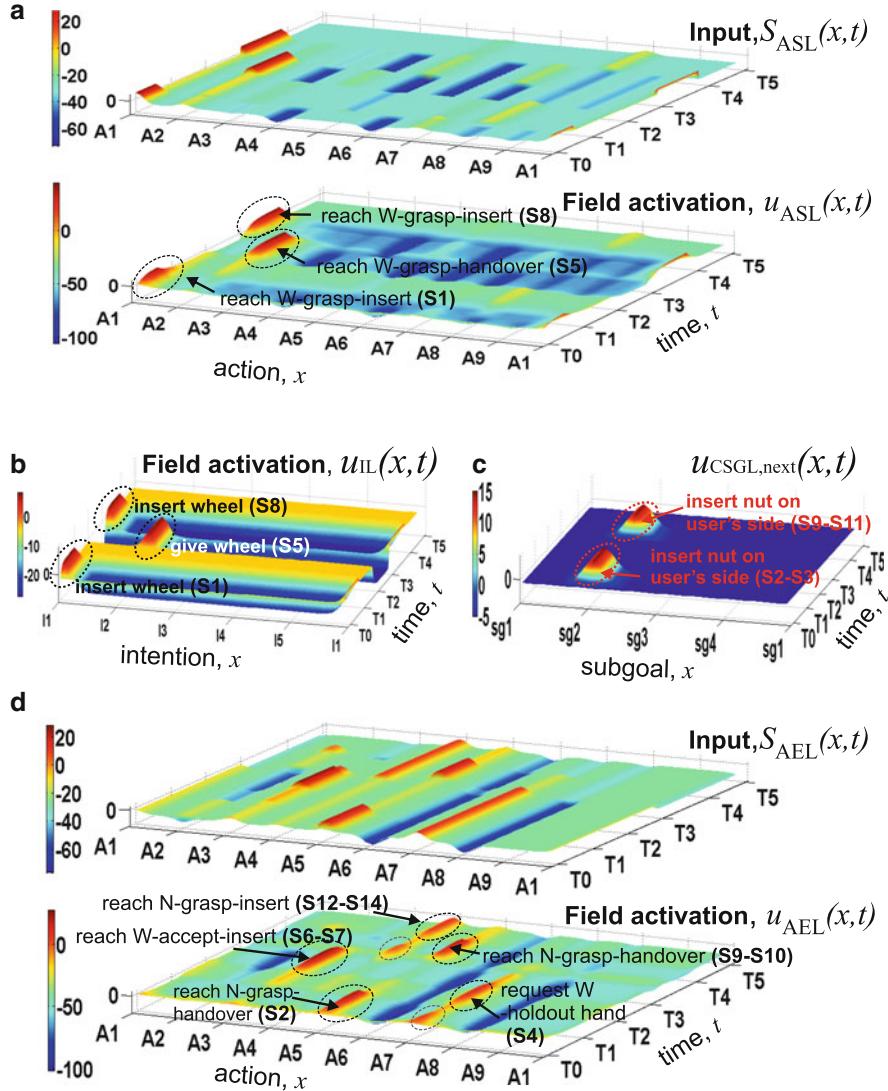


Fig. 13.5 Field activities in layers ASL, IL, CSGL and AEL for the experiment in Fig. 13.4. (a) Temporal evolution of input to ASL (top) and field activity in ASL (bottom). (b) Temporal evolution of field activity in IL. (c) Updating of CSGL layer representing future subgoals based on the inferred motor intention of the user (in IL). (d) Temporal evolution of input to AEL (top) and of activity in AEL (bottom)

suprathreshold activation exists at that time in ASL (see panel a, Fig. 13.5, time interval T1–T2) and activity below threshold in IL indicates that the robot has currently not attributed any action goal to the co-actor (see panel b, Fig. 13.5, time interval T1–T2). The robot now takes initiative and decides to request a wheel to

mount it on its side of the platform (snapshot S4, Fig. 13.4). This change in decision is possible because the population representing the previously selected (but not yet executed) behavior to transfer a nut is not supported anymore by input from IL. On the other hand, information about currently possible subgoals and the location of parts in the two working areas create sufficiently strong input to AEL to trigger a self-stabilized activation of the population representing the ‘request-wheel’ gesture (panel d, Fig. 13.5, time interval T1–T2).

Subsequently, the human grasps the wheel with a side grip (snapshot S5, Fig. 13.4). This information coded in AOL (not shown) together with information about currently active subgoals trigger a bump in ASL that represents the simulation of the corresponding ‘reach-grasp-handover’ chain (panel a, Fig. 13.5, time interval T2–T3), which in turn automatically activates the underlying goal representation ‘give wheel’ in IL (panel b in Fig. 13.5, time interval T2–T3). The evolving suprathreshold activation in AEL (panel d, Fig. 13.5, time interval T2–T3) shows the robot’s decision to receive the wheel and attach it (see also snapshots S6–S7 in Fig. 13.4). When the robot has attached the wheel, the vision input updates the task representations in CSGL and a new bump encoding the subsequent subgoal ‘insert nut on robot’s side’ evolves (panel c, Fig. 13.5, time interval time T2–T3). The second possible subgoal ‘insert wheel on user’s side’ remains active.

Next, the user grasps again a wheel from above, ARoS predicts as before that the user will attach the wheel on his side (panel b in Fig. 13.5, time interval T3–T4) and decides to hand over a nut to fix the wheel (snapshots S8–S9 in Fig. 13.4; see panel d in Fig. 13.5, time interval T3–T4). Note that an alternative decision in AEL could be to ‘grasp and attach a nut on the robot’s side’. The input from OML (not shown) indicating that the two nuts are located in the workspace of the robot together with the input from CSGL support the two action alternatives in AEL. As explained above, the decision process appears to be biased toward serving the human first due to the difference in input strengths from suprathreshold population activity in CSGL. As can be seen in the snapshots S9–S11 (Fig. 13.4), the user attaches the wheel, and subsequently grasps the nut from the robot’s hand to plug it on the axle. As the vision system detects the change in the target object, the representations of already achieved subgoals in the memory layer of CSGL are updated accordingly and the subgoal ‘insert nut on robot’s side’ becomes active (not shown). As a consequence, a bump in AEL evolves that represents the decision of the robot to grasp and attach a nut on its side of the platform (see panel d in Fig. 13.5, time interval T4–T5). The overt robot behavior is depicted in snapshots S12–S14 (Fig. 13.4).

13.5.2 Understanding Partially Occluded Actions

In the previous example, we have seen that the robot could infer through motor simulation the co-actor’s motor intention from the way the object is grasped. But what happens when the robot cannot directly observe the hand-object interaction? In natural environments with multiple objects and occluding surfaces this is a



Fig. 13.6 Snapshots of a video showing action understanding of partially occluded actions. Snapshot S1 shows the view of the vision system of the robot

common scenario. The capacity to discern the user’s motor intention and to select an appropriate complementary behavior should of course not be disrupted by missing information about the grip type used. The firing of mirror neurons in similar occluder paradigms suggests that working memory about objects in the scene and shared task information about what the user should do in a specific situation may sustain the motor simulation process. This is illustrated in the following interaction scenario in which only the reaching part of the user’s action sequence can be observed (see Fig. 13.6).

In this experiment, one wheel and the two nuts are located within the working area of the robot while the second wheel is located in the user’s workspace. Initially all objects are visible for the robot and their locations can thus be memorized in OML. Then a box is introduced into the scene. The robot sees the user’s hand disappearing behind the occluding surface but remembers that there is a wheel behind it. Figure 13.7³ illustrates the goal inference mechanism in this situation.

The corresponding population in AOL (not shown) codes only the reaching behavior. The currently possible subgoals represented in CSGL are ‘insert wheel

³For the video see <http://www.youtube.com/watch?v=7t5DLgH4DeQ>.

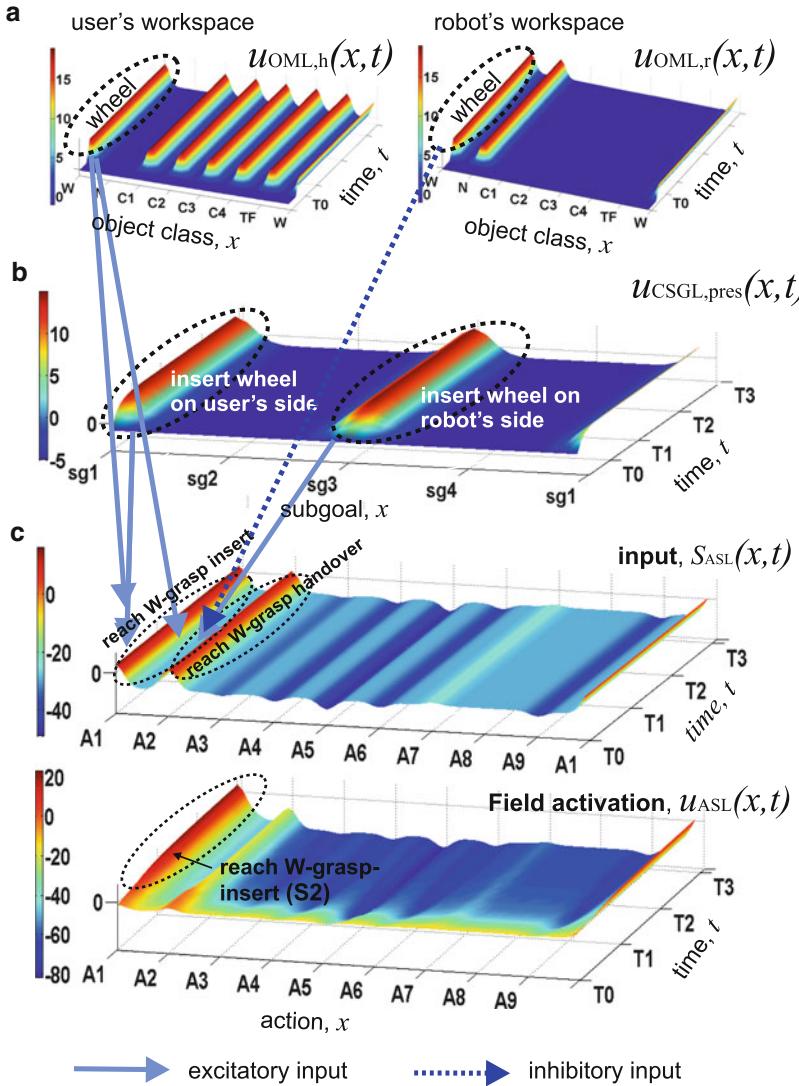


Fig. 13.7 Field activities for the experiment in Fig. 13.6. (a) Temporal evolution of activity in OML. (b) Temporal evolution of CSGL activity representing currently possible subgoals. (c) Temporal evolution of input to ASL (top) and of activity in ASL (bottom)

on user's side' and 'insert wheel on robot's side' (panel b in Fig. 13.7). The inputs from AOL and CSGL to ASL thus pre-activate the representations of two competing action chains associated with two possible motor intentions. The additional input necessary for goal inference comes from the information about the memorized location of the wheels in the two workspaces represented in OML (see panel a in Fig. 13.7). These inputs trigger the evolution of a self-stabilized activation peak in

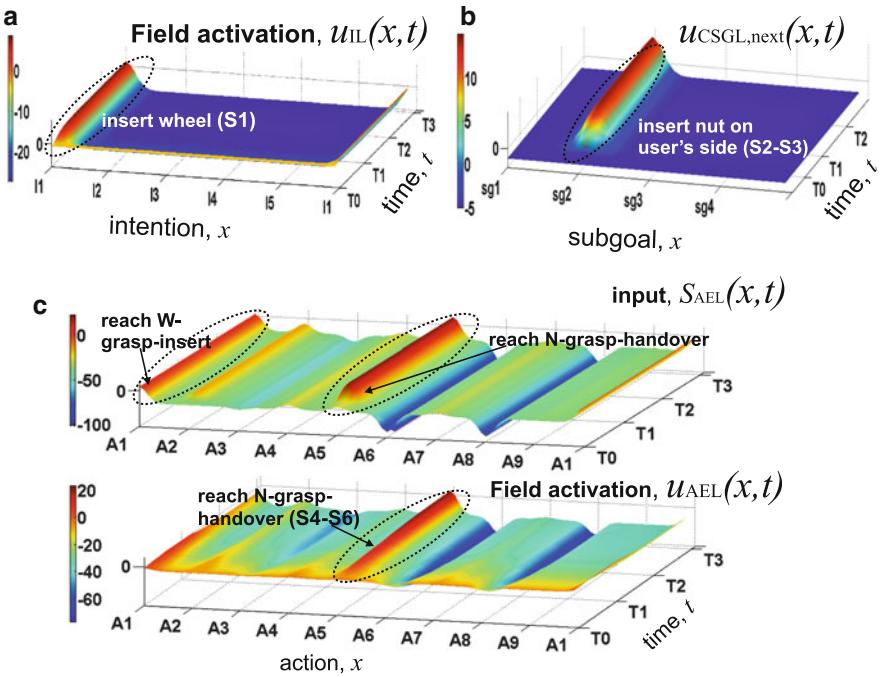


Fig. 13.8 Field activities in IL, CSGL and AEL for the experiment in Fig. 13.6. (a) Temporal evolution of activity in IL. (b) Updating of future subgoals for the user in CSGL based on input from IL. (c) Temporal evolution of input to AEL (top) and of activity in AEL (bottom)

ASL representing the action sequence ‘reach wheel-grasp-insert’ (see panel c in Fig. 13.7; see also snapshot S2 in Fig. 13.6). This suprathreshold activation in turn induces the evolution of a bump in IL representing the inferred goal of the human to insert the wheel (see panel a in Fig. 13.8). Input from IL triggers a dynamic updating process in the second layer of the CSGL, representing the next possible subgoal(s) for the user (see panel b in Fig. 13.8). This allows the robot, as explained in the previous example, to select a complementary action that serves the user’s future needs. As can be seen when comparing the pattern of localized activation that evolves in AEL, the robot decides to serve the human by grasping a nut for handing it over (see panel c in Fig. 13.8 and snapshots S3–S5 in Fig. 13.6).

Note that the simplification for the current robotics work to represent an entire action sequence like reaching-grasping-attaching in a single population does not affect the mechanisms supporting the simulation of partially occlude actions in ASL. A chain of coupled populations of mirror neurons representing individual motor acts [19] may become sequentially activated above threshold by assuming that all individual population of the chain are pre-activated by input from OML and CSGL, and the initial “reaching” population gets additional input from the corresponding neuronal pool in the action observation layer [16].

13.6 Discussion

This work showed that Dynamic Neural Fields provide a powerful theoretical framework for designing autonomous robots able to naturally interact with humans in challenging real-word environments. Flexible and intelligent robot behavior in a social context cannot be purely explained by a stimulus-reaction paradigm in which the system merely maps in a pre-determined manner current environmental inputs onto overt behavior. Dynamic neural fields explain the emergence of persistent neural activation patterns that allow a cognitive agent to initiate and organize behavior informed by past sensory experience, anticipated future environmental inputs and distal behavioral goals. The DNF architecture for joint action reflects the notion that cognitive representations, that is, all items of memory and knowledge, consist of distributed, interactive, and overlapping networks of cortical populations ([21]). Network neurons showing suprathreshold activity are participating in the selection of actions and their associated consequences. Since the decision-making normally involves multiple, distributed representations of potential actions that compete for expression in overt performance, the robot's goal-directed behavior is continuously updated for the current environmental context. Important for decision making in a collaborative setting, inferring others' goals from their behavior is realized by internal motor simulation based on the activation of the same joint representations of actions and their environmental effects ("mirror mechanism", [36]). Through this automatic motor resonance process, the observer becomes aligned with the co-actor in terms of actions and goals. This alignment allows the robot to adjust its behavior without explicit communication to those of the human co-actor in space and time (for an integration of verbal communication in the DNF architecture see [4]).

The implementation of aspects of real-time social cognition in a robot based on continuously changing patterns of neuronal activity in a distributed, interactive network strongly contrasts with traditional AI approaches. They realize the underlying cognitive processes as the manipulation of discrete symbols that are qualitatively distinct and entirely separated from sensory and motor information. We do not deny that the sequence of decisions shown in our robotics experiments could be implemented by symbolic planning as well. In fact, similar joint assembly tasks have been used in the past to test AI-style control architectures for human-robot interactions [1, 28, 42]. Typically, these architectures include a dedicated module that organize the high-level task of intention coordination using rule-based logic. However, the additional planning step which is needed to link the representation of every high-level decision to the level of action preparation for the robot's actuators greatly reduces the efficiency of those representations. This makes it hard or even impossible to achieve the impressive flexibility and fluency of human team performance.

The functional neural field architecture shares many features with neural network models that also take the known connectivity and functional role of different brain areas into account to explain cognitive behavior. For instance, Eliasmith and

colleagues [11] have recently tested a large-scale network of populations of spiking neurons in various tasks ranging from pattern recognition and completion to basic semantic reasoning. The network is presented with sequences of simple visual characters and controls a simulated robot arm drawing these characters. Like in the DNF model of joint action, task and context information is able to autonomously change the information flow between subsystems of the brain-inspired network architecture to generate goal-directed behavior. However, the central role of stable attractor states of the population dynamics for higher level cognition (e.g., internal simulation, decision making, working memory) postulated and analyzed by the dynamic neural field approach (see also the discussion in Chap. 12) is much less emphasized. Different to our current robotics implementations, the large-scale neural network performs the compression operations from higher-dimensional input spaces to lower-dimensional functional spaces that we assume for the DNF representations. We are currently testing learning algorithms that may explain for instance how neural populations in layers AOL and IL representing specific grasping categories develop from observing the trajectories of different grasping behaviors.

In the experiments reported here, the robot-human team executed the individual assembly steps without errors and in the correct temporal order. It is important to keep in mind, however, that decisions based on noisy or incomplete sensory information and anticipated environmental inputs may fail. It is thus no surprise that execution and prediction errors occur with some probability in complex real-world scenarios such as the joint assembly task. To work efficiently as a team, it is important that these errors are detected and compensated by one or both team members before success is compromised. Neurophysiological and behavioral findings suggest that similar neural mechanisms are involved in monitoring one's own and other's task performance [40]. We have described in detail elsewhere how the basic DNF model of joint action coordination can be extended to include also an action monitoring function [6]. The key idea is that specific populations integrate activity from connected neural pools or external sensory signals that carry the conflicting information. For instance, the user might want to transfer a nut to the robot but a nut has been already attached at the robot's construction side. To detect the conflict between the inferred intention of the user and the state of the construction it is sufficient to postulate that input from IL and CSGL may drive the target population beyond threshold. This suprathreshold activity may then produce (inhibitory) biasing effects for the competition between action representations in AEL. In the example, the prepotent complementary behavior of receiving the nut has to be suppressed to favor a correct response like a communicative pointing at the attached object. As integral part of the distributed network, the action monitoring thus provides just another input to the dynamic action selection process.

The applications in the domain of cognitive robotics provide new challenges for the theoretical analysis of dynamic neural fields. Most current mathematical studies are exclusively concerned with the existence and stability of characteristic patterns like bumps or traveling waves [9]. They do not address the spatio-temporal properties that external inputs must satisfy to generate those patterns when applied to a field at rest or in a pre-activated state. For instance, multi-bump solutions that

we and others apply as a memory model for multiple items or sequential events [17] are known to exist when a coupling function with oscillatory decay is used [31]. From an application point of view, analyzing the spatial properties of the inputs (e.g., width, relative distance etc.) that may generate multi-bump solutions when they are presented simultaneously or in sequential order is of highest importance (Ferreira, Erlhagen and Bicho, in preparation).

The present robotics implementations with hand-coded synaptic links between connected populations are based on the seminal analytical studies of Amari and co-workers on the formation of patterns with stationary localized stimuli. For the robotics domain, it would be highly desirable to combine the field dynamics with a learning dynamics that would allow us to establish the inter-field connections in the distributed network during training and practice. According to the principle first enunciated by Hebb [27], memory is formed by associative synaptic modulations of connections between neuronal assemblies simultaneously excited. Important for cognitive control, persistent population activity allows the learning system to establish associations between transient events separated in time. In previous simulation studies, we have shown for instance that a rate-based Hebbian learning rule (for review of mathematical formulations see [23]) can be applied to establish the goal-directed mappings for action simulation in the mirror circuit [15, 16]. A more rigorous understanding of the field dynamics with the weighted, self-stabilized activity from connected populations as non-stationary input would be an important contribution for the design of an autonomous learning system.

Dynamic approaches to robotics and cognition have been often criticized to address mainly lower-level cognitive phenomena like sensory-motor coordination, path planning or perception and not the high-level cognitive capacities which are characteristics of human beings [46]. Being able to synthesize in an embodied artificial agent the cognitive demands of real-time cooperative interactions with a human co-actor shows that dynamic neural field theory provides a promising research program for bridging this gap.

Acknowledgements The present research was conducted in the context of the fp6-IST2 EU-IP Project JAST (proj. nr. 003747) and partly financed by fp7-people-2011-itn NETT (proj. nr. 289146) and FCT grants POCI/V.5/A0119/2005 and CONC-REEQ/17/2001. We would like to thank Luis Louro, Emanuel Sousa, Flora Ferreira, Eliana Costa e Silva, Rui Silva and Toni Machado for their assistance during the robotic experiments.

References

1. Alami, R., Clodic, A., Montreuil, V., Sisbot, E.A., Chatila, R.: Task planning for human-robot interaction. In: Proceedings of the 2005 Joint Conference on Smart Objects and Ambient Intelligence, Genoble. ACM International Conference Proceeding Series, vol. 121, pp. 81–85 (2005)
2. Amari, S.: Dynamics of pattern formation in lateral-inhibitory type neural fields. *Biol. Cybern.* **27**, 77–87 (1977)

3. Bicho, E., Mallet, P., Schöner, G.: Target representation on an autonomous vehicle with low-level sensors. *Int. J. Robot. Res.* **19**, 424–447 (2000)
4. Bicho, E., Louro, L., Erlhagen, W.: Integrating verbal and nonverbal communication in a dynamic neural field architecture for human-robot interaction. *Front. Neurorobotics* (2010). doi:10.3389/fnbot.2010.0005
5. Bicho, E., Erlhagen, W., Louro, L., Costa e Silva, E.: Neuro-cognitive mechanisms of decision making in joint action: a human-robot interaction study. *Hum. Mov. Sci.* **30**, 846–868 (2011a)
6. Bicho, E., Erlhagen, W., Louro, L., Costa e Silva, E., Silva, R., Hipolito, N.: A dynamic field approach to goal inference, error detection and anticipatory action selection in human-robot collaboration. In: Dautenhahn, K., Saunders, J. (eds.) *New Frontiers in Human-Robot Interaction*, pp. 135–164. John Benjamins, Amsterdam (2011b)
7. Cisek, P.: Cortical mechanisms of action selection: the affordance competition hypothesis. *Philos. Trans. R. Soc. B* **362**, 1585–1599 (2007)
8. Cohen, P., Levesque, H.J.: Intention is choice with commitment. *Artif. Intell.* **42**, 213–261 (1990)
9. Coombes, S.: Waves, bumps, and patterns in neural field theories. *Biol. Cybern.* **93**, 91–108 (2005)
10. Costa e Silva, E., Costa, F., Bicho, E., Erlhagen, W.: Nonlinear optimization for human-like movements of a high degree of freedom robotics arm-hand system. In: Murante, B. (ed.) *Computational Science and Its Applications – ICCSA 2011*. Lecture Notes in Computer Science, vol. 6794, Part III, pp. 327–342. Springer, Berlin/Heidelberg (2011)
11. Eliasmith, C., Stewart, T.C., Choo, X., Bekolay, T., DeWolf, T., Tang, Y., Rasmussen, D.: A large scale model of the functioning brain. *Science* **338**, 1202–1205 (2012)
12. Erlhagen, W., Bicho, E.: The dynamic neural field approach to cognitive robotics. *J. Neural Eng.* **3**, R36–R54 (2006)
13. Erlhagen, W., Schöner, G.: Dynamic field theory of movement preparation. *Psychol. Rev.* **109**, 545–572 (2002)
14. Erlhagen, W., Bastian, A., Jancke, D., Riehle, A., Schöner, G.: The distribution of neuronal population activation as a tool to study interaction and integration in cortical representations. *J. Neurosci. Methods* **94**, 53–66 (1999)
15. Erlhagen, W., Mukovskiy, A., Bicho, E.: A dynamic model for action understanding and goal-directed imitation. *Brain Res.* **1083**, 174–188 (2006)
16. Erlhagen, W., Mukovskiy, A., Chersi, F., Bicho, E.: On the development of intention understanding for joint action tasks. In: 6th IEEE International Conference on Development and Learning, Imperial College London, pp. 140–145 (2007)
17. Ferreira, F., Erlhagen, W., Bicho, E.: A dynamic field model of ordinal and timing properties of sequential events. In: Honkela, T., Duch, W., Girolami, M., Kaski, S. (eds.) *Artificial Neural Networks and Machine Learning – ICANN 2011*. Lecture Notes in Computer Science vol. 6792, Part II, pp. 325–332. Springer, Berlin/Heidelberg (2011)
18. Flanagan, J.R., Bowman, M.C., Johansson, R.S.: Control strategies in object manipulation tasks. *Curr. Opin. Neurobiol.* **16**, 650–659 (2006)
19. Fogassi, L., Ferrari, P.F., Gesierich, B., Rozzi, S., Chersi, F., Rizzolatti, G.: Parietal lobe: from action organization to intention understanding. *Science* **308**, 662–667 (2005)
20. Fong, T., Nourbakhsh, I., Dautenhahn, K.: A survey of socially interactive robots. *Robot. Auton. Syst.* **42**, 143–166 (2003)
21. Fuster, J.M.: A cognit: a network model of cortical representation. *Int. J. Psychophysiol.* **60**, 125–132 (2006)
22. Genovesio, A., Brasted, P.J., Wise, P.: Representation of future and previous spatial goals by separate neural populations in prefrontal cortex. *J. Neurosci.* **26**(27), 7305–7316 (2006)
23. Gerstner, W., Kistler, W.M.: Mathematical formulations of Hebbian learning. *Biol. Cybern.* **87**, 404–415 (2002)
24. Gibson, J.J.: *The Ecological Approach to Visual Perception*. Houghton Mifflin, Boston (1979)
25. Gold, J.I., Shadlen, M.: The neural basis of decision making. *Ann. Rev. Neurosci.* **30**, 535–574 (2007)

26. Haazebroek, P., van Dantzig, A., Hommel, B.: A computational model of perception and action for cognitive robots. *Cognit. Process* **12**, 355–365 (2011)
27. Hebb, D.O.: *The Organization of Behavior*. Wiley, New York (1949)
28. Hoffman, G., Breazeal, C.: Cost-based anticipatory action selection for human-robot fluency. *IEEE Trans. Robot.* **23**, 952–961 (2007)
29. Kishimoto, K., Amari, S.: Existence and stability of local excitations in homogeneous neural fields. *J. Math. Biol.* **7**, 303–318 (1979)
30. Kozma, R.: Intentional systems: review of neurodynamics, modelling, and robotics implementations. *Phys. Life Rev.* **5**, 1–21 (2008)
31. Laing, C.R., Troy, W.C., Gutkin, B., Ermentrout, G.B.: Multiple bumps in a neuronal model of working memory. *SIAM J. Appl. Math.* **63**, 62–97 (2002)
32. Levesque, H., Lakemeyer, G.: Cognitive robotics. In: van Harmelen, F., Lifschitz, V., Porter, B. (eds.) *Handbook of Knowledge Representation*, pp. 869–886. Elsevier B. V., Amsterdam (2008)
33. Newman-Norlund, R.D., van Schie, H.T., van Zuijlen, A.M.J., Bekkering, H.: The mirror neuron system is more active during complementary compared with imitative action. *Nat. Neurosci.* **10**, 817–818 (2007)
34. Pinheiro, M., Bicho, E., Erlhagen, W.: A dynamic neural field architecture for a pro-active assistant robot. In: Proceedings of the 3rd IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics (IEEE BioRob 2010), Tokyo, pp. 777–784 (2010)
35. Rizzolatti, G., Luppino, G.: The cortical motor system. *Neuron* **31**, 889–901 (2001)
36. Rizzolatti, G., Sinigaglia, C.: The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nat. Rev. Neurosci.* **11**, 264–274 (2010)
37. Sandamirskaya, Y., Schöner, G.: An embodied account for serial order: how instabilities drive sequence generation. *Neural Netw.* **23**, 1164–1179 (2010)
38. Schaal, S.: The new robotics: towards human-centered machines. *HFSP J.* **1**, 115–126 (2007)
39. Schöner, G.: Dynamical systems approaches to cognition. In: Sun, R. (ed.) *The Cambridge Handbook of Computational Psychology*, pp. 101–125. Cambridge University Press, Cambridge (2008)
40. Sebanz, N., Bekkering, H., Knoblich, G.: Joint action: bodies and minds moving together. *Trends Cognit. Sci.* **10**, 70–76 (2006)
41. Shepard, R.N.: Toward a universal law of generalization for psychological science. *Science* **237**, 1317–1323 (1997)
42. Steil, J.J., Röthling, F., Haschke, R., Ritter, H.: Situated robot learning for multi-modal instruction and imitation of grasping. *Robot. Auton. Syst.* **47**, 129–141 (2004)
43. Tanji, J., Shima, K., Mushiake, H.: Concept-based behavioural planning and the lateral prefrontal cortex. *Trends Cognit. Sci.* **11**, 528–534 (2007)
44. Trappenburg, T., Standage, D.I.: Multi-packet regions in stabilized continuous attractor networks. *Neurocomputing* **65–66**, 617–625 (2005)
45. Umiltà, M.A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., Rizzolatti, G.: I know what you are doing: a neurophysiological study. *Neuron* **31**, 155–165 (2001)
46. Vernon, D., Metta, G., Sandini, G.: A survey of artificial cognitive systems: implications for the autonomous development of mental capabilities in computational agents. *IEEE Trans. Evol. Comput.* **2**, 151–181 (2007)
47. Westphal, G., von der Malsburg, C., Würtz, R.P.: Feature-driven emergence of model graphs for object recognition and categorization. In: Bunke, H., Kandel, A., Last, M. (eds.) *Applied Pattern Recognition. Studies in Computational Intelligence*, vol. 91, pp. 155–199. Springer, Berlin (2008)
48. Wilson, H.R., Cowan, J.D.: A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* **13**, 55–80 (1973)

Chapter 14

Neural Field Modelling of the Electroencephalogram: Physiological Insights and Practical Applications

David T. J. Liley

Abstract The aim of this chapter is to outline a mean field approach to modelling brain activity that has been particularly successful in articulating the genesis of rhythmic electroencephalographic activity in the mammalian brain. In addition to being able to provide a physiologically consistent explanation for the genesis of the alpha rhythm, as well as expressing an array of complex dynamical phenomena that may be of relevance to understanding cognition, the model is also capable of accounting for many of the macroscopic electroencephalographic effects associated with anaesthetic action, a feature often missing in similar formulations. This chapter will then conclude with an example of how the physiological insights afforded by this mean field modelling approach can be translated into improved methods for the clinical monitoring of depth of anaesthesia.

14.1 Introduction

In recent years there has been a resurgence of interest in utilising the electroencephalogram (EEG) to understand brain function. While it was the first functional measure of brain function [1, 32], unlike the subsequently developed blood oxygen level dependent functional magnetic resonance (fMRI) and radionuclide imaging techniques, it has generally been viewed as too coarse in its spatial field of view to reveal anything meaningful about the inner workings of the brain. However with the limitations of fMRI becoming all too apparent and recent advances in our understanding of the anatomical and physiological organization of cortex challenging simplistic views of cortex being just an axo-synaptically coupled network of neurons, EEG, together with its electromagnetic counterpart the

D.T.J. Liley (✉)

Brain and Psychological Sciences Research Centre, Swinburne University of Technology, P.O. Box 218, Hawthorn VIC 3122, Australia
e-mail: dliley@swin.edu.au

magnetoencephalogram (MEG), is re-emerging not only as an important functional measure of brain activity, but also as the basis about which physiologically meaningful mesoscopic formulations of cortical dynamics can be formulated [46].

The EEG is a sensitive measure of behavioural and cognitive state [53]. Spontaneous EEG reveals characteristic and systematic changes during sleep and anaesthesia [34], whereas time locked and/or averaged EEG, has been shown to be a sensitive indicator of cognitive performance and function [63]. In disease it can exhibit features of singular diagnostic importance – from spike and wave activity characteristic of epilepsy to the wicket rhythms pathognomonic for the transmissible spongiform encephalopathy known as Creutzfeldt-Jakob disease [55]. Nevertheless despite our detailed empirical knowledge regarding the patterns and features of EEG our understanding of the physiological genesis of such patterns is comparatively meagre.

While the biophysical origins of the EEG are relatively well established [57] the mechanisms responsible for its dynamical genesis, despite decades of investigation, remain uncertain. For example it was previously believed that the alpha rhythm, a characteristic waxing and waning oscillation of between 8 and 13 Hz, was restricted to the occipital lobe and was due to cortical tissue being driven by oscillatory activity arising from the thalamus [5] or to the existence of distributed subpopulations of pyramidal neurons having some form of intrinsic rhythmicity [47]. However we now believe that dynamical activity in the EEG *emerges* from a panoply of interactions between neuronal and non-neuronal cell populations in cortex [48]. How then do we theoretically instantiate such a view so that we can use it to explain existing electroencephalographic phenomena and make the predictions necessary for its ontological justification?

Two broad theoretical approaches declare themselves as frameworks for understanding the genesis of the EEG. The most obvious is to assume that cortex is a network of neurons and model the individual neurons and their interactions. Apart from the obvious problem of dealing with the computational tractability of simulating the hundreds of thousands of neurons, and their connections, that underly a typical scalp EEG electrode, is the issue of how much physiological detail to include and how to meaningfully parameterise it. We now know that the functional structure of cortex extends well beyond neurons and their axosynaptic interactions. Glial cells, originally thought to only provide structural and biochemical support to the neuronal parenchyma, have been shown to regulate neuronal activity through a “tripartite” synapse – a complex involving an astrocyte and the pre- and post-synaptic terminals of a pair of neurons [31]. Add to this the suggestions of new modes of neuronal interaction (e.g. the axo-myelinic synapse) [73], potential ephaptic (local field) [74] and diffusive (gap junction) neuronal coupling [69], volume (extrasynaptic) neurotransmission [23] and non-synaptic plasticity [52], let alone the known complexities of single neuronal function, then not only does a network motivated approach to understanding the EEG seem daunting, it comes with considerable uncertainties as to how much detail should be included.

A preferable approach then to articulating the genesis of the EEG will be one that (i) has spatiotemporal scales commensurate with EEG and ECoG (ii) is able to deal with the uncertainties of structural, and therefore functional, composition

of cortical tissue and (iii) handles the sparseness of neuronal firing [8] and the unreliability of neuronal interconnectivity that arise as a consequence of low in vivo synaptic release probabilities [12]. The framework of mean field modelling, in which interactions between individual elements are replaced by *effective* averages, has emerged as one powerful way to address these requirements while at the same time remaining physiologically and anatomically pertinent [22, 46].

Commencing with the pioneering works of Beurle [10], Freeman [26], Nunez [56], Wilson and Cowan [77, 78], and Amari [4], mean field models of the EEG have evolved from being relatively abstract biomathematical formulations to frameworks that will be central for the analysis, organization and integration of large volumes of high dimensional functional imaging data [21]. The aim of this chapter is to describe one mean field modelling approach that has been developed to account for electrorhythmogenesis of the mammalian EEG and to illustrate its relevance to understanding the modulation of cortical activity during health and disease.

14.2 A Mean Field Model of Electrocortical Rhythmogenesis

Despite the avowed advantages of a mean field approach over a network approach in understanding the genesis of rhythmic activity in the EEG, almost all mean field approaches take as their starting point the physiological and anatomical properties of axo-synaptically coupled networks of neurons. Commencing with the work of Beurle [10], in which cortex was modelled as a continuous network of spatially uniform, fixed firing threshold, excitatory neurons devoid of any membrane or synaptic dynamics, mean field models have evolved to include many of the most significant anatomical and physiological features immanent to cortical tissue. Of these models one of the more successful in generating dynamics consistent with that of human EEG is that of Liley et al. [11, 40–42]. This model differs from other well known mean-field formulations in (i) explicitly separating the synaptic kinetics of cortical excitatory and inhibitory neuronal activity (cf. [66]) (ii) not needing to explicitly model cortico-thalamic feedback in order to generate cortical rhythms (cf. [66]) (iii) separating out intracortical (short-range) and cortico-cortical (long range) fibre connectivity (cf. [37]) (iv) including the full panoply of local feedforward and feedback excitatory and inhibitory coupling (cf. [35]) and (vi) incorporating synaptic reversal potentials such that a conductance-based mean neuron (see below) is defined (cf. [35, 66, 76]).

In essence the model of Liley et al. [11, 40–42] is constructed at the scale of the cortical macrocolumn – an approximately barrel shaped region extending through the entire thickness of the cortical sheet that has a lateral extent within the cortical sheet of the order of the characteristic scale of pyramidal neuron recurrent collaterals. Within this column, extending across all cortical layers, are distributed populations of excitatory and inhibitory neurons interacting with each other by all possible feedforward and feedback axo-dendritic connections. Macrocolumns then interact with each other by the axons of the excitatory pyramidal neurons

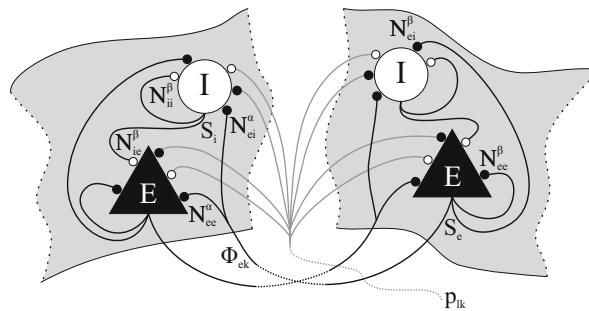


Fig. 14.1 Model topology (Figure reproduced with permission from [25])

that pierce through the bottom layer of cortex to form the long range cortico-cortical conduction system. The topological organization of this model is well known and is depicted in Fig. 14.1 in which the interactions within and between two cortical macrocolumns are shown. In this model “equations-of-motion” for the soma membrane potentials of excitatory and inhibitory neurons, averaged over the spatial extent of the macrocolumn, are defined. Cortical activity is then described by the spatiotemporal evolution of these mean excitatory, h_e , and inhibitory, h_i , membrane potentials. The connection with electrophysiological experiment is through h_e , which is assumed to be linearly related to the EEG. Excitatory and inhibitory neurons are modelled as a single passive resistor-capacitor circuit in which all synaptically induced postsynaptic currents(I_{lk}) flow. On this basis the following *conductance-based* mean neuron can be defined:

$$\tau_k \partial_t h_k = h_k^r - h_k(r, t) + \sum_{l=e,i} \frac{h_{lk}^{eq} - h_k(r, t)}{|h_{lk}^{eq} - h_k|} I_{lk}(r, t) , \quad (14.1)$$

where $r \in \mathbb{R}^2$ is position on the cortical sheet and double subscripts represent first source and then target.¹ Postsynaptic “currents”² (I_{lk}) are weighted by the ionic driving forces, which are defined to be unity at the resting membrane potential h_k^r such that a unitary postsynaptic potential can be simply parameterised. The remaining parameters are defined in Table 14.1.

The dynamics of the postsynaptic “currents” (I_{lk}) are described by a critically damped oscillator driven by the mean rate (i.e. the *mean field*) of incoming excitatory or inhibitory axonal pulses, A_{lk} :

$$(\partial_t + \gamma_{lk})^2 I_{lk}(r, t) = e \gamma_{lk} \Gamma_{lk} \cdot A_{lk}(r, t) \quad (14.2)$$

¹Where in contrast to other authors we have adopted a “anatomical” index ordering.

²We are being quite sloppy with our terminology here as these “currents” are more correctly identified as being conductances but have units of volts as a consequence of being weighted by ionic driving forces normalised to be unity at rest.

Table 14.1 Liley model parameters (Table reproduced with permission from [45])

	Definition	Min., Max.	Units
h_k^r	Resting membrane potential	-80, -60	mV
τ_k	Passive membrane decay time	5, 150	ms
h_{ek}^{eq}	Excitatory reversal potential	-20, 10	mV
h_{ik}^{eq}	Inhibitory reversal potential	-90, $h_k^r - 5$	mV
Γ_{ek}	EPSP peak amplitude	0.1, 2.0	mV
Γ_{ik}	IPSP peak amplitude	0.1, 2.0	mV
$1/\gamma_{ek}$	EPSP rise time to peak	1, 10	ms
$1/\gamma_{ik}$	IPSP rise time to peak	2, 100	ms
N_{ek}^{α}	No. of excitatory cortico-cortical synapses	$k = e: 2,000, 5,000$ $k = i: 1,000, 3,000$	-
N_{ek}^{β}	No. of excitatory intracortical synapses	2,000, 5,000	-
N_{ik}^{β}	No. of inhibitory intracortical synapses	100, 1,000	-
v_{ek}	Axonal conduction velocity	0.1, 1	$\frac{\text{mm}}{\text{ms}}$
$1/\Lambda_{ek}$	Decay scale of cortico-cortical connectivity	10, 100	mm
S_k^{\max}	Maximum firing rate	0.05, 0.5	ms^{-1}
μ_k	Firing threshold	-55, -40	mV
σ_k	Standard deviation of firing threshold	2, 7	mV
p_{ek}	Extracortical synaptic input rate	0, 10	ms^{-1}

In general the A_{lk} is comprised of cortically local, cortically distant and extracortical/subcortical axonal pulses. Because subcortical and cortically distant axonal pulses arise exclusively from excitatory neurons, A_{ek} and A_{ik} are defined as

$$A_{ek}(r, t) = N_{ek}^{\beta} S_e[h_e(r, t)] + \phi_{ek}(r, t) + p_{ek}(r, t) , \quad (14.3)$$

$$A_{ik}(r, t) = N_{ik}^{\beta} S_i[h_i(r, t)] , \quad (14.4)$$

where $N_{lk}^{\beta} S_l$, the mean number of connections from local neuronal population l times their mean firing rate S_l , models local inputs to target population k , p_{ek} represents extracortical (thalamic) excitatory sources and ϕ_{ek} pulses arriving across larger distances via the excitatory cortico-cortical fibre system.

The lynchpin of the mean field formulation is the closure of the macroscopic equations (14.1)–(14.4) by the definition of S_l . In rate based models it is typically assumed that mean population firing rates are an instantaneous function of the respective mean soma membrane potential. One very general form for S_l in which mean firing rates monotonically increase with h_k , are bounded below by zero and above by a maximal firing rate and has a flexible shape is [41]

$$S_l[h_l(r, t)] = S_l^{\max} - S_l^{\max}(1 + \exp\{\sqrt{2}[h_l(r, t) - \mu_l]/\sigma_l\})^{-\kappa_l} \quad (14.5)$$

which for $\kappa_l = 1$ reduces to the well-known symmetric sigmoid function.

Axonal pulses that propagate locally through intracortical fibre systems are assumed to result in conduction delays that are negligible in comparison to the

delays induced by neurotransmitter activation and the passive electrical properties of dendrites. However such delays cannot be ignored for axonal pulses that are propagated by the long range cortico-cortical fibre systems. In the simplest case of a single cortico-cortical conduction velocity v_{ek} and an exponential fall off in the strength of cortico-cortical connectivity with increasing distance between source and target neuronal populations of characteristic scale $1/\Lambda_{ek}$ it can be shown that the propagation of ϕ_{ek} can be approximately described by the following two-dimensional telegraph equation

$$\left[(\partial_t + v_{ek} \Lambda_{ek})^2 - \frac{3}{2} v_{ek}^2 \nabla^2 \right] \phi_{ek} (r, t) = v_{ek}^2 \Lambda_{ek}^2 N_{ek}^\alpha S_e [h_e (r, t)] , \quad (14.6)$$

where N_{ek}^α is the total number of excitatory synaptic connections formed by long-range cortico-cortical axons on local population k . Robinson et al. [65] and Jirsa and Haken [37] have both defined similar long-range propagators.

Equations (14.1)–(14.6) typically define the model of Liley et al. and in addition to being able to reproduce the main features of spontaneous human EEG gives rise to a rich repertoire of interesting and/or novel dynamical activity.

14.2.1 Model Extensions

While the physiological specificity of the model of Liley can be easily extended by the addition of sub-populations of excitatory and inhibitory neurons or by the inclusion of ancillary axonal conduction systems of differing characteristic scales and conduction velocities, the resulting formulation, while arguably of greater biological veracity, will have a substantially augmented phase space. In non-linear systems larger phase spaces make it more difficult to characterise system dynamics. It is therefore fortunate that there are a number of modifications and extensions that can be made to the model of Liley et al. that further supplement its physiological relevance without causing its phase space to expand. Further such modifications can be utilised by other mean field approaches aimed at modelling cortical electrophysiology.

A simple modification of the equation describing the dynamics of the postsynaptic currents enables independent adjustments of the rise and decay times of the unitary postsynaptic potential so defined [11]. By defining I_{Ik} to satisfy

$$[\partial_t + \gamma(\epsilon)][\partial_t + \tilde{\gamma}(\epsilon)]I(r, t) = \tilde{\gamma}(\epsilon)e^{\gamma(\epsilon)/\gamma_0} \Gamma \cdot A(r, t) , \quad (14.7)$$

$$\gamma(\epsilon) = \epsilon\gamma_0/(e^\epsilon - 1) , \quad \tilde{\gamma} = \gamma(\epsilon)e^\epsilon \quad (14.8)$$

where we have dropped the subscripts for clarity, $1/\gamma_0$ defines the time to peak of, and $\epsilon > 0$ controls the decay of, the unitary postsynaptic potential. It is worth noting that Eq. (14.7) reduces to Eq. (14.2) as $\epsilon \rightarrow 0$.

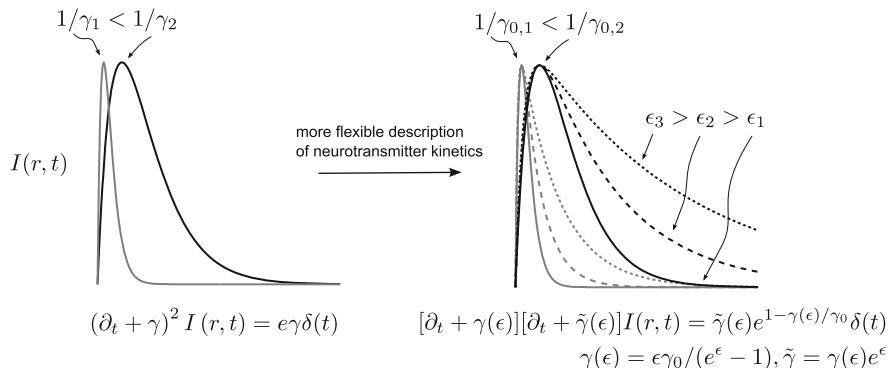


Fig. 14.2 The independent adjustment of the rise and decay times of the modelled unitary postsynaptic potential enables the more faithful modelling of physiologically and pharmacologically induced alterations in bulk neurotransmitter kinetics and dendritic cable properties

Why might such a formulation be useful? Firstly, isoflurane, a volatile halogenated anaesthetic, has been shown to prolong the decay time of the unitary inhibitory postsynaptic potential as well as reducing the peak amplitude of both excitatory and inhibitory postsynaptic potentials [7, 49]. Based on experiment γ_0 , ϵ and Γ can then be defined to be functions of extracellular anaesthetic concentration such that the effects of isoflurane on the EEG can be modelled. By taking this approach Bojak and Liley [11] have been able to account for the increases in low frequency power in the human EEG induced by isoflurane action. Secondly, the bulk voltage-dependence of excitatory postsynaptic potential amplitude and time course, that arises as a consequence of a N-methyl-d-aspartate (NMDA)-mediated synaptic component, can plausibly be approximated by such a formulation by allowing Γ , γ_0 and ϵ to be functions of the mean soma membrane potential h i.e. $\Gamma(h)$, $\gamma_0(h)$ and $\epsilon(h)$. At present no work has been performed in this regard however it may offer a fruitful way forward towards modelling the bulk effects of behaviourally or pharmacologically induced alterations in NMDA-mediated receptor activity (Fig. 14.2).

In the model of Liley et al., and other related mean field formulations [37, 56, 65, 72], it is typically assumed that activity propagated between distant cortical areas by the cortico-cortical conduction system is by fibres of relatively uniform conduction velocity. However empirical measurement, either by direct physiological measurement of conduction latencies or indirectly via histological measurement of axonal diameter and the subsequent mapping to conduction velocity, suggests that propagation velocities of the cortico-cortical fibres are instead rather broadly distributed [60, 64]. While such broad distributions are easily incorporated in integral mean field formulations, at least until recently, it has not been possible to include them in the computationally more efficient and tractable mean field partial differential formulations. However as shown in Bojak and Liley [11] by defining ϕ_{ek} to satisfy

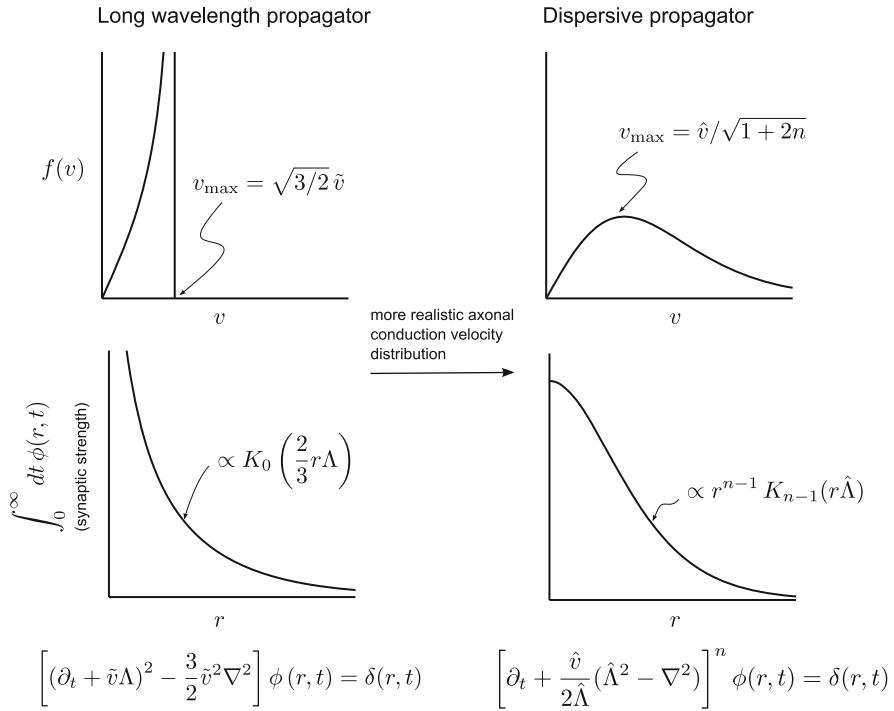


Fig. 14.3 It is computationally convenient to describe pulse propagation by the long range cortico-cortical conduction system by a partial differential formulation. The long range propagator defined by Eq. (14.6) is simple to implement but assumes that the axonal velocity distribution is sharply peaked (*top left*) about a central value with the synaptic connectivity having an integrable infinite divergence at $r = 0$ (*bottom left*). Empirically however, cortico-cortical axonal velocities are found to be quite broadly distributed. Fortunately a partial differential propagator can be found that assumes a broad distribution of axonal conduction velocities (*top right*) while retaining a monotonic decay in axonal fibre density as a function of distance (*bottom right*)

$$\left[\partial_t + \frac{\hat{v}_{ek}}{2\hat{\Lambda}_{ek}}(\hat{\Lambda}_{ek}^2 - \nabla^2) \right]^n \phi_{ek}(r, t) = 2^{-n} \hat{v}_{ek}^n \hat{\Lambda}_{ek}^n N_{ek}^\alpha S_e[h(r, t)] \quad (14.9)$$

a marginal velocity distribution, $f_{ek}(v)$, of the form

$$f_{ek}(v) = \frac{2nv\hat{v}_{ek}^{2n}}{(v^2 + \hat{v}_{ek}^2)^{n+1}} \quad (14.10)$$

is implied for the propagation of axonal pulses by long-range fibres. Based on fits to callosal fibre data obtained from humans it is estimated that $n = 4$ and $\hat{v}_{ek} \approx 19 \text{ m s}^{-1}$. Such a formulation preserves monotonically decreasing connectivity between source and target regions as a function of increasing separation (Fig. 14.3).

14.3 Dynamical Features

Despite the structural simplicity of the defining mean field equations, they reveal a rich repertoire of dynamics in the physiologically admissible parameter space. While much of this is of physiological relevance we will in this section restrict our discussion to the dynamical properties and features that have some degree of mathematical novelty.

14.3.1 A Novel Route to Chaos

The brain is undoubtably a structured and highly complex dynamical system. Attempts to characterise and explain the structured emergence of such complex activity face many hurdles not the least of which is experimental. How do we measure the state of a complex system in the presence of substantial physiological measurement noise? Attempts to determine whether the brain supports the existence of deterministic dynamical macroscopic brain states are inevitably frustrated by the noisy and non-stationary time series data obtained from EEG, MEG or resting state fMRI. Might the existence of theoretical evidence for such complexity help guide and motivate such empirical explorations? In this regard the chaotic dynamical behaviour of a macrocolumnar reduction of the Liley model, and its parametric organization, might be relevant to this quest. By ignoring long range connectivity the Liley model's phase space can be dramatically reduced in size – yet retain considerable dynamical complexity. For example by assuming $\gamma_{ik} \equiv \gamma_i$, $\gamma_{ek} \equiv \gamma_e$ $N_{ek}^\alpha = 0$, and under some weak assumptions of convergence, Eqs. (14.1)–(14.4) can be rewritten as [15]

$$\tau_k \partial_t h_k = h_k^r - h_k + \sum_{l=e,i} e \Gamma_{lk} \frac{h_{lk}^{eq} - h_k}{|h_{lk}^{eq} - h_k^r|} \{ \gamma_l N_{lk}^\beta I_l + (1 - \delta_{lk}) [p_{lk} - N_{lk}^\beta p_l] / \gamma_l \} \quad (14.11)$$

$$(\partial_t + \gamma_k)^2 I_k = S_k(h_k) + p_k \quad (14.12)$$

where $p_e = p_{ee}/N_{ee}^\beta$, $p_i = p_{ii}/N_{ii}^\beta$ and δ_{lk} is the Kronecker delta. It is worth noting that our “synaptic currents” have been trivially rescaled to have units of s^{-1} . For a range of physiologically admissible parameter values these simplified equations are, not surprisingly, capable of producing aperiodic behaviour characteristic of deterministic chaos. However what is perhaps surprising is that such chaotic activity arises through a number of routes, one of which is quite unusual. Initial numerical explorations involving the full set of equations, in which long range connectivity was ignored, revealed extensive chaos in a parameter plane defined by p_{ee} and p_{ei} . Subsequently it was shown that such chaos was spawned by a co-dimension one homoclinic bifurcation, known as a Shil'nikov saddle-node

bifurcation³ located in an unphysiological region of the (p_{ee}, p_{ei}) parameter plane [75]. However, even though this co-dimension one bifurcation is located in an unphysiological region of parameter space it nevertheless organizes the qualitative behaviour of the emerging dynamics in the physiological meaningful parameter space. Interestingly this route to chaos occurs in the vicinity of a codimension three, focus type, degenerate Bogdanov-Takens point, suggesting that there might exist an *organising centre* for the qualitative organization of dynamics in an even larger region of parameter space. Importantly for *the same* parameter set the reduced Eqs. (14.11) also give rise to chaos [15] in the (p_{ee}, p_{ei}) parameter plane spawned by an identical Shil'nikov saddle-node bifurcation at unphysiological (i.e. negative) values of p_{ee} and p_{ei} . Such a reduction may therefore aid in the efficient characterisation of high co-dimension organising centres.

14.3.2 Metabifurcations

One of the limitations in performing a bifurcation analysis on a high dimensional system is the difficulty in establishing a canonical parameter set from which to explore the qualitative organization of the parameter space and to relate it to physiologically meaningful or significant behaviour. Often parameter sets are degenerate, in the sense that parametrically widely separated sets can produce similar, physiologically relevant, behaviour. For example [25] numerically generated over 70,000 parameter sets for the spatially homogeneous (i.e. $\nabla^2\phi_{ek} = 0$) Liley model that gave rise to electroencephalographically and physiologically plausible behaviour: parameters within empirically established ranges, alpha band oscillatory activity, ' $1/f$ ' low frequency activity and modelled mean neuronal firing rates $\lesssim 20 \text{ s}^{-1}$. A subsequent principal components analysis revealed that the first ten principal components cumulatively accounted for less than 50 % of the total parametric variance i.e. the structure of the parameter space could not be appreciably simplified by assuming linear combinations of parameters. Therefore how might we investigate the qualitative dynamical properties of the model's physiological admissible parameter space? One possible solution is to attempt to partition the parameter space based on a classification of the patterns of bifurcation diagrams obtained by continuing in one or more appropriately chosen parameters, and to determine the conditions under which, if any, inter-bifurcation pattern transition occurs. We refer to such a general method as a "*metabifurcation analysis*".

³If a saddle node has a single homoclinic orbit, then a unique limit cycle will form when the equilibrium disappears. This is often referred to as a saddle-node on an invariant circle bifurcation. If a saddle node has two or more homoclinic orbits then infinitely many saddle limit cycles (i.e. chaos) appear when the equilibrium disappears. In general this is referred to as a Shil'nikov saddle-node bifurcation. These bifurcations should not be confused with either saddle node (fold) or saddle-homoclinic bifurcations.

How might such a metabifurcation analysis proceed? It is known that in the Liley model inhibition is found to be a very sensitive locus of dynamical control. Small alterations in modelled inhibitory coupling strengths and neurotransmitter kinetics sensitively induce changes in model stability, excitability and frequencies of driven and autonomous oscillatory modes. It has been theorised that such sensitivity can explain many of the electroencephalographic features of γ -amino butyric acid (GABA)-ergic anaesthetic action (see Sect. 14.4.1.1). On this basis [25] constructed two-dimensional parameter continuations in an (R, k) plane defined by $\Gamma_{ik} \rightarrow R\Gamma_{ik}$ and $N_{ii}^\beta \rightarrow kN_{ii}^\beta$. As $p_{ik} \equiv 0$ variations in R and k enable the independent specification of global and individual population changes in inhibitory input coupling strength. Two dimensional continuations in (R, k) were then performed for a large number of randomly chosen parameter sets, for the spatially homogeneous Liley model, selected to exhibit electroencephalographically and physiological plausible dynamical behaviour.

Based on the analysis of 405 randomly chosen parameter sets it was found that topologically the bifurcation diagrams conformed to two broad patterns or families (Fig. 14.4). For one of the families (F1) two, almost parallel, lines of saddle-nodes (equilibrium) partition the (R, k) plane into three major regions. For the other family (F2) the (R, k) plane is characterised by the presence of two cusp points, such that the region containing three equilibria is the union of two separated wedge-shaped areas with the cusps as their vertices. In both families emergent Hopf bifurcations interact with the saddle node bifurcations by so-called fold-Hopf points. The two families could be distinguished by differences in the distribution of certain parameters, with the parameter distributions accounting for the greatest dissimilarity being in τ_e , the mean excitatory neuronal membrane time constant, and σ_e , the standard deviation in the excitatory neuron mean firing threshold. In particular $\tau_e^{F1}, \sigma_e^{F1} < \tau_e^{F2}, \sigma_e^{F2}$. In general it is found that parameter sets belonging to F1 are associated with a more restricted dynamical repertoire than parameter sets belonging to F2.

Topological transitions between the two types of families can be induced by changes in p_{ee} and p_{ei} . In general reductions in p_{ei} in parameter sets belonging to F1 result in the topological metamorphosis of the (R, k) bifurcation diagram to that of F2. Specifically decreasing p_{ei} results in the appearance of two cusp points via a so-called *swallow tail* bifurcation until the bifurcation diagram resembles that of F2. A similar topological transition is induced from F1 to F2 if a metabifurcation parameter p_{ee} is instead increased.

What, if at all, might be the physiological significance of such “*metabifurcations*”? Two speculations present themselves. Firstly the metabifurcation parameters p_{ee} and p_{ei} model thalamic input, and therefore suggest an alternative role for thalamus other than its classically defined character as a relay station for peripherally derived sensory information. Because changes in p_{ee} and p_{ei} transfigure the topological organization of bifurcations in the (R, k) plane we might hypothesise that thalamic activity modulates, and in a sense selects, the cortical dynamical landscape. Viewed from this perspective sensory input may be conceived as configuring the possible domain of the cortical response in addition to initiating it. Further by

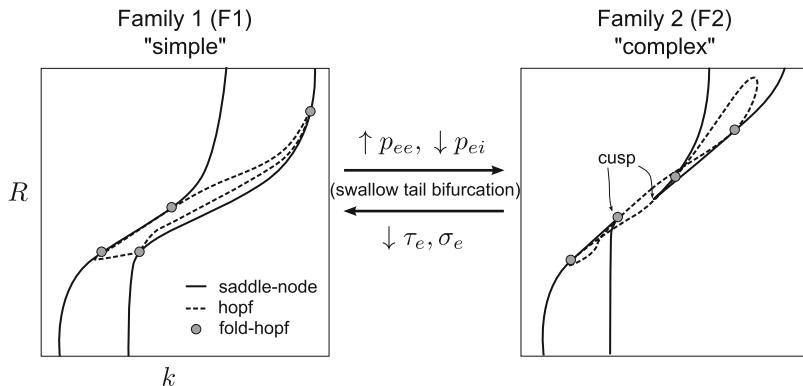


Fig. 14.4 Schematic diagram illustrating two topological families of bifurcation diagrams obtained by the classification of continuations in the (R, k) parameter plane of ≈ 400 parameter sets chosen to exhibit physiologically and electroencephalographically admissible behaviour. Also illustrated are the parameters identified to most sensitively effect topological transitions between the two families. Numerical analysis reveals that parameter sets belonging to family 1 have a more restricted repertoire of dynamical behaviour than those belonging to family 2. For further details see [25] (Figure adapted from [25])

considering thalamocortical feedback the possibility is opened up for some form of auto-regulation of cortical dynamics: cortical feedback through thalamus could initiate a sequence of transitions between topologically distinct bifurcation patterns and thus the cortical dynamical repertoire could be reconfigured “*on the fly*”, and on a time scale quite distinct to activity dependent synaptic plasticity. The second speculation concerns activity dependent changes in τ_e and σ_e . The widely held view is that learning principally involves modifications of synaptic strength. However there exists an alternative, though less well known, view in which learning may also involve non-synaptic processes, such as modulations in voltage dependent membrane conductances, that manifest themselves in alterations in neuronal excitability [52]. For example widely identified non-synaptic changes observed during learning include changes in neuronal input resistance and alterations in neuronal burst/spike threshold. The mean field correlates of these single neuron properties include τ_e and σ_e . Thus we might speculate that activity dependent changes in τ_e and σ_e cause long lasting changes in the cortical dynamical landscape, and that these alterations contribute to the behavioural changes observed during learning.

14.3.3 Multistability

Many systems in nature exhibit multistability: when starting from different initial conditions the system can evolve into different attractors with quite different long term behaviour. The term “*generalised multistability*” was coined in order to

distinguish it from “*trivial multistability*” which arises from the co-existence of multiple stable fixed points. Subsequent to its initial theoretical and experimental delineation it is now a well described phenomenon in neuroscience, optics and condensed matter physics.

In neuroscience one of the most extensive examples of generalised multistability is found in the R15 *Aplysia* neuron model in which five different limit cycles and two chaotic attractors are found to co-exist [16]. Functionally multistability might provide a mechanism whereby transient changes in neural activity or sensory input induce persistent changes in oscillatory activity. Such oscillatory mode shifts may therefore directly initiate changes in behaviour or perception, or act as a dynamical substrate from which further activity dependent modulations in dynamics arise. While relatively well studied in the context of single neuron dynamics, multistability at the cortical population level has been little appraised either experimentally or theoretically. However emerging evidence does suggest that resting alpha (8–13 Hz) band activity can be decomposed into distinct high and low amplitude modes [27], and that this can be interpreted as evidence for cortical population level multistability [28]. It is therefore natural to ask whether such multistable activity can be found in the Liley model.

It is shown in [20] that a macrocolumnar version (i.e. $N_{ek}^\alpha = 0$) of the Liley model is able to support the co-existence of two limit cycle attractors and one chaotic attractor in an initial condition space. The limit cycle attractors consist of (i) a high amplitude, high firing rate ($\approx 300 \text{ s}^{-1}$), limit cycle with a dominant frequency of $\approx 5 \text{ Hz}$ and a strong first harmonic $\approx 10 \text{ Hz}$, and (ii) a low amplitude, low firing rate ($\approx 20 \text{ s}^{-1}$), limit cycle with a dominant frequency of $\approx 10 \text{ Hz}$. The set of initial conditions which gives rise to the low amplitude limit cycle is embedded in a sea of initial conditions which gives rise to a small amplitude chaotic attractor (largest Lyapunov exponent = 3.4 s^{-1} ; Kaplan-Yorke dimension = 2.086 ± 0.003) having a dominant frequency in the alpha band. Surrounding the initial conditions of these low amplitude dynamics is an extensive region of large-amplitude limit cycle dynamics.

Unlike multistable dynamics observed in a similar mean field model [28], the multistable dynamics in this model does not arise due to noise driving in the vicinity of a sub-critical Hopf bifurcation. Parametric continuations in p_{ee} instead reveals that a high amplitude limit cycle, born from a subcritical Hopf bifurcation at large p_{ee} , surrounds chaos born through a period doubling cascade at small p_{ee} in which is embedded a low amplitude limit cycle orbit that appears to arise through a homoclinic bifurcation at intermediate values of p_{ee} .

14.4 Physiological Relevance

One of the strengths of the model of Liley is the physiological relevance of its parameterisation. All model parameters correspond to quantities that can be physiologically and anatomically independently measured, and thus the important

question can be asked: to what extent does the physiologically admissible model parameter space produce behaviour that is both physiologically and electroencephalographically tenable? For example can the model of Liley account for the electroencephalographically observed alpha rhythm in the context of a physiologically meaningful parameterisation? Being able to account for the alpha rhythm would signal the suitability of this model as one basis to account for bulk perturbations in dynamical brain activity that are observed to occur in health, disease and during drug administration.

14.4.1 *The Resting Alpha Rhythm*

Between 1926 and 1929 Hans Berger laid the foundations for the development of electroencephalography in humans [32]. While canine EEG had been discovered many decades earlier [17], it was Berger who first described the alpha rhythm (8–13 Hz), its occipital dominance and its attenuation in response to mental effort and opening of the eyes. The intervening years have revealed that alpha band activity is not restricted to occipital cortex. Alpha band activity is recordable over much of the cortical surface and is reactive (i.e. enhanced or attenuated) in response to a much wider variety of cognitive activity than just opening and closing the eyes. For this reason it is often preferable to refer to 8–13 Hz electroencephalographic activity as *alpha band activity* rather than as the *alpha rhythm*.

Despite amassing a great deal of knowledge regarding the phenomenology of alpha band activity we remain comparatively ignorant regarding the physiological basis for its genesis: does it (i) arise from intrinsic oscillatory activity in individual cortical neurons (ii) stem from oscillatory thalamic activity directly driving populations of cortical neurons or (iii) emerge through the reverberant activity generated by reciprocal interactions of synaptically connected neuronal populations in cortex, and/or through such reciprocal interactions between cortex and thalamus? Theoretically the last of these is the most interesting and the one best addressed by the mean field modelling approach. Specifically, for a range of physiologically admissible parameter values the model of Liley reveals a wide array of deterministic and noise-driven dynamics that includes alpha band activity [11, 41]. In particular, physiologically plausible alpha band activity can appear in three distinct dynamical scenarios: linear noise driven, limit-cycle, and chaotic dynamics. For appropriate parameterisations linearisations of the defining equations about a stable singular point reveals alpha band oscillatory activity in h_e and h_i at physiologically plausible firing rates ($0.1\text{--}20\text{ s}^{-1}$), as well as rhythmic activity in other bands of electroencephalographic interest. In the case of electroencephalographically plausible alpha band activity (full-width-half-maximum of the peak alpha band frequency $\gtrsim 5$) linearisation reveals model activity to be essentially determined by conjugate pairs of weakly damped poles at alpha band frequencies. The physiological plausibility of these linearisations suggests that resting EEG may be viewed as a filtered random linear process. Indeed empirical analysis has found that, except for short bursts

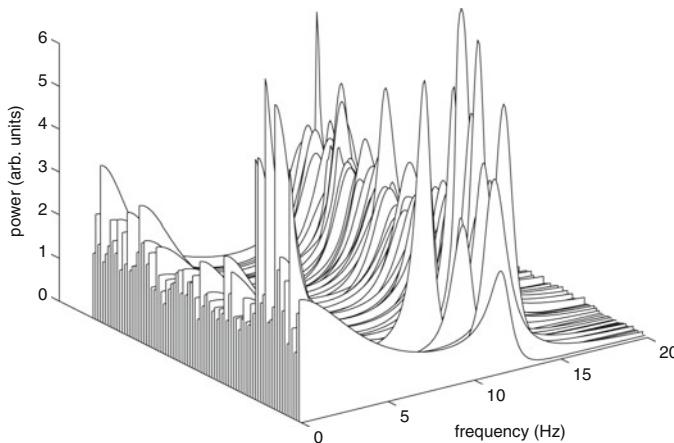


Fig. 14.5 Analytical fluctuation spectra for 86 physiologically admissible parameter sets chosen to exhibit plausible mean resting neuronal firing rates and resting alpha band activity, as well as a biphasic surge during simulated anaesthetic action. For details of parameters and method of calculation refer to [11]

of subdural and scalp-recorded EEG, the alpha rhythm is indistinguishable from linearly filtered white noise [70, 71]. On this basis we might reasonably assert that “resting” cortex is dynamically in a state of marginal linear stability.

In the Liley model, based on a range of heuristic search strategies [11], physiologically and electroencephalographically plausible alpha band activity is found to be widely, but sparsely, distributed over the whole biologically valid parameter space without easily discernable structure in most parameter dimensions. For example by randomly searching the physiologically admissible parameter space [11] found that of 7×10^9 randomly generated parameter sets 73,454 ($\approx 0.001\%$) produced electroencephalographically plausible alpha band activity (Fig. 14.5). However as it is likely there are biological co-dependencies between various model parameters such sparseness may be artificial.

By analysing the response of model dynamics to small parameter perturbations it is found that emergent alpha band activity is particularly sensitive to alterations in those parameters that characterise inhibitory action. This in turn suggests a novel mechanism for alpha band rhythmogenesis: alpha band oscillatory activity arises from reverberant activity between populations of inhibitory interneurons. To see this consider the following sequence of events – (i) initially (basally) excited inhibitory neurons, following a delay related to the characteristic time of inhibitory neurotransmitter kinetics, are inhibited due to negative feedback and thus inhibitory neuronal firing rates decrease (ii) because the activity of inhibitory neurons has decreased feedback inhibition is reduced and thus inhibitory neuronal firing rates increase again and return to basal levels on a time scale related to the characteristic time of inhibitory neurotransmitter kinetics (iii) once mean inhibitory firing rates return to basal levels feedback inhibition between inhibitory

neurons is again strengthened and mean inhibitory firing rate again decrease and the cycle then repeats. Such oscillatory activity then “slaves” population excitatory neuronal activity and thus gives rise to alpha band variations in scalp recorded electroencephalographic activity [43].

Thus the model of Liley hypothesises that (i) resting alpha band activity is a marginally stable rhythm and (ii) inhibition is a sensitive locus for the dynamical control of alpha band oscillations. These hypotheses have important implications for the functional role of alpha band oscillations during cognition and their physiological control and pharmacological modulation. For example parameter sets chosen to produce electroencephalographically plausible linear noise driven alpha, as well as a surge in total EEG power during modelled anaesthetic induction, can under small parametric perturbations produce autonomous gamma band ($>30\text{ Hz}$) oscillatory activity. Gamma band oscillations are thought to be the *sine qua non* of cognitive functioning. Indeed there exists much evidence to suggest that the emergence of synchronised gamma-band activity (local field potential or EEG) functional underpins perceptual binding and subserves the processes of cortical computation [29]. Thus the existence of weakly damped, noise-driven, linear alpha activity may be a dynamical precursor to gamma band electroencephalographic activity, and more generally as a physiologically meaningful state from which transitions can be made from or to. From this perspective the alpha rhythms may be better viewed as readiness rhythms and not idling or resting rhythms as is often asserted.

What physiological factors may drive such transitions? Given that cortical population dynamics are hypothesised to be particularly sensitive to variations in inhibitory activity it may represent a target for control by the relatively sparse thalamocortical afferents. Averaged over cortex, less than 2–3 % of all synapses can be attributed to thalamocortical projections [13]. While thalamocortical afferents synapse onto both excitatory (pyramidal) and inhibitory layer IV cortical neurons the strength of such synaptic connections may be far from uniform across the respective target neuronal populations. For example studies in rat barrel cortex show that cortical inhibitory neurons receive thalamocortical synapses that are on average five-fold stronger (in terms of evoked inhibitory postsynaptic amplitude) than those received by nearby pyramidal neurons [50]. In this way relatively weak excitatory thalamocortical input to inhibitory cortical neuronal populations (p_{ei}) may be able to precipitate transitions in cortical state. Thus thalamic input arising from either first-order thalamic relay neurons being driven directly by incoming sensory/subcortical input, or from higher-order thalamic relay neurons driven by feedback from cortex, may, rather than only “driving” cortex act also to “modulate” cortical activity (in the sense of [68]).

While we have speculated that cortical inhibition may be a sensitive target for the control of alpha band activity by thalamocortical afferents at present the empirical evidence for such an assertion is weak. Fortunately stronger evidence for the hypothesised role of inhibitory modulation in the regulation and control of alpha band activity exists. Of particular interest and relevance is the fact that the

endogenous and exogenous pharmacological modulation of GABAergic activity is known to perturb the resting alpha rhythm.

14.4.1.1 Endogenous Pharmacological Modulation

Surprisingly the EEG is known to undergo systematic changes in women during their menstrual cycle. In particular it has been observed that during the late and mid-luteal phases alpha band activity is enhanced. For example [19] observed that the mean occipital alpha band activity increased by 0.3 Hz during the luteal phase and that the average time course of acceleration followed the time course of increases in blood progesterone levels. Similar changes in alpha band activity during oestrus are observed in other studies [6, 9, 61]. How do such changes implicate modulations in GABAergic activity?

Progesterone, a steroid hormone involved in the female menstrual cycle, is metabolised to a high degree to the neurosteroids allopregnanolone and pregnanolone which are potent positive allosteric modulators of GABA subtype A (GABA_A) receptor activity such that GABA action is potentiated [33, 62]. These neurosteroids bind to discrete sites on the GABA_A receptors that are distinct to those that bind ethanol, benzodiazepine, barbiturates and a range of general anaesthetic agents. During the mid and late luteal phases progesterone concentrations are highest and thus their effect in modulating GABAergic function is maximal. The model of Liley predicts that the antagonism of GABAergic activity (increases in Γ_{ik}) should alter the spectral features of resting alpha band activity. In particular the model predicts that increases in Γ_{ii} will increase the frequency, and reduce the damping (i.e. reduce the full-width-half-maximum) of the alpha-band linear resonance, whereas increasing Γ_{ie} will produce the opposite effect. On this basis the model of Liley would predict that the neurosteroids allopregnanolone and pregnanolone potentiate GABAergic activity in cortical inhibitory neurons to a greater degree than in cortical excitatory neurons. Is such a prediction supported by pharmacological differences in the properties of GABA_A receptors in excitatory and inhibitory neurons?

Structurally GABA_A receptors are composed of 5 membrane spanning protein subunits that are assembled from a family of at least 18 subunits ($\alpha_1\text{--}6$, $\beta_1\text{--}3$, $\gamma_1\text{--}3$, $\sigma_1\text{--}3$, δ , ϵ , θ) that determine, among other properties, their pharmacological profiles [59]. A range of studies have established that they are heterogeneously distributed across brain areas and neuronal subtypes [30, 58]. In cortex the most abundant receptor isoforms are $\alpha_1\beta_{2/3}\gamma_2$ and $\alpha_2\beta_{2/3}\gamma_2$ differentially localised to inhibitory and excitatory cortical neurons respectively. In general the $\alpha_1\beta_{2/3}\gamma_2$ and $\alpha_2\beta_{2/3}\gamma_2$ isoforms exhibit differential binding affinities for benzodiazepines. Indeed the presence of the α subunit isoform exerts a major effect on the affinity and efficacy of ligands at the benzodiazepine binding site. It is thought that the neurosteroids allopregnanolone and pregnanolone evince similar differential binding affinities based on the demonstrated importance of the α subunit for the binding of neurosteroids [33]. Thus the prediction of the Liley model regarding the

predominant target of neurosteroid action is consistent with the known molecular pharmacology. As the next section will discuss, the differential ligand-binding affinity of GABA receptor isoforms may be relevant to understanding the actions that anaesthetics and sedatives have on the EEG.

14.4.1.2 Exogenous Pharmacological Modulation

General anaesthetic agents induce profound reversible alterations in brain activity and behaviour. While positron emission tomography and fMRI have revealed a range of non-uniform reductions in inferred cerebral neuronal activity during anaesthetic drug action [2, 3, 39], to date only changes in the EEG have been reliably correlated with the clinically documented effects of anaesthesia [14]. In general during the progression to deep anaesthesia the EEG undergoes a series of well described quantitative changes: (i) the EEG is transiently activated such that beta band (13–30 Hz) oscillatory activity is increased and alpha band activity is decreased (the so-called “*beta buzz*”) (ii) the EEG is slowed (reduction in median and spectral edge frequencies), the alpha rhythm is abolished, and total EEG power transiently increases (the “*biphasic response*”) (iii) the appearance of isoelectric (defined as <5 µV peak-peak amplitude) periods lasting many seconds separated by short bursts of high amplitude slow, sharp or spiking activity – a phenomenon known as *burst suppression*. While not all anaesthetic agents produce these changes (notable exceptions being nitrous oxide, xenon and ketamine – agents often collectively referred to as “*dissociatives*”) they are sufficiently general to motivate systematic processed EEG approaches to the clinical monitoring of depth of anaesthesia.

To what extent can these features be explained by known anaesthetic molecular pharmacology? Countless studies have revealed the synaptic GABA_A receptor to be one of the most important molecular targets mediating the action of anaesthetic and sedative agents [24, 67]. It has been established that a variety of anaesthetic agents reduce the peak amplitude, and selectively prolong the decay, of the inhibitory postsynaptic potential [7]. On this basis [11] have shown that many of the electroencephalographic features of “*typical*” anaesthesia can be accounted for by the model of Liley by utilising a description that enables the independent adjustment of rise and decay times of the inhibitory postsynaptic potential i.e. Eq. (14.7). By parameterising Γ_{ik} and ϵ_{ik} as a function of anaesthetic concentration (c) on the basis of experimental measurement, modelled EEG slows, and exhibits a transient increase in total power, with increasing c for appropriately chosen “*base*” parameter sets. While such bulk effects of anaesthetic action can be produced by homogeneous variations of Γ_{ik} and ϵ_{ik} , other well known electroencephalographic effects might only be explained by assuming that inhibitory postsynaptic potentials in inhibitory neurons are differentially modulated by anaesthetic action compared to inhibitory postsynaptic potentials in excitatory neurons.

In [43] it was theorised, on the basis of empirical EEG evidence involving the benzodiazepine alprazolam, that the well documented benzodiazepine “*beta*

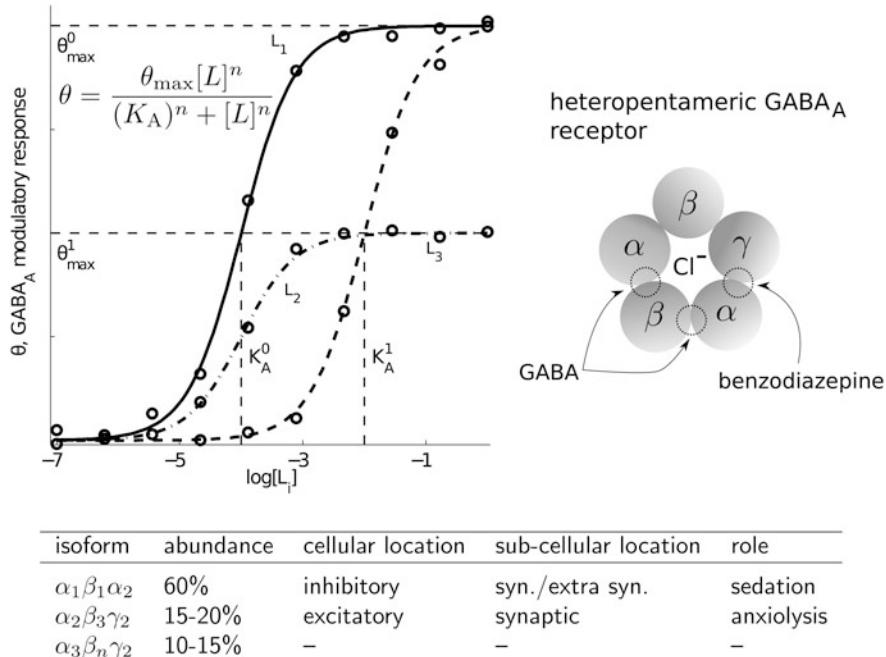


Fig. 14.6 *Top left:* differences in binding affinity, K_A , and maximal modulatory response, θ^{\max} , for three hypothetical positive allosteric modulators of the $GABA_A$ receptor. *Top right:* the $GABA_A$ receptor is a co-assembly (oligomer) of five protein subunits, the interfaces of which contain binding sites for GABA and benzodiazepines. *Bottom:* summary of the major $GABA_A$ receptor isoforms found in cortical neurons [30, 51]. The type of α subunits present determines the pharmacological properties of the given isoform. Isoforms containing the α_1 subunit have a BZ₁-type pharmacology (bind zolpidem and CL218,872 with high affinity), whereas those co-assembled with the α_2 isoform have a BZ₂-type pharmacology (bind zolpidem and CL218,872 with low affinity)

“buzz” could be explained by assuming that benzodiazepines acted with greater efficacy at GABAergic synapses on inhibitory neurons than at GABAergic synapses on excitatory neurons (see Fig. 14.6). Such differential potency accords with the empirically established cellular distribution of $GABA_A$ receptor isoforms that exhibit ligand-based differences in the potentiation of GABA induced activity. As discussed previously the high affinity (to zolpidem and CL218,872) $GABA_A$ receptor isoform $\alpha_1\beta_{2/3}\gamma_2$ is found predominantly in cortical inhibitory neurons whereas the low affinity $GABA_A$ receptor isoform $\alpha_2\beta_{2/3}\gamma_2$ is chiefly localised to cortical excitatory neurons [30]. The differential effects of modulating feed-forward and feed-back may also be relevant to understanding the electroencephalographic actions of propofol.

Propofol (2,6-di-isopropylphenol) is a widely used intravenous anaesthetic agent that is distinguished in that, rather than uniformly attenuating alpha band activity, instead elicits strong increases in frontal alpha band (10–13 Hz) activity in addition

to the typical increases in slow wave activity seen with other anaesthetic agents. While it has been suggested that this “alpha” rhythm emerges because propofol enhances feed-forward GABA_A conductances in cortical pyramidal neurons such that thalamocortical feedback is strengthened [18], another possibility suggests itself. Propofol, like the benzodiazepines and the neurosteroids, allosterically enhances GABA-mediated activation of GABA_A receptor activity to a degree depending on its isoform. For example [38] found that in recombinant GABA_A receptors expressed in *Xenopus* oocytes, that those consisting of the isoform $\alpha_1\beta_{2/3}\gamma_2$ were potentiated to a much greater degree (maximum potentiation compared to baseline $\approx 1,400\%$) by propofol than those of the $\alpha_2\beta_{2/3}\gamma_2$ (maximum potentiation compared to baseline $\approx 500\%$) isoform. From a theoretical perspective feed-back disinhibition of cortical inhibitory neuronal activity would then be favoured over the feed-forward inhibition of cortical excitatory neuronal activity and thus alpha band activity would be promoted.

14.4.2 Mass Action and the Monitoring of Anaesthetic Action

To date depth of anaesthesia monitoring has relied on a range of heuristic data driven approaches to objectively define optimal levels of hypnosis such that intraoperative awareness is minimised. The most successful of these approaches are arguably those that are based upon the analysis of spontaneous or time locked electroencephalographic activity. Of these approaches the Bispectral Index in particular has become commonplace in clinical anaesthesia [14]. However its use occurs in the context of a number of well documented limitations (i) not all hypnotic agents are reliably detected (e.g. nitrous oxide and the short acting synthetic opioids being quintessential examples), and (ii) the index admits of no clear physiological interpretation as it has been constructed to act as a quantitative surrogate for an ostensibly subjective state.

Given that the model of Liley can offer potential explanation for the electrorhythmogenesis of the resting EEG and its perturbation by a range of factors that include sedative and anaesthetic agents, it may have some utility in monitoring the cerebral effects of general anaesthesia and thus resolve some of the uncertainties associated with the use of the Bispectral Index. While in principle it is possible to estimate parameters of the Liley model on the basis of real data, and to investigate how they correlate with anaesthesia, practically the difficulties are substantial given the model’s non-linear partial differential formulation and its high dimensional parameter space. Fortunately many of the model’s important qualitative properties can be understood through a linearisation of the form [41, 43]

$$H_e(k, \omega) = g(h_e^*, q') \frac{\omega^M + \sum_{m=1}^M b_m(k; h_e^*, q)\omega^{M-k}}{\omega^N + \sum_{n=1}^N a_n(k; h_e^*, q)\omega^{N-n}} P_{ee}(k, \omega) \quad (14.13)$$

$$= g(h_e^*, q) G_e(k, \omega; h_e^*, q) P_{ee}(k, \omega) \quad (14.14)$$

where it has been assumed that only the excitatory input to the excitatory cortical neuronal population is non-zero. $H_e(k, \omega)$ and $P_{ee}(k, \omega)$ are the Fourier transforms of $h_e(r, t)$ and $p_{ee}(r, t)$ respectively. $G_e(k, \omega; q)$ is defined as the *electrocortical transfer function* and arises from the linearisation about a spatially homogeneous stable singular point h_e^* for a given set of model parameters q , and $g(h_e^*, q')$ represents a factored out common term depending on a subset of model parameters $q' \in q$. By assuming $\gamma_{lk} \equiv \gamma_l$ and $\Lambda_{ek} \equiv \Lambda_e$, N and M can be set to 8 and 5 respectively. Such a linearisation reveals that under physiologically plausible parameterisations $G_e(k, \omega)$ gives resonances corresponding to all the major EEG frequency bands [41]. Thus such a linearisation implies that resting EEG can theoretically be understood as arising from a filtered spatio-temporal random process. This is of particular physiological relevance when it is considered that EEG during rest and anaesthesia is typically found to be indistinguishable from a white-noise process [36, 70, 71]. On this basis Eq. (14.13) suggests a quite specific signal processing strategy by which to estimate $G_e(k, \omega)$ and changes in P_e . By assuming (i) a matching of poles and zeros in transforming from the continuous to discrete time domains and (ii) a restricted range of wavenumbers k over which physiologically relevant model linear EEG activity occurs Eq. (14.13) can be rewritten in the discrete time domain as [43]

$$H_e(z) = k_d q(q') \frac{1 + \sum_{k=1}^{k=5} b_k(q) z^{-k}}{1 + \sum_{k=1}^{k=8} a_k(q) z^{-k}} P(z) \quad (14.15)$$

where k_d is a constant required to match the gain in going from continuous to discrete time and $z = e^{i\omega/f_s}$ with f_s being the sampling frequency. By assuming $P(z)$ describes a band-limited white-noise process Eq. (14.15) can be written as the following fixed-order autoregressive moving average (ARMA) process

$$h_e[n] = - \sum_{k=1}^{k=8} a_k h_e[n-k] + \sum_{k=0}^{k=5} b_k u[n-k], \text{ or} \quad (14.16)$$

$$A(z)h_e[n] = B(z)u[n] \quad (14.17)$$

where $u[n] \equiv k_d q(q') p[n]$ is a stationary uncorrelated random process and $A(z) = 1 + \sum_{k=1}^{k=8} a_k z^{-k}$ and $B(z) = 1 + \sum_{k=1}^{k=5} b_k z^{-k}$. Tracking the state of this estimated electrocortical filter and its innovating input will provide one possible measure for characterising the cortical effects of anaesthesia. One easily calculated measure of the *state* of the estimated electrocortical filter is the scaled mean pole location, a_1 . The innovating *input* can be estimated, by assuming that the factor $k_d g(q')$ remains invariant to any intervention, as $\sqrt{\text{Var}[Z^{-1}\{A(z)S(z)/B(z)\}]}$ where Z^{-1} is the inverse Z -transform. These respective measures are referred to as the *Cortical State* (CS) and *Cortical Input* respectively, and have been useful in differentiating the electroencephalographic effects of the hypnotic propofol and the analgesic remifentanil (an ultra-short acting synthetic opioid). Figure 14.7 illustrates the

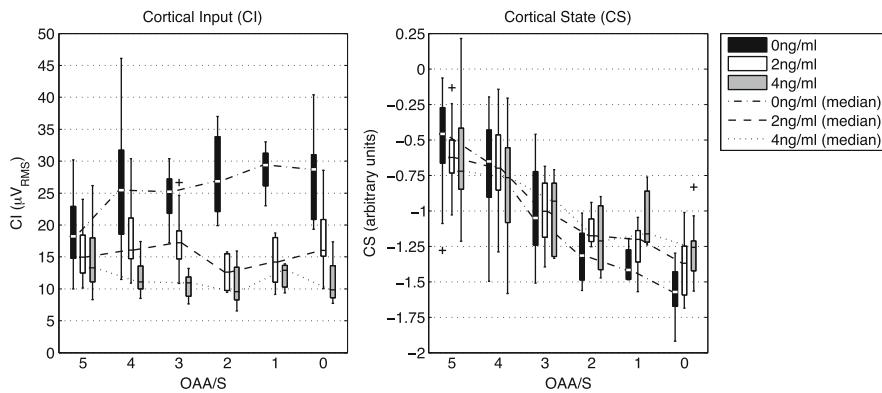


Fig. 14.7 Box-and-whisker plots for the derived electroencephalographic measures CI and CS, during propofol-remifentanil anaesthesia, as a function of the modified Observer's Assessment of Alertness/Sedation (OAA/S) score for 0, 2 and 4 ng/ml effect site remifentanil concentrations. The modified OAA/S scale is a measure of alertness and sedation: 5 = responds readily to name, 4 = responds lethargically to name, 3 = responds to name called loudly and repeatedly, 2 = responds to mild prodding/shaking, 1 = responds only after painful stimulus, 0 = completely unresponsive. Boxes represent interquartile ranges, lines enclosed within boxes median values, whiskers represent largest/smallest values and crosses outliers (Figure adapted from [44])

differences in the response of CS and CI as a function of the Observers Assessment of Alertness/Sedation (OAA/S) (OAA/S = 5 is fully responsive, OAA/S = 0 is completely unresponsive) and the level of analgesia (0, 2 or 4 ng/ml remifentanil). On this basis we can speculate that CS is a measure of *hypnosis* and CI is a measure of *nociception*.

14.5 Conclusion

The aim of this chapter has been to give an account of a relatively simple neural field model of the resting EEG and to briefly illustrate some of its more interesting dynamical features as well as speculating on its physiological relevance in accounting for resting alpha band activity and its perturbation by a range of endogenous and exogenous pharmacological factors. In particular we discussed the predicted sensitivity of model dynamics to differential perturbations in cortical inhibition and how this might account for the electroencephalographic effects of anesthetics that act principally through GABAergic agonism. We argued that to first approximation anesthetics alter noise-driven linear properties of the resting EEG. However it is known that anesthetics are also able to induce quite profound qualitative alterations in the EEG. At high levels of many sedative and anaesthetic agents the EEG can exhibit burst suppression. Typically the burst suppression pattern consists of bursts of high amplitude slow, sharp or spiking activity separated

by periods of near iso-electricity (suppression) [54]. As anaesthetic depth increases the periods of burst become shorter. While it is often assumed that the burst-suppression pattern arises from slow thalamic oscillations driving cortex, the fact that the pattern survives following cortical deafferentation suggests that it arises as a consequence of intrinsic dynamical properties of cortical tissue. Therefore the challenge of ours, and similar models, is to account for this phenomenon in the context of a plausible physiological framework and the known molecular and cellular targets associated with anaesthetic action.

There is an emerging practical utility for mean field models as evinced by our simple fixed order ARMA approach. While this single-electrode approach can, at least in principle, be easily extended to the multi-electrode case by the suitable definition of a multivariate (vector) ARMA model the real challenge is to estimate actual model parameters from empirical data and to see if they accord with known physiology. At present the action of anaesthetic agents would appear to provide the most robust context in which to estimate model parameters as they can be directly associated with known cellular and molecular targets of action.

References

1. Adrian, E., Matthews, B.: The berger rhythm, potential changes from the occipital lobe in man. *Brain* **57**, 355–385 (1934)
2. Alkire, M.T.: Probing the mind: anesthesia and neuroimaging. *Clin. Pharmacol. Ther.* **84**(1), 149–152 (2008)
3. Alkire, M.T., Haier, R.J., Barker, S.J., Shah, N.K., Wu, J.C., Kao, Y.J.: Cerebral metabolism during propofol anesthesia in humans studied with positron emission tomography. *Anesthesiology* **82**(2), 393–403 (1995)
4. Amari, S.: Homogeneous nets of neuron-like elements. *Biol. Cybern.* **17**, 211–220 (1975)
5. Andersen, P., Andersson, S.: *Physiological Basis of the Alpha Rhythm*. Appelton-Century-Crofts, New York (1968)
6. Baker, F.C., Colrain, I.M.: Daytime sleepiness, psychomotor performance, waking EEG spectra and evoked potentials in women with severe premenstrual syndrome. *J. Sleep Res.* **19**(1 Pt 2), 214–227 (2010)
7. Banks, M.I., Pearce, R.A.: Dual actions of volatile anesthetics on GABA(A) IPSCs: dissociation of blocking and prolonging effects. *Anesthesiology* **90**(1), 120–134 (1999)
8. Barth, A.L., Poulet, J.F.: Experimental evidence for sparse firing in the neocortex. *Trends Neurosci.* **35**(6), 345–355 (2012)
9. Becker, D., Creutzfeldt, O.D., Schwibbe, M., Wuttke, W.: Changes in physiological, EEG and psychological parameters in women during the spontaneous menstrual cycle and following oral contraceptives. *Psychoneuroendocrinology* **7**(1), 75–90 (1982)
10. Beurle, R.: Properties of a mass of cells capable of regenerating pulses. *Phil. Trans. R. Soc. B* **240**, 55–94 (1956)
11. Bojak, I., Liley, D.: Modeling the effects of anesthesia on the electroencephalogram. *Phys. Rev. E* **71**, 041902 (2005)
12. Borst, J.G.: The low synaptic release probability in vivo. *Trends Neurosci.* **33**(6), 259–266 (2010)
13. Braintenberg, V., Schüz, A.: *Cortex: Statistics and Geometry of Neuronal Connectivity*, 2nd edn. Springer, New York (1998)

14. Bruhn, J., Myles, P.S., Sneyd, R., Struys, M.M.: Depth of anaesthesia monitoring: what's available, what's validated and what's next? *Br. J. Anaesth.* **97**(1), 85–94 (2006)
15. Buente, D., Frascoli, F., Liley, D.: Complex dynamics for a reduced model of human eeg: implications for the physiological basis of brain activity. *BMC Neurosci.* **12**, 198 (2011)
16. Canavier, C.C., Baxter, D.A., Clark, J.W., Byrne, J.H.: Nonlinear dynamics in a model neuron provide a novel mechanism for transient synaptic inputs to produce long-term alterations of postsynaptic activity. *J. Neurophysiol.* **69**(6), 2252–2257 (1993)
17. Caton, R.: The electric currents of the brain. *Br. Med. J.* **2**, 278 (1875)
18. Ching, S., Cimenser, A., Purdon, P.L., Brown, E.N., Kopell, N.J.: Thalamocortical model for a propofol-induced alpha-rhythm associated with loss of consciousness. *Proc. Natl. Acad. Sci. U.S.A.* **107**(52), 22665–22670 (2010)
19. Creutzfeldt, O.D., Arnold, P.M., Becker, D., Langenstein, S., Tirsch, W., Wilhelm, H., Wuttke, W.: EEG changes during spontaneous and controlled menstrual cycles and their correlation with psychological performance. *Electroencephalogr. Clin. Neurophysiol.* **40**(2), 113–131 (1976)
20. Dafilis, M., Frascoli, F., Cadusch, P., Liley, D.: Chaos and generalised multistability in a mesoscopic model of the electroencephalogram. *Physica D* **13**, 1056–1060 (2009)
21. David, O., Kiebel, S., Harrison, L., Mattout, J., Kilner, J., Friston, K.: Dynamic causal modeling of evoked responses in EEG and MEG. *NeuroImage* **30**, 1255–1272 (2006)
22. Deco, G., Jirsa, V., Robinson, P., Breakspear, M., Friston, K.: The dynamic brain: from spiking neurons to neural masses and cortical fields. *PLoS Comput. Biol.* **4**, e1000092 (2008)
23. De-Miguel, F.F., Fuxé, K.: Extrasynaptic neurotransmission as a way of modulating neuronal functions. *Front. Physiol.* **3**, 16 (2012)
24. Franks, N.P.: General anaesthesia: from molecular targets to neuronal pathways of sleep and arousal. *Nat. Rev. Neurosci.* **9**, 370–386 (2008)
25. Frascoli, F., van Veen, L., Bojak, I., Liley, D.: Metabifurcation analysis of a mean field model of the cortex. *Physica D* **240**, 949–62 (2011)
26. Freeman, W.: *Mass Action in the Nervous System: Examination of the Neurophysiological Basis of Adaptive Behavior Through the EEG*, 1st edn. Academic, New York (1975). Also electronic edn.: <http://sulcus.berkeley.edu/MANSWW/MANSWW.html> (2004)
27. Freyer, F., Aquino, K., Robinson, P.A., Ritter, P., Breakspear, M.: Bistability and non-Gaussian fluctuations in spontaneous cortical activity. *J. Neurosci.* **29**(26), 8512–8524 (2009)
28. Freyer, F., Roberts, J.A., Becker, R., Robinson, P.A., Ritter, P., Breakspear, M.: Biophysical mechanisms of multistability in resting-state cortical rhythms. *J. Neurosci.* **31**(17), 6353–6361 (2011)
29. Fries, P.: Neuronal gamma-band synchronization as a fundamental process in cortical computation. *Annu. Rev. Neurosci.* **32**, 209–224 (2009)
30. Fritschy, J.M., Mohler, H.: GABA_A-receptor heterogeneity in the adult rat brain: differential regional and cellular distribution of seven major subunits. *J. Comp. Neurol.* **359**(1), 154–194 (1995)
31. Giaume, C., Koulakoff, A., Roux, L., Holcman, D., Rouach, N.: Astroglial networks: a step further in neuroglial and gliovascular interactions. *Nat. Rev. Neurosci.* **11**, 87–99 (2010)
32. Gloor, P.: Hans Berger on the electroencephalogram of man. *Electroencephalogr. Clin. Neurophysiol.* **S28**, 350 (1969)
33. Hosie, A.M., Clarke, L., da Silva, H., Smart, T.G.: Conserved site for neurosteroid modulation of GABA A receptors. *Neuropharmacology* **56**(1), 149–154 (2009)
34. Hutt, A. (ed.): *Sleep and Anesthesia: Neural Correlates in Theory and Experiment*. Springer Series in Computational Neuroscience. Springer, New York (2011)
35. Jansen, B.H., Rit, V.G.: Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biol. Cybern.* **73**(4), 357–366 (1995)
36. Jeleazcov, C., Fechner, J., Schwilden, H.: Electroencephalogram monitoring during anesthesia with propofol and alfentanil: the impact of second order spectral analysis. *Anesth. Analg.* **100**(5), 1365–1369 (2005)

37. Jirsa, V., Haken, H.: Field theory of electromagnetic brain activity. *Phys. Rev. Lett.* **77**, 960–963 (1996)
38. Lam, D.W., Reynolds, J.N.: Modulatory and direct effects of propofol on recombinant GABA_A receptors expressed in xenopus oocytes: influence of alpha- and gamma2-subunits. *Brain Res.* **784**(1–2), 179–187 (1998)
39. Langsjo, J.W., Alkire, M.T., Kaskinoro, K., Hayama, H., Maksimow, A., Kaisti, K.K., Aalto, S., Aantaa, R., Jaaskelainen, S.K., Revonsuo, A., Scheinin, H.: Returning from oblivion: imaging the neural core of consciousness. *J. Neurosci.* **32**(14), 4935–4943 (2012)
40. Liley, D., Cadusch, P., Wright, J.: A continuum theory of electro-cortical activity. *Neurocomputing* **26–27**, 795–800 (1999)
41. Liley, D., Cadusch, P., Dafilis, M.: A spatially continuous mean field theory of electrocortical activity. *Netw.: Comput. Neural Syst.* **13**, 67–113 (2002)
42. Liley, D., Cadusch, P., Dafilis, M.: Corrigendum. *Netw.: Comput. Neural Syst.* **14**, 369 (2003)
43. Liley, D., Cadusch, P., Gray, M., Nathan, P.: Drug-induced modification of the system properties associated with spontaneous human electroencephalographic activity. *Phys. Rev. E* **68**, 051906 (2003)
44. Liley, D., Sinclair, N., Lipping, T., Heyse, B., Vereecke, E., Struys, M.: Propofol and remifentanil differentially modulate frontal electroencephalographic activity. *Anesthesiology* **113**, 1–13 (2010)
45. Liley, D., Foster, B., Bojak, I.: A mesoscopic modelling approach to characterising anaesthetic action on brain electrical activity. In: Hutt, A. (ed.) *Sleep and Anesthesia: Neural Correlates in Theory and Experiment*. Springer Series in Computational Neuroscience, pp. 139–66. Springer, New York (2011)
46. Liley, D., Foster, B., Bojak, I.: Co-operative populations of neurons: mean field models of mesoscopic brain activity. In: Le Novère, N. (ed.) *Computational Systems Neurobiology*, pp. 317–64. Springer, New York (2012)
47. Llinas, R.: The intrinsic electrophysiological properties of mammalian neurons: insights into central nervous system function. *Science* **242**, 1654–1664 (1988)
48. Lopes da Silva, F.: Dynamics of EEGs as signals of neuronal populations: models and theoretical considerations. In: Niedermeyer, E., Lopes da Silva, F. (eds.) *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, pp. 85–106, 5th edn. Lippincott Williams & Wilkins, Philadelphia (2005)
49. MacIver, M.B., Mikulec, A.A., Amagasu, S.M., Monroe, F.A.: Volatile anesthetics depress glutamate transmission via presynaptic actions. *Anesthesiology* **85**(4), 823–834 (1996)
50. Miller, K.D., Pinto, D.J., Simons, D.J.: Processing in layer 4 of the neocortical circuit: new insights from visual and somatosensory cortex. *Curr. Opin. Neurobiol.* **11**(4), 488–497 (2001)
51. Mohler, H., Fritschy, J.M., Rudolph, U.: A new benzodiazepine pharmacology. *J. Pharmacol. Exp. Ther.* **300**(1), 2–8 (2002)
52. Mozzachiodi, R., Byrne, J.H.: More than synaptic plasticity: role of nonsynaptic plasticity in learning and memory. *Trends Neurosci.* **33**(1), 17–26 (2010)
53. Niedermeyer, N.: The normal EEG of the waking adult. In: Niedermeyer, E., Lopes da Silva, F. (eds.) *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, pp. 167–192, 5th edn. Lippincott Williams & Wilkins, Philadelphia (2005)
54. Niedermeyer, E.: The burst-suppression electroencephalogram. *Am. J. Electroneurodiagnostic Technol.* **49**(4), 333–341 (2009)
55. Niedermeyer, E., Lopes da Silva, F. (eds.): *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, 5th edn. Lippincott Williams & Wilkins, Philadelphia (2005)
56. Nunez, P.: The brain wave equation: a model for the EEG. *Math. Biosci.* **21**, 279–297 (1974)
57. Nunez, P., Srinivasan, R.: *Electric Fields of the Brain: The Neurophysics of EEG*, 2nd edn. Oxford University Press, New York (2005)
58. Nutt, D.: GABA_A receptors: subtypes, regional distribution, and function. *J. Clin. Sleep Med.* **2**(2), 7–11 (2006)
59. Olsen, R.W., Sieghart, W.: International Union of Pharmacology. LXX. Subtypes of gamma-aminobutyric acid(A) receptors: classification on the basis of subunit composition, pharmacology, and function. *Update. Pharmacol. Rev.* **60**(3), 243–260 (2008)

60. Partadiredja, G., Miller, R., Oorschot, D.: The number, size, and type of axons in rat subcortical white matter on left and right sides: a stereological, ultrastructural study. *J. Neurocytol.* **32**, 1165–1179 (2003)
61. Pitot, M., Gastaut, H.: [Electroencephalographic modifications during the estrus cycle]. *Rev. Neurol. (Paris)* **89**(5), 427–430 (1953)
62. Reddy, D.S.: Neurosteroids: endogenous role in the human brain and therapeutic potentials. *Prog. Brain Res.* **186**, 113–137 (2010)
63. Regan, D.: Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic Fields in Science and Medicine. Elsevier, New York (1989)
64. Ritchie, J.: Physiology of axons. In: Waxman, S., Kocsis, J., Stys, P. (eds.) *The Axon: Structure, Function and Pathophysiology*, pp. 68–96. Oxford University Press, New York (1995)
65. Robinson, P., Rennie, C., Wright, J.: Propagation and stability of waves of electrical activity in the cerebral cortex. *Phys. Rev. E* **56**, 826–840 (1997)
66. Robinson, P.A., Rennie, C.J., Rowe, D.L., O'Connor, S.C.: Estimation of multiscale neurophysiologic parameters by electroencephalographic means. *Hum. Brain Mapp.* **23**(1), 53–72 (2004)
67. Rudolph, U., Antkowiak, B.: Molecular and neuronal substrates for general anaesthetics. *Nat. Rev. Neurosci.* **5**, 709–720 (2004)
68. Sherman, S., Guillery, R.: Exploring the thalamus and its role in cortical function. MIT, Cambridge (2005)
69. Sohl, G., Maxeiner, S., Willecke, K.: Expression and functions of neuronal gap junctions. *Nat. Rev. Neurosci.* **6**(3), 191–200 (2005)
70. Stam, C.: Nonlinear dynamical analysis of EEG and MEG: review of an emerging field. *Clin. Neurophysiol.* **116**, 2266–2301 (2005)
71. Stam, C., Pijn, J., Suffczyński, P., Lopes da Silva, P.: Dynamics of the human alpha rhythm: evidence for non-linearity? *Clin. Neurophysiol.* **110**, 1801–1813 (1999)
72. Steyn-Ross, D., Steyn-Ross, M., Sleigh, J., Wilson, M.: Progress in modelling EEG effects of general anesthesia: biphasic response and hysteresis. In: Hutt, A. (ed.) *Sleep and Anesthesia: Neural Correlates in Theory and Experiment*. Springer Series in Computational Neuroscience, pp. 167–94. Springer, New York (2011)
73. Stys, P.K.: The axo-myelinic synapse. *Trends Neurosci.* **34**(8), 393–400 (2011)
74. Terzuolo, C.A., Bullock, T.H.: Measurement of imposed voltage gradient adequate to modulate neuronal firing. *Proc. Natl. Acad. Sci. U.S.A.* **42**(9), 687–694 (1956)
75. van Veen, L., Liley, D.: Chaos via Shilnikov's saddle-node bifurcation in a theory of the electroencephalogram. *Phys. Rev. Lett.* **97**, 208101 (2006)
76. Wendling, F., Bartolomei, F., Bellanger, J.J., Chauvel, P.: Epileptic fast activity can be explained by a model of impaired GABAergic dendritic inhibition. *Eur. J. Neurosci.* **15**(9), 1499–1508 (2002)
77. Wilson, H., Cowan, J.: Excitatory and inhibitory interactions in localized populations of model neuron. *Biophys. J.* **12**, 1–24 (1972)
78. Wilson, H., Cowan, J.: A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* **13**, 55–80 (1973)

Chapter 15

Equilibrium and Nonequilibrium Phase Transitions in a Continuum Model of an Anesthetized Cortex

D. Alistair Steyn-Ross, Moira L. Steyn-Ross, and Jamie W. Sleigh

Abstract In this chapter we investigate a range of dynamic behaviors accessible to a continuum model of the cerebral cortex placed close to the anesthetic phase transition. If the anesthetic transition from the high-firing (conscious) to the low-firing (comatose) state can be modeled as a jump between two equilibrium states of the cortex, then we can draw an analogy with the vapor-to-liquid phase transition of the van der Waals gas of classical thermodynamics. In this analogy, specific volume (inverse density) of the gas maps to cortical activity, with pressure and temperature being the analogs of anesthetic concentration and subcortical excitation. It is well known that at the thermodynamic critical point, large fluctuations in specific volume are observed; we find analogous critically-slowed fluctuations in cortical activity at its critical point. Unlike the van der Waals system, the cortical model can also exhibit nonequilibrium phase transitions in which the homogeneous equilibrium can destabilize in favor of slow global oscillations (Hopf temporal instability), stationary structures (Turing spatial instability), and chaotic spatiotemporal activity patterns (Hopf–Turing interactions). We comment on possible physiological and pathological interpretations for these dynamics. In particular, the turbulent state may correspond to the cortical slow oscillation between “up” and “down” states observed in nonREM sleep and clinical anesthesia.

D.A. Steyn-Ross (✉) • M.L. Steyn-Ross

School of Engineering, University of Waikato, P.B. 3105, Hamilton 3240, New Zealand
e-mail: asr@waikato.ac.nz; msr@waikato.ac.nz

J.W. Sleigh

Waikato Clinical School, University of Auckland, Hamilton 3204, New Zealand
e-mail: Jamie.Sleigh@waikatodhb.health.nz

15.1 Introduction

Over the past decade the authors have presented a series of articles [26–28, 31] describing a theoretical model for the bulk electrodynamical behavior of the cortex undergoing a general anesthetic-induced phase transition. We picture the cortex as a continuum mass of excitable tissue in which populations of neurons are densely interlinked not only by *chemical* synapses but also by direct *electrical* connections known as *gap junctions*. The time evolution of the spatially-averaged excitatory and inhibitory soma voltages are governed by a set of eight coupled partial differential equations; full details are provided in the Appendix.

general anesthetic drugs, such as the commonly-used agent propofol, act by prolonging the duration of the inhibitory postsynaptic potential (IPSP), thus boosting the strength of chemical inhibition in a concentration-dependent manner [7, 14], and, at higher concentrations, can reduce the strength of inhibitory gap-junction coupling [33]. The drug-induced boost in chemical inhibition tends to lower the overall level of cortical activity. When incorporated into our cortical model, we find that this reduction in cortical tone can either be smooth and continuous, or abrupt and step-like, depending on the relative balance of inhibitory drive versus background excitation from subcortical sources.

These alternative descents into anesthesia—either smooth or abrupt—can be visualized by tracing them onto a manifold of equilibrium states obtained by solving the cortical equations for their homogenous steady-state values. This means that we set all time- and space-dependencies to zero, so for the moment there is no noise, no gap-junction diffusion, no wave propagations through the cortical continuum, and no point-to-point long-range axonal connections. Although highly idealized, this homogeneous steady-state cortex provides a useful reference which guides our analysis by allowing us to investigate the conditions under which interesting spatiotemporal dynamics—generated by interacting spatial and temporal instabilities—can emerge spontaneously when the time- and space-dependencies are re-enabled.

Figure 15.1a shows a representative manifold of equilibrium states for our idealized homogeneous cortex; these equilibria are distributed across a two-dimensional “anesthesia domain” with inhibition (anesthetic drug effect) and excitation (subcortical tone) displayed on mutually-perpendicular axes. The presence of a reentrant “fold” indicates that, for certain combinations of inhibition and excitation, the cortex has access to three alternative states of cortical activity: a high-firing “up” state, a low-firing “down” state, and an intermediate unstable mode that serves as a separatrix between these activated and quiescent states.

The form of the equilibrium manifold for the cortex (Fig. 15.1a) invites an analogy with thermodynamic states of matter (Fig. 15.1b). For a pure substance, the gaseous, liquid, and solid states can be located on a pressure–volume–temperature (P – v – T) thermodynamic phase diagram of equilibrium states: the specific volume v (volume per unit mass; inverse density) is determined by both temperature T and applied pressure P . The isotherms (lines of constant temperature) drawn on

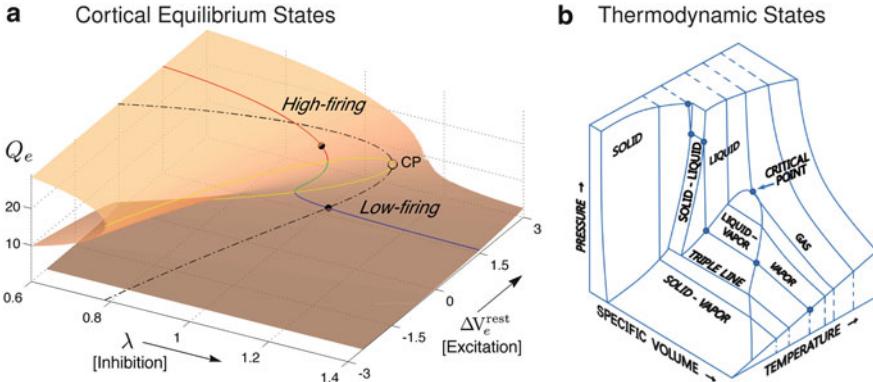


Fig. 15.1 Analogy between cortical equilibrium states and thermodynamic states of matter. **(a)** Distribution of equilibrium firing rates for a homogeneous noiseless cortex as a multi-valued function of inhibitory and excitatory drives. Control parameter λ sets the anesthetic effect; ΔV^{rest} is an additive offset representing background cortical excitation. Yellow curve marks the edge of the reentrant fold; dashed-black curve shows the projection of this edge onto the lower and upper surfaces, bounding the zone of multiple steady states. Red-green-blue curve shows distribution of steady states for varying anesthetic inhibition at constant cortical excitation $\Delta V^{\text{rest}} = 1.5 \text{ mV}$ (see green curve of Fig. 15.2a). CP marks the critical point at which high- and low-firing states become indistinguishable. **(b)** Classical pressure-temperature-volume (p - v - T) thermodynamic phase diagram showing solid, liquid and gas phases of a pure substance. CP marks the critical point at which gas and liquid become indistinguishable (Adapted from Fig. 2–6 of Sears and Salinger [23])

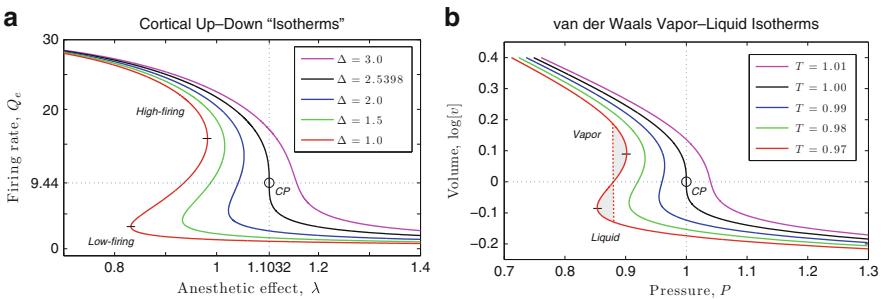


Fig. 15.2 Comparison between **(a)** up/down equilibrium states of the cortex, and **(b)** vapor/liquid states for a van der Waals gas. **(a)** Firing rate Q_e versus anesthetic inhibition λ for stepped values of ΔV^{rest} (abbreviated as Δ in the legend). CP marks the critical point of inflection at $(\lambda, Q_e/\text{s}^{-1}, \Delta V^{\text{rest}}/\text{mV}) \approx (1.1032, 9.44, 2.5398)$. **(b)** Specific volume v versus pressure P for stepped values of temperature T for a van der Waals gas. The three state variables have been scaled so that the van der Waals critical point CP is located at $(P_c, v_c, T_c) = (1, 1, 1)$; a \log_{10} scale has been used for v . The red-dashed vertical line for the $T = 0.97$ isotherm illustrates the kind of Maxwell construction used to “flatten” the vapor-liquid coexistence isotherms in Fig. 15.1b by equalizing the areas above and below the Maxwell pressure line

Fig. 15.1b show how a change of phase from vapor to liquid, and then from liquid to solid, can be organized by increasing pressure while maintaining constant temperature. Thus changes in pressure can result in phase transitions between distinct states of matter. If we make the identifications

$$\begin{aligned} \text{pressure, } P &\leftrightarrow \text{anesthetic inhibition, } \lambda \\ \text{temperature, } T &\leftrightarrow \text{subcortical excitation, } \Delta V_e^{\text{rest}} \\ \text{specific volume, } v &\leftrightarrow \text{cortical firing rate, } Q_e \end{aligned}$$

then compression along a cortical “isotherm” corresponds to suppression of cortical activity via increase in anesthetic inhibition (at constant subcortical excitation), causing a phase transition from the high-firing “up” state to the low-firing “down” state; this transition can be either abrupt and first-order (e.g., $\Delta = 1.5 \text{ mV}$) in Fig. 15.2a, or smoothly continuous (e.g., $\Delta = 3.0 \text{ mV}$) and second-order. Thus one might identify the low-density state of thermodynamics (vapor) with the activated up-state of the cortex,¹ and the high-density thermodynamic state (liquid) with the quiescent down-state,

$$\begin{aligned} \text{vapor} &\leftrightarrow \text{high-firing up-state of cortex} \\ \text{liquid} &\leftrightarrow \text{low-firing down-state} \end{aligned}$$

The reentrant fold in the Fig. 15.1a distribution of cortical equilibria demarcates the region in which the cortex can exist in either the high- or low-firing state; the mid-branch equilibrium states are unstable. The analogous vapor–liquid coexistence region in the Fig. 15.1b thermodynamic diagram has been simplified by collapsing what would have been reentrant p -vs- v isotherms onto horizontal lines of constant pressure. This simplification is achieved by applying a “Maxwell construction”² to determine an effective pressure that equalizes the area deviations of the curved isotherm above and below the constructed line (see Fig. 15.2b for an illustration).

The indicative vapor–liquid isocones in Fig. 15.1b can be made quantitative using van der Waals 1873 gas model which provides a reasonable description of the condensation behavior of near-ideal gases such as hydrogen and helium,

$$\left(P + \frac{a}{v^2} \right) (v - b) = RT \quad (15.1)$$

where a and b are constants for any given gas, and R is the universal gas constant. The (P_c, v_c, T_c) coordinate of the critical point of Fig. 15.1b marks the location at

¹Of course, the (specific volume) \equiv (firing rate) analogy is not perfect: the volume of a gas can increase without limit, but cortical firing rate is limited by biological constraints, implemented in the model by imposing a maximum firing rate Q_e^{\max} (see Table 15.1).

²See http://en.wikipedia.org/wiki/Maxwell_construction

Table 15.1 Symbol definitions and standard values for cortical model

Symbol	Description	Value	Unit
$\tau_{e,i}$	Neuron time constant	0.040, 0.040	s
$V_{e,i}^{\text{rev}}$	Reversal potential at dendrite	0, -70	mV
$V_{e,i}^{\text{rest}}$	Neuron resting potential	-64, -64	mV
$\Delta V_{e,i}^{\text{rest}}$	Offset to resting potential	1.5, 0	mV
ρ_e	Excitatory synaptic gain	1.00×10^{-3}	mV s
ρ_i^0	Inhibitory synaptic gain at zero anesthetic	-1.05×10^{-3}	mV s
γ_e	Excitatory rate-constant	170	s^{-1}
γ_i^0	Inhibitory rate-constant at zero anesthetic	90–100	s^{-1}
D_2	$i \leftrightarrow i$ gap-junction diffusive coupling strength	0–0.26	cm^2
D_1	$e \leftrightarrow e$ gap-junction diffusive coupling strength	$D_2/100$	cm^2
N_{eb}^α	Longer-range $e \rightarrow b$ axonal connectivity	2000	—
$N_{eb,ib}^\beta$	Local $e \rightarrow b, i \rightarrow b$ axonal connectivity	800, 600	—
$\langle \phi_{eb}^{\text{sc}} \rangle$	$e \rightarrow b$ tonic flux entering from subcortex	300	s^{-1}
α	Subcortical noise scale-factor	0.2	—
v	Axonal conduction speed	140	cm s^{-1}
Λ_{eb}	Inverse-length scale for $e \rightarrow b$ axonal connections	4	cm^{-1}
$Q_{e,i}^{\max}$	Maximum firing rate	30, 60	s^{-1}
$\theta_{e,i}$	Sigmoid threshold voltage	-58.5, -58.5	mV
$\sigma_{e,i}$	Standard deviation for threshold	3, 5	mV
$L_{x,y}$	Length and width of cortical sheet	25, 25	cm

Subscript label b means destination cell can be either of type e (excitatory) or i (inhibitory)

which liquid and vapor states become indistinguishable. After rescaling pressure, volume, and temperature as multiples of their respective critical values,

$$P' = P/P_c, \quad v' = v/v_c, \quad T' = T/T_c$$

the van der Waals equation of state can be restated in the dimensionless form [23],

$$\left(P' + \frac{3}{v'^2} \right) (3v' - 1) = 8T'. \quad (15.2)$$

This reduced form is known as the *law of corresponding states* that applies to any van der Waals gas. To simplify notation, we now drop the primes. Figure 15.2b plots Eq. (15.2) for a family of five van der Waals isotherms bracketing the $(P_c, v_c, T_c) = (1, 1, 1)$ critical point. When plotted as a P -vs- v graph, the slope of the T_c isotherm at the critical point is zero, and this is also a point of inflection,

$$\left(\frac{\partial P}{\partial v} \right)_T \Big|_{(P_c, v_c, T_c)} = 0, \quad \text{and} \quad \left(\frac{\partial^2 P}{\partial v^2} \right)_T \Big|_{(P_c, v_c, T_c)} = 0.$$

Conversely, the slope of the v -vs- P graph is infinite here, $(\partial v / \partial P)_T \rightarrow \infty$, implying that the specific volume (and its inverse, density) of the van der Waals gas will be exquisitely sensitive to minor fluctuations in pressure. Indeed, the critical point is characterized by the phenomenon of *critical opalescence* in which molecular-level fluctuations between vapor and liquid phases grow to micron length scales, causing strong scattering of light and a cloudy appearance to the mixture of phases [24].

The analogous critical point for the model cortex can be identified on the $\Delta = 2.5398$ cortical “isotherm” of Fig. 15.2a. This special point corresponds to the apex of the fold of the Fig. 15.1a manifold of steady states: here, the three firing-rate solutions merge to the single value $Q_e/s^{-1} \approx 9.44$. By analogy with the van der Waals gas, because $(\partial Q_e / \partial \lambda)_\Delta \rightarrow \infty$ here, we expect a divergent increase in cortical fluctuations as anesthetic effect approaches its critical value, $\lambda \rightarrow \lambda_c$. In fact, such divergences in cortical response are not confined to trajectories passing through the critical point: as we will demonstrate in the next section, any transition that approaches a fold in the steady-states manifold will exhibit divergent growth of critically-slowed fluctuations.

The remainder of the chapter is structured as follows. We simulate an anesthetic-induced phase transition from the up- (high-firing) to the down- (low-firing) state along a ΔV_e^{rest} “isotherm” as a possible representation of induction of anesthesia. We describe this up \rightarrow down state change as an *equilibrium phase transition* because the cortex is transiting from one homogeneous equilibrium state to another.

We then place the cortex at the critical point, and alter two parameters to demonstrate emergence of two distinct classes of symmetry-breaking *nonequilibrium* phase transitions in the cortical model. By reducing the inhibitory rate constant γ_i that governs the timeliness of the inhibitory postsynaptic potential, we can generate a temporal or *Hopf* instability that causes whole-of-cortex oscillations. Alternatively, by increasing the strength of diffusive gap-junction coupling D_2 between inhibitory neurons, we can precipitate a spatial or *Turing* instability that forms structured patterns of cortical activity. Finally, if both instabilities are allowed to interact, then slow chaotic spatiotemporal oscillations can emerge. We conclude the chapter with a discussion of possible physiological interpretations of these distinct phase transition types. In the Appendix, we list the cortical equations and tabulate the parameter values used in the model, describe the parameterization of anesthetic drug effect, and explain how the linear stability analysis was performed.

15.2 Induction of Anesthesia as an Equilibrium Phase Transition

We simulate induction of anesthesia by iterating the cortical equations (15.3)–(15.7) (listed in section “Model Equations” of Appendix) for slowly incrementing steps in drug inhibitory effect starting from $\lambda = 0.94$ on the up-state branch of the $\Delta V_e^{\text{rest}} = 2.0$ mV “isotherm” (the “ $\Delta = 2.0$ ” curve in Fig. 15.2a). We deliberately

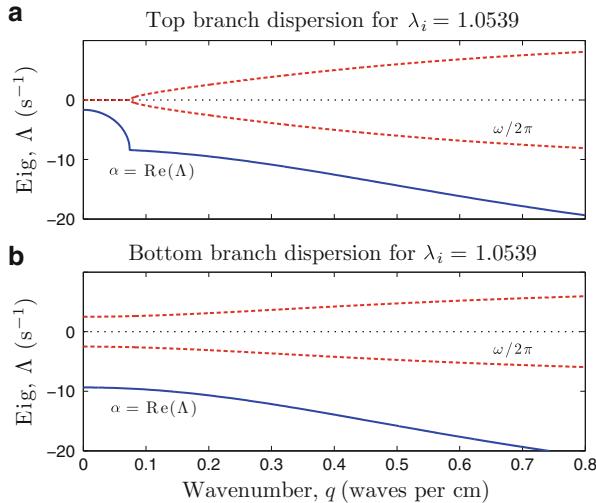


Fig. 15.3 Linear stability predictions for homogeneous cortex close to turning point near $\lambda_i = 1.0539$. Solid and dashed curves show respectively the real and imaginary parts of the dominant eigenvalue Λ as a function of wavenumber q . **(a)** Top branch is marginally stable: $\Lambda(q = 0)$ approaches zero from below, implying globally-slowed decay times $1/\Lambda$ and exaggerated low-frequency responsiveness. **(b)** Bottom branch is strongly stable: perturbations will be rapidly damped. *Settings:* $\Delta V_e^{\text{rest}} = 2.0 \text{ mV}$, $\gamma_i^0 = 100 \text{ s}^{-1}$, $D_2 = 0$

suppress the possibility of emergent temporal and spatial instabilities with appropriate choices for inhibitory postsynaptic potential (IPSP) rate constant ($\gamma_i^0 = 100 \text{ s}^{-1}$), and for inhibitory gap-junction diffusion strength ($D_2 = 0$) respectively. These parameter selections permit direct observation of the critical slowing and growing of noise-induced fluctuations about homogeneous equilibrium as the cortex gradually approaches the upper-branch turning point.

These selections were guided by analysis of the linear stability dispersion graphs for the cortical equations (see section “Linear Stability Analysis for Homogenous Stationary States” of Appendix for details). We plot the real and imaginary parts of the dominant eigenvalue, $\Lambda(q) = \alpha(q) + i\omega(q)$ as a function of wavenumber q , looking for regions for which the growth term, α , changes sign from *negative* (exponential damping: equilibrium is stable) to *positive* (predicting exponential growth of perturbations: equilibrium state has become unstable).

For our default parameter settings ($\gamma_i^0 = 100 \text{ s}^{-1}$, $D_2 = 0$), we find that all points along the $\Delta = 2.0$ “isotherm” are strongly stable (i.e., $\alpha(q)$ is significantly negative across the full range of wavenumbers), except for coordinates approaching the turning points marking the extremum for the up → down (*induction*) or the down → up (*emergence*) state changes. As illustrated in Fig. 15.3a, a homogenous equilibrium state close to the upper-branch turning point is only marginally stable, since the dominant eigenvalue is trending towards zero from below, becoming precisely zero at the turning point. This indicates that noise-induced perturbations

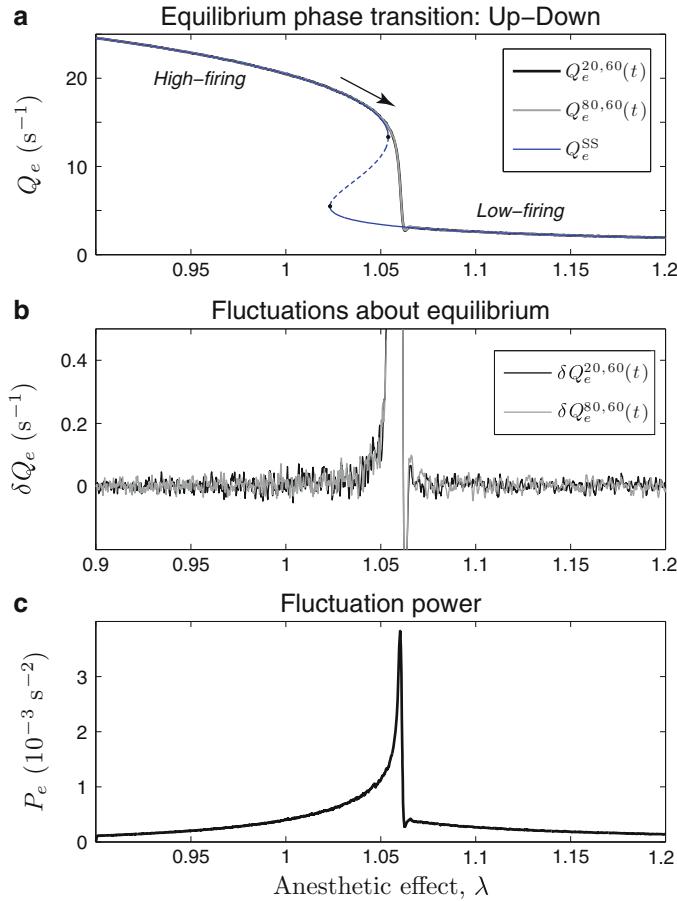


Fig. 15.4 Numerical simulation of the cortical equations showing induction of anesthesia as a phase transition between high- and low-firing equilibrium states. The cortex is represented as 25×25 -cm square of tissue mapped onto a 120×120 grid. (a) Time-series of excitatory firing rates Q_e for two positions on the cortical grid (black and gray traces) superimposed on the distribution of homogeneous steady-states Q_e^{SS} (thin-black trace). (b) Firing-rate fluctuations $\delta Q_e = Q_e - Q_e^{SS}$ for the two time-series displayed in panel-a. (c) Fluctuation power P_e computed from the spatial variance across the 120×120 cortical grid at each time-step. *Simulation settings:* Duration: 20 s; time-step $\Delta t = 0.4$ ms; stimulus: continuous low-intensity subcortical white noise

are predicted to produce fluctuation responses whose lifetimes and amplitudes will become more pronounced as increasing drug effect moves the cortex inexorably towards the turning point for anesthetic induction. Fluctuations about the up-state will grow in scale and spatial extent until eventually the entire cortex fluctuates into the low-firing down-state.

These predictions are verified in the numerical simulations reported in Fig. 15.4 showing induction of anesthesia along the $\Delta V_e^{\text{rest}} = 2$ mV “isotherm”

with $\gamma_i^0 = 100 \text{ s}^{-1}$, $D_2 = 0$. As expected, Fig. 15.4a shows that the Q_e excitatory firing rates at two representative grid-points follow closely the Q_e^{ss} homogeneous equilibrium curve right up until the turning point; then activity “falls” onto the low-firing bottom branch corresponding to the unconscious comatose state. Of greater interest are the δQ_e firing-rate *fluctuations* of Fig. 15.4b, since changes in the spectral content and amplitude of these deviations from equilibrium provide an early warning of the impending state change. The continuous low-intensity subcortical white noise provides a subtle background signal that constantly probes cortical susceptibility to random stimuli. It is evident from Fig. 15.4b that the noise-induced fluctuations simultaneously become stronger and slower as the turning point is approached, so that the transition itself can be viewed as an “infinite period” (i.e., zero frequency) giant fluctuation into the new (lower-branch) state.

We can also track the divergent growth in fluctuation power by computing the spatial variance $P_e = \text{var}[Q_e(x, y)]$ across 120×120 cortical grid: Fig. 15.4c shows that the intensity of spatial fluctuations increases by about a factor of 40 during the course of the induction. Previously [26], we have suggested that this surge in critical fluctuations might provide a natural explanation for the counterintuitive *biphasic drug effect* observed across a range of inductive anesthetic agents [16, 17]: a drug intended to suppress brain activity can provoke an *excitatory* EEG response as the point of induction is approached.

15.3 Critical Fluctuations at the “Opalescent” Point

We now place the cortex “exactly” at the critical cusp at which the three-root region collapses to a single root. This is the point labeled “CP” on Figs. 15.1a and 15.2a, and corresponds to the vapor–liquid equivalence point at coordinate $(P, v, T) = (1, 1, 1)$ for a reduced van der Waals gas (see Fig. 15.2b and Eq. (15.2)).

Unlike the van der Waals gas, the critical point for the cortical equations cannot be located analytically. Instead we must perform a two-dimensional numerical search within the three-root region of the equilibrium manifold, marching alternately along the λ and ΔV_e^{rest} orthogonal directions, eventually converging to the point where the three roots smoothly coalesce to a single solution. In fact, it is not possible to identify this critical coordinate exactly. Using MATLAB’s IEEE-754 double-precision floating-point arithmetic, the best estimate we can achieve (after several thousand binary-search iterations) for the critical coordinate is $(\lambda; \Delta; Q_e)_c \approx (1.1032\,3341\,0877\,578; 2.5397\,7874\,0236\,721; 9.4385 \pm 0.0024)$, with respective units 1, mV, s^{-1} . At this particular $(\lambda, \Delta)_c$ coordinate, we find three equilibrium firing rates: $Q_e = [9.4409, 9.4399, 9.4362] \text{ s}^{-1}$, and this spread of values provides the uncertainty estimate in $(Q_e)_c$. We find that further iterations do not improve convergence. Thus $(Q_e)_c$ is only known to a precision of 3 significant figures, 13 orders of magnitude less precise than the values for λ_c and Δ_c . This failure to locate the critical firing rate more precisely is a natural consequence of the

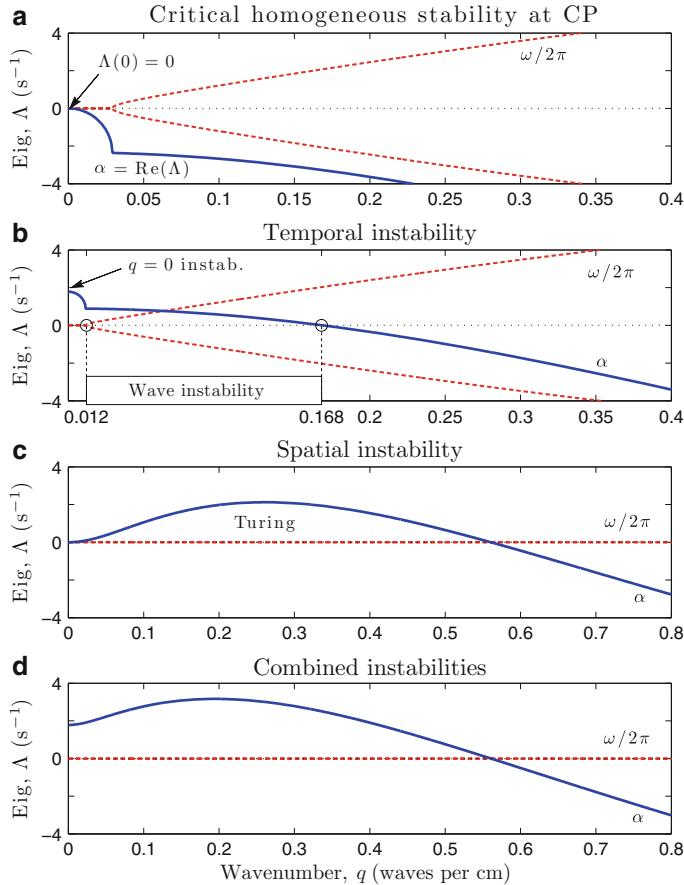


Fig. 15.5 Stability predictions at critical point CP (Fig. 15.2a) showing effect of variations in inhibitory rate-constant γ_i^0 (s^{-1}) and inhibitory diffusion D_2 (cm^2): (a) Critical homogeneous stability $(\gamma_i^0, D_2) = (100, 0)$. (b) Temporal instability $(\gamma_i^0, D_2) = (90, 0)$: all wavenumbers $q < 0.168$ cm $^{-1}$ are unstable, with frequencies ranging from 0 to 2 Hz. Outcome is a \sim 2-Hz Hopf oscillation (Fig. 15.7). (c) Spatial instability $(\gamma_i^0, D_2) = (100, 0.26)$: A band of spatial frequencies is unstable, with peak instability near 0.26 waves/cm. A stationary Turing pattern is expected (Fig. 15.8). (d) Combined $q = 0$ and Turing instabilities $(\gamma_i^0, D_2) = (90, 0.26)$. Outcome is chaotic spatiotemporal oscillations (Fig. 15.9). Solid and dashed curves show real and imaginary parts of the dominant eigenvalue respectively

manifold topology: at the critical point, the gradients $(\partial Q_e / \partial \lambda)_c$ and $(\partial Q_e / \partial \Delta)_c$ both become infinite, so equilibrium firing rates become exquisitely sensitive to tiny variations in either anesthetic pressure λ or subcortical excitation Δ .

Figure 15.5 shows the linear stability predictions for the cortical model placed “at” the $(\lambda, \Delta, Q_e)_c$ critical point. Leaving the IPSP rate-constant and inhibitory diffusion strength unchanged from their default values ($\gamma_i^0 = 100$ s $^{-1}$, $D_2 = 0$; same as for Figs. 15.3 and 15.4) produces the dispersion curves of Figs. 15.5a. The

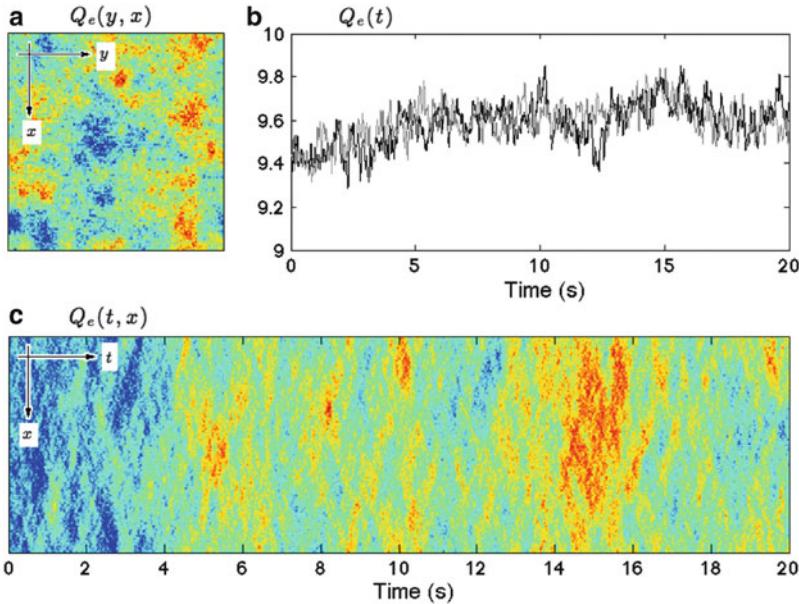


Fig. 15.6 Noise-perturbed simulations at cortical critical point (λ_i , $\Delta V_e^{\text{rest}}/\text{mV} \approx (1.1032, 2.5398)$ marked CP in Fig. 15.2). With $\gamma_i^0 = 100 \text{ s}^{-1}$ and $D_2 = 0$, Fig. 15.5a predicts that no temporal or spatial instabilities are expected. Cortical equations are iterated for 20 s. (a) Bird's-eye view of excitatory firing-rate activity Q_e across the 120×120 cortical grid at final time $t = 20$ s (red = raised activity; blue = reduced activity). (b) Time-series of cortical activity at two points $x = 20, x = 80$ down the grid midline $y = 60$. (c) Space-time strip-chart of cortical activity along the $y = 60$ midline as a function of time. Critical fluctuations show activity that is correlated over space and time

top-branch dominant eigenvalue has a real part that is almost precisely zero at zero wavenumber ($\Lambda(q = 0) \approx -5 \times 10^{-6} \text{ s}^{-1}$), and more negative elsewhere. This means that the homogeneous equilibrium state is marginally stable: noise-induced fluctuations will be very long-lived and correlated throughout the cortex.

The simulation results of Fig. 15.6 illustrate the emergence over 20 s of large, slow firing-rate fluctuations whose spatial variance, $P_e \gtrsim 4 \times 10^{-3} \text{ s}^{-2}$, exceeds the power recorded at the biphasic induction peak of Fig. 15.4c. The space–time record in Fig. 15.6c shows formation of correlated “cloudy” structures that persist over space and time that we might identify with the critical opalescence phenomenon characterizing the liquid–gas equivalence point of a van der Waals gas.

15.4 Nonequilibrium Phase Transitions in the Cortical Model

We have characterized the anesthetic state change of Sect. 15.2 as an *equilibrium* phase transition because the model cortex transits from one homogeneous stationary phase—the activated up-state—to another homogeneous phase—the comatose

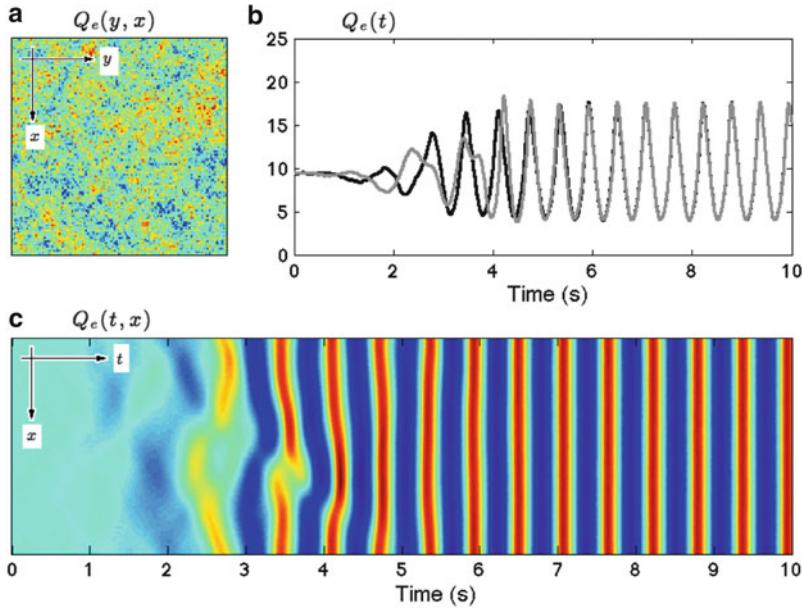


Fig. 15.7 Stochastic simulation run at critical point CP with inhibitory rate-constant lowered to $\gamma_i^0 = 90 \text{ s}^{-1}$ to precipitate a temporal instability (diffusion remains unchanged at $D_2 = 0$). (a) Bird's-eye snapshot of excitatory activity Q_e across the cortical grid at $t = 20$ s. (b) Time-series of cortical activity at two points ($x = 20, 80$ down $y = 60$ midline). (c) Space-time chart of cortical activity along the $y = 60$ midline. Figure 15.5b predicts emergence of long-wavelength low-frequency waves; these rapidly organize into synchronous ~ 1.6 -Hz whole-of-cortex Hopf oscillations

down-state; significantly, both the source and the destination states lie on the manifold of equilibrium states. We now examine three examples of cortical transitions to *nonequilibrium* states in which the cortex (i) oscillates coherently in time via emergence of a Hopf instability, or (ii) breaks up into a stationary spatial pattern via a Turing instability, or (iii) generates a complex spatiotemporal wave dynamics arising from interactions between the Hopf and Turing instabilities. Although these latter examples can be described as “far from equilibrium”, Figs. 15.7–15.9 show that the emergent dynamics tends to organize itself into periodic or chaotic excursions about the destabilized homogeneous equilibrium state.

15.4.1 Emergence of Hopf Instability

We now make a small adjustment to γ_i^0 , the rate-constant governing the time-course of the inhibitory alpha-function impulse response at the postsynaptic membrane: we reduce the rate-constant from its default value of $\gamma_i^0 = 100$ to 90 s^{-1} , delaying

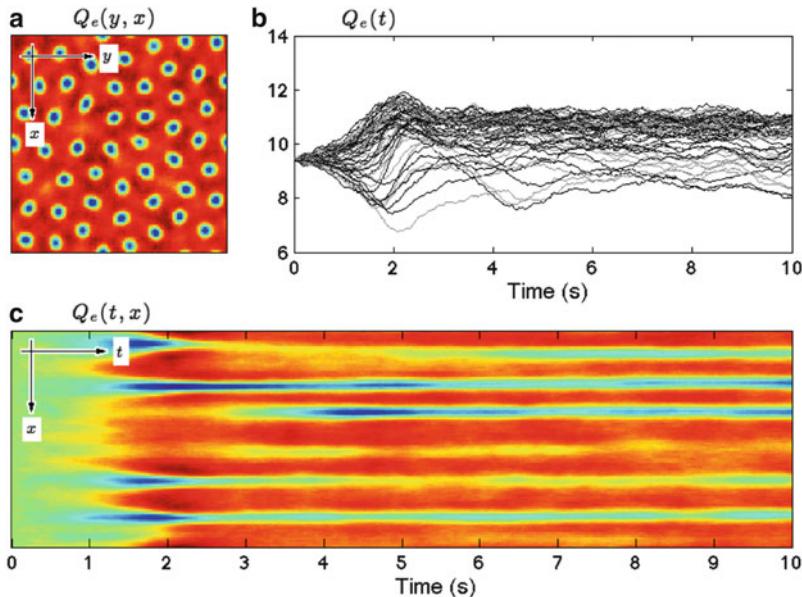


Fig. 15.8 Stochastic simulation run at critical point CP with inhibitory diffusion raised to $D_2 = 0.26 \text{ cm}^2$ to provoke a spatial (Turing) instability (IPSP rate-constant has been restored to default, $\gamma_i^0 = 100 \text{ s}^{-1}$). (a) Bird's-eye snapshot of grid activity Q_e at $t = 20$ s. (b) Time-series of cortical activity at 60 sample points evenly spaced down the $y = 60$ midline. (c) Space-time chart of cortical activity along the $y = 60$ midline. Figure 15.5c predicts emergence of a stationary pattern at wavenumber $q \approx 0.26$ waves/cm, consistent with the emergent pattern of ~ 6 blue (low-firing) islands per 25 cm length of cortex

the time to IPSP peak from 10.0 to 11.1 ms. Applying this apparently minor change at the CP critical equilibrium point ($\lambda_i, \Delta V_e^{\text{rest}}/\text{mV} \approx (1.1032, 2.5398)$ of Figs. 15.1a and 15.2a causes a significant change to the stability properties at the critical point—comparing the linear dispersion graphs of Fig. 15.5b against Fig. 15.5a (default), we see that a range of slow spatial frequencies $0 \leq q \lesssim 0.168 \text{ cm}^{-1}$, and their associated temporal frequencies $0 \leq \omega/2\pi \lesssim 2.0 \text{ Hz}$, are predicted to become unstable, with peak instability predicted at $q = 0$ (i.e., infinite wavelength).

The simulation result in Fig. 15.7 demonstrates that the homogeneous steady state is no longer stable, and, within a few seconds, the competing but weaker long-wavelength unstable modes rapidly quench in favor of a coherent whole-of-cortex temporal oscillation of frequency $\sim 1.6 \text{ Hz}$ that we classify as a Hopf mode.

15.4.2 Emergence of Turing Instability

We can precipitate a spatial instability in cortical firing-rate activity by restoring the IPSP rate-constant to its $\gamma_i^0 = 100 \text{ s}^{-1}$ default value (to suppress the Hopf

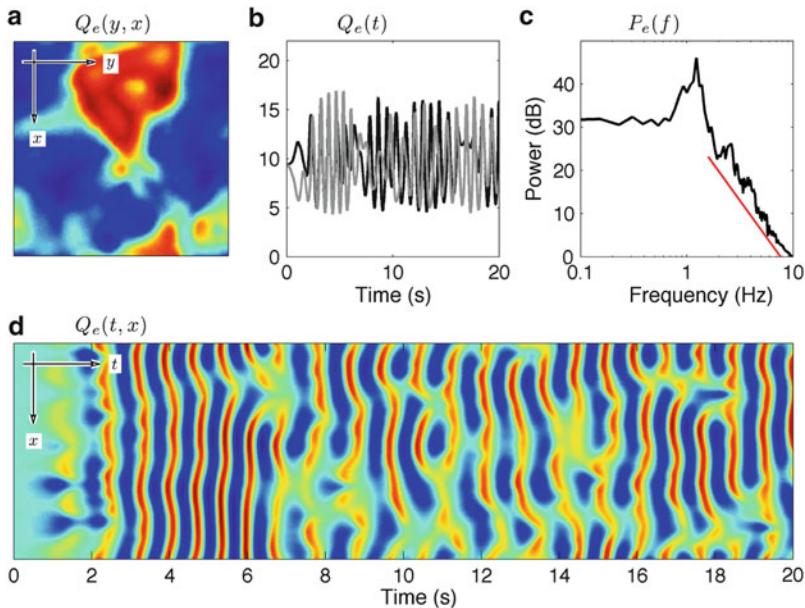


Fig. 15.9 Critical-point simulation with inhibitory PSP rate-constant and diffusion strength set to $\gamma_i^0 = 90 \text{ s}^{-1}$ and $D_2 = 0.26 \text{ cm}^2$ respectively. This combination allows Hopf (Fig. 15.7) and Turing (Fig. 15.8) instabilities to interact. (a) Bird's-eye snapshot of grid activity Q_e at $t = 20 \text{ s}$. (b) Time-series of cortical activity at two points ($x = 20, 80$ along the $y = 60$ midline). Oscillations are now incoherent. (c) Power spectrum for Q_e fluctuations over final 13 s of simulation, expressed in dB relative to 10 Hz. Reference line shows a power law $f^{-3.4}$. (d) Space-time chart of cortical activity along the $y = 60$ midline. Figure 15.5d predicts competition between an unstable $q = 0$ homogeneous mode and a Turing mode at $q \approx 0.2 \text{ waves/cm}$; the outcome is chaotic traveling wavefronts that generate slow ($\sim 1 \text{ Hz}$) oscillations in firing rate

instability), then boosting the strength of inhibitory gap-junction diffusion from $D_2 = 0$ to 0.26 cm^2 . Because excitatory diffusion D_1 is presumed to be much weaker than inhibitory diffusion ($D_1 = D_2/100$), conditions are now favorable for spontaneous emergence of spatial patterning. This is evident from the dispersion panels of Fig. 15.5: the low-frequency wave-instability of panel-B has been suppressed and replaced in panel-c by a broad-spectrum Turing pattern (stationary wave) with a wavenumber peak near $q = 0.26 \text{ waves/cm}$. For a square cortex of side-length 25 cm, we expect the resulting Turing pattern to have a wavelength of $\sim 4 \text{ cm}$, but this linear analysis cannot provide more detailed information about the final pattern structure.

Figure 15.8 shows the results of a stochastic simulation with inhibitory diffusion set to $D_2 = 0.26 \text{ cm}^2$. The cortex is initialized at the CP homogeneous critical point, and immediately starts evolving towards a two-dimensional patterned configuration. The dominant inhibitory diffusion causes the cortex to clump into islands of lowered activity (colored blue), surrounded by seas of relatively higher activity (red), with roughly symmetric excursions above and below the homogeneous equilibrium firing

rate of $\sim 9.44 \text{ s}^{-1}$. As seen in panels-b (time-series) and -c (space–time chart), the Turing pattern is almost fully developed after about 2 s.

15.4.3 Mixed Modes: Hopf–Turing Interactions

We can investigate interaction between the Hopf (temporal) and Turing (Turing) instabilities by simultaneously lowering the inhibitory PSP rate-constant to $\gamma_i^0 = 90 \text{ s}^{-1}$ while raising the inhibitory diffusion strength to $D_2 = 0.26 \text{ cm}^2$. The linear dispersion graph of Fig. 15.5d predicts a competition between a stationary Turing mode at wavenumber $q \approx 0.2 \text{ cm}^{-1}$ and an instability at $q = 0$, but there is no indication of any temporal oscillations in the dominant eigenvalue spectrum.

The simulation result of Fig. 15.9 confirms that the $q = 0$ homogeneous mode is indeed unstable, but as the cortex “falls away” from the uniform state, slow Hopf-like $\sim 1.2 \text{ Hz}$ temporal oscillations develop which, after a few seconds, break up into turbulent spatiotemporal patterns. By running paired noise-free simulations from closely similar initial conditions (not shown here), we find exponential divergence in trajectories (i.e., positive dominant Lyapunov exponent) indicating that Hopf–Turing interactions can produce patterns that are chaotic in space and time.

15.5 Discussion

It is self-evident that the brain has access to a multitude of distinct activity states, with wake, sleep, and anesthetic-induced coma being three of the more grossly obvious examples. Although EEG and fMRI sensing modalities reveal that these brain states are essentially dynamical and unsteady, it does not seem unreasonable to suggest that, to first approximation, one can map a given brain state to a particular coordinate within a suitably chosen domain of homogenous equilibrium states, and to consider the spatiotemporal dynamics for that state either as a small fluctuation—or as a large oscillation in space and time—about the equilibrium reference. For the wake-to-anesthesia transition we selected anesthetic inhibition (λ_i) and subcortical excitation (ΔV_e^{rest}) to represent orthogonal wake–sleep control parameters, with excitatory firing rate (Q_e) representing the level of cortical activation.

The significant feature of the Fig. 15.1a manifold of equilibrium states is the presence of a multirroot region, suggesting that, for certain regions of the sleep–wake domain, the brain can switch activity states by moving from one equilibrium point (e.g., the “up” state of wakefulness) to another (e.g., the “down” state of comatose unconsciousness). We characterize this change as an *equilibrium phase transition* because both states can be located on the stable branches of the equilibrium manifold.

For default parameter settings, the model predicts a first-order jump transition between states at the point of loss of consciousness (LOC); this should be associated

with a pronounced growth in low-frequency fluctuation power as the saddle-node bifurcation point is approached. In earlier work [26, 28, 30], we have suggested that this fluctuation power surge might provide an explanation for the paradoxical boost in delirium-like brain activity on approach to the point of anesthetic induction [16, 17].

A similar boost in fluctuation power is expected for the *recovery* of consciousness (ROC), but because this second bifurcation point generally occurs at a *lower* drug concentration (except for trajectories passing through the CP critical point), the EEG power peak for recovery will be hysteretically displaced from that for induction. This prediction of *two* power surges per induction–recovery cycle is consistent with clinical reports by Kuizenga et al. [16, 17]. This hysteretic separation between the points of anesthetic induction and recovery will provide a form of “neural inertia” that acts to preserve the current behavioral state. Strong evidence for anesthetic neural inertia in insects (fruit flies) and mammals (mice) has been reported in a recent paper by Friedman et al. [8].

15.5.1 Critical Fluctuations

We highlighted an analogy between our cortical model and the vapor \leftrightarrow liquid phase transition of classical thermodynamics (more accurately: *thermostatics*). If gas pressure P and gas temperature T can be mapped to anesthetic “pressure” λ_i and subcortical “temperature” ΔV_e^{rest} , then specific volume v (inverse density) could be analogous to cortical firing rate Q_e such that the “up” and “down” states of the cortex become the vapor and liquid states of a van der Waals substance. The van der Waals model predicts the existence of a critical point CP at which liquid and vapor states are indistinguishable, and which is characterized by the phenomenon of critical slowing: noise-evoked divergent density fluctuations that become slower and larger on close approach to CP. The corresponding critical point on the cortical manifold also displays divergent behavior: extreme sensitivity to small perturbations about the equilibrium point (Fig. 15.6).

We find that it is not necessary to traverse the CP critical point in order to observe critically-slowed fluctuations: it is sufficient to closely approach any of the saddle-node turning points on the anesthetic induction “isotherms” of Fig. 15.2a, so in this sense, the locus of saddle-node bifurcations locates a set of “secondary” critical points. The discontinuous phase transition behavior associated with a saddle-node annihilation was demonstrated in Fig. 15.4.

15.5.2 Nonequilibrium Phase Transitions in the Cortex

Unlike the van der Waals gas, the cortical model can also exhibit *nonequilibrium phase transitions* in which the homogeneous equilibrium state can destabilize in

favor of either (i) a whole-of-cortex temporal oscillation (Hopf instability), or (ii) a symmetry-breaking stationary spatial pattern of cortical activity (Turing instability), or (iii) a complex dynamical spatiotemporal pattern resulting from interacting Hopf–Turing instabilities.

For the particular set of cortical parameters listed in Table 15.1, a low-frequency (~ 1.6 -Hz) Hopf instability can be evoked by imposing a 10 % reduction to γ_i , the rate-constant for the inhibitory postsynaptic potential (Fig. 15.7), while a stationary Turing structure will spontaneously precipitate from homogeneous rest if the inhibitory gap-junction diffusion strength is boosted from $D_2 = 0$ (all gap junctions closed) to $D_2 = 0.26 \text{ cm}^2$ (Fig. 15.8). And if both parameters are changed simultaneously, the result is a dynamic interaction between Hopf and Turing instabilities giving a turbulent mix of slow traveling waves of cortical activity that permeates the entire cortex (Fig. 15.9).

Several authors [15, 19, 34] have identified the pure-Hopf oscillatory mode with the pathological state of *grand mal* seizure. In the present model, emergence of this temporal instability is associated with delayed responsiveness at the inhibitory chemical synapses (i.e., reduction in γ_i rate-constant \Rightarrow prolonged time-to-peak for inhibitory postsynaptic potential) paired with weak gap-junction diffusion between inhibitory neurons (i.e., $D_2 \approx 0$). If this is the case, then boosting inhibitory diffusion—by opening gap junctions to allow competitive interaction between Hopf and Turing modes—should be protective against seizure; conversely, blocking inhibitory gap junctions will tend to increase propensity to seizure. While there is clinical evidence in support of this idea [12, 32, 36], other workers have reported the opposite effect (e.g., [9, 13, 21]), namely that gap-junction blockers have an anticonvulsant effect. More clinical and modeling work will be required to resolve this discrepancy.

Strong inhibitory diffusion tends to suppress oscillations in favor of a stationary Turing pattern. The extreme case of a frozen Turing “crystal” of cortical activations and inactivations is probably pathological since there is no opportunity for information transfer across the cortex. However, Turing structuring may be relevant to early brain development when the immature brain is richly endowed with gap junctions [3] and the conditions for formation of permanent Turing patterns are most favorable. Cartwright [4] has proposed that the morphogenesis of the maze-like features of the gyri and sulci of the cortex is a Turing solution, but his speculation posits an axonal guidance competition between (unknown) inhibitor and activator chemical species rather than a gap-junction process dominated by inhibitory diffusion.

Suitable tuning of the inhibitory synaptic rate-constant γ_i and inhibitory gap-junction diffusivity D_2 locates a codimension-2 Turing–Hopf bifurcation point at which Turing and Hopf instabilities are delicately balanced, giving rise to oscillating Turing patterns and traveling waves. These interactions can produce a chaotic dynamics typified by large-amplitude traveling wavefronts whose passage generates broad-spectrum (0.1–3-Hz) slow oscillations between “up” and “down” states (Fig. 15.9). This chaotic dynamics may correspond to the cortical slow oscillations

seen in NREM sleep and anesthesia [20, 25]. This putative Turing–Hopf dynamical mechanism for the slow oscillation is complementary to, and quite distinct from, the classical explanations that assume slow modulations of ionic currents (e.g., [2]).

15.5.3 Effect of Anesthesia on Sub-brain Structures

Our model assumes that anaesthetic drugs have no effect on either the subcortical tone ϕ^{sc} entering the cortex, or on the cortical resting potential V_e^{rest} . In practice, sedative concentrations of propofol and isoflurane have been found to inhibit the aminergic brain-stem centers which should result in a decrease in resting potential. Thus a more realistic induction-of-anesthesia trajectory on the manifold would be along the “diagonal” (towards the reader) in Fig. 15.1a—rather than following an exact ΔV_e^{rest} “isotherm”. However, the degree of direct anesthetic suppression (or even activation) of aminergic activity is uncertain, so the actual trajectories are unknown. Another unresolved question is the effect of intense structured input into the (anesthetised) cortex from the thalamus: in particular, the oscillatory bursting that gives rise to sleep-like spindle (10–14 Hz) patterns in the EEG. We plan to incorporate a thalamic structure into the cortical model to permit exploration of thalamocortical interactions during wake, sleep, and anesthesia.

There is accumulating experimental evidence that the loss of consciousness that occurs with both GABAergic and nonGABAergic general anesthesia is strongly associated with loss of long-range functional connectivity in the brain [1, 10, 18]; it seems that this long-range connectivity has to be associated with information-rich activity in order for consciousness to be present. Although there is a paradoxical anesthetic boost in long-range spindle synchrony, these 10–14-Hz waveforms are of stereotypical shape and are information-poor [5].

It should be possible to test and refine our model by applying an appropriate form of Bayesian methodology (such as dynamic causal modeling) to clinical data in order to infer the likely trajectories across the manifold, as has been done recently for sleep staging [6].

15.5.4 Multiple Equilibria and Cortical Self-Organization

Implicit in our model construction is the idea that, in order to maximise flexibility and speed of response, the healthy cortex operates in regions that are never far from points of bifurcation and controlled instability. Referring to Fig. 15.1a, this means that we expect the cortical operating point to be located within—or close to—the reentrant fold, allowing access to multiple equilibrium states. Induction of anesthesia can be viewed as an enforced traversal of the manifold leading to a phase transition-like switching between high- and low-firing states of the cortex and

an abrupt loss of consciousness. Similarly, the observation of drug hysteresis—the awakening of a patient at a lower level of anesthetic concentration than that required to put her to sleep—can be visualized as traversals of distinct induction and recovery trajectories.

We find that when the cortical operating point is located within—or close to—the multiroot domain, the model can exhibit a wide and diverse range of dynamical behaviors. But at locations far from the fold (not shown here), the response to stimulus is typically weak and stereotyped—the cortex becomes “too stable” and is no longer “fit for purpose” in an interactive biological sense. If diversity of response is biologically desirable, then we propose that there exists an (as yet unknown) self-organizing principle that biases the cortex towards its multiroot region, permitting both equilibrium and nonequilibrium phase transition-like changes in brain state. If this idea is correct, then there must exist multiple redundant neurobiological mechanisms that work to ensure that the brain remains close to its bifurcation points. The grand challenge will be to identify the nature of this self-organizing learning rule.

Appendix

Model Equations

The cortex is modeled as a 2-D continuum of excitatory and inhibitory neurons, interconnected via resistive gap junctions and neurotransmitter-mediated chemical synapses. The spatially-averaged excitatory (inhibitory) soma potentials V_e (V_i) obey partial differential equations,

$$\tau_e \frac{\partial V_e}{\partial t} = V_e^{\text{rest}} + \Delta V_e^{\text{rest}} - V_e + [\rho_e \psi_{ee} \Phi_{ee} + \rho_i \psi_{ie} \Phi_{ie}] + D_1 \nabla^2 V_e , \quad (15.3)$$

$$\tau_i \frac{\partial V_i}{\partial t} = V_i^{\text{rest}} - V_i + [\rho_e \psi_{ei} \Phi_{ei} + \rho_i \psi_{ii} \Phi_{ii}] + D_2 \nabla^2 V_i , \quad (15.4)$$

with chemical-synaptic inputs [...] enclosed in square brackets, and gap-junction inputs entering as diffusion terms $D \nabla^2 V_{e,i}$. Here, τ_b ($b = e, i$) is the soma time-constant; V_b^{rest} is the soma resting voltage; ρ_b is the chemical synaptic strength with $\rho_e > 0$ (EPSP) and $\rho_i < 0$ (IPSP). These strengths are scaled by dimensionless reversal-potential functions ψ_{ab} ($a = e, i$),

$$\psi_{ab}(t) = \frac{V_a^{\text{rev}} - V_b(t)}{V_a^{\text{rev}} - V_b^{\text{rest}}} ,$$

that are normalized to unity when the neuron is at its resting voltage, and are zero when the membrane voltage reaches the relevant reversal potential (see Table 15.1

for values). The $\Phi_{eb,ib}$ functions in Eqs. (15.3, 15.4) are chemical-synaptic input fluxes obeying second-order differential equations,

$$\left(\frac{d}{dt} + \gamma_e\right)^2 \Phi_{eb} = \gamma_e^2 [N_{eb}^\alpha \phi_{eb}^\alpha(t) + N_{eb}^\beta Q_e(t) + \phi_{eb}^{sc}(t)], \quad (15.5)$$

$$\left(\frac{d}{dt} + \gamma_i\right)^2 \Phi_{ib} = \gamma_i^2 N_{ib}^\beta Q_i(t), \quad (15.6)$$

with dendritic rate constants $\gamma_{e,i}$. The cortico-cortical and local connectivities N^α and N^β scale their respective incoming fluxes ϕ^α , $Q_{e,i}$ respectively; these fluxes are supplemented by an unstructured subcortical stimulation ϕ^{sc} modeled as a small white-noise variation $\xi(t)$ about a constant tone $\langle\phi^{sc}\rangle$,

$$\phi_{eb}^{sc}(t) = \langle\phi_{eb}^{sc}\rangle + \alpha \sqrt{\langle\phi_{eb}^{sc}\rangle} \xi_{eb}(t),$$

where α is a dimensionless noise-amplitude scale-factor. The local fluxes $Q_{e,i}$ in Eqs. (15.5, 15.6) are defined by a sigmoidal mapping from soma voltage to firing rate,

$$Q_{e,i}(t) = \frac{Q_{e,i}^{\max}}{1 + \exp[-C(V_{e,i}(t) - \theta_{e,i})/\sigma_{e,i}]},$$

with $C = \pi/\sqrt{3}$. Here, θ is the population-average threshold for firing, σ is its standard deviation, and Q^{\max} is the maximum firing rate.

The cortico-cortical flux ϕ^α in Eq. (15.5) is generated by excitatory sources $Q_e(\mathbf{r}, t)$, and obeys a 2-D damped wave equation [22],

$$\left[\left(\frac{\partial}{\partial t} + v \Lambda_{eb} \right)^2 - v^2 \nabla^2 \right] \phi_{eb}^\alpha(\mathbf{r}, t) = (v \Lambda_{eb})^2 Q_e(\mathbf{r}, t), \quad (15.7)$$

where Λ_{eb} is the inverse-length scale for $e \rightarrow b$ axonal connections, and v is the axonal conduction speed.

The ∇^2 diffusion terms in Eqs. (15.3, 15.4) describe the voltage contribution arising from gap-junction currents between adjacent neurons. Gap-junction i -to- i coupling between inhibitory interneurons is substantially more abundant than e -to- e coupling between excitatory neurons [3], and in layer-1 of cortex, over 90% of the neural density is inhibitory [35], suggesting the existence of a syncytium of interneuron-to-interneuron diffusive scaffolding that spans the cortex. In view of the relative dominance of i -to- i diffusion, we set the e -to- e diffusion strength D_1 to be small fraction of inhibitory diffusion D_2 (viz. $D_1 = D_2/100$), with D_2 being an adjustable parameter (see [29] for detailed derivation and estimation of D_2 diffusive coupling).

Modeling Propofol Anesthesia

Inductive anesthetic agents, such as propofol, suppress neural activity by prolonging the opening of GABA (*gamma*-aminobutyric acid) channels on the postsynaptic neuron [7], allowing increased influx of chloride ions leading to hyperpolarization. We model propofol effect by simultaneously scaling both the inhibitory synaptic strength ρ_i (in Eqs. (15.3, 15.4)) and the dendritic rate-constant γ_i (Eq. (15.6)) by a dimensionless scale-factor λ that is set to unity in the absence of propofol, and which grows proportionately to propofol concentration,

$$\rho_i \rightarrow \lambda \rho_i^0, \quad \gamma_i \rightarrow \gamma_i^0 / \lambda$$

where γ_i^0 and ρ_i^0 are the anesthetic-free default values. This scaling prolongs the duration of the inhibitory postsynaptic potential (IPSP) without altering its peak amplitude [14] so that the area of the IPSP response (representing total charge transfer) increases linearly with drug concentration. We note that at very high propofol concentrations—well above the clinically relevant range—the charge-transfer versus drug-concentration curve shows saturation effects [14], but the assumption of linearity is accurate at low concentrations, and has been used by Hutt and Longtin [11] in their anesthesia modeling.

Linear Stability Analysis for Homogenous Stationary States

Equations (15.3–15.7) define the cortical model in terms of two first-order ($V_{e,i}$ soma voltages), and six second-order ($\Phi_{ee,ei}; \Phi_{ie,ii}; \phi_{ee,ei}$ firing-rate fluxes) partial differential equations (DEs). If we disable the subcortical noise and take note of the parameter symmetries evident in Table 15.1 (viz., $N_{ee}^\alpha = N_{ei}^\alpha; N_{ee}^\beta = N_{ei}^\beta; N_{ie}^\beta = N_{ii}^\beta$), the cortical system reduces to a set of two first-order and three second-order DEs, equivalent to eight first-order equations. We locate the homogenous equilibrium states by eliminating all space- the time-derivatives in differential equations (15.3–15.7) ($\nabla^2 = 0; \partial/\partial t = \partial^2/\partial t^2 = 0$), then solving (numerically) the resulting set of nonlinear coupled algebraic equations for the steady-state firing rates (Q_e, Q_i) of the excitatory and inhibitory neural populations as a function of anesthetic effect λ and resting potential offset ΔV_e^{rest} . The resulting distribution of homogeneous stationary states are displayed in Figs. 15.1a and 15.2a.

We define an eight-variable state vector $\mathbf{X} = [V_e, V_i, \Phi_{eb}, \dot{\Phi}_{eb}, \Phi_{ib}, \dot{\Phi}_{ib}, \phi_{eb}, \dot{\phi}_{eb}]^T$ with homogeneous equilibrium value $\mathbf{X}^{(0)}$. We examine the linear stability of this stationary state by imposing a small spatiotemporal disturbance $\delta\mathbf{X}$ about $\mathbf{X}^{(0)}$,

$$\mathbf{X}(t, \mathbf{r}) = \mathbf{X}^{(0)} + \delta\mathbf{X}(t, \mathbf{r})$$

with $\delta\mathbf{X}$ being a plane-wave perturbation

$$\delta\mathbf{X}(t, \mathbf{r}) = \delta\mathbf{X}(t) e^{i\mathbf{q}\cdot\mathbf{r}} = \delta\mathbf{X}(0) e^{\Lambda t} e^{i\mathbf{q}\cdot\mathbf{r}},$$

where \mathbf{q} is the wavevector with wavenumber $|\mathbf{q}| = q$, and Λ is an eigenvalue whose real part gives the growth rate of the $\delta\mathbf{X}(0)$ initial perturbation: if $\text{Re}(\Lambda) > 0$, an instability is predicted. Substituting $\mathbf{X} = \mathbf{X}^{(0)} + \delta\mathbf{X}$ into Eqs. (15.3–15.7) and retaining only linear terms results in the matrix equation,

$$\frac{d}{dt} \delta\mathbf{X} = \mathbf{J}(q) \delta\mathbf{X},$$

where \mathbf{J} is an 8×8 Jacobian matrix in which the ∇^2 Laplacians for excitatory and inhibitory diffusion (Eqs. 15.3, 15.4), and wave propagation (Eq. 15.7) appear as $-q^2$ terms. The eight eigenvalues owned by \mathbf{J} describe the linearized dynamics of the homogeneous cortex. For each wavenumber q , we extract and plot the *dominant* eigenvalue—i.e., that eigenvalue whose real part is most positive (or least negative)—since this describes the most strongly growing (or most long-lived) mode at a given spatial frequency. The resulting Λ vs q dispersion curves are shown in Figs. 15.3 and 15.5.

Although the linear dispersion curve provides valuable guidance regarding the onset of instability (i.e., when the real part of the dominating eigenvalue will approach zero), it cannot predict accurately the new dynamics that will emerge once the homogeneous steady state has lost stability and the nonlinear terms can no longer be ignored. This mismatch between linear dispersion prediction and actual simulation outcome is nicely illustrated in Fig. 15.5b which suggests a zero-frequency instability at zero wavenumber will interact with a low frequency wave instability, while the simulation of Fig. 15.7 shows that the actual outcome is a 1.6-Hz Hopf oscillation at $q = 0$. Similarly, Fig. 15.5d indicates a zero-frequency instability at $q = 0$ competing with a stationary Turing, but the Fig. 15.9 simulation shows destabilization in favor of strongly turbulent unsteady interactions between Hopf and Turing instabilities.

References

1. Alkire, M.T., Hudetz, A.G., Tononi, G.: Consciousness and anesthesia. *Science* **322**(5903), 876–880 (2008)
2. Bazhenov, M., Timofeev, I., Steriade, M., Sejnowski, T.J.: Model of thalamocortical slow-wave sleep oscillations and transitions to activated states. *J. Neurosci.* **22**(19), 8691–8704 (2002)
3. Bennett, M.V., Zukin, R.S.: Electrical coupling and neuronal synchronization in the mammalian brain. *Neuron* **41**, 495–511 (2004)
4. Cartwright, J.H.E.: Labyrinthine turing pattern formation in the cerebral cortex. *J. Theor. Biol.* **217**(1), 97–103 (2002)

5. Casali, A.G., Gosseries, O., Rosanova, M., Boly, M., Sarasso, S., Casali, K.R., Casarotto, S., Bruno, M.A., Laureys, S., Tononi, G., Massimini, M.: A theoretically based index of consciousness independent of sensory processing and behavior. *Sci. Transl. Med.* **5**(198), 198ra105 (2013)
6. Dadok, V., Szeri, A.J., Sleigh, J.W.: A probabilistic framework for a physiological representation of dynamically evolving sleep state. *J. Comput. Neurosci.* (2013). doi:10.1007/s10827-013-0489-x
7. Franks, N.P., Lieb, W.R.: Molecular and cellular mechanisms of general anaesthesia. *Nature* **367**, 607–613 (1994)
8. Friedman, E.B., Sun, Y., Moore, J.T., Hung, H.T., Meng, Q.C., Perera, P., Joiner, W.J., Thomas, S.A., Eckenhoff, R.G., Sehgal, A., Kelz, M.B.: A conserved behavioral state barrier impedes transitions between anesthetic-induced unconsciousness and wakefulness: evidence for neural inertia. *PLoS One* **5**(7), e11,903 (2010)
9. Gajda, Z., Gyengesi, E., Hermesz, E., Ali, K.S., Szente, M.: Involvement of gap junctions in the manifestation and control of the duration of seizures in rats *in vivo*. *Epilepsia* **44**(12), 1596–1600 (2003)
10. Guldenmund, P., Demertzi, A., Boveroux, P., Boly, M., Vanhaudenhuyse, A., Bruno, M.A., Gosseries, O., Noirhomme, Q., Brichant, J.F., Bonhomme, V., Laureys, S., Soddu, A.: Thalamus, brainstem and salience network connectivity changes during propofol-induced sedation and unconsciousness. *Brain Connect* **3**(3), 273–285 (2013)
11. Hutt, A., Longtin, A.: Effects of the anesthetic agent propofol on neural populations. *Cog. Neurodyn.* **4**(1), 37–59 (2010)
12. Jacobson, G.M., Voss, L.J., Melin, S.M., Mason, J.P., Cursons, R.T., Steyn-Ross, D.A., Steyn-Ross, M.L., Sleigh, J.W.: Connexin36 knockout mice display increased sensitivity to pentylenetetrazol-induced seizure-like behaviors. *Brain Res.* **1360**, 198–204 (2010)
13. Jahromi, S.S., Wentlandt, K., Piran, S., Carlen, P.L.: Anticonvulsant actions of gap junctional blockers in an *in vitro* seizure model. *J. Neurophysiol.* **88**(4), 1893–1902 (2002)
14. Kitamura, A., Marszałek, W., Yeh, J.Z., Narahashi, T.: Effects of halothane and propofol on excitatory and inhibitory synaptic transmission in rat cortical neurons. *J. Pharmacol.* **304**(1), 162–171 (2002)
15. Kramer, M.A., Kirsch, H.E., Szeri, A.J.: Pathological pattern formation and cortical propagation of epileptic seizures. *J. R. Soc. Lond.: Interface* **2**, 113–207 (2005)
16. Kuizenga, K., Kalkman, C.J., Hennis, P.J.: Quantitative electroencephalographic analysis of the biphasic concentration–effect relationship of propofol in surgical patients during extradural analgesia. *Br. J. Anaesth.* **80**, 725–732 (1998)
17. Kuizenga, K., Wierda, J.M.K.H., Kalkman, C.J.: Biphasic EEG changes in relation to loss of consciousness during induction with thiopental, propofol, etomidate, midazolam or sevoflurane. *Br. J. Anaesth.* **86**, 354–360 (2001)
18. Lee, U., Ku, S., Noh, G., Baek, S., Choi, B., Mashour, G.A.: Disruption of frontal-parietal communication by ketamine, propofol, and sevoflurane. *Anesthesiology* **118**(6), 1264–1275 (2013)
19. Liley, D.T.J., Bojak, I.: Understanding the transition to seizure by modeling the epileptiform activity of general anesthetic agents. *Clin. Neurophysiol.* **22**(5), 300–313 (2005)
20. Murphy, M., Bruno, M.A., Riedner, B.A., Boveroux, P., Noirhomme, Q., Landsness, E.C., Brichant, J.F., Phillips, C., Massimini, M., Laureys, S., Tononi, G., Boly, M.: Propofol anesthesia and sleep: a high-density EEG study. *Sleep* **34**(3), 283–291A (2011)
21. Nilsen, K.E., Kelso, A.R., Cock, H.R.: Antiepileptic effect of gap-junction blockers in a rat model of refractory focal cortical epilepsy. *Epilepsia* **47**(7), 1169–1175 (2006)
22. Robinson, P.A., Rennie, C.J., Wright, J.J.: Propagation and stability of waves of electrical activity in the cerebral cortex. *Phys. Rev. E* **56**, 826–840 (1997)
23. Sears, F.W., Salinger, G.L.: Thermodynamics, Kinetic Theory, and Statistical Thermodynamics, 3rd edn. Addison-Wesley, Reading (1975)
24. Stanley, H.E.: Introduction to Phase Transitions and Critical Phenomena. Clarendon Press, Oxford (1971)

25. Steriade, M., Nuñez, A., Amzica, F.: A novel slow (<1 Hz) oscillation of neocortical neurons in vivo: depolarizing and hyperpolarizing components. *J. Neurosci.* **13**, 3252–3265 (1993)
26. Steyn-Ross, M.L., Steyn-Ross, D.A., Sleigh, J.W., Liley, D.T.J.: Theoretical electroencephalogram stationary spectrum for a white-noise-driven cortex: evidence for a general anesthetic-induced phase transition. *Phys. Rev. E* **60**, 7299–7311 (1999)
27. Steyn-Ross, M.L., Steyn-Ross, D.A., Sleigh, J.W., Wilcocks, L.C.: Toward a theory of the general anesthetic-induced phase transition of the cerebral cortex: I. A statistical mechanics analogy. *Phys. Rev. E* **64**, 011,917 (2001)
28. Steyn-Ross, M.L., Steyn-Ross, D.A., Sleigh, J.W.: Modelling general anaesthesia as a first-order phase transition in the cortex. *Prog. Biophys. Mol. Biol.* **85**(2–3), 369–385 (2004)
29. Steyn-Ross, M.L., Steyn-Ross, D.A., Wilson, M.T., Sleigh, J.W.: Gap junctions mediate large-scale turing structures in a mean-field cortex driven by subcortical noise. *Phys. Rev. E* **76**, 011,916 (2007)
30. Steyn-Ross, D.A., Steyn-Ross, M.L., Wilson, M.T., Sleigh, J.W.: Phase transitions in single neurons and neural populations: Critical slowing, anesthesia, and sleep cycles. In: Steyn-Ross, D.A., Steyn-Ross, M.L. (eds.) *Modeling Phase Transitions in the Brain*. Springer Series in Computational Neuroscience, vol. 4, chap. 1, pp. 1–26. Springer, New York (2010)
31. Steyn-Ross, D.A., Steyn-Ross, M.L., Sleigh, J.W., Wilson, M.T.: Progress in modeling EEG effects of general anesthesia: Biphasic response and hysteresis. In: Hutt, A. (ed.) *Sleep and Anesthesia: Neural Correlates in Theory and Experiment*. Springer Series in Computational Neuroscience, vol. 15, chap. 8, pp. 167–194. Springer, New York (2011)
32. Voss, L.J., Jacobson, G., Sleigh, J.W., Steyn-Ross, D.A., Steyn-Ross, M.L.: Excitatory effects of gap junction blockers on cerebral cortex seizure-like activity in rats and mice. *Epilepsia* **50**(8), 1971–1978 (2009)
33. Wentlandt, K., Samoilova, M., Carlen, P.L., El Beheiry, H.: General anesthetics inhibit gap junction communication in cultured organotypic hippocampal slices. *Anesth. Analg.* **102**(6), 1692–1698 (2006)
34. Wilson, M.T., Sleigh, J.W., Steyn-Ross, D.A., Steyn-Ross, M.L.: General anesthetic-induced seizures can be explained by a mean-field model of cortical dynamics. *Anesthesiology* **104**, 588–593 (2006)
35. Wozny, C., Williams, S.R.: Specificity of synaptic connectivity between layer 1 inhibitory interneurons and layer 2/3 pyramidal neurons in the rat neocortex. *Cereb. Cortex* **21**(8), 1818–1826 (2011)
36. Yang, L., Ling, D.S.F.: Carbenoxolone modifies spontaneous inhibitory and excitatory synaptic transmission in rat somatosensory cortex. *Neurosci. Lett.* **416**, 221–226 (2007)

Chapter 16

Large Scale Brain Networks of Neural Fields

Viktor Jirsa

Abstract Neural fields describe neural activations continuous in space and time. Neurons at a particular location in the brain receive input from its local neighbors and from far distant neuronal populations. Both types of connectivity, local and global, contribute approximately equally to the complete connectivity, but differ qualitatively in their connection topology. The local connectivity is characterized by a connection density that monotonously decreases with the distance, typically independent of the location in the brain, whereas the global connectivity is characterized by sparse long-range connections (Connectome) between brain areas. In this chapter I discuss some developments of local-global descriptions of neural fields culminating in the international neuroscience project The Virtual Brain.

16.1 Introduction

The binome Structure-Function has intrigued human mankind across all disciplines of science for many centuries. Paradoxically, it is nowhere as exciting and as little understood as in system neuroscience. The human brain – on the system level – is effectively a network. A network is composed of network nodes and network links. Traditionally researchers choose a level of observation for the network nodes, such as neurons, neuronal populations or complexes composed of neurons, astrocytes and glial cells. Less commonly discussed, and more difficult to justify, is the selection of network links. About half of all synaptic connections in a given area stem from intracortical fibers and are hence of short length and local, typically smaller than 1 cm [7]. The other half of incoming fibers is corticocortical (global) and originates in distant regions with length distributions of 1–20 cm and a mean value of about

V. Jirsa (✉)

Institute de Neurosciences des Systèmes, UMR Inserm 1106, Aix-Marseille University,
Marseille, France

e-mail: viktor.jirsa@univ-amu.fr

8 cm [45]. The two fiber systems are differentiated also regarding time delays via signal transmission and topology. The local fibers are typically not myelinated (with the exception of fibers from interneurons that tend to be longer) and the signal transmission is slow with speeds of 0.1–1 m/s. The associated time delays are on the order of 1–10 ms. The connectivity of the short fiber system is translationally invariant and the connection density falls off exponentially with increasing distance. The global fiber system is differently organized. Its axons are all myelinated and make up the white matter with characteristic transmission speeds of 6–9 m/s and transmission delays on the order of 10–100 ms. The key difference to the local fiber system is its sparse connectivity, which is not translationally invariant and has become to be known under the expression connectome, the finite and countable set of connections of corticocortical white matter fibers [12]. In the following I will discuss some of the developments over the last 20 years that have dealt with the interplay of local and global connectivity and have lead to the conception of The Virtual Brain, a large scale modeling and simulation platform [38, 53].

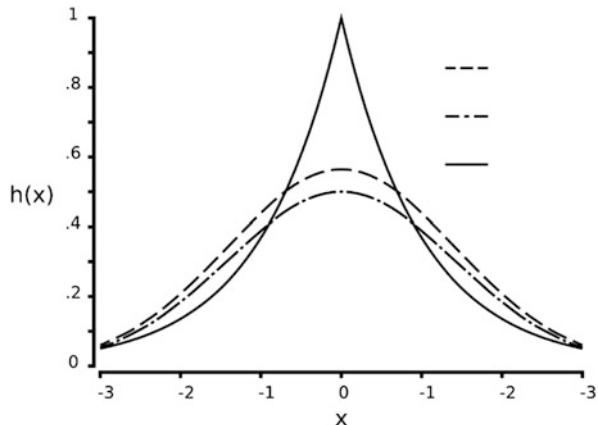
16.2 Mathematical Formulation of a Large-Scale Brain Network of Neural Fields

Let $\psi(x, t)$ be the neural field capturing the population activity at time point t and position x . Unless explicitly stated I will consider scalar neural fields for reasons of simplicity. A full vectorial formulation is provided in [54]. The dynamics of the neural field can then be described by the following integro-differential equation:

$$\begin{aligned} \dot{\psi}(x, t) = & N(\psi(x, t)) + \int_T G(t - s) \left(\int_{\Gamma} W_{hom}(|x - y|) S[\psi(y, t - |x - y|/c)] dy \right. \\ & \left. + \int_{\Gamma} W_{het}(x, y) S[\psi(y, t - |x - y|/v)] dy \right) ds \end{aligned} \quad (16.1)$$

where N denotes the nonlinearity of the neural mass model and the dot indicates the first time derivative. The transfer function $G(t - s)$ is identified with a mean synaptic transfer process. The spatial domain of the neural field is denoted by Γ where $x \in \Gamma = [0, L]$ and L is the spatial length of the neural field. The local homogeneous connectivity function, $W_{hom}(|x - y|)$, is translationally invariant (see Fig. 16.1). The global connectivity function is not translationally invariant and is called heterogeneous, $W_{het}(x, y) \neq W_{het}(|x - y|)$. The parameters c and v are the propagation velocities through the homogeneous and heterogeneous connections, respectively. The general network of neural fields in (16.1) represents one of the most general formulations of brain network dynamics. Derivatives thereof are used in many applications and will be discussed in the following as a function of the degree, to which they incorporate realistic connectivity.

Fig. 16.1 Local fiber system is homogeneous: The distribution of synaptic footprints as a function of distance is translationally invariant. The horizontal axis shows the distance x for three example distributions $h(x)$



16.3 Homogeneous Approximations of Neural Fields

A key approximation allowing the development of detailed neural field models applied to real world problems was the assumption of homogeneous connectivity. This assumption was motivated by the fact that the connection density decreases with the corticocortical fiber length. Mathematically this is captured through translationally invariant integral kernels $W_{hom}(|x - y|)$ (see Fig. 16.1), which finds its roots in classic works including [1, 44, 57, 58]. At that time more detailed large-scale connectivity of the full primate brain was unavailable routinely, hence the introduction of homogeneous connectivity scaled up to the full brain was proposed as a first approximation [44]. The approach proved successful for the study of certain phenomena as observed in large-scale brain systems including spontaneous activity [8, 23, 42, 49, 51, 59, 60], evoked potentials [50], anesthesia [3, 29], epilepsy [9] and sensorimotor coordination [34, 35]. In Fig. 16.2 two connectivity matrices are plotted: on the left, a realistic connectivity matrix of the two hemispheres of the human brain, the Connectome, is plotted. Brain areas are ordered along the horizontal and vertical axes, separating the hemispheres. The grey values code for the strength of a connection. On the right of Fig. 16.2, the homogeneous approximation of the connectivity matrix is plotted.

After having made the homogeneous connectivity approximation, Jirsa and Haken introduced the Green's function method into neural field theory [34, 35] to link the integral field formulation with the partial differential equation formulation of neural fields. Various researchers have since then explored the pattern forming capacities of neural fields as a general tool to study principles of spatiotemporal pattern formation [41]. Since synaptic processes act typically on a short time scale τ compared to the dynamics of $\psi(x, t)$, most neural field approaches also neglect or truncate them at first order of τ (but see [30] for treatments of gamma-distributed synaptic footprint). Linearizing the nonlinear population dynamics N around its fixed point solution, the neural field equation can be written as

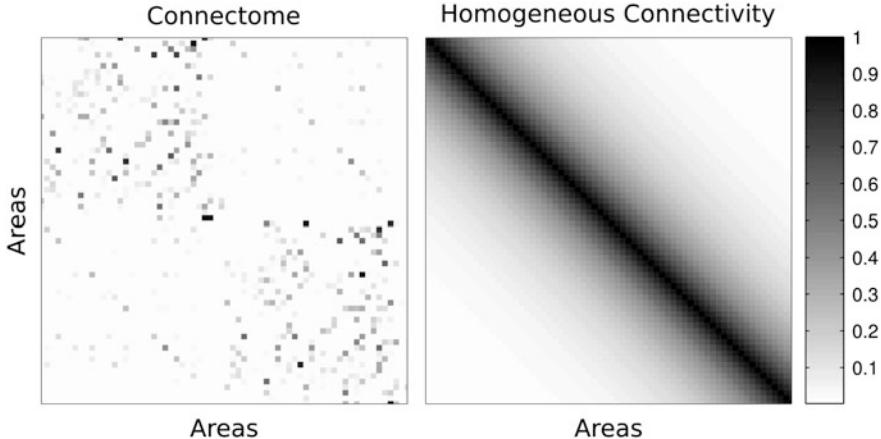


Fig. 16.2 Large-scale brain connectivity. *Left:* the human connectome is represented as a connectivity matrix. The brain areas are ordered along its two axes and the entries of the matrix quantify the strength of a link between two areas. The connectivity strengths are coded in *grey scale*. Intrahemispheric connections can be easily recognized in the distributions within two quadrants. *Right:* the connectivity matrix is shown after applying a homogeneous translationally invariant approximation

$$\hat{K}^{-1}\psi(x, t) = \dots + \tau\dot{\psi}(x, t) + \psi(x, t) = \int_{\Gamma} w(|x - y|) \cdot S[\psi(y, t - \frac{|x - y|}{c})] dy, \quad (16.2)$$

where \hat{K}^{-1} is the inverse of the integral operator linked to the synaptic transfer $G(t - s)$ and $w(|x - y|)$ is the novel integral kernel as shown in Fig. 16.2. The neural field equation (16.2) is a nonlinear retarded differential-integral equation with a spatially invariant integral kernel. To gain a first intuition of the neural field's dynamics, the following simplifications may be made: we set $\tau = 0$ and assume the cortical surface Γ to be one-dimensional with its connectivity $w(|x - y|)$ as

$$w(|x - y|) = \frac{1}{2\sigma} e^{-|x - y|/\sigma} \quad (16.3)$$

Then the method of Green's functions [34, 35, 45, 49] may be applied which transforms (16.2) into the Fourier-space of physical space-time; we reshuffle the terms in the equation and perform a back-transformation into physical space-time. The resulting equation is a local nonlinear partial differential equation of the form

$$\ddot{\psi}(x, t) + (\omega_0^2 - c^2 \Delta) \psi(x, t) + 2\omega_0 \dot{\psi}(x, t) = (\omega_0^2 + \omega_0 \frac{\partial}{\partial t}) \cdot S[\psi(x, t)] \quad (16.4)$$

where $\omega_0 = c/\sigma > 0$ and the one-dimensional Laplacian $\Delta = \frac{\partial^2}{\partial x^2}$. Without any input the left-hand-side is a damped wave equation and allows for oscillatory dynamics. The spatially uniform pattern is generally stable, if the slope of the

sigmoid function $S[.]$ on the right-hand-side of (16.4) is sufficiently small, though transient wave propagation may occur following small discrete perturbations [49]. If the slope increases beyond a threshold, then the spatially uniform state is destabilized and wave propagation may occur. Coombes et al. [15] discuss the effects of connectivity strengths, which do not decrease with increasing distance, but rather remain constant within a finite regime. In this case, it is not sufficient to describe the spatiotemporal dynamics by a local partial differential equation as in (16.4), but non-local delayed terms arise.

To achieve better realism [37] expanded this equation to two dimensions and defined the neural field on the biologically realistically folded cortical hemisphere employing EEG and MEG forward solutions to obtain simulation imaging signals that could be compared directly to empirical data for the first time.

In Fig. 16.3 the stimulus excites the neural sheet at site A , $x = x_0$, and initiates wave propagation by means of a circular traveling wave front undergoing attenuation in space and in time. The time courses of the neural ensembles are plotted for the sites A and B, where the latter is more distant to the stimulus site. For several selected time points the spatiotemporal activity patterns on the sphere are plotted in the top row of Fig. 16.3. Here and in the following the color code represents $-\max$ to $+\max$ as blue goes through black to red and yellow. In the rows below, the same neural activity patterns are represented on the unfolded cortex and on the folded cortex for the same time points. Here the circular traveling wave structure is preserved in all three representations, which may be understood by the fact that the symmetry of the connectivity imposes constraints upon the range of the observable spatiotemporal dynamics. Effectively, an introduction of heterogeneity into the connectivity will break this symmetry and reorganize the dynamics spatiotemporally. From this perspective, it is not surprising that almost all modeling efforts based on this homogeneous approximation focus on the temporal organization of brain dynamics rather than the organization of brain topographies and its temporal evolution. This was pointed out early on [37] and a suggestion was made to integrate biologically realistic diffusion tensor weighted imaging (DTI) based connectivity into full brain modeling efforts, which is nowadays at the core of connectome based modeling efforts. Recent studies [46] have systematically investigated the degree, to which homogeneous approximations may serve to understand realistic connection topologies and have concluded that homogeneous approximations are more appropriate for mesoscopic descriptions of brain activity, but less well suited to address full brain network dynamics. These insights underwrite the need to incorporate realistic connectivity into large scale brain network models and integrate the models with neuroinformatics tools for common brain imaging signatures such as electroencephalography (EEG), magnetoencephalography (MEG) and functional Magnetic Resonance Imaging (fMRI) [4, 5, 53].

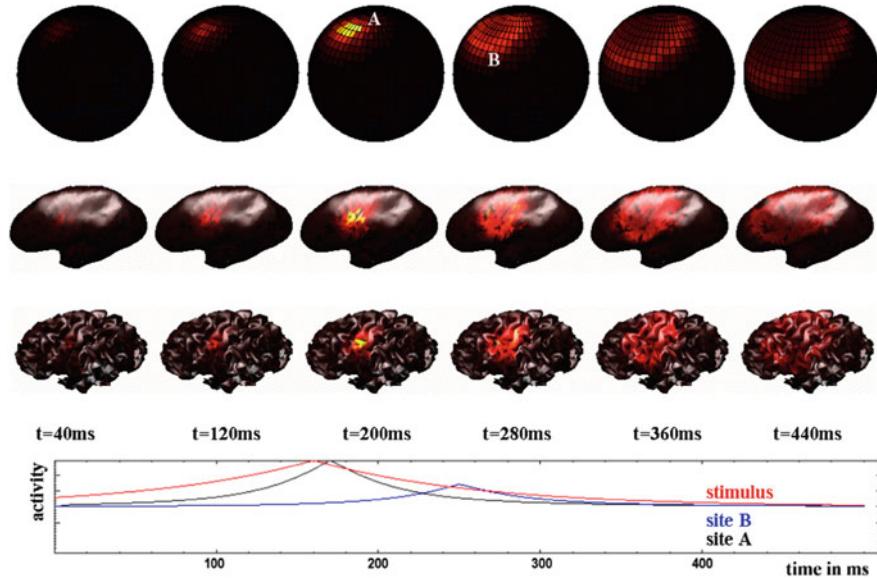


Fig. 16.3 Stimulus evoked activity of neural fields on the sphere (top row), inflated cortex (second row) and folded cortex (third row). The time evolution is plotted along the rows for six separate time points. The bottom panel shows the time course of the stimulus (red line) and the activation pattern for two individual sites on the spherical surface (Modified after [37])

16.4 Homogeneous Modeling of Neural Fields and First Heterogeneous Extensions

The neural field dynamics with local and global couplings is determined by Eq.(16.1). Now consider a neural field with short range local connectivity and a single heterogeneous connection at two points of the field. The latter connection has been termed a two-point connection [36]. Neural field architectures with two-point connections have been investigated in theoretical and computational detail as a function of connection parameters (signal transmission speeds, connection strengths) [32, 47, 48] or distributions of two-point connections [6]. Random matrix theory has been used to study the effects of random connectivity on network dynamics, but only with one constant time delay for signal propagation [22, 31]. Much insight can be gained on the effects of heterogeneities by considering the limit case of zero local coupling, when the system then reduces to the abstract model of two coupled oscillators described by the following delay differential equation [32]:

$$\dot{x}_i(t) = -\epsilon x_i(t) + \mu_{ij} x_j(t - \tau) \quad (16.5)$$

with $i, j = 1, 2; i \neq j$. The time-dependent amplitude of the i th oscillator is defined by $x_i \in \mathbb{R}$ and the discrete time delay by $\tau \in \mathbb{R}_0^+$. Letting $\mathbf{B} = (\mu_{ij})$

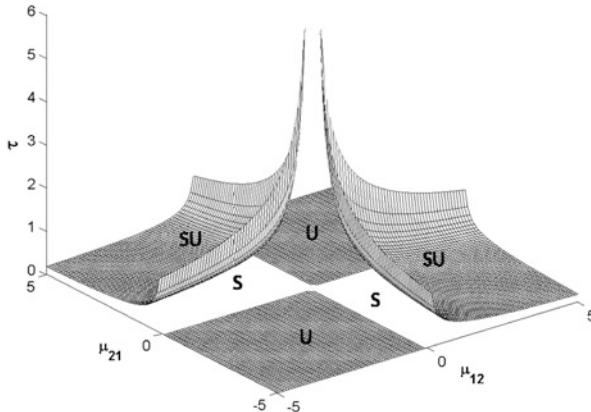


Fig. 16.4 Stability surface for the system of two coupled oscillators with delay at $\epsilon = 0.1$. The surface defines the minimal delay value as a function of the coupling strengths of the two heterogeneous connections. The region denoted by S shows the stable regions while U shows the regions where the system is always unstable. SU shows the regions where the system is unstable above the critical surface and stable below it

be the coupling matrix, the system can be written in the vector/matrix form as:

$$\dot{\mathbf{x}}(t) = -\epsilon \mathbf{x}(t) + \mathbf{B} \mathbf{x}(t - \tau) \quad (16.6)$$

Following [33], we can decompose \mathbf{B} according to $\mathbf{B} = \mathbf{E}^{-1} \boldsymbol{\lambda} \mathbf{E}$ where $\boldsymbol{\lambda}$ is the Jordan form with eigenvalues $\beta = \pm \sqrt{\mu_{12}\mu_{21}}$ and matrix \mathbf{E} contains the corresponding eigenvectors $\hat{\mathbf{e}}$. Multiplying (16.6) from the left with \mathbf{E} we obtain:

$$\mathbf{E} \dot{\mathbf{x}}(t) = -\epsilon \mathbf{E} \mathbf{x}(t) + \boldsymbol{\lambda} \mathbf{E} \mathbf{x}(t - \tau) \quad (16.7)$$

If we set the eigenmodes to $u(t) = \hat{\mathbf{e}} \mathbf{x}(t)$, Eq.(16.7) reduces to a decoupled representation of the dynamics of the system in terms of the eigenmodes as follows:

$$\dot{u}(t) = -\epsilon u(t) + \beta u(t - \tau) \quad (16.8)$$

The stability properties of the above one-dimensional delay differential equation have been widely studied. Assuming a solution of the form $u(t) = e^{\lambda t}$, $\lambda \in \mathbb{C}$, the stability condition is determined by the following characteristic equation:

$$H(\lambda) = \lambda + \epsilon - \beta e^{-\lambda \tau} = 0 \quad (16.9)$$

Notice that any solution of (16.8) satisfying $\Re[\lambda] < 0$ is stable and when $\Re[\lambda]$ becomes positive, the system destabilizes. On the critical surface parameterized by $\epsilon, \mu_{12}, \mu_{21}$ and τ , the solution λ of the characteristic equation is purely imaginary, i.e., $\lambda = \Im[\lambda] = i\omega$, $\omega \in \mathbb{R}_0^+$. This surface represents the boundary at which the system changes stability and is shown in Fig. 16.4. A stability change through

$\Re[\lambda] = \infty$ is impossible [16], hence every instability must occur at the surface defined by $\lambda = i\omega$. Using the characteristic polynomial $H(\lambda)$ for $\lambda = i\omega$ with some algebra we can show that:

$$\frac{d\Re[\lambda]}{d\tau} = -\operatorname{Re}\left[\frac{\partial H/\partial\tau}{\partial H/\partial\lambda}\right] = \frac{\omega^2}{|K|^2} > 0 \quad , \quad \lambda = i\omega \quad (16.10)$$

where $K = \beta\tau + e^{i\omega\tau}$. The above result implies that as τ increases across the critical surface from below, the equilibrium state always destabilizes and remains unstable for all larger τ . It is noteworthy that manipulations of a variety of parameters including excitability ϵ , connection strength μ_{ij} and transmission speeds (via time delay τ) do not alter the topology of the parameter space in Fig. 16.4, but rather translate and rotate the existing hypersurfaces [32]. We do not know however to what degree this property translates to more complex architectures as described by [6] or below in Sect. 16.6.

16.5 Full Brain Network Modeling Using the Connectome

With the increasing availability of structural information on the global fiber system, the Connectome, it has become possible to investigate the connectivity-specific dynamic characteristics of the brain network. I wish to emphasize this aspect here: Connectome-based modeling finds its primary usefulness in exploring the effects of connectivity upon network dynamics. So far it remains an open question to what degree the inclusion of biophysical detail of neural mass models aids in improving our understanding of full brain network dynamics. Generally there are three ingredients in full-brain network modeling. The first ingredient of these models is the anatomical structure. In all models the structural information is extracted from databases compiling different type of tracing studies (e.g. for the macaque cortex this information is provided by the CoCoMac neuroinformatics [40]), or from DTI/DSI (diffusion tensor Imaging/ Diffusion Spectrum Imaging) techniques [27]. The second ingredient is the signal transmission time delays. Generally time delays do not alter the stationary spatial patterns in a network, but they do change their stability [14]. Though time delays play an essential role in the synchronization behaviour of oscillatory patterns, it is less clear to what extent they influence patterns of spiking neurons. The third ingredient of the models is the type of intrinsic dynamics of the local nodes. It is here where typically a large range of models has been employed. Models considered simple oscillatory dynamics [13, 19, 24, 25], chaotic dynamics [28], and detailed realistic neural mass models; the latter include, for instance, ensembles of excitatory and inhibitory populations of spiking neurons coupled through N-Methyl-D-aspartic acid (NMDA), alpha-amino-3-hydroxy-5-methylisoxazole-4-propionic acid (AMPA) and gamma-Aminobutyric acid (GABA) synaptic dynamics [17]. All these modeling approaches have discretized the neural fields into network nodes and only focussed on the global connections via the

Connectome, ignoring the local intracortical connectivity. Through this approach, the large-scale network models began to display spatiotemporal pattern formation with biologically realistic spatial patterns. Spieglér and Jirsa [54] systematically investigated in this context the relevance of the local contributions to the overall connectivity and provided guidelines for choosing the right spatial distribution function for given anatomical and geometrical constraints.

This type of full brain network modeling has been applied extensively to the exploration of the intrinsic activity of the brain network at rest. Experimentally, the first demonstration of a resting state network (RSN) using functional Magnetic Resonance Imaging (fMRI) was given in the seminal paper of Biswal and colleagues [2] where temporal fluctuations in the blood oxygen level dependent (BOLD) signal were exploited to examine the cross-correlation (that is, functional connectivity) between activity in the primary motor cortex (M1) and other brain regions independent of any overt task. Spatially, the functional connectivity pattern seemed to mimic the pattern of activation seen when subjects executed an overt motor response. This observation led to a veritable explosion of empirical work exclusively focused on the identification and characterization of these resting state networks [26, 43, 52]. Theoretical modeling efforts of RSN dynamics focussed on full brain network modeling using the connectome (see for a review [20]). In particular a new hypothesis emerged regarding the working point of the full brain network at rest: [25] demonstrated that the resting state emerges from noise-induced transient fluctuations around the stable equilibrium state of a full brain network of coupled FitzHugh-Nagumo oscillators just below the edge of a bifurcation. The noise provokes fluctuations between different multi-stable oscillatory brain states in a network of Wilson-Cowan oscillators, again in the neighborhood of the critical point, but above the edge of instability [19]. In these models, it is the underlying anatomical structure that shapes the structure of the dynamical landscape, which then the fluctuations of the network model can explore. The time delays may alter the location of the instability, but will not change the actual structure of spatial patterns in the dynamical landscape. These are the insights we gained when discussing the neural field models with embedded heterogeneous fibers in Sect. 16.2. At the edge of the critical instability of any model the spatial correlations of the noisy excursions are mainly shaped by the anatomical structure. The degree to which the RSNs are expressed depends on the proximity to the instability, which is determined by the space-time structure of the couplings, that is the topology of the connectivity and its associated time delays. The global dynamics of a brain working at a critical point amplifies the underlying structure of the anatomical connections and of its interactions with the local dynamics.

These results are best illustrated in Fig. 16.5 where we follow the lines of thought as described in [21]. The network node model of a local brain area consists of a population of integrate-and-fire spiking neurons with excitatory (AMPA and NMDA) and inhibitory (GABA-A) synaptic receptor types. The nodes are connected using a large-scale anatomical connection matrix obtained from human subjects via DTI/DSI tractography [17, 27]. Let us assume that the control parameter is the coupling strength G across different brain areas. The global dynamical states of

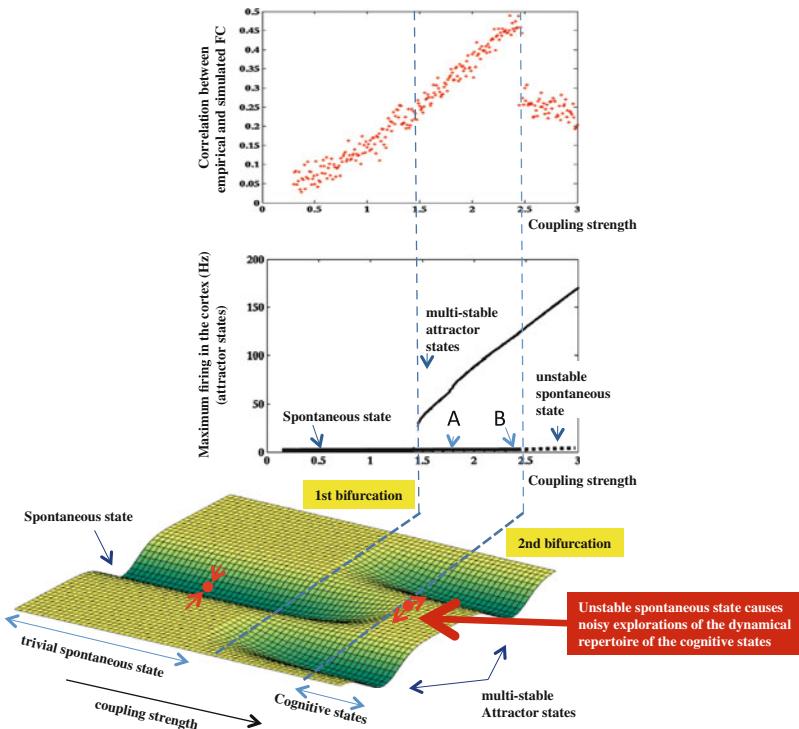


Fig. 16.5 Resting State Network dynamics is critical. *Top*: The fit between both the empirical and simulated functional connectivity is measured by the Pearson correlation coefficient. In the same plot, the second bifurcation line obtained below is also shown. The best fit of the empirical data is observed at the brink of the second bifurcation. *Middle*: Bifurcation diagrams characterizing the stationary states of the brain system as a function of the scaling of the connectivity strength, which acts as a control parameter. The number of different possible stable states is plotted. For small values of the global connectivity strength, only one stable state exists, namely the spontaneous state characterized by low firing activity in all cortical areas. For a critical value of connectivity strength a first bifurcation emerges where at least one new multi-stable state appears while the spontaneous state is still stable. For even larger values of connectivity strength, a second bifurcation appears where the spontaneous state becomes unstable. *Bottom*: The two bifurcations are illustrated in a landscape diagram on the bottom, and in the bifurcation diagram on top of that, where the maximal rate activity in the brain is plotted. After the

the system can be described by a so-called bifurcation diagram, which captures the stationary states (attractors) of the system as a function of the coupling strength. For low coupling strength, only a trivial stable spontaneous state, characterized by an asynchronous low level of firing activity in all brain areas, is stable. By increasing the coupling strength at the first bifurcation point, new stable attractors emerge. These new attractors reflect increased activity in some brain areas, defining the emergence and stabilization of specific networks of brain areas. In Fig. 16.5, this is shown in the landscape cartoon on the bottom, and in the bifurcation diagram on top of that, where the maximal rate activity in the brain is plotted. After the

first bifurcation a new branch appears evidencing the emergence of new attractors. Nevertheless, after this first bifurcation, the spontaneous state is still stable. Only after a second bifurcation the spontaneous state loses stability. It is precisely at the edge of this second bifurcation point, where the brain seems to operate. In the top panel, we show the fitting between the empirical and simulated resting BOLD functional connectivity. The best fitting (maximal correlation) occurs exactly at the edge of the second bifurcation, where the trivial spontaneous state loses its stability and therefore the noisy fluctuations of the dynamics are able to explore and reflect the structure of the other “cognitive” attractors shaped by the underlying anatomy.

16.6 The Virtual Brain

The research described so far has led to the development of the international research project The Virtual Brain (see www.thevirtualbrain.org and [53] for details). The Virtual Brain (TVB) is a neuroinformatics platform, which simulates large scale brain network dynamics using neural fields with the embedded Connectome. In line with previous studies, TVB incorporates a biologically realistic, large scale connectivity of brain regions in the primate brain. The long-range connectivity is mediated by white matter fibre tracts as identified by tractography based methods [27], or obtained from CoCoMac neuroinformatics database [40]. Two types of brain connectivity are distinguished in TVB, that is region-based and surface-based connectivity. In the former case, the networks comprise discrete nodes and connectivity, in which each network node models the neural population activity of a brain region connected via interregional fibres, the Connectome, as described in the previous Sect. 16.5. The connectivity matrix defines the connection strengths and time delays via signal transmission between all network nodes. In the latter case of surface-based connectivity, cortical and subcortical areas are modeled on a much finer and more realistic scale. The connectivity comprises the local intracortical homogeneous and the global corticocortical fibres, the Connectome. Due to the fine spatial sampling, in particular of the cortical surface, the neural activity of coupled neural masses approximates the spatially continuous neural field [18, 54]. There is also a wide choice of neural mass models in TVB that are placed on each network node (for instance [10, 55, 56, 58]).

In TVB, our main interest is to provide the researcher with a neuroinformatics platform that allows asking questions on how connectivity influences and drives the processes on the macroscopic brain network scale. Researchers from different backgrounds can benefit from an integrative software platform including a supporting framework for data management (generation, organization, storage, integration and sharing) and a simulation core written in Python. TVB allows the reproduction and evaluation of personalized configurations of the brain by using individual subject or patient data. This facilitates an exploration of the consequences of pathological changes in the system as well as permitting the investigation of potential ways to counteract such unfavorable processes. In particular, neural field

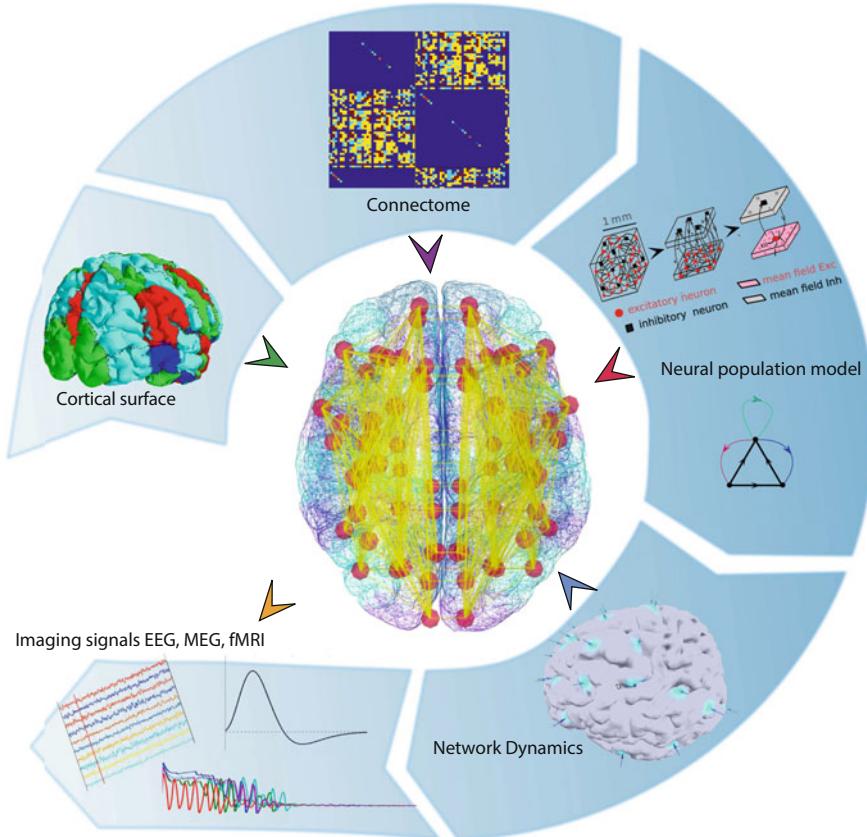


Fig. 16.6 Work flow in The Virtual Brain (TVB). From top left, clockwise: TVB is a neuroinformatics platform integrating biologically realistic structural constraints on connectivity and geometry with simulations of neural fields. The geometry of the cortical surface in 3D is integrated with the short-range connectivity (grey matter) to build neural fields on the folded cortical surface. The set of white matter fibers (long range) comprising the Connectome is embedded into the total connectivity matrix. Each network node holds a neural mass model, that receives input from all the network links. The signal transmission via the Connectome is time-delayed due to finite transmission speeds. As a whole the elements define the full brain network and allow for large-scale simulations under manipulations of the connectivity, neural mass model parameters or external stimulations. The output of the simulations are the BOLD, and electro- and magnetoencephalographic (EEG, MEG) signatures as observed in non-invasive human brain imaging

simulations performed on the folded cortical surface generate the biologically realistic electrographic and magnetographic (EEG, MEG) time series, as well as the volumetric fMRI time series that researchers are used to from human brain imaging. The architecture of TVB supports interaction with MATLAB packages, for example, in order to perform graph theoretic network analyses, the well known Brain Connectivity Toolbox has been integrated into TVB. TVB can be used in

a client-server configuration, such that it can be remotely accessed through the Internet thanks to its web based HTML5, JS and WebGL graphical user interface. Otherwise, TVB is also accessible as a standalone cross-platform Python library and application, and users can interact with the scientific core through the Python interactive scripting interface IDLE, enabling easy modeling, development and debugging of the scientific kernel. This second interface makes it possible to extend TVB by combining it with other libraries and modules developed by the Python scientific community. TVB is available from <http://thevirtualbrain.org>.

16.7 Conclusions and Final Remarks

Traditional neural field models are characterized by translationally invariant integral kernels reflecting the connectivity of the brain network. Such models are dominated though by either short range connectivity and are not suited to describe large scale brain models; or they suffer from the low pass filtering properties in the spatial domain (Laplacian approximation of a realistic large-scale connectivity) and do not capture the spatial details of the spatiotemporal organization. To overcome this limitation the suggestion was made to integrate biologically realistic DTI based connectivity into full brain modeling efforts. Large scale brain dynamics are basically expected to reflect the underlying anatomical connectivity between brain areas [11, 20], even though structural connectivity is not the only constraint. The transmission delays also play an essential role in shaping the brain network dynamics [19, 25, 36, 38, 39]. Toy models of neural fields with embedded heterogeneous fibers allow to study and illustrate the effects of the interplay between the local and global fiber systems, as well as the effect of the time delays. These studies underwrite the need to incorporate realistic connectivity into large scale brain network models. In this sense The Virtual Brain (TVB) has evolved from a research program seeking to identify and reproduce realistic full brain network dynamics, on the basis of empirical connectivity and neural field models.

Acknowledgements The work reported here was supported by the Brain Network Recovery Group through the James S. McDonnell Foundation and the FP7-ICT BrainScales. The author thanks Paula Sanz Leon and Enrique Hansen for help with the preparation of figures.

References

1. Amari, S.: Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.* **27**, 77–87 (1977)
2. Biswal, B., Yetkin, F.Z., Haughton, V.M., Hyde, J.S.: Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn. Reson. Med.* **34**(4), 537–541 (1995)

3. Bojak, I., Liley, D.: Modeling the effects of anesthesia on the electroencephalogram. *Phys. Rev. E* **71**, 041,902 (2005)
4. Bojak, I., Oostendorp, T., Reid, A., Kötter, R.: Connecting mean field models of neural activity to eeg and fmri data. *Brain Topogr.* **23**(2), 139–149 (2010)
5. Bojak, I., Oostendorp, T., Reid, A., Kötter, R.: Towards a model-based integration of co-registered electroencephalography/functional magnetic resonance imaging data with realistic neural population meshes. *Phil. Trans. R. Soc. Lond. A.* **369**(1952), 3785–3801 (2011)
6. Brackley, C., Turner, M.: Two-point heterogeneous connections in a continuum neural field model. *Biol. Cybern.* **100**, 371–383 (2009)
7. Braintenberg, V., Schüz, A.: Anatomy of the cortex. Statistics and geometry. Springer, Berlin/Heidelberg/New York (1991)
8. Breakspear, M., Terry, J., Friston, K.: Modulation of excitatory synaptic coupling facilitates synchronization and complex dynamics in a biophysical model of neuronal dynamics. *Netw. Comput. Neural Syst.* **14**, 703–732 (2003)
9. Breakspear, M., Roberts, J.A., Terry, J.R., Rodrigues, S., Mahant, N., Robinson, P.A.A.: Unifying explanation of primary generalized seizures through nonlinear brain modeling and bifurcation analysis. *Cereb. Cortex* **16**, 1296–1313 (2006)
10. Brunel, N., Wang, X.J.: What determines the frequency of fast network oscillations with irregular neural discharges? I. Synaptic dynamics and excitation-inhibition balance. *J. Neurophysiol.* **90**, 415–430 (2003)
11. Bullmore, E., Sporns, O.: Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* **10**(3), 186–198 (2009)
12. Bullmore, E., Sporns, O.: The economy of brain network organization. *Nat. Rev. Neurosci.* **13**(5), 336–349 (2012)
13. Cabral, J., Hugues, E., Sporns, O., Deco, G.: Role of local network oscillations in resting-state functional connectivity. *Neuroimage* **57**, 130–139 (2011)
14. Campbell, S.: Time delays in neural systems. In: Jirsa, V.K., McIntosh, A.R. (eds.) *Handbook of Brain Connectivity* pp. 65–90. Springer, Berlin/New York (2007)
15. Coombes, S., Lord, G., Owen, M.: Waves and bumps in neuronal networks with axo-dendritic synaptic interactions. *Physica D* **178**, 219–241 (2003)
16. Datko, R.: A procedure for determination of the exponential stability of certain differential-difference equations. *Q. Appl. Math.* **36**, 279–292 (1978)
17. Deco, G., Jirsa, V.: Ongoing cortical activity at rest: criticality, multistability and ghost attractors. *J. Neurosci.* **32**, 3366–3375 (2012)
18. Deco, G., Jirsa, V., Robinson, P.A., Breakspear, M., Friston, K.: The dynamic brain: from spiking neurons to neural masses and cortical fields. *PLoS Comput. Biol.* **4**(8), e1000,092 (2008)
19. Deco, G., Jirsa, V., McIntosh, A., Sporns, O., Kötter, R.: Key role of coupling, delay, and noise in resting brain fluctuations. *Proc. Natl. Acad. Sci. U.S.A.* **106**(25), 10,302–10,307 (2009)
20. Deco, G., Jirsa, V., McIntosh, A.: Emerging concepts for the dynamical organization of resting-state activity in the brain. *Nat. Rev. Neurosci.* **12**(1), 43–56 (2011)
21. Deco, G., Ponce-Alvarez, A., Mantini, D., Romani, G.L., Hagmann, P., Corbetta, M.: Resting-state functional connectivity emerges from structurally and dynamically shaped slow linear fluctuations. *J. Neurosci.* **33**(27), 11,239–11,252 (2013)
22. Feng, J., Jirsa, V., Ding, M.: Synchronization in networks with random interactions: theory and applications. *Chaos* **16**, 015,109 (2006)
23. Freyer, F., Roberts, J.A., Becker, R., Robinson, P.A., Ritter, P., Breakspear, M.: Biophysical mechanisms of multistability in resting-state cortical rhythms. *J. Neurosci.* **31**(17), 6353–6361 (2011)
24. Ghosh, A., Rho, Y., McIntosh, A.R., Kötter, R., Jirsa, V.K.: Cortical network dynamics with time delays reveals functional connectivity in the resting brain. *Cogn. Neurodyn.* **2**, 115–120 (2008)
25. Ghosh, A., Rho, Y., McIntosh, A.R., Kötter, R., Jirsa, V.K.: Noise during rest enables the exploration of the brain's dynamic repertoire. *PLoS Comput. Biol.* **4**(10), e1000196 (2008)

26. Greicius, M.D., Krasnow, B., Reiss, A.L., Menon, V.: Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 253–258 (2003)
27. Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C.J., Wedeen, V.J., Sporns, O.: Mapping the structural core of human cerebral cortex. *PLoS Biol.* **6**(7), e159 (2008)
28. Honey, C.J., Kötter, R., Breakspear, M., Sporns, O.: Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proc. Natl. Acad. Sci. U.S.A.* **104**(24), 10,240–10,245 (2007)
29. Hutt, A.: The anaesthetic propofol shifts the frequency of maximum spectral power in eeg during general anaesthesia: analytical insights from a linear model. *Front. Comput. Neurosci.* **7**(2), 1–15 (2013)
30. Hutt, A., Atay, F.M.: Analysis of nonlocal neural fields for both general and gamma-distributed connectivities. *Physica D* **203**(1–2), 30–54 (2005)
31. Jirsa, V.K.: Dispersion and time delay effects in synchronized spike-burst networks. *Cogn. Neurodyn.* **2**, 29–38 (2008)
32. Jirsa, V.K.: Neural field dynamics with local and global connectivity and time delay. *Phil. Trans. R. Soc. A* **367**, 1131–1143 (2009)
33. Jirsa, V.K., Ding, M.: Will a large complex system with time delays be stable? *Phys. Rev. Lett.* **93**, 070,602:1–4 (2004)
34. Jirsa, V.K., Haken, H.: Field theory of electromagnetic brain activity. *Phys. Rev. Lett.* **77**, 960–963 (1996)
35. Jirsa, V.K., Haken, H.: A derivation of a macroscopic field theory of the brain from the quasi-microscopic neural dynamics. *Physica D* **99**, 503–526 (1997)
36. Jirsa, V.K., Kelso, J.A.: Spatiotemporal pattern formation in neural systems with heterogeneous connection topologies. *Phys. Rev. E: Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.* **62**(6 Pt B), 8462–8465 (2000)
37. Jirsa, V.K., Jantzen, K.J., Fuchs, A., Kelso, J.A.S.: Spatiotemporal forward solution of the EEG and MEG using network modelling. *IEEE Trans. Med. Imaging* **21**, 493–504 (2002)
38. Jirsa, V.K., Sporns, O., Breakspear, M., Deco, G., McIntosh, A.R.: Towards the virtual brain: network modeling of the intact and the damaged brain. *Arch. Ital. Biol.* **148**(3), 189–205 (2010)
39. Knock, S., McIntosh, A., Sporns, O., Kötter, R., Hagmann, P., Jirsa, V.K.: The effects of physiologically plausible connectivity structure on local and global dynamics in large scale brain models. *J. Neurosci. Methods* **183**(1), 86–94 (2009)
40. Kötter, R.: Online retrieval, processing, and visualization of primate connectivity data from the cocomac database. *Neuroinformatics* **2**, 127–144 (2004)
41. Laing, C.: PDE methods for two-dimensional neural fields. In: Coombes S., Beim Graben P., Potthast R., Wright J. (eds.) *Neural Fields: Theory and Applications*. Springer, Berlin/Heidelberg (2014)
42. Liley, D.T., Alexander, D.M., Wright, J.J., Aldous, M.D.: Alpha rhythm emerges from large-scale networks of realistically coupled multicompartmental model cortical neurons. *Network* **10**(1), 79–92 (1999)
43. Lowe, M.J., Mock, B.J., Sorenson, J.A.: Functional connectivity in single and multislice echoplanar imaging using resting-state fluctuations. *Neuroimage* **7**, 119–132 (1998)
44. Nunez, P.L.: The brain wave equation: A model for EEG. *Mathematical Biosciences* **21**, 279–297 (1974)
45. Nunez, P.: *Neocortical Dynamics and Human EEG Rhythms*. Oxford University Press, New York/Oxford (1995)
46. Pinotsis, D., Hansen, E., Friston, K., Jirsa, V.: Anatomical connectivity and the resting state activity of large cortical networks. *Neuroimage* **65**, 127–138 (2013)
47. Qubaj, M., Jirsa, V.K.: Neural field dynamics with heterogeneous connection topology. *Phys. Rev. Lett.* **93**, 238,102 (2007)
48. Qubaj, M., Jirsa, V.K.: Neural field dynamics under variation of local and global connectivity and finite transmission speed. *Physica D* **238**, 2331–2346 (2009)

49. Robinson, P.A., Rennie, C.A., Wright, J.J.: Propagation and stability of waves of electrical activity in the cerebral cortex. *Phys. Rev.E* **56**, 826–840 (1997)
50. Robinson, P.A., Loxley, P.N., O'Connor, S.C., Rennie, C.J.: Modal analysis of corticothalamic dynamics, electroencephalographic spectra, and evoked potentials. *Phys. Rev. E Stat. Nonlinear Soft Matter Phys.* **63**(4 Pt 1), 041,909 (2001)
51. Robinson, P.A., Rennie, C.J., Wright, J.J., Bahramali, H., Gordon, E., Rowe, D.L.: Prediction of electroencephalographic spectra from neurophysiology. *Phys. Rev. E Stat Nonlinear Soft Matter Phys.* **63**(2 Pt 1), 021,903 (2001)
52. Rogers, B.P., Morgan, V.L., Newton, A.T., Gore, J.C.: Assessing functional connectivity in the human brain by fMRI. *Magn. Reson. Imaging* **25**, 1347–1357 (2007)
53. Sanz Leon, P., Knock, S., Woodman, M., Domide, L., Mersmann, J., McIntosh, A., Jirsa, V.: The virtual brain: a simulator of primate brain network dynamics. *Front. Neuroinfl.* **10**, 1–23 (2013)
54. Spiegler, A., Jirsa, V.K.: Systematic approximations of neural fields through networks of neural masses in the virtual brain. *Neuroimage* **83C**, 704–725 (2013)
55. Stefanescu, R., Jirsa, V.K.: A low dimensional description of globally coupled heterogeneous neural networks of excitatory and inhibitory. *PLoS Comput. Biol.* **4**(11), 26–36 (2008)
56. Stefanescu, R., Jirsa, V.K.: Reduced representations of heterogeneous mixed neural networks with synaptic coupling. *Phys. Rev. E: Stat. Nonlinear Soft Matter Phys.* **83**(2), 026204 (2011)
57. Wilson, H.R., Cowan, J.D.: Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.* **12**, 1–24 (1972)
58. Wilson, H.R., Cowan, J.D.: A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetic* **13**, 55–80 (1973)
59. Wright, J.J., Liley, D.: Dynamics of the brain at global and microscopic scales: Neural networks and the eeg. *Behav. Brain Sci.* **19**, 285–320 (1996)
60. Wright, J.J., Liley, D.T.J.: Simulation of electrocortical waves. *Biol. Cybern.* **72**, 347–356 (1995)

Chapter 17

Neural Fields, Masses and Bayesian Modelling

Dimitris A. Pinotsis and Karl J. Friston

Abstract This chapter considers the relationship between neural field and mass models and their application to modelling empirical data. Specifically, we consider neural masses as a special case of neural fields, when conduction times tend to zero and focus on two exemplar models of cortical microcircuitry; namely, the Jansen-Rit and the canonical microcircuit model. Both models incorporate parameters pertaining to important neurobiological attributes, such as synaptic rate constants and the extent of lateral connections. We describe these models and show how Bayesian inference can be used to assess the validity of their field and mass variants, given empirical data. Interestingly, we find greater evidence for neural field variants in analyses of LFP data but fail to find more evidence for such variants, relative to their neural mass counterparts, in MEG (virtual electrode) data. The key distinction between these data is that LFP data are sensitive to a wide range of spatial frequencies and the temporal fluctuations that these frequencies contain. In contrast, the lead fields, inherent in non-invasive electromagnetic recordings, are necessarily broader and suppress temporal dynamics that are expressed in high spatial frequencies. We present this as an example of how neuronal field and mass models (hypotheses) can be compared formally.

17.1 Introduction

This chapter reviews recent developments in the modelling of brain imaging data that exploits neural field theory. We focus on the Bayesian optimization of model parameters, called *Dynamic Causal Modelling* (DCM) and its application in the context of neural fields [18, 46]. This framework is part of the academic freeware

D.A. Pinotsis (✉) • K.J. Friston

The Wellcome Trust Centre for Neuroimaging, University College London, Queen Square,
London, WC1N 3BG, UK

e-mail: d.pinotsis@ucl.ac.uk

*Statistical Parametric Mapping*¹ (SPM), which is a popular platform for analyzing neuroimaging data, used by several neuroscience communities worldwide. DCM allows for a formal (Bayesian) statistical analysis of cortical network connectivity, based upon realistic biophysical models of brain responses. It is this particular feature of DCM—the unique combination of generative models with optimization techniques based upon (variational) Bayesian principles—that furnishes a novel way to characterize brain organization. In particular, it provides answers to questions about how the brain is wired and how it responds in different situations. In this chapter, we first present the general framework with an emphasis on the role of neural fields. We then consider particular applications, in the context of LFP [39] and MEG data [56] and show how DCM allows one to adjudicate between alternative models of brain imaging data, such as neural masses and fields. This comparison is based upon the free energy approximation to Bayesian model evidence that comprises both accuracy and complexity (see e.g. [19]). This allows one to adjudicate among competing models, in terms of explanations for data that are both accurate and parsimonious.

17.1.1 Neural Masses and Fields

Neural mass models are a particular case of neural fields, where the (hidden neuronal) states of populations of neurons are functions of time only. Such models can generate temporal responses from one or several interconnected populations and have been used successfully to explain empirical electrophysiological data from local field potentials (LFP) and EEG/MEG (see e.g. [28,31,34,35,37,49,57–59]). To date, neural mass models have been largely based upon point sources and formulated using ordinary differential equations (ODEs). A key challenge in this area has been to model observed signals as being generated by continuous and spatially distributed neuronal activity, of the sort observed directly using high density multi-electrode arrays and optical imaging. Here, we consider how one can address this challenge using neural field models.

Neural field models model current fluxes as continuous processes on a cortical manifold, using integro-differential or partial differential equations (PDEs) (see [13] for a review and also [2,4–6,8–10,16,17,21,25,26,32,33,41,42,47,50–53]). These models have also been used to study anaesthetic action and associated macroscopic electroencephalographic effects (for a review, see [30]). The key advance that neural field models offer, over conventional neural mass models, is that they embody spatial parameters (like the density and extent of lateral connections). This means, in principle, one can estimate the spatial parameters of cortical infrastructures generating electrophysiological signals (and infer changes in those parameters over different levels of an experimental factor) from empirical data. This rests on

¹<http://www.fil.ion.ucl.ac.uk/spm/>.

modelling responses not just in time but also over space. Clearly, to exploit this sort of model, one needs to measure the temporal dynamics of observed cortical responses over different spatial scales; for example, with high-density recordings, at the epidural or intracortical level. However, the impact of spatially extended dynamics is not restricted to expression over space but can also have profound effects on temporal (e.g., spectral) responses at one point (or averaged locally over the cortical surface). This means that neural field models can also play a key role in the modelling of non-invasive electrophysiological data that does not resolve spatial activity directly.

17.1.2 *Neural Fields as Models for Empirical Data*

The modelling of electrophysiological signals depends upon models of how they are generated in source space and how the resulting (hidden) neuronal states are detected by sensors. At the source level we consider two neural field models that are inspired by biophysical considerations: an extension of a widely used mass model introduced by Jansen and Rit comprising three neural populations [24]; and a canonical microcircuit field model, where the pyramidal cell population of the previous model is split into two subpopulations. This separates the sources of forward and backward connections in cortical hierarchies that have proven useful to explain several aspects of distributed cortical computations in theoretical neurobiology.

In terms of the mapping from source to sensor space, we use a conventional lead field formulation (that is expanded in terms of spatial basis functions) and focus on the ensuing observations of power spectra of the sort considered in conventional time-series analysis. There is a long history of modelling steady-state (or ongoing) activity spectra, associated with neural fields, usually in models of the whole cortex; e.g., [25]. Robinson and colleagues [50] have developed a neurophysiologically grounded model of corticothalamic activity, which reproduces many properties of empirical EEG signals; such as the spectral peaks seen in various sleep states and seizure activity. Technically, the spectra summarizing the response of cortical sources can be defined in terms of transfer functions, mapping endogenous neuronal fluctuations to observed responses [15, 40, 51–53]. We will derive transfer functions and expressions for the spectral responses for sources that comprise multiple layers. This allows us to model the spectral activity of cortical fields as measured on the cortical surface and to also compute the corresponding spectral responses, as measured by invasive (LFP) and non-invasive (MEG) electrophysiological sensors in an efficient manner.

The resulting scheme can be regarded as inverting population models of the Amari type—using real data—with a particular focus on Bayesian model inversion. Previous work in a similar vein includes the use of Kalman filters to develop estimation schemes for both neural mass [49, 58] and neural field models of a

single population [20, 55]. In a related approach, [11] replaced the standard dipole source—used in neural mass models—with the principal Fourier mode of a neural field, for the particular case of exponentially decaying synaptic density over the cortical surface. In [36], the authors used a combination of linear and nonlinear optimisation methods to invert a two-layered neural field model of voltage-sensitive dye data, describing inhibitory and excitatory populations (without conduction delays). Finally, the inverse problem for the neural field equation was rigorously studied in [44].

17.2 Generative Models for Cross Spectral Densities

Dynamic Causal Modelling (DCM) allows for the comparison and estimation of biophysically plausible models of fMRI, EEG, MEG and LFP data [12, 18, 37–39]. DCM calls on an underlying generative model to predict observed data. As with other state-space models, Dynamic Causal Models are based on a combination of evolution equations for the hidden neuronal states with static observer equations:

$$\begin{aligned}\dot{V}(t) &= f(V, U, \theta) \\ Y(t) &= L(V, \theta)\end{aligned}\tag{17.1}$$

where $V(t) \in \mathbb{R}$ and $U(t) \in \mathbb{R}$ are vectors of hidden state variables and inputs respectively, $Y(t)$ are the predicted time series and we use θ to denote the parameters of the model. Many generative models can be cast in the form of Eq. (17.1): these include neural mass models, such as the Jansen and Rit model [24] considered below. Most DCMs for electrophysiological data have been based on neural mass models, which use point sources (e.g., equivalent current dipoles) and preclude spatially extended dynamics. The DCM introduced in [46] allows for an explicit modelling of the spatiotemporal aspects of cortical activity. This means one can make inferences about parameters pertaining to the topographic distribution of cortical sources from LFP and M/EEG data; like the extent of connectivity and the conduction velocity of axonal propagation. Furthermore, including spatial parameters allows one to explain some effects that other models such as neural masses attribute to time parameters. In particular, neural field models characterize the propagation of electrical activity on the cortex to provide a more complete parameterization of the mechanisms generating cortical responses. In the models considered here, the postsynaptic convolution of presynaptic inputs is described by a second-order ODE or two first-order ODEs pertaining to voltage and current. This means that the left hand side of Eq. (17.1) is augmented with the second derivative of hidden (depolarization) states. For cortical sources comprising of n layers, our neural field formulation is based on the following extension of Eq. (17.1)

$$\begin{aligned}
\ddot{V} &= -2B\dot{V} - B^2V + D \otimes F \circ V + G \circ U \\
Y(t) &= \int L(x, \varphi) \cdot V(x, t) \, dx \\
D \otimes Q &= \iint D(x - x', t - t') \cdot Q(x', t') \, dx' \, dt' \\
V(x, t) &= \begin{bmatrix} v_1(x, t) \\ \vdots \\ v_n(x, t) \end{bmatrix},
\end{aligned} \tag{17.2}$$

where the $n \times 1$ vector $V(t) \in \mathbb{R}^n$ of hidden neuronal states in each layer in Eq. (17.1) is replaced by $V(x, t) \in \mathbb{R}^n$; both this vector and the input $U(x, t) \in \mathbb{R}^n$ are explicit functions of both space and time. The dynamics of cortical sources now conform to integrodifferential equations, as implied by the spatiotemporal convolution denoted by \otimes on the right hand side of Eq. (17.2).

In the above equation, $D(x, t)$ is a $n \times n$ smooth (analytic) matrix-valued connectivity function or kernel; $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a nonlinear mapping from postsynaptic depolarization to presynaptic firing rates at each point on the cortical manifold and B is a $n \times n$ matrix encoding average synaptic decay rates. In short, Eq. (17.2) says that the rate of change of voltage in each layer comprises three terms; the first is a simple decay, the second is due to presynaptic inputs from other parts of the cortical manifold and the final part is due to external inputs, where $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$ maps the inputs to the motion of hidden states. It is the second component, involving the convolution with the connectivity kernel $D(x, t)$ that embodies lateral interactions over the cortical manifold. In terms of the observer function, the linear mapping from hidden states to observed signal now becomes a $m \times n$ matrix function of source space $L(x, \varphi)$ encoding the contribution of the n hidden states to each of m sensors.

17.2.1 The Jansen and Rit Model

In this section, we provide a brief review of the well-known Jansen and Rit (JR) neural mass model [24]. In the JR model, each cortical source is modelled with three subpopulations: excitatory spiny stellate input cells, inhibitory interneurons and deep excitatory output pyramidal cells (for classical approaches to modelling such populations with neural fields, see e.g. [1, 15, 40, 62]). For simplicity, we consider a single source, noting that extensions to multiple sources involve adding extrinsic (between-source) connections or kernels. The JR model is a particular instance of Eq. (17.1) (where the convolution of presynaptic input is second-order):

$$\begin{aligned}
\ddot{V} + 2B\dot{V} &= -2B^2V - ABF \circ V + G \circ U \\
Y &= L \cdot V.
\end{aligned} \tag{17.3}$$

Here, A and B are the 3×3 matrices of synaptic parameters controlling the maximum postsynaptic responses and the rate-constants of postsynaptic filtering and G is a column vector:

$$\begin{aligned} A &= \text{diag}(m_e, m_i, m_e) \\ B &= \text{diag}(\kappa_e, \kappa_i, \kappa_e) \\ G &= \begin{bmatrix} \kappa_e m_e \\ 0 \\ 0 \end{bmatrix} \end{aligned} \quad (17.4)$$

As above $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is a nonlinear mapping from depolarization to firing and $U(x, t) \in \mathbb{R}^3$ is the external input to each population. Note that there is only one input and this input enters the first (spiny stellate) population. Based on cortical microcircuitry of intrinsic connections, the JR model prescribes the mapping $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ in terms of nonlinear firing rate functions of the depolarization in the three populations. Writing out Eq. (17.3) in full we have

$$\begin{aligned} \ddot{v}_1 + 2\kappa_e \dot{v}_1 + \kappa_e^2 v_1 &= \kappa_e m_e (d_{13} \cdot \sigma(v_3) + U) \\ \ddot{v}_2 + 2\kappa_i \dot{v}_2 + \kappa_i^2 v_2 &= \kappa_i m_i d_{23} \cdot \sigma(v_3) \\ \ddot{v}_3 + 2\kappa_e \dot{v}_3 + \kappa_e^2 v_3 &= \kappa_e m_e (d_{31} \cdot \sigma(v_1) - d_{32} \cdot \sigma(v_2)) \end{aligned} \quad (17.5)$$

where $v_a(t) : a = 1, 2, 3$ denotes the expected depolarization in the a -th population (excitatory stellate, inhibitory population and excitatory pyramidal respectively) and the sigmoid firing rate function is

$$\sigma(v_a) = \frac{1}{1 + \exp(r(\eta - v_a))}. \quad (17.6)$$

Here, r and η are parameters that determine the shape of this sigmoid activation function. In particular, r is synaptic gain and η is the postsynaptic potential at which the half of the maximum firing rate is elicited. In Eq. (17.5), $d_{ab} \cdot \sigma(v_b)$ is (endogenous) presynaptic input to the a -th population from the b -th and corresponds to the mapping $F \circ V$. See Fig. 17.1 for a schematic of this model.

17.2.2 The Canonical Microcircuit Model

Recent work suggests that superficial layers of visual cortex oscillate preferentially at gamma frequencies, while deep layers primarily oscillate at lower alpha/beta frequencies [7]. Since forward connections originate largely from superficial layers and backward connections primarily originate from deep layers, these spectral asymmetries suggest that forward connections use faster (gamma) temporal frequencies, while backward connections may employ lower (beta) frequencies—a suggestion that has experimental support [54]. These asymmetries mandate

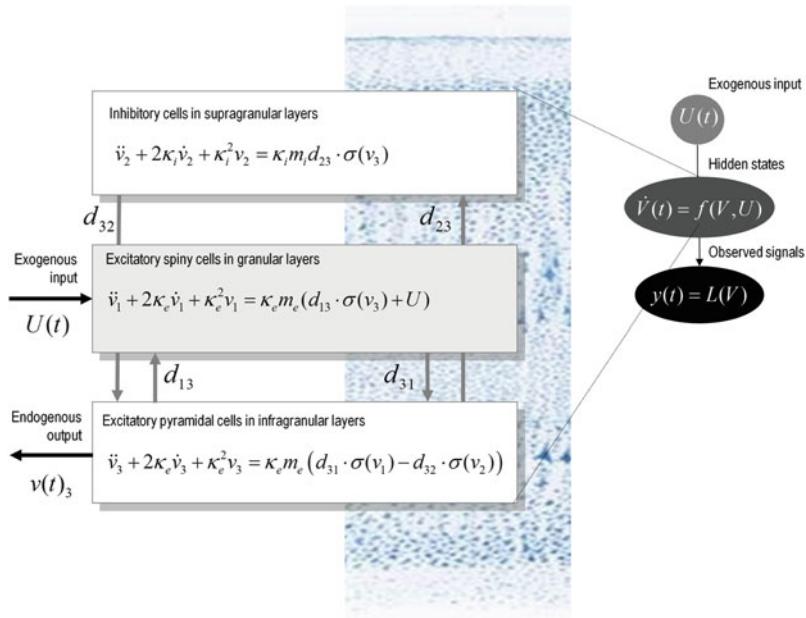


Fig. 17.1 This schematic summarizes the equations of motion or state equations that specify a Jansen and Rit neural mass model of a single source. This model contains three populations, each loosely associated with a specific cortical subpopulation or layer. The second-order differential equations describe changes in hidden states (e.g., voltage) that subtend observed local field potentials or EEG signals. These differential equations effectively implement a linear convolution of presynaptic activity to produce postsynaptic depolarization. Average firing rates within each subpopulation are then transformed through a nonlinear (sigmoid) voltage-firing rate function to provide inputs to other populations. These inputs are weighted by connection strengths

something quite remarkable: namely, a synthesis and segregation of forward and backward output from afferent input. This segregation can only arise from local neuronal computations that are formally structured and precisely interconnected. The canonical microcircuit is a detailed proposal for such a laminar-specific intracortical architecture that describes how information flows through the cortical column. This model is based on findings in the primary visual cortex [14] but recent work [29, 61] indicates that similar microcircuits exist in other regions—such as somatosensory and motor cortex.

Douglas and Martin recorded intracellular potentials from cells in area 17 of the cat, while they stimulated cortical afferents, and noticed a strong compartmentalization of the superficial and deep cell properties—reflected in slow superficial responses and fast input layer responses. The authors created a conductance-based model that reproduced the evolution of excitation and inhibition through the cortical circuit with great precision. This model contained three groups of cells: superficial and deep pyramidal cells, and a common pool of inhibitory cells. All three pools of neurons receive thalamic drive—although the thalamic drive to deep layer cells

was weaker than the other inputs. All neuronal populations had self-connectivity, and are interconnected with the other populations. Finally, inhibition was stronger onto the deep pyramidal population. This model was able to reproduce the features observed in their electrophysiological recordings—including the latency difference between superficial and input layer neurons—and has served to establish several basic properties that are now believed to be replicated in other cortical areas: first, although superficial and deep compartments are strongly interconnected, their response properties are also segregated. Second, cortex is not under tonic inhibition, rather, both excitation and inhibition are generated by afferent thalamic input and both shape ongoing cortical responses. Third, the canonical microcircuit can amplify thalamic inputs to generate self-sustaining activity, while also maintaining a delicate balance between excitation and inhibition—so as to prevent runaway excitation.

Previous computational modelling studies indicate that this circuitry allows the cortex to optimally organize and integrate bottom-up, lateral, and top-down information [48]. Douglas and Martin suggest that the rich anatomical connectivity of superficial layer 2/3 pyramidal cells allows them to collect information from all relevant top-down, lateral, and bottom-up inputs, and—through processing in the dendritic tree—select the most likely interpretation of its subcortical inputs. For a discussion and more details about the canonical microcircuit and its potential role in predictive coding we refer the reader to [3].

Haeusler and Maass used Hodgkin and Huxley neurons to build a realistic microcircuit model and showed that a cortical column—whose connectivity conforms to the canonical microcircuit—can perform various computations efficiently, in relation to a column with random connectivity [23]. By collapsing two pairs of cell types in the Haeusler and Maass model, while preserving the topology of the connectivity, one obtains the canonical microcircuit (CMC) depicted in Fig. 17.2: this circuit comprises four populations: excitatory spiny stellate input cells (17.1), inhibitory interneurons (17.2), deep excitatory pyramidal cells (17.3) and superficial excitatory pyramidal cells (17.4). The corresponding evolution equations for the neuronal states (the analogues of Eq. 17.5) are (see also Fig. 17.2):

$$\begin{aligned}\ddot{v}_1 + 2\kappa_e \dot{v}_1 + \kappa_e^2 v_1 &= \kappa_1 m_e (-d_{14} \cdot \sigma(v_4) + d_{11} \cdot \sigma(v_1) - d_{12} \cdot \sigma(v_2) + U) \\ \ddot{v}_2 + 2\kappa_i \dot{v}_2 + \kappa_i^2 v_2 &= \kappa_2 m_i (d_{21} \cdot \sigma(v_1) + d_{22} \cdot \sigma(v_2) + d_{23} \cdot \sigma(v_3)) \\ \ddot{v}_3 + 2\kappa_e \dot{v}_3 + \kappa_e^2 v_3 &= \kappa_3 m_e (-d_{32} \cdot \sigma(v_2) + d_{33} \cdot \sigma(v_3)) \\ \ddot{v}_4 + 2\kappa_e \dot{v}_4 + \kappa_e^2 v_4 &= \kappa_4 m_e (d_{41} \cdot \sigma(v_1) + d_{44} \cdot \sigma(v_4))\end{aligned}\tag{17.7}$$

17.2.3 Neural Field Extensions of the Canonical Microcircuit and Jansen-Rit Models

In this section, we transcribe the neural mass models described above into neural fields. In the case of fields, we consider spatially extended sources occupying

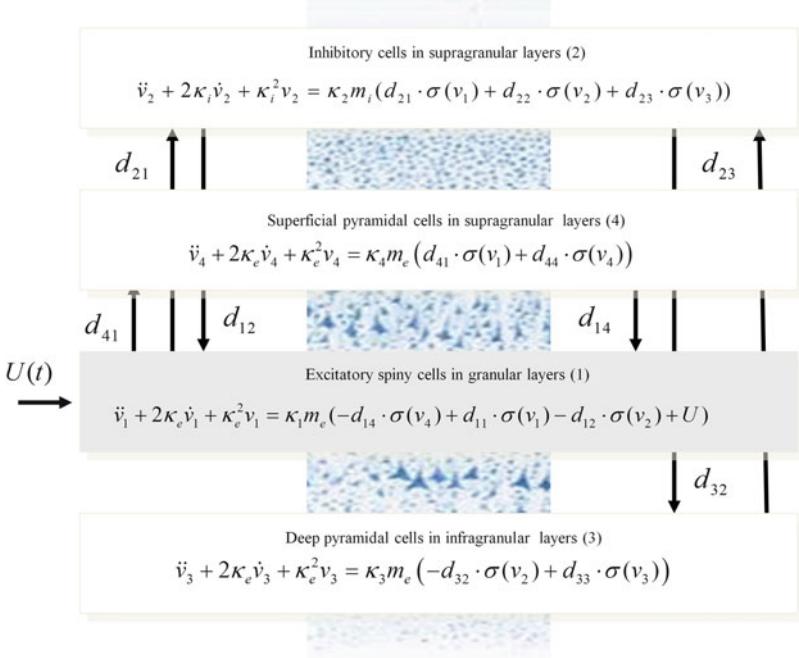


Fig. 17.2 This figure shows the evolution equations that specify a Canonical Microcircuit neural mass model of a single source. This model contains four populations occupying different cortical layers: the pyramidal cell population of the JR model is here split into two subpopulations allowing a separation of the sources of forward and backward connections in cortical hierarchies. As with the JR model, second-order differential equations mediate a linear convolution of presynaptic activity to produce postsynaptic depolarization. This depolarization gives rise to firing rates within each sub-population that provide inputs to other populations

bounded manifolds (patches) in different layers that lie beneath the cortical surface. In this setting, each subpopulation of a neural mass model now becomes a layer in the cortical sheet. The dynamics of cortical sources conform to integrodifferential equations, such as the Wilson-Cowan or Amari equations, where coupling is parameterised by matrix-valued coupling kernels—namely, smooth (analytic) connectivity matrices that also depend on time and space. Assuming that the connectivity kernels $d_{ij}(x, t)$ appearing in Eq. (17.2) factorize into $d_{ij}(x, t) = \kappa_{ij}(|x|)\delta(t - |x - x'|v)$, Eq. (17.2) becomes

$$(\ddot{V} + 2B\dot{V} + B^2 V)(x, t) = AB \int K(x - x') F \circ V(x', t - |x - x'|v) dx' + GU \quad (17.8)$$

where v is the inverse speed at which spikes propagate along connections and interactions among populations—within and across macrocolumns—are described by the connectivity kernel $K = K^{(i)} + K^{(e)}$. This form provides an explicit

parameterization of conduction delays that will be exploited later, when using the field model as an observation model. One can see that in the infinite speed limit $v = 0$ the spatial convolution in Eq. (17.8) disappears (to within a scaling constant) and we recover the neural mass Eqs. (17.3).

For both the canonical microcircuit and the Jansen and Rit models, the intrinsic connectivity $K^{(i)}$ is an exponentially decaying kernel, commonly used in the literature to account for excitatory and inhibitory interactions (see e.g. [46]). The extrinsic connectivity kernel $K^{(e)}$ was introduced in [22,45] to model patchy lateral (horizontal) connections and is characterized by non-central peaks allowing for differences in (and estimation of) the range and dispersion of lateral connections, summarized in terms of the parameters h_a and c_{aa} , respectively. The canonical microcircuit (CMC) field model we consider here is described by Eq. (17.8), where the connectivity kernel is given by

$$K = K^{(i)} + K^{(e)}$$

$$K^{(i)} = \begin{bmatrix} k_{11}^{(i)} & k_{12}^{(i)} & 0 & k_{14}^{(i)} \\ k_{21}^{(i)} & k_{22}^{(i)} & k_{23}^{(i)} & 0 \\ 0 & k_{32}^{(i)} & k_{33}^{(i)} & 0 \\ k_{41}^{(i)} & 0 & 0 & k_{44}^{(i)} \end{bmatrix}, \quad K^{(e)} = \begin{bmatrix} k_{11}^{(e)} & 0 & 0 & 0 \\ 0 & k_{22}^{(e)} & 0 & 0 \\ 0 & 0 & k_{33}^{(e)} & 0 \\ 0 & 0 & 0 & k_{44}^{(e)} \end{bmatrix} \quad (17.9)$$

$$k_{ab}^{(i)}(x) = \frac{1}{2}a_{ab}e^{-c_{ab}|x|}$$

$$k_{aa}^{(e)}(x) = \frac{1}{2}c_{aa}(e^{-c_{aa}|x-h_a|} + e^{-c_{aa}|x+h_a|}).$$

Here, the parameters a_{ab} and c_{ab} encode the strength (analogous to the number of synaptic connections) and extent of intrinsic connections between cortical layers. The intrinsic connections can be regarded as inter-laminar connections within a macrocolumn, while the extrinsic—between macrocolumn connections—correspond to horizontal connections and connect layers of the same type at a distance h_a , see Fig. 17.3.

This kernel models sparse intrinsic connections in cortical circuits that mediate both local and non-local interactions and allows one to estimate properties of lateral interactions that are particularly relevant in the context of data obtained from retinotopically mapped visual cortex. We will see an example of this using an attention task and MEG data below. On the other hand, to illustrate the Jansen and Rit field model in non-mapped cortex, we will use data from the auditory cortex under anaesthesia and neglect extrinsic connectivity; that is $K^{(e)} = 0$. In this example, we will use the same exponential decaying form, as for the CMC field model above, namely the function $K(x) = K^{(i)}(x)$, where

$$K^{(i)} = \begin{bmatrix} 0 & 0 & k_{13}^{(i)} \\ 0 & 0 & k_{23}^{(i)} \\ k_{31}^{(i)} & k_{32}^{(i)} & 0 \end{bmatrix} \quad (17.10)$$

and $k_{ab}^{(i)}$ are given by Eq. (17.9).

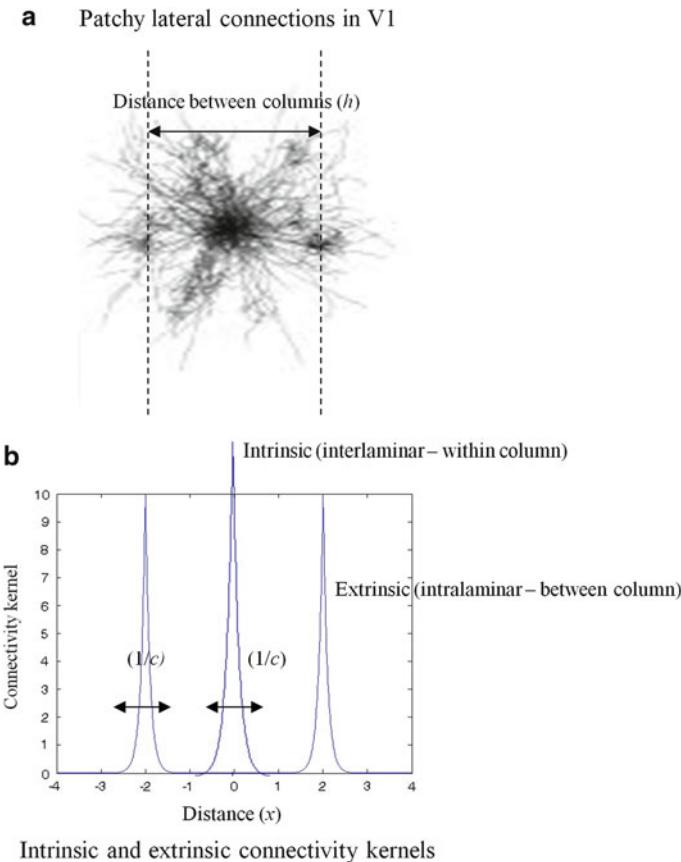


Fig. 17.3 Connectivity kernel describing a combination of patchy but isotropic distributions by using connectivity kernels with non-central peaks. This kernel models sparse intrinsic connections in cortical circuits that mediate both local and non-local interactions. In other words, neurons talk both to their immediate neighbours and receive input from remote populations who share the same functional selectivity; see Eq. (17.9) (The *insert* is modified from www.ini.uzh.ch/node/23776)

17.2.4 Power Spectra of Neural Fields

In what follows, we describe the generative or forward mapping from external inputs (exogenous neuronal fluctuations) to observed spectral responses for a single cortical source. This allows one to compare the predictions of our model with real data and requires a mapping of neuronal states (the depolarisation fields above) to sensors—called the lead field. The lead field allows one to infer hidden parameters characterizing the deployment of sources on the cortical surface, even when there is no explicit spatial information in the data; see also [50]. The lead field samples particular spatiotemporal frequencies, depending on the sensitivity profile of the

sensors used. For example, if the lead field has a narrow spatial support (e.g., when using LFP electrodes), its spatial Fourier transform will be broad and it will be sensitive to a wide range of spatial frequencies. Conversely, when the lead field sees a large part of the cortical surface (e.g., non-invasive EEG sensors), the spatial Fourier transform will be narrow and only fluctuations in low spatial frequencies will contribute to the observed cross-spectra.

The lead field is parameterised by φ as a continuous gain function $L(x, \varphi)$ over the cortical patch that is applied to a mixture of neuronal (depolarisation) states at each point on the patch. In the case of the CMC field model, this mixture is determined by four coefficients $Q = [q_1 q_2 q_3 q_4]$, while the gain function is parameterised in terms of the coefficients $L(k, \varphi)$ of a spatial Fourier basis set:

$$L(x, \varphi) = \sum_k L(k, \varphi) e^{ikx}. \quad (17.11)$$

The predicted response at an LFP or virtual electrode—for a given set of neural and lead field parameters θ is obtained by integrating over the cortical patch

$$y(t, \theta) = \int L(x, \varphi) (Q \cdot V(x, t)) dx, \quad (17.12)$$

which leads to a spectral response of the form

$$Y(\omega, \theta) = \sum_k L(k, \varphi) (Q \cdot T(k, \omega) U(k, \omega)). \quad (17.13)$$

The predicted spectral response measured by the sensor is therefore

$$\begin{aligned} g(\omega, \theta) &= Y(\omega, \theta) Y^*(\omega, \theta) \\ &= \sum_k |L(k, \varphi)|^2 Q T(k, \omega) g_u(k, \omega) T(k, \omega)^* Q^T, \end{aligned} \quad (17.14)$$

where $g_u(k, \omega) = |U(k, \omega)|^2$ is the auto-power spectrum of external input. The above expression depends upon the transfer function $T(k, \omega)$ associated with the evolution Equations of the model, which we will derive below.

For the case of MEG virtual electrodes and LFP data we consider here, the gain function has a simple Gaussian form, which we parameterize in terms of its dispersion φ such that $L(x, \varphi) = e^{-x^2/2\varphi^2}$ —noting that the amplitude is fixed to avoid redundancy with the parameters Q . This leads to Fourier coefficients of the form $L(k, \varphi) = e^{-2\pi^2\varphi^2k^2}$ and Eq. (17.14) becomes

$$g(\omega, \theta) = \sum_k \left| e^{-2\pi^2\varphi^2k^2} \right|^2 Q T(k, \omega) g_u(k, \omega) T(k, \omega)^* Q^T. \quad (17.15)$$

The function $T(k, \omega)$ can be derived by assuming that the neural field defined by Eq. (17.2) is perturbed around a spatially homogeneous steady-state attained in the absence of external or exogenous perturbations, (see also [45, 46]):

$$V_0 = B^{-1} A \cdot F(V_0) \int K(x) dx. \quad (17.16)$$

Using linear systems analysis, we define the transfer function of a field model with the following relation

$$T(k, \omega) = \frac{P(k, \omega)}{U(k, \omega)}, \quad (17.17)$$

where $U(k, \omega)$ is the two dimensional Fourier transform of external input:

$$\begin{aligned} U(k, \omega) &= FT(U(x, t)) \\ &= \iint U(x, t) e^{-ikx+i\omega t} dt dx \end{aligned} \quad (17.18)$$

and $P(k, \omega)$ is the Fourier transform of the perturbations around the steady-state solution. Given the transfer function, we can characterise the spectral response of the system to any external input, in terms of the underlying connectivity kernel, propagation velocities and post-synaptic response function: Substituting $V(x, t) = V_0 + P(x, t)$ into Eq. (17.2) and expanding $F \circ V$ around V_0 , we obtain a second-order expression for the perturbations $P(x, t)$

$$\begin{aligned} \ddot{P} + 2B\dot{P} &= -B^2 P + AB\gamma D \otimes P + GU \\ \gamma_{ab} &= \frac{\partial \sigma(v_a=0)}{\partial v_b} = \begin{cases} \frac{re^{r\eta}}{(1+e^{r\eta})^2}, & a = b \\ 0, & a \neq b, \end{cases} \end{aligned} \quad (17.19)$$

where γ is the gain of the nonlinear mapping between depolarisation and firing rate. Equations (17.17) and (17.19) provide the transfer function of our canonical microcircuit neural field model. Taking the Fourier transform of Eq. (17.19) and substituting into Eq. (17.17) gives:

$$\begin{aligned} T(k, \omega) &= (-\omega^2 I_4 - 2i\omega B + B^2 - J(k, \omega))^{-1} G \\ J(k, \omega) &= ABD(k, \omega)\gamma, \end{aligned} \quad (17.20)$$

where $J(k, \omega)$ is a 4×4 matrix incorporating the synaptic parameters, connectivity parameters and gain matrix and $D(k, \omega)$ is the Fourier transform of the spatiotemporal connectivity: see Eq. (17.22) below.

In summary, Eq. (17.20) provides a transfer function mapping from exogenous inputs or fluctuations acting upon each neuronal layer and the resulting

spatiotemporal response in source space. This transfer function is specified completely by synaptic and connectivity parameters implicit in the neural field model.

Substituting Eq. (17.20) into Eq. (17.15), we obtain an expression for the predicted spectra as a mixture of contributions from each population weighted by $q_a (a = 1, \dots, 4)$:

$$\begin{aligned} g(\omega, \theta) &= \sum_{a,k} q_a W_a(k, \omega) \\ W_a(k, \omega) &= \left| e^{-2\pi^2\varphi^2 k^2} \kappa_1 m_e S_a(k, \omega) R^{-1}(k, \omega) \right|^2 g_u(k, \omega). \end{aligned} \quad (17.21)$$

The term $S_a(k, \omega) R^{-1}(k, \omega)$ in (17.21) expresses the relative contribution of each population to the predictions at the source level and depends upon the particular form of the connections among these populations. It can be seen from Eq. (17.9), that this ratio depends upon the (Fourier transforms of) intrinsic and extrinsic connectivity;

$$\begin{aligned} D_{ab}^{(i)}(k, \omega) &= \frac{a_{ab}(c_{ab} - i\nu\omega)}{c_{ab}^2 - \nu_{ab}^2\omega^2 - 2i\nu c_{ab}\omega + k^2} \\ D_{aa}^{(e)}(k, \omega) &= \frac{c_{aa}}{2} \left[\frac{e^{-h_a c_{aa} c^-} - e^{i\nu h_a \omega - 2h_a c_{aa}} (c^- \alpha + \beta)}{4k^2\pi^2 + c^- c^-} \right. \\ &\quad \left. + \frac{e^{i\nu h_a \omega} c^+ \alpha - e^{-h_a c_{aa} c^+} + e^{i\nu h_a \omega - h_a c_{aa}} \beta}{4k^2\pi^2 + c^+ c^+} \right] \\ \alpha &= \cos(2h_a k \pi), \quad \beta = 2k\pi \sin(2h_a k \pi) \\ c^+ &= c_{aa} + i\nu\omega, \quad c^- = c_{aa} - i\nu\omega. \end{aligned} \quad (17.22)$$

In particular, $R(k, \omega)$ and $S_a(k, \omega)$ are given by

$$\begin{aligned} R(k, \omega) &= -V_{14}(k, \omega)(-V_{23}(k, \omega) + Q_2(k, \omega)Q_3(k, \omega)) \\ &\quad + Q_4(k, \omega)[-V_{23}(k, \omega)Q_1(k, \omega) \\ &\quad + Q_3(k, \omega)(-V_{12}(k, \omega) + Q_1(k, \omega)Q_2(k, \omega))] \\ S_1(k, \omega) &= -Q_4(k, \omega)(-V_{23}(k, \omega) + Q_2(k, \omega)Q_3(k, \omega)) \\ S_2(k, \omega) &= D_{21}^{(i)}(k, \omega)\gamma\kappa_2 m_i Q_3(k, \omega)Q_4(k, \omega) \\ S_3(k, \omega) &= -D_{21}^{(i)}(k, \omega)D_{32}^i(k, \omega)\gamma^2\kappa_2\kappa_3 m_e m_i Q_4(k, \omega) \\ S_4(k, \omega) &= D_{41}^{(i)}(k, \omega)\gamma\kappa_4 m_e (-V_{23}(k, \omega) + Q_2(k, \omega)Q_3(k, \omega)) \end{aligned} \quad (17.23)$$

where the functions $Q_a(k, \omega)$ and $V_{ab}(k, \omega)$ depend on the Fourier transforms $D_{ab}^{(i)}(k, \omega)$ and $D_{aa}^{(e)}(k, \omega)$ as follows:

$$\begin{aligned} Q_a(k, \omega) &= -\kappa_a^2 + \gamma(D_{aa}^{(i)}(k, \omega) + D_{aa}^{(e)}(k, \omega))\kappa_a m_a + 2i\kappa_a \omega + \omega^2 \\ V_{ab}(k, \omega) &= D_{ab}^{(i)}(k, \omega)D_{ba}^{(i)}(k, \omega)\gamma^2\kappa_a\kappa_b m_a m_b. \end{aligned} \quad (17.24)$$

In summary, the predicted spectral response at the sensor for the CMC field model is given by:

$$g(\omega, \theta) = \sum_k \left| e^{-2\pi^2 \varphi^2 k^2} \kappa_1 m_e \sum_a q_a S_a(k, \omega) \right|^2 g_u(k, \omega), \quad (17.25)$$

where the functions $S_a(k, \omega)$ and $R(k, \omega)$ are defined by (17.23). Equation (17.25) reflects the fact that the predicted spectral responses of the system are coupled to its spatial as well as its temporal properties; these properties are encoded in the transfer functions $S_a(k, \omega)$ and $R(k, \omega)$ through the underlying connectivity functions $D_{ab}(k, \omega)$. In turn, these are specified by the synaptic parameters associated with the canonical microcircuit $\theta \in \{m_i, m_e, \kappa_i, \kappa_e, r, \eta\}$ and the spatial parameters $\theta \in \{a_{ab}, c_{ab}, h_a, v_{ab}\}$ that encode intrinsic and extrinsic connections among different layers and neighbouring columns or points on the cortical circuits.

The predicted spectral responses for the JR field model obeying Eq. (17.8) with connectivity determined by Eq. (17.10) are also given by an equation of the form of Eq. (17.25) above (with $a = 1, 2, 3$) where the functions $S_a(k, \omega)$ and $R(k, \omega)$ are now given by:

$$\begin{aligned} S_1(k, \omega) &= -\frac{D_{32}^{(i)}(k, \omega)}{D_{31}^{(i)}(k, \omega)} S_2(k, \omega) + (\kappa_e + i\omega)^2 (\kappa_i + i\omega)^2 \\ S_2(k, \omega) &= D_{23}^{(i)}(k, \omega) D_{31}^{(i)}(k, \omega) \gamma^2 \kappa_e \kappa_i m_e m_i \\ S_3(k, \omega) &= D_{31}^{(i)}(k, \omega) \gamma \kappa_e m_e (\kappa_i + i\omega)^2 \\ R(k, \omega) &= -D_{23}^{(i)}(k, \omega) D_{32}^{(i)}(k, \omega) \gamma^2 \kappa_e \kappa_i m_e m_i (\kappa_e + i\omega)^2 \\ &\quad + (\kappa_i + i\omega)^2 \left(\kappa_e^4 + 4i\kappa_e^3 \omega - 4i\kappa_e \omega^3 + \omega^4 \right. \\ &\quad \left. - \kappa_e^2 (D_{13}^{(i)}(k, \omega) D_{31}^{(i)}(k, \omega) \gamma^2 m_e^2 + 6\omega^2) \right). \end{aligned} \quad (17.26)$$

17.3 Neural Fields as Dynamic Causal Models

17.3.1 Probabilistic Models of Empirical Data and Their Inversion

To complete our specification of a generative model, we assume that the observed cross-spectra g_y are a mixture of predicted spectra, channel and Gaussian observation noise

$$\begin{aligned} g_y(\omega) &= g(\omega, \theta) + g_n(\omega, \theta) + \epsilon_y \\ g_u(\omega, \theta) &= \alpha_u + \frac{\beta_u}{\omega}, \quad g_n(\omega, \theta) = \alpha_n + \frac{\beta_n}{\omega}, \\ \text{Re}(\epsilon) &\sim \mathcal{N}(0, \Sigma(\omega, \lambda)), \quad \text{Im}(\epsilon) \sim \mathcal{N}(0, \Sigma(\omega, \lambda)). \end{aligned} \quad (17.27)$$

Here, $g(\omega, \theta) + g_n(\omega, \theta)$ are the predictions of the data features $g_y(\omega)$ and ϵ_y are the corresponding prediction errors with covariance $\Sigma(\omega, \lambda)$. The spectra of the neuronal fluctuations or input $g_u(\omega, \theta)$, are assumed to be spatially white; namely,

they are independent of spatial frequency. However both input and noise spectra are modelled as a mixture of white and coloured fluctuations over time. In particular, we have introduced extra free parameters $\theta \subset \{\alpha_n, \alpha_u, \beta_n, \beta_u\}$ controlling the spectra of the inputs and channel noise.

Equation (17.27) provides the basis for our generative model and entails free parameters controlling the spectra of the inputs and channel noise as well as the amplitude of observation error. Gaussian assumptions about the observation error mean that we have a probabilistic mapping from all of the unknown parameters to observed (spectral) data features. Inversion of this model means estimating, probabilistically, the free parameters from the data.

Having prescribed the generative model of our DCM, we can now turn to its inversion via Bayesian techniques. Almost universally, the fitting or inversion of Dynamic Causal Models optimizes variational free energy. Variational free energy serves as a bound approximation to the log-evidence $\ln p(g_y|M)$ for a model M . This optimization is carried out with respect to a variational density $q(\theta)$ on the unknown model parameters. By construction, the free energy bound ensures that when the variational density maximizes free energy, it approximates the true posterior density over parameters, $q(\theta) \approx p(\theta|g_y, M)$. At the same time, the free energy itself approximates the log-evidence (log-marginal likelihood of the data). The (approximate) conditional density and (approximate) log-evidence are used for inference on parameters and models respectively. In other words, one first compares different models (e.g. neural fields and masses) using their log-evidence and then turns to inferences on parameters, under the model selected. One usually assumes the conditional density has a Gaussian form $q(\theta) = \mathcal{N}(\mu, C)$. This is known as the Laplace assumption. The conditional density is quantified by the most likely value of the parameters, μ and their conditional covariance C that encodes uncertainty about the estimates and their conditional dependencies. Under this assumption about the variational density and Gaussian observation noise, the free energy has a very simple form:

$$\begin{aligned} F &= G(\mu) + \frac{1}{2} \ln |\partial_{\mu\mu} G| \\ G &= H - \frac{1}{2} \text{Re}(\epsilon(\mu))^T \sigma^{-1} \text{Re}(\epsilon(\mu)) - \frac{1}{2} \text{Im}(\epsilon(\mu))^T \sigma^{-1} \text{Im}(\epsilon(\mu)) \\ H &= -\frac{1}{2} \rho(\mu)^T \Omega^{-1} \rho(\mu) - \frac{1}{2} \ln |\sigma| - \frac{1}{2} \ln |\Omega| \\ \rho &= \mu - \phi. \end{aligned} \quad (17.28)$$

Here, $\rho(\mu) \in \mathbb{R}$ are prediction errors on the parameters, in relation to their prior density $p(\theta|m) = \mathcal{N}(\phi, \Omega)$. Model complexity in Eq.(17.28) corresponds to the $-\frac{1}{2} \rho^T \Omega^{-1} \rho$ term: This reports the deviation of the estimated parameters from their prior expectations and effectively penalizes the free energy objective function in proportion to the degrees of freedom used to explain the data. A full description of the resulting Variational Laplace scheme can be found in [19].

The underlying generative model generally admits a unique solution during model inversion; this follows from the use of biophysically plausible priors over the biophysical parameters. Tables 17.1 and 17.2 describe the priors over synaptic

Table 17.1 Prior expectations of parameters of the CMC field model

Parameter	Physiological interpretation	Prior mean
m_e, m_i	Maximum postsynaptic depolarisation	8, 32 (mV)
$\kappa_1, \kappa_2, \kappa_3, \kappa_4$	Postsynaptic rate constants	1/2, 1/2, 1/16, 1/28 (ms^{-1})
a_{22}, a_{33}, a_{41}	Amplitude of intrinsic connectivity kernels	3,200
$a_{12}, a_{44}, a_{23}, a_{32}$		800,800,1,600,1,600
a_{11}, a_{14}, a_{21}		9,600,4,000,4,800
c_{ab}	Spatial decay of connectivity kernels	$\begin{cases} 0.6 & a \neq b \\ 2 & a = b \end{cases} (\text{mm}^{-1})$
h_a	Separation between columns	4.5 (mm)
r, η	Parameters of the postsynaptic firing rate function	0.54, 0
v	Inverse conduction speed	1.5 s/m
ϕ	Dispersion of the lead field	$\sqrt{2}/20$

Table 17.2 Priors of parameters of the JR field model

Parameter	Physiological interpretation	Prior mean
$\alpha_{13}, \alpha_{23}, \alpha_{31}, \alpha_{32}$	Amplitude of intrinsic connectivity kernels	2,000, 8,000, 2,000, 1,000

parameters for the CMC and JR field models; as well as parameters pertaining to the spatial structure of cortical sources. These priors are based on the modelling literature, while others come from the experimental literature [27]. In general, priors are chosen to restrict parameter estimates to a physiologically meaningful range. However, it should be noted that the precise values of the priors are not important: the inversion scheme has the latitude to accommodate deviations from these values to optimise model evidence.

In what follows, we apply the theory of preceding sections to real electrophysiological data to illustrate the sorts of questions and quantitative characterisations that are enabled by combining neural field models with dynamic causal modelling.

17.3.2 LFP Auditory Cortex Data

We will first formulate neural mass models as a special case of neural field models by simply setting the conduction times to zero. This provides a useful perspective on the relationship between these two models, in terms of the implicit assumptions we make when modelling observed data. A pragmatic advantage of emulating neural mass models, with a transit time of zero, is that we can apply precise shrinkage priors to conduction times to facilitate model comparison. In other words, it provides a simple means of comparing models with and without spatial dynamics (with and without prior constraints on conduction or transit times). In particular, we first consider the mass and field variants of the JR model and assume that output comes primarily from pyramidal cells. The corresponding predictions for neural fields and masses are shown in Fig. 17.4. These predictions are based on model

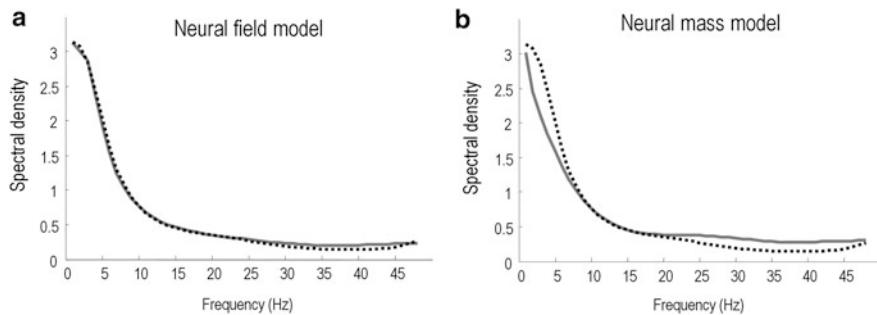


Fig. 17.4 Real data (dashed line) and model predictions (full line) showing a spectral profile that is typically seen under anaesthesia. We observe a better fit of the field, relative to the mass model, in particular for low frequencies

fits or inversions using local field potentials recorded from primary (A1) auditory cortex in the Lister hooded rat, following the application of the anaesthetic agent Isoflurane (see [39] for details) under acoustic white noise stimuli at a level of 83 dB. In brief, 10 min of recordings were extracted from the continuous time domain data and down-sampled to a sampling rate of 125 Hz. Frequency domain data-features were obtained from this epoch using a vector autoregression model of order eight. The model predictions of Fig. 17.4 illustrate nicely the difference between the field and mass models: one can see that the neural field model has approximated the preponderance of low frequencies more accurately than the neural mass model. This is because it has extra degrees of freedom; namely conduction velocity and the extent of lateral connections. These extend the repertoire of predictions to include those afforded by spatial dynamics. Crucially, the log-evidence for the neural field model was 1271 above the log-evidence for the neural mass model. This suggests that there is a very strong evidence for spatial dynamics over the cortical manifold in these auditory cortex data [43]. Recall that the model fit is based on optimising the free energy bound to model-evidence. The free energy is just the difference between a term quantifying accuracy (goodness of fit) and a term quantifying complexity. This means the inversion provides explanations for empirical data that are both accurate and parsimonious.

Clearly, the choice of an appropriate model depends upon the question of interest; in particular, neural fields are appropriate for addressing questions about the deployment of sources on the cortical surface and induced spatial dynamics. However, the above example highlights that neural field models can be more appropriate than mass models, from a Bayesian perspective, even if the spatial parameters of a neuronal model are not the focus of study: In the context of our Bayesian scheme, each model is scored using a free energy bound on model-evidence, where better models have a higher free energy. This provides a principled way to compare (score) different modes or hypotheses about how neuronal time-series are generated.

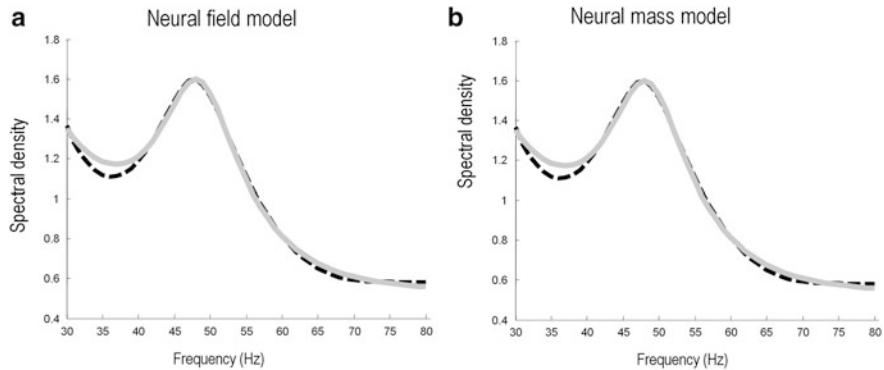


Fig. 17.5 Real data (dashed line) and model predictions (full line) for spectra in the gamma band obtained from the human visual cortex during an attention task [56]. We observe that the fits of both the field and mass models are equally good with no manifest differences

17.3.3 MEG Data from the Visual Cortex

We now turn to MEG data that summarise the spectral expression of endogenous activity in the visual cortex of (twelve) human subjects described in detail in [56]. We used an adaptive spatial filter or beamformer [60] to obtain estimates of ongoing neuronal activity in primary visual cortex. This provides estimates of electrical cortical activity based on a weighted combination of sensors—sometimes referred to as a *virtual electrode*. We then used the CMC model describing inter-laminar and lateral intra-laminar interactions and inverted its mass and field variants. In these illustrative inversions, the synaptic and spatial parameters were optimized and the intermodal distance h_a was fixed to a physiologically plausible value. We computed the log evidence ratio (using the free energy approximation) comparing field and mass variants at the group level. As above, the neural mass model was formulated as a special case of the neural field model by shrinking the conduction times to zero. Contrary to our earlier result, we found no evidence in favour of the neural field model (the relative log evidence between mass and field models was on average 2.62): see Fig. 17.5 for spectral fits of an exemplar subject.

This highlights the fact that the best model depends upon the data modelled. It also underlines the importance of combining a neuronal model with a spatial forward model: although both auditory and visual cortices are thought to conform to the local homogeneity constraints implicit in neural field models, the loss of spatial frequency resolution—with non-invasive data—might render neural field models unnecessary, in relation to neural mass models. Our failure to establish a significantly greater evidence for neural field models, in the present model comparison, is intuitively sensible because non-invasive MEG data have much lower spatial resolution than the LFP data we used in the previous model comparison. This observation speaks to the potential importance of using spatially resolved data

to take full advantage of neural field models. Data with high signal to noise ratio and wide brain coverage—such as those afforded by ECoG sensors or multi-array grids—can, in principle, disclose a full spectrum of spatiotemporal dynamics at different scales, which may be important for an informed (efficient) estimate of spatial parameters in neural field models.

17.4 Conclusions

By exploiting a combination of neural field modelling and Bayesian inference, we have shown that Dynamic Causal Modelling can help appraise biophysical models for explaining electrophysiological data. We have focused on two main classes of biophysical models of brain activity, the so-called neural field and mass models and have considered their application to modelling empirical LFP and MEG data. Bayesian model comparison—using a variational free energy approximation to log model evidence—suggests neural field models provide a better explanation of empirical data if, and only if, there is sufficient spatial frequency information in the data. In other words, we found greater evidence for neural field models in analyses of LFP data but failed to find more evidence for neural field models, relative to neural mass models, in MEG (virtual electrode) data. The key distinction between these different modalities is that the LFP data is sensitive to a wide range of spatial frequencies and the temporal fluctuations that these frequencies contain. In contrast, the lead fields inherent in non-invasive electromagnetic recordings are necessarily broader and suppress temporal dynamics that are expressed in high spatial frequencies. This is a nice example of modelling with neural mass and field models that highlights the key role of both data and biophysically informed models in hypothesis testing and model comparison.

References

1. Amari, S.: Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.* **27**, 77–87 (1977)
2. Atay, F.M., Hutt, A.: Neural fields with distributed transmission speeds and long-range feedback delays. *SIAM J. Appl. Dyn. Syst.* **5**, 670–698 (2006)
3. Bastos, A.M., Usrey, W.M., Adams, R.A., Mangun, G.R., Fries, P., Friston, K.J.: Canonical microcircuits for predictive coding. *Neuron.* **76**(4), 695–711 (2012). doi:10.1016/j.neuron.2012.10.038
4. Breakspear, M., Roberts, J.A., Terry, J.R., Rodrigues, S., Mahant, N., Robinson, P.A.: A unifying explanation of primary generalized seizures through nonlinear brain modeling and bifurcation analysis. *Cereb. Cortex* **16**, 1296–1313 (2006)
5. Bressloff, P.C.: New mechanism for neural pattern formation. *Phys. Rev. Lett.* **76**, 4644–4647 (1996)
6. Bressloff, P.C.: Traveling fronts and wave propagation failure in an inhomogeneous neural network. *Phys. D Nonlinear Phenom.* **155**, 83–100 (2001)

7. Buffalo, E.A., Fries, P., Landman, R., Buschman, T.J., Desimone, R.: Laminar differences in gamma and alpha coherence in the ventral stream. *Proc. Natl. Acad. Sci.* **108**, 11262 (2011)
8. Coombes, S.: Waves, bumps, and patterns in neural field theories. *Biol. Cybern.* **93**, 91–108 (2005)
9. Coombes, S., Lord, G.J., Owen, M.R.: Waves and bumps in neuronal networks with axo-dendritic synaptic interactions. *Phys. D Nonlinear Phenom.* **178**, 219–241 (2003)
10. Coombes, S., Venkov, N.A., Shiu, L., Bojak, I., Liley, D.T.J., Laing, C.R.: Modeling electrocortical activity through improved local approximations of integral neural field equations. *Phys. Rev. E* **76**, 051901 (2007)
11. Daunizeau, J., Kiebel, S.J., Friston, K.J.: Dynamic causal modelling of distributed electromagnetic responses. *Neuroimage* **47**, 590–601 (2009)
12. David, O., Kiebel, S.J., Harrison, L.M., Mattout, J., Kilner, J.M., Friston, K.J.: Dynamic causal modeling of evoked responses in EEG and MEG. *Neuroimage* **30**, 1255–1272 (2006)
13. Deco, G., Jirsa, V.K., Robinson, P.A., Breakspear, M., Friston, K.: The dynamic brain: from spiking neurons to neural masses and cortical fields. *PLoS Comput. Biol.* **4**, e1000092 (2008)
14. Douglas, R.J., Martin, K.: A functional microcircuit for cat visual cortex. *J. Physiol.* **440**, 735 (1991)
15. Freeman, W.J.: Linear analysis of dynamics of neural masses. *Ann. Rev. Biophys. Bioeng.* **1**, 225–256 (1972)
16. Freeman, W.J.: A neurobiological theory of meaning in perception. In: Proceedings of the international joint conference on neural networks, Portland, vols. 1–4, pp. 1373–1378 (2003)
17. Freeman, W.J.: A field-theoretic approach to understanding scale-free neocortical dynamics. *Biol. Cybern.* **92**, 350–359 (2005)
18. Friston, K.J., Harrison, L., Penny, W.: Dynamic causal modelling. *Neuroimage* **19**, 1273–1302 (2003)
19. Friston, K., Mattout, J., Trujillo-Barreto, N., Ashburner, J., Penny, W.: Variational free energy and the Laplace approximation. *Neuroimage* **34**, 220–234 (2007)
20. Galka, A., Ozaki, T., Muhle, H., Stephani, U., Siniatchkin, M.: A data-driven model of the generation of human EEG based on a spatially distributed stochastic wave equation. *Cognit. Neurodyn.* **2**, 101–113 (2008)
21. Ghosh, A., Rho, Y., McIntosh, A.R., Kotter, R., Jirsa, V.K.: Cortical network dynamics with time delays reveals functional connectivity in the resting brain. *Cognit. Neurodyn.* **2**, 115–120 (2008)
22. Grindrod, P., Pinotsis, D.A.: On the spectra of certain integro-differential-delay problems with applications in neurodynamics. *Phys. D Nonlinear Phenom.* **240**, 13–20 (2011)
23. Haeusler, S., Maass, W.: A statistical analysis of information-processing properties of lamina-specific cortical microcircuit models. *Cereb. Cortex* **17**, 149 (2007)
24. Jansen, B.H., Rit, V.G.: Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biol. Cybern.* **73**, 357–366 (1995)
25. Jirsa, V.K.: Neural field dynamics with local and global connectivity and time delay. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **367**, 1131 (2009)
26. Jirsa, V.K., Haken, H.: Field theory of electromagnetic brain activity. *Phys. Rev. Lett.* **77**, 960–963 (1996)
27. Kandel, E.R., Schwartz, J.H., Jessell, T.M.: *Principles of Neural Science*, 4th edn. McGraw-Hill, New York (2000)
28. Kiebel, S.J., Garrido, M.I., Moran, R., Chen, C.C., Friston, K.J.: Dynamic causal modeling for EEG and MEG. *Hum. Brain Mapp.* **30**, 1866–1876 (2009)
29. Lefort, S., Tomm, C., Floyd Sarria, J.-C., Petersen, C.C.H.: The excitatory neuronal network of the C2 barrel column in mouse primary somatosensory cortex. *Neuron* **61**, 301–316 (2009)
30. Liley, D.T.J.: Neural field modelling of the electroencephalogram: physiological insights and practical applications. In: Coombes, S., beim Graben, P., Potthast, R., Wright, J. (eds.) *Neural Fields: Theory and Applications*. Springer, Berlin/Heidelberg (2014)

31. Liley, D.T.J., Alexander, D.M., Wright, J.J., Aldous, M.D.: Alpha rhythm emerges from large-scale networks of realistically coupled multicompartmental model cortical neurons. *Netw. Comput. Neural Syst.* **10**, 79–92 (1999)
32. Liley, D.T.J., Cadusch, P.J., Dafilis, M.P.: A spatially continuous mean field theory of electrocortical activity. *Netw. Comput. Neural Syst.* **13**, 67–113 (2002)
33. Lopes da Silva, F.H., Storm van Leeuwen, W.: The cortical alpha rhythm in dog: the depth and surface profile of phase. Raven Press, New York (1978)
34. Lopes da Silva, F.H., Hoeks, A., Smits, H., Zetterberg, L.H.: Model of brain rhythmic activity. *Biol. Cybern.* **15**, 27–37 (1974)
35. Lumer, E.D., Edelman, G.M., Tononi, G.: Neural dynamics in a model of the thalamocortical system. I. Layers, loops and the emergence of fast synchronous rhythms. *Cereb. Cortex* **7**, 207 (1997)
36. Markounikau, V., Igel, C., Grinvald, A., Jancke, D.: A dynamic neural field model of mesoscopic cortical activity captured with voltage-sensitive dye imaging. *PLoS Comput. Biol.* **6**(9), e1000919 (2010). doi:10.1371/journal.pcbi.1000919
37. Moran, R.J., Kiebel, S.J., Stephan, K.E., Reilly, R.B., Daunizeau, J., Friston, K.J.: A neural mass model of spectral responses in electrophysiology. *Neuroimage* **37**, 706–720 (2007)
38. Moran, R.J., Stephan, K.E., Seidenbecher, T., Pape, H.C., Dolan, R.J., Friston, K.J.: Dynamic causal models of steady-state responses. *Neuroimage* **44**, 796–811 (2009)
39. Moran, R.J., Jung, F., Kumagai, T., Endepols, H., Graf, R., Dolan, R.J., Friston, K.J., Stephan, K.E., Tittgemeyer, M.: Dynamic causal models and physiological inference: a validation study using isoflurane anaesthesia in rodents. *PLoS One* **6**(8), e22790 (2011). doi:10.1371/journal.pone.0022790
40. Nunez, P.L.: *Neocortical Dynamics and Human EEG Rhythms*. Oxford University Press, New York (1995)
41. Nunez, P.L.: Multiscale neocortical dynamics, experimental EEG measures, and global facilitation of local cell assemblies. *Behav. Brain Sci.* **19**, 305 (1996)
42. Nunez, P.L., Srinivasan, R.: *Electric Fields of the Brain*, vol. 1, Oxford University Press, p. i–612 (2006)
43. Penny, W.D., Stephan, K.E., Daunizeau, J., Rosa, M.J., Friston, K.J., Schofield, T.M., Leff, A.P.: Comparing families of dynamic causal models. *PLoS Comput. Biol.* **6**, e1000709 (2010)
44. Potthast, R., beim Graben, P.: Inverse problems in neural field theory. *SIAM J. Appl. Dyn. Syst.* **8**(4), 1405–1433 (2009)
45. Pinotsis, D.A., Friston, K.J.: Neural fields, spectral responses and lateral connections. *Neuroimage* **55**, 39–48 (2011)
46. Pinotsis, D.A., Moran, R.J., Friston, K.J.: Dynamic causal modeling with neural fields. *Neuroimage* **59**, 1261–1274 (2012)
47. Qubbaj, M.R., Jirsa, V.K.: Neural field dynamics under variation of local and global connectivity and finite transmission speed. *Phys. D Nonlinear Phenom.* **238**, 2331–2346 (2009)
48. Raizada, R.D.S., Grossberg, S.: Towards a theory of the laminar architecture of cerebral cortex: computational clues from the visual system. *Cereb. Cortex* **13**, 100–113 (2003)
49. Riera, J.J., Jimenez, J.C., Wan, X., Kawashima, R., Ozaki, T.: Nonlinear local electrovascular coupling. II: from data to neuronal masses. *Hum. Brain Mapp.* **28**, 335–354 (2007)
50. Robinson, P.A.: Patchy propagators, brain dynamics, and the generation of spatially structured gamma oscillations. *Phys. Rev. E* **73**, 041904 (2006)
51. Robinson, P.A., Loxley, P.N., O'Connor, S.C., Rennie, C.J.: Modal analysis of corticothalamic dynamics, electroencephalographic spectra, and evoked potentials. *Phys. Rev. E* **63**(4), 041909 (2001)
52. Robinson, P.A., Rennie, C.J., Rowe, D.L.: Dynamics of large-scale brain activity in normal arousal states and epileptic seizures. *Phys. Rev. E* **65**, 041924 (2002)
53. Robinson, P.A., Rennie, C.J., Rowe, D., O'Connor, S.C., Wright, J.J., Gordon, E., Whitehouse, R.W.: Neurophysical modeling of brain dynamics. *Neuropsychopharmacology* **28**, S74–S79 (2003)

54. Roopun, A.K., Kramer, M.A., Carracedo, L.M., Kaiser, M., Davies, C.H., Traub, R.D., Kopell, N.J., Whittington, M.A.: Period concatenation underlies interactions between gamma and beta rhythms in neocortex. *Front Cell Neurosci.* **2**, 1 (2008)
55. Schiff, S., Sauer, T.: Kalman filter control of a model of spatiotemporal cortical dynamics. *BMC Neurosci.* **9**, O1 (2008)
56. Schwarzkopf, D.S., Robertson, D.J., Song, C., Barnes, G.R., Rees, G.: The frequency of visually induced gamma-band oscillations depends on the size of early human visual cortex. *J. Neurosci.* **32**, 1507–1512 (2012)
57. Steriade, M., Deschenes, M.: The thalamus as a neuronal oscillator. *Brain Res. Rev.* **8**, 1–63 (1984)
58. Valdes, P.A., Jimenez, J.C., Riera, J., Biscay, R., Ozaki, T.: Nonlinear EEG analysis based on a neural mass model. *Biolog. Cybern.* **81**, 415–424 (1999)
59. Van Rotterdam, A., Lopes da Silva, F.H., Van den Ende, J., Viergever, M.A., Hermans, A.J.: A model of the spatial-temporal characteristics of the alpha rhythm. *Bull. Math. Biol.* **44**, 283–305 (1982)
60. Van Veen, B.D., Van Drongelen, W., Yuchtman, M., Suzuki, A.: Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Trans. Biomed. Eng.* **44**, 867–880 (1997)
61. Weiler, N., Wood, L., Yu, J., Solla, S.A., Shepherd, G.M.G.: Top-down laminar organization of the excitatory network in motor cortex. *Nat. Neurosci.* **11**, 360–366 (2008)
62. Wilson, H.R., Cowan, J.D.: Mathematical theory of functional dynamics of cortical and thalamic nervous-tissue. *Kybernetik* **13**, 55–80 (1973)

Chapter 18

Neural Field Dynamics and the Evolution of the Cerebral Cortex

James J. Wright and Paul D. Bourke

Abstract We describe principles for cortical development which may apply both to the evolution of species, and to the antenatal development of the cortex of individuals. Our account depends upon the occurrence of synchronous oscillation in the neural field during embryonic development, and the assumption that synchrony is linked to cell survival during apoptosis. This leads to selection of arrays of neurons with ultra-small-world characteristics. The “degree of separation” power law is supplied by the combination of neuron sub-populations with differing exponential axonal tree distributions, and consequently, in the visual cortex, connections emerge in anatomically realistic patterns, with an ante-natal arrangement which projects signals from the surrounding cortex onto each macrocolumn, in a form analogous to the projection of a Euclidean plane onto a Möbius strip. Simulations of signal flow explain cortical responses to moving lines as functions of stimulus velocity, length and orientation. With the introduction of direct visual inputs, under the operation of Hebbian learning, development of mature selective response “tuning” to stimuli “features” then takes place, overwriting the earlier ante-natal configuration. Further assuming similar development principles apply to inter-areal interactions in the developing cortex, a general principle for the evolution of increasingly complicated sensory-motor sequences, at both species-evolution and individual time-scales, is implicit.

J.J. Wright (✉)

Faculty of Medicine, Department of Psychological Medicine, University of Auckland, Auckland,
New Zealand

e-mail: jadawright@gmail.com

P.D. Bourke

iVEC@UWA, University of Western Australia, Perth, WA, Australia

18.1 Introduction

This chapter outlines the wider biological motivation of recent work from our group, in which we have applied neural field theory to the embryological development of the primary visual cortex. The embryogenesis of brains appears to mirror the phylogenetic history of the brains of antecedent species, and neurodevelopment and later learning must, throughout life, take place hand-in-hand. So, perhaps it will be of value to consider neuron dynamics within this evolutionary and developmental context? This idea is hardly new—in relation to neural networks it can be traced back through Hebb [42] to William James [50], and beyond. In its anatomical aspects, it is given its strongest evolutionary context in the works of Papez [76], Yakovlev [110], Sperry [88], and MacLean [57, 69]. In these latter works, the process of encephalization was explained in terms of the drive toward ever more neurons, and of the advantages of envelopment of the “older” (species’) brain within the “newer” brain, thus providing centripetal/centrifugal control and supervisory functions, so that the function of hard-wiring circuits was not lost as progressively flexible “new” circuits were added, at paleo-cortical, and then neo-cortical level. Two corollary aspects of this evolutionary sequence have been less emphasised, but seem also to be important. The first aspect seems almost too obvious to require stating—the developing neural organization must retain, as cortical size increases, a primary capacity to convert information delivered to the sensory cortices into motor outputs, beginning from simple sensory-motor systems exemplified by the tadpole tectum [45]. Perhaps less obviously, it seems that there must be a modular principle for sensory-motor signal conversions, such that new pieces of cortex can be “inserted”, without disruption of antecedent functions. The latter aspect has gained in importance since the classic works of MacLean and his precursors. As encephalization increases, there are corollary demands to minimize information transfer times and physical size, while maximizing total synaptic connectivity and total information storage capacity, all the while minimizing metabolic demand as much as possible. In approaching an optimum neuronal assembly, there is a synergy between the need to maximize connectivity, minimize connection distances, and maximize information storage capacity, for the following reason: as encephalization increases, the small neurons of small, primitive creatures give way to long, attenuated neurons of large, advanced creatures. This increases the connectivity of each neuron, and is compatible with an efficient connection system among the neurons, for which some “ultra-small-world” arrangement [21] would be optimal. The tendency toward attenuation of neurons has a limit at which the neuron, described as a fractal object, approaches a dimension of three—i.e., as large a surface area of synaptic contacts as possible, for as small a cell volume as possible. Assuming the supply of metabolites is subject to some upper practical limit, there must also result an increasing competition for crucial metabolites among synapses. There is good evidence that competition between synapses for resource takes place at a number of anatomical sites (e.g., [7, 40, 51, 61, 74]). If there was only enough critical metabolite for half the synapses to operate at maximum capacity, then, as well as generating a maximum

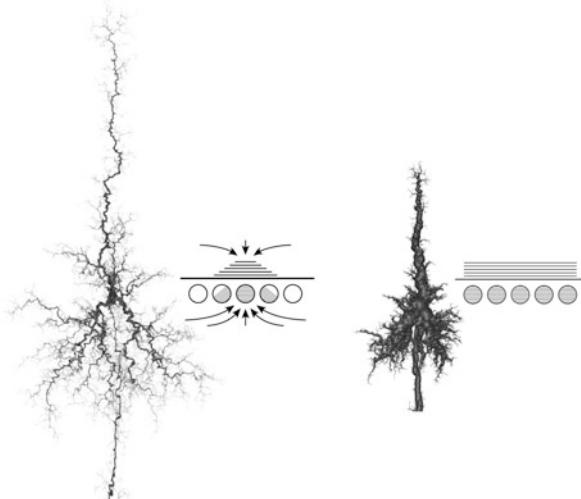


Fig. 18.1 Surface-to-volume ratios of dendrites, and competition for critical metabolites. Attenuated dendrites of a neuron with fractal dimension approaching three (*left*), are contrasted with less attenuated dendrites of a neuron with fractal dimension markedly less than three (*right*). To the right of each neuron, synapses and post-synaptic membranes are represented schematically. For the more “developed” neuron, where cell surface to volume is high, *arrows* indicate the flow of a critical metabolite away from inactive synapses to active synapses, induced by demand—a competition that need not take place when dendritic surface area is relatively reduced. Model neurons constructed by diffusion-aggregation [103]

of synapses, attenuation plus competition could maximize the possible Shannon entropy of the synaptic states, by maximizing the complexity of possible neural signal pathways among the neurons (See Fig. 18.1). If variation of the supply of metabolites fluctuates with firing states of the network, and there are a multiplicity of critical factors, and consequently of time-scales of their supply, conditional Markov processes of great complexity are possible, from which ensemble those beneficial to survival must be selected.

Since “ontogeny recapitulates phylogeny”, and the considerations above have determined the pathway followed by species-level natural selection, then what is their analogue during individual development? During embryogenesis there is a further happy convergence of effects, which at first seem antagonistic. The active firing of the neurons, which seems to add a burden of metabolic demand to the developing cells may assist the avoidance of cell-death by apoptosis, and may do so in a way which leads to an efficient primary organizational underpinning for the learning of ever-more complicated sensory-motor sequences in post-natal life. We next sketch relevant background findings, before presenting application of these principles to problems of development of the primary visual cortex (V1). We have concentrated on the primary visual cortex because of the wealth of experimental data that has been gathered in that cortical area, but we intend our treatment to be more general, and applicable to the cortex more widely, as is later described.

18.1.1 Genetic Expression, Cell Firing and Apoptosis in Cortical Development

The emergence of functional neuronal organization and connectivity in the developing cerebral cortex depends on differentiation, proliferation and migration of neurons [44]. Early, thalamus-independent (i.e., sensory-pathway-independent) steps in the process of cortical arealization take place on the basis of information intrinsic to the cells, as proposed by Rakic in his protomap hypothesis [78]. It is these genetic programmes that lead to the characteristic cellular shapes of different populations of neurons during the stages of cell differentiation. However, factors not simply explained by direct gene expression in the cells seem to be important. Action potential generation is present from early embryonic development (e.g., [5]) and plays a part in the development of cortical microcircuitry [111]. As cell differentiation proceeds, programmed cell death plays a major role. Fragmented nuclear DNA markers suggest that the bulk of differentiating neurons die soon after they are generated, and the majority of the cells that die are in the fastest proliferating regions [12]. Cell firing itself is not essential to synaptic development, since cultured neurons blocked from generating action potentials by xylocaine continue to develop synapses [65]. Yet, although the generation of action potentials must greatly increase metabolic demands, synchronous action potential generation appears to protect against apoptosis, since neurons in neonatal cerebral cortical slices show increased apoptosis when their capacity to enter into synchronous firing is disrupted by pharmacological means [43]. Embryonic neurons developing in vitro develop synchronous firing, and as their growth proceeds, also show self-organization into “small world” networks [22].

We propose that synchronous firing and protection from apoptosis are related because competition among developing neurons and synapses, although mediated by trophic factors [39, 94, 97, 98], is ultimately a competition for available metabolic energy, and that pulse synchrony increases uptake of critical metabolic resources, perhaps by some collective pumping action. Consequently, cell groups interlinked in such a way as to fire in maximum synchrony can supply themselves with sufficient resource to survive, while others cannot.

18.1.2 Cell Firing and Synchronous Oscillation

Synchronous oscillation of pulses and local field potentials is a ubiquitous aspect of cortical activity [16, 26, 27, 37, 87] and has been proposed as a solution to the “binding problem” of perceptual grouping and cognitive processing [27, 87]. Synchrony is not absolute, but refers to occurrence of maximum cross-correlation at zero lag, and is a broadband phenomenon in the temporal frequency domain [16]. Detailed models of synchronous firing in specific cell assemblies [26, 83, 87, 89, 95, 100] do not explain the synchrony seen in neuron cultures, brain slices, or the

early foetal brain. A more fundamental mechanism that is a universal property of networks with summing junctions including dendrites [20, 80, 107] is applicable, however, and appears in simulations that also accurately reproduce spectra, cross-correlations and excitatory/inhibitory timings characteristic of activated cortex [104, 105]. In these simulations synchrony results from interaction of waves travelling in opposite directions, and increases in amplitude toward an ideal steady-state in which there is sustained symmetrical exchange of signals between all excitatory neurons [105], associated with concurrent local excitatory/inhibitory oscillation. That is, synchrony reflects an oscillatory steady-state with bidirectional equality of signal exchange. Unidirectional traveling waves are transient deviations from that equilibrium of exchange.

18.1.3 Unresolved Issues in the Development of V1

18.1.3.1 The Geometry of Response Organization

Since the discovery that individual cells in V1 respond with an orientation preference (OP) to visual lines of differing orientation [48], attempts to analyze the response organization and explain its relationship to cortical function [92, 99, 102] have played a pivotal role in neuroscience. The surface organization of OP in V1 has recently been compared with appropriate random surrogates, and shown, in some species at least, to approximate an hexagonal rotational periodicity in which each roughly delineated macrocolumnar unit exhibits all values of OP arrayed around a pinwheel [68, 75]. Varying chirality and orientation of the pinwheels achieves continuity of OP at the columnar margins, thus producing linear zones and saddles. In any individual, irregular variation from the average periodicity occurs, and some species—particularly those with smaller brains and hence visual cortices—exhibit little or no sign of this ordering. Because of this marked interspecies variation, serious doubt has been expressed that the pattern is of functional significance at all, since response maps are absent in some species without those species having any apparent deficit in vision [47].

18.1.3.2 The Superficial Patch System

A further puzzle of intracortical V1 organization is posed by the superficial patch system. This system, composed of relatively long-range, largely excitatory [46, 59] patchy connections [35, 81] is ubiquitous in cortex [67] and has a functional relationship to OP. Patchy connections develop before sensory afferents reach the cortex [18, 23, 77, 82] and do not arise or terminate in the vicinity of OP singularities. They link areas of common OP (“like-to-like”) over distances several times the diameter of a macrocolumn [17, 36, 63, 68], are periodic on roughly the same interval as OP, and are largely patch-reciprocal [4, 81]. Just as for maps of response properties, there is variation of patchy connection orderliness between species [68].

18.1.3.3 Model Characterization of Primary Feature Responses

Explanation of organization of OP has been attempted in a group of now-classical theories, which we will refer to as “standard models”, following the comparative description of Swindale [92]. Dimension reduction methods [24, 25, 52] show that the response maps of OP, eye preference (OC), direction preference (DP) and spatial frequency preference (SF) are consequences of requiring continuity and completeness of representation of each response property, in a two-dimensional representation in which every type of response property occurs within any small area on the surface of V1 [19, 92]. The same ordering can also be explained as a consequence of competitive Hebbian learning among small neighborhood assemblies of excitatory neurons [38]. All standard models depend on seeding with oriented lines, in one way or another [24, 38, 64, 70, 71, 90, 91, 93, 99] and otherwise similar models avoiding this limitation do not accurately reproduce response maps [54–56, 60]. Initial belief that response to simple oriented lines in the visual field formed the basis of OP maps has been undermined in two ways. Firstly, maps of OP appear in the cortex prior to visual experience [11, 85, 101], and although it is argued that structured stimuli may arise from retinal inputs in the absence of visual experience [1, 75, 79], the absence of particular visual stimuli in the post-natal environment eliminates subsequent neural response to those stimuli [10] indicating that direct visual experience is essential at some stage. Secondly, and more recently, Basole and colleagues, who tested OP using stimulus lines moving at different speeds, and oriented at differing angles to the line of movement of the stimulus, found OP to be a function of these variables to such a degree that for lines oriented non-orthogonally to the direction of movement, OP could vary progressively with increments of speed to an asymptotic limit of 90° [8, 9]. This effect was attenuated for lines of progressively greater length. Standard models could not account for these effects, and to salvage the standard models in essence, if not specifics, subsequent workers explained these results by considering the temporal and spatial frequencies associated with the moving stimuli. Issa and colleagues [6, 49] showed responses to specific features could be explained by fitting six parameters—OP, SF preference, and temporal frequency preference, and the tuning bandwidths of all three. This description is referred to as the spatio-temporal filter model.

18.2 Developmental Synergy of Apoptosis and Synchrony, Applied to V1

Wright and Bourke [106, 109] used a generic form of neural field equations for an idealised, isotropic, neural field, within which individual neurons are embedded. This represents the developing cortex’s potential isotropic connections, from which actual connections are selected during development, by the combined unfolding of genetic cascades, and of apoptosis. The scale of the field is that of a cortical

area such as V1, representing intracortical connections rather than cortico-cortical. Thus, the density of connection between neurons declines with increasing separation of their cell bodies [15]. The high non-linearity of synapto-dendritic summations are linearized at the field level, and axonal conduction speed is considered single-valued. Subject to these strictures, these general equations include the minimum relevant features:

$$\varphi_p^{\mathbf{qr}'}(t) = f_p^{\mathbf{qr}'} \times Q_p \left(\mathbf{r}', t - \frac{|\mathbf{q} - \mathbf{r}'|}{v} \right) \quad (18.1)$$

$$\psi_p^{\mathbf{qr}'}(t) = M_p^{\mathbf{qr}'}(t) * \varphi_p^{\mathbf{qr}'}(t) \quad (18.2)$$

$$\Psi_p(\mathbf{q}, t) = \int_D \psi_p^{\mathbf{qr}'}(t) d\mathbf{r}' \quad (18.3)$$

$$V_p(\mathbf{q}, t) = G_e(t) * \Psi_e(\mathbf{q}, t) + G_i(t) * \Psi_i(\mathbf{q}, t) \quad (18.4)$$

$$Q_p(\mathbf{q}, t) = f_\Sigma(V_p(\mathbf{q}, t)) + E_p(\mathbf{q}, t). \quad (18.5)$$

Subscript $p \in \{e, i\}$ refers to excitatory or inhibitory neurons; superscript \mathbf{qr}' refers to synaptic connection from \mathbf{r}' to \mathbf{q} where \mathbf{q}, \mathbf{r}' are cortical positions in domain D , occupied by single neurons. $\varphi_p^{\mathbf{qr}'}(t)$ is the flux of pulses reaching presynapses at the neuron at \mathbf{q} , from the neuron at \mathbf{r}' . $\psi_p^{\mathbf{qr}'}(t)$ is the synaptic current generated by $\varphi_p^{\mathbf{qr}'}(t)$. $\Psi_p(\mathbf{q}, t)$ is the aggregate synaptic current of type p generated at \mathbf{q} . $V_p(\mathbf{q}, t)$ is the soma membrane potential (relative to the resting potential) generated at \mathbf{q} . $Q_p(\mathbf{q}, t)$ is the pulse emission rate at \mathbf{q} . $f_p^{\mathbf{qr}'}$ is the probability density of occurrence of presynapses generated by axons of the neuron at \mathbf{r}' terminating at \mathbf{q} . v is axonal conduction speed. $M_p^{\mathbf{qr}'}(t)$ is the impulse response function transforming presynaptic flux to synaptic current. $G_p(t)$ is the impulse response function transforming synaptic current into dendritic potentials. $f_\Sigma(V_p(\mathbf{q}, t))$ is a sigmoid function describing the local conversion of dendritic potentials into the rate of generation of action potentials. $E_p(\mathbf{q}, t)$ is a driving noise, arising from intrinsic random cell action potentials.

Restriction of the field to the scale of a cortical area carries several implications, all because the probability of connections between any two neurons declines with distance of separation. Firstly, descriptively we can consider “reciprocal couplings” as an idealization/representation of field coupling symmetry, and in some instances reciprocal couplings will in fact exist. Secondly because of more generally dense connections among near neighbours, smoothing at dendritic summation requires that $Q_p(\mathbf{q}, t)$ is spatially and temporally “brown”—i.e., has high correlation at short distances and times of separation. Thirdly, in the sparsely connected network, the average “degree” of separation—i.e., the average number of neighboring cells traversed by synaptic connections linking one cell to another—will also increase in proportion to physical distance of separation.

A further crucial property upon which our results depend is the occurrence of gamma oscillation in the cortical field, when the cortex is sufficiently excited, as occurs in the developing mammalian cortex in later foetal development [58, 62]. Experimental observations [32, 33, 41] show intrinsic cortical oscillation arises from alternating excitatory cell and inhibitory cell firing at lags 1/4 of the period of oscillation. Simulations of the oscillations [104, 105] show that travelling waves are thus generated, the intersection of which produces broadband synchrony. In conditions of uniform cortical excitation without strong perturbation from external inputs the exchange of pulses between all cells reaches an equilibrium—that is, a steady-state of symmetrical exchange of signals between excitatory cells at any two positions on the cortex, so that in the oscillating field over sufficient intervals, T ,

$$\frac{1}{T} \int_0^T \varphi_e(\mathbf{q}, t) - \bar{\varphi}_e \, dt = \frac{1}{T} \int_0^T \varphi_e(\mathbf{r}', t) - \bar{\varphi}_e \, dt \quad (18.6)$$

where $\bar{\varphi}_p$ is the time-average presynaptic flux, uniform throughout the cortical field. Since conduction delays are short compared to the period of oscillation, the equality of Eq.(18.6) is generally approached even when T is smaller than the period of oscillation [20], and because there are equal time-lags in both directions of conduction excitatory pulse trains throughout the cortex have maximum correlation at zero lag.

Zero-lag synchronous oscillation thus entails presynaptic pulse synchrony, with a magnitude of presynaptic flux variation that can be defined respectively for individual synapses, and in aggregate, as

$$J^{qr'} = \left[\frac{1}{T} \int_0^T (\varphi_e^{qr'}(t) - \bar{\varphi}_e)^2 \, dt \right]^{1/2} \quad (18.7)$$

$$J = \left[\frac{1}{T} \int_0^T \int_D \int_D (\varphi_e^{qr'}(t) - \bar{\varphi}_e)^2 \, d\mathbf{q} \, d\mathbf{r}' \, dt \right]^{1/2} \quad (18.8)$$

$J^{qr'}$ is RMS presynaptic flux variation between \mathbf{q} and \mathbf{r}' , and J is the aggregate of $J^{qr'}$ over the cortex. The assumption that selection of neurons that survive apoptosis depends on maximization of J has a series of important consequences.

18.2.1 Selection of Scale-Free Small-World Configurations of Neurons

For any given level of cortical excitation, J is greatest for that ensemble of C connected neurons, in which excitatory pulses arrive at dendrites, from all sources at differing distances of separation, as closely in-phase as possible, so as to maximize their summation. Axonal delays, small compared to the period of gamma oscillation,

contribute a phase difference between cell firing at \mathbf{r}' and the arrival of presynaptic pulses at \mathbf{q} , of

$$\Delta\phi^{\mathbf{qr}'} = 2\pi \frac{|\mathbf{q} - \mathbf{r}'|}{Pv} \quad (18.9)$$

where P is the period of oscillation. Therefore that ensemble selected by its capacity to maximize presynaptic synchrony must approach minimal total axonal length, $L = \int_D \int_D |\mathbf{q} - \mathbf{r}'| d\mathbf{q} d\mathbf{r}'$, and minimization of this length also minimizes the metabolic requirements of the axons.

It has been shown generally [21] for all systems of connected elements, the path length in a topological sense is at a minimum where degree distribution follows a power law. As was pointed out in conjunction with Eqs. (18.1)–(18.5), in our idealised neural field, average degree of separation, in the topological sense, increases linearly as metric distance of separation of the cell bodies, so that if L , their total length of axonal connections, is minimal, then the path length in the topological sense is also minimal, and the degree distribution is that of a scale-free, or ultra-small world. Therefore, the connection density between cells versus their metric distance of separation should also be approximated by a power-law distribution. Further, according to Cohen and Havlin [21]

$$L \sim \log \log C \quad (18.10)$$

so the metabolic efficiency of the connection system is further enhanced if the surviving cells are linked into a continuum, as opposed to separate pools of neurons. The number of neighbouring excitatory cells connected to a given excitatory neuron, as a function of distance of separation, is proportional to $2\pi \times f_e^{\mathbf{qr}'}(|\mathbf{q} - \mathbf{r}'|)$ and intracortical axonal trees have approximately exponential density/range relations [15, 84], therefore, because a power function can be fitted exactly by a sum of exponential functions, an ultra-small-world connectivity can be achieved by sets of populations of cells with differing axonal characteristic lengths. During embryogenesis primal cells divide sequentially by layer [78, 86] with differences in growth pattern and characteristic axonal length programmed in sequential cell divisions. For simplicity, we consider only two populations of excitatory cells, with cell bodies partially separated by layer, but with intermingled axonal and dendritic trees, and axonal tree connection probabilities described by

$$f_\alpha^{\mathbf{qR}} = \frac{N_\alpha}{N} 2\pi \lambda_\alpha \exp[-\lambda_\alpha 2\pi |\mathbf{q} - \mathbf{R}|] \quad (18.11)$$

$$f_\beta^{\mathbf{qr}} = \frac{N_\beta}{N} 2\pi \lambda_\beta \exp[-\lambda_\beta 2\pi |\mathbf{q} - \mathbf{r}|] \quad (18.12)$$

$$f_e^{\mathbf{qr}'} = f_\alpha^{\mathbf{qR}} + f_\beta^{\mathbf{qr}}$$

$f_\alpha^{\mathbf{qR}}$ refers to the axonal trees with longest axonal extensions, and $f_\beta^{\mathbf{qR}}$ refers to the axonal trees with short axonal extension, thus $\lambda_\alpha < \lambda_\beta$. $N = N_\alpha + N_\beta$ is the number of synapses received/generated by each cell. Distances from \mathbf{r}' to \mathbf{q} are substituted as $\mathbf{r} - \mathbf{R}$ to indicate equal distances, $\mathbf{q} - \mathbf{r}$ and $\mathbf{q} - \mathbf{R}$, measured along the axonal trees of the respective populations.

The further defining characteristic of small-world connectivity—the occurrence of connection nodes—emerges as a consequence of the formation of the superficial patch system, as follows.

18.2.2 The Origin of the Superficial Patch System

The two populations of cells and the synapses they give rise to can be referred to as α -cells and synapses, and β -cells and synapses. We first make a provisional assumption (later justified on a species-specific basis) that $N_\beta \gg N_\alpha$, so that α -cells with long-range axons are embedded among much more numerous β -cells. Applying Eqs. (18.11) and (18.12) via Eq. (18.1) to find values of $J^{\mathbf{q}\mathbf{r}'}$ in Eq. (18.7) as functions of $|\mathbf{q} - \mathbf{r}, \mathbf{R}|$, shows that

$$\begin{aligned} J^{\mathbf{q}\mathbf{r}} &= J^{\mathbf{qR}} && \text{if } |\mathbf{q} - \mathbf{r}, \mathbf{R}| = x \\ J^{\mathbf{q}\mathbf{r}} &> J^{\mathbf{qR}} && \text{if } |\mathbf{q} - \mathbf{r}, \mathbf{R}| < x \\ J^{\mathbf{q}\mathbf{r}} &< J^{\mathbf{qR}} && \text{if } |\mathbf{q} - \mathbf{r}, \mathbf{R}| > x \end{aligned} \quad (18.13)$$

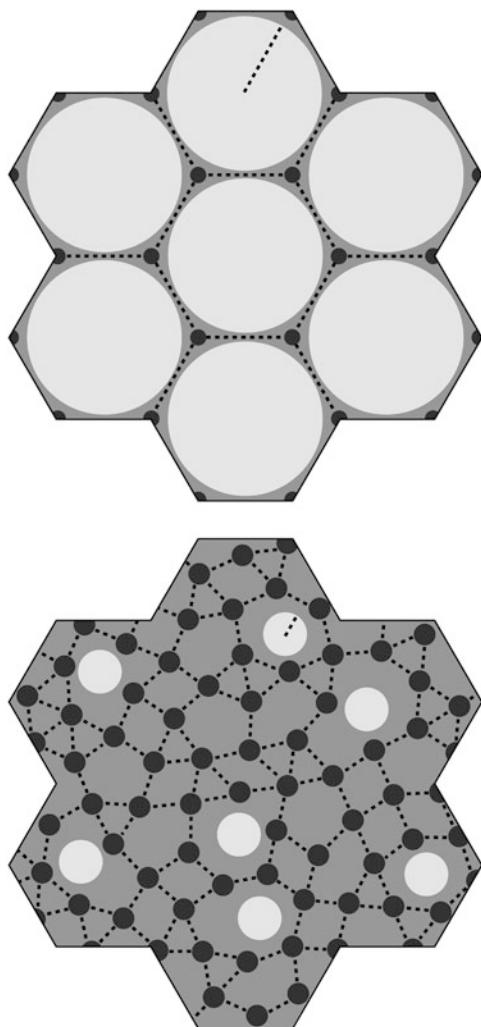
$$\ln \frac{N_\alpha \lambda_\alpha}{N_\beta \lambda_\beta}$$

where $x = -\frac{2\pi(\lambda_\beta - \lambda_\alpha)}{N_\alpha \lambda_\alpha}$.

Consequently [106, 109] it can be shown that J (Eq. (18.8)) is at a maximum if β -cells are clustered so they make reciprocal connections at minimum distance and maximum density (β -clusters), and α -cells also form clusters (α -clusters) making reciprocal synaptic connections at distances greater than x , so that they may form multiple patches of synaptic connections, skipping from α -cluster to α -cluster. Also, α -clusters are necessarily placed at the vertices of hexagons tiling the cortical surface, with each hexagon embracing a β -cluster, while reciprocal connections between α - and β -cells occur at cluster margins, over distances approximate to x . Analogy to the superficial patch system in larger-brained species is apparent. See Fig. 18.2.

As noted earlier, hexagonal symmetry of OP and the superficial patch system is an idealization that is roughly approached in some species, while in others it is effectively absent [47]. Since approximation of a power law distribution by two populations of neurons requires $N_\alpha \ll N_\beta$ if $\lambda_\alpha \ll \lambda_\beta$, this case is more closely approached for larger cortical sizes, and the patchy connection system will have higher orderliness and hexagonal rotational symmetry. If $\lambda_\alpha < \lambda_\beta$ by

Fig. 18.2 *Pale circles:* neurons coupled by short-range connections. *Black circles:* neurons coupled by long-range patchy connections. *Grey background:* neurons receiving connections of both types. *Dashed lines* are of length $x = -\frac{\ln \frac{N_\alpha \lambda_\alpha}{N_\beta \lambda_\beta}}{2\pi(\lambda_\beta - \lambda_\alpha)}$. *Top:* $N_\beta \gg N_\alpha$. *Bottom:* $N_\alpha > N_\beta$



a small amount, as in animals with small cortical size, then N_β is not necessarily greater than N_α , and an ordered hexagonal structure need not be apparent. Such reduction of the apparent orderliness does not imply the absence of “small world” connectivity, nor imply impairment of function. As a corollary, the same principle of development may apply widely throughout the cortex, as the emergence of clearly defined macrocolumns is determined by the availability of cell types with marked differences in axonal length. This appears to be the case for V1 in particular, whereas elsewhere, resolution into clear macrocolumns is not so apparent [47].

18.2.3 Self-Organization of Pre-vision Response Properties

Turning from optimization of energy demand of axons, to that of dendrites, we can modify Eq. (18.2) [106, 109] to

$$\psi_e^{qr'}(t) = \Gamma^{qr'} M_e^{qr'}(t) * \varphi_e^{qr'}(t) \quad (18.14)$$

where $\Gamma^{qr'}$ is the available fraction of the metabolic supply rate needed to attain maximum current flow, and $M_e^{qr'}(t)$ includes terms for synaptic adaptation and impulse decay, and, most importantly, for presynaptic synergy [96].

Since we have assumed increasing synaptic current in synchronously activated synapses increases the available metabolic supply, the value of $\Gamma^{qr'}$ must follow that of $\psi_e^{qr'}(t)$, and as well as inter-cellular competition between assemblies of neurons, we assume competition takes place between adjacent individual synapses arising from the same neuron. Therefore those neurons that survive apoptosis must have found an efficient deployment of resource to the synapses best positioned to maximize the magnitude of synchrony. Since any two adjacent synapses arising from the same pre-synaptic neuron may terminate on the same, or different, post-synaptic neurons, then if they terminate on the same neuron their conditions are essentially identical. If they terminate on different neurons, then the relevant values of J^q —their respective synaptic cooperativity with other synapses terminating on the same cell—need not identical—and their competition for resources would lead, via the feedback between $\psi_e^{qr'}(t)$ and $\Gamma^{qr'}$, to low synaptic current at one synapse, and high current at the other. Just what the physiological corollary of these opposite high and low-activity states is, and the critical metabolic component for which the synapses compete, we do not specify. A likely, but by no means unique contributing factor is the supply of extracellular calcium [66]. Whatever the critical component(s), the important consequence is that, at synchronous equilibrium, closely situated neurons each receiving synapses from the same cell, must have either high, or low, pulse correlations with each other.

We can term those synapses that are transmitting impulses more strongly near equilibrium “saturated” synapses, and those which are more quiescent, but potentially able to be activated, “sensitive” synapses, and can consider what spatial patterns of saturated connections would best meet the requirement to maximize synchrony. Here a further property of the neural field commented on in relation to Eqs. (18.1)–(18.5)—higher spatial cross-correlation of pulses and field potentials at shorter range—has a decisive impact on the equilibrium pattern of synaptic saturations, in concert with the need for saturated and sensitive synapses to be generated on adjacent post-synaptic neurons. Then, for reasons further argued in [106, 109], the emergent patterns, diagrammed in Fig. 18.3, have the following properties:

- (a) Saturated connections within each β -cluster form a re-entrant network analogous to a Möbius strip.
- (b) Saturated connections between the α -cluster system and each of the β -clusters form a projection between scales which is homeomorphic, preserving topological identity between scales, and thus mapping a disk to a Möbius strip, and imposing an orientation and chirality on each β -cluster.
- (c) Cells in the α -system are linked by saturated synapses.
- (d) Saturated connections between β -clusters must project to each of their six neighbors as closely as possible to mirror symmetry, with both saturated and sensitive synapses linking points homologous with respect to position in the α -system—that is to say, points with similar OP as classically measured with low object speeds. The necessarily broken symmetry permits the particular pattern generated to be one of a large set of possible combinations.

Further analogy between the hypothetical α - and β -systems and real anatomical structures can now be drawn. As well as the α -system's congruence with the superficial patch system, the β -systems, each with a dense system of local connections that are centrally spared from patchy connections, are analogous to macrocolumns each centred about an OP singularity. The distribution of OP for lines of orientation $0 - \pi$ to angles $0 - 2\pi$ in pinwheels about a singularity finds analogy in the wrapping of a Euclidean plane onto a Möbius strip. It has also been earlier shown that arrangements of adjacent pinwheels in broken mirror symmetry match classical OP maps [108]. These relations are shown in Figs. 18.3 and 18.4.

Just as OP organization in some species is apparent before eye opening, so too is the organization into OD columns [11, 31]. Explanation of this can be included in the present model by an argument similar to that of Erwin and Miller, who suppose the correlation of cell firing at short distances of separation of V1 cells to be greater than the correlation of visual inputs over a similar distance. This forces a columnar OD organization because of instability—in the present model's terms, the resulting disruption of the synchronous field at equilibrium produced by binocular inputs to the same cells—resolved by formation of columns in Turing patterns.

18.2.4 *Consequently, Following Eye-Opening...*

After eye opening, visual inputs will provoke ordered departures from the average equilibrium condition.

The emergent map at equilibrium, by which the patchy connections over a part of V1 link to positions within each macrocolumn, can be expressed as 1:1 projection from a disk on a Euclidean plane (the global map), P, to a Möbius strip (the local map), p^[2]—the square brackets [2] indicating the map's resemblance, if viewed from

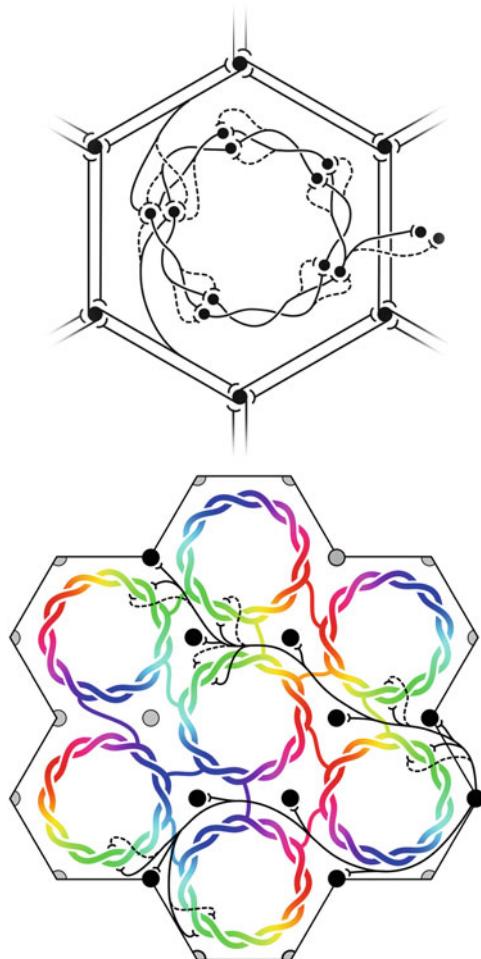
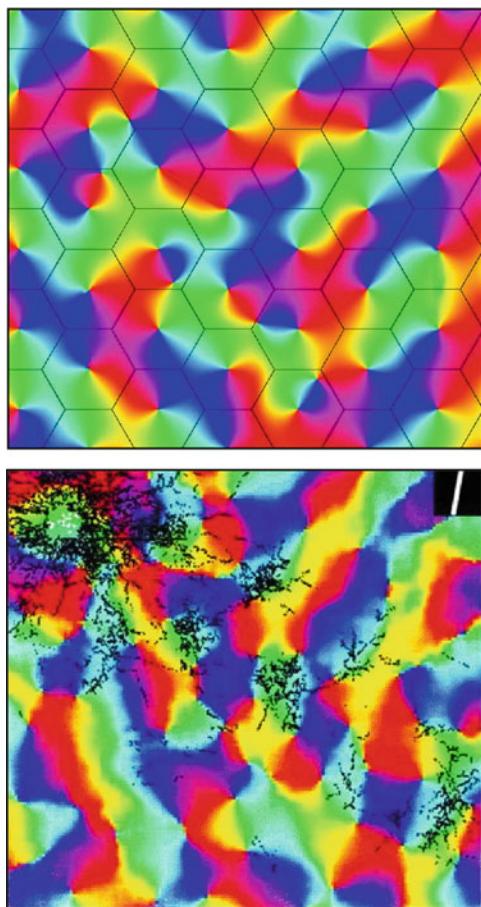


Fig. 18.3 *Top* Equilibrium disposition of saturated and sensitive synapses. Black circles represent cell bodies and dendrites. Synapses are indicated as saturated (solid) or sensitive (dashed) terminations of axons. Reciprocal connections between α -patches (patchy connections) form an hexagonal array. (Other connections, although shown as unidirectional, are also reciprocal.) A representative pair of connections from α -cells to the β -patch is displayed in the upper-and lower aspects of the figure. At the centre of the figure, saturated and sensitive synapses show the network's analogy to a Möbius-strip within a β -patch (macrocolumn). To the right, representative links from the central macrocolumn to cells at homologous positions in neighboring macrocolumns are indicated. *Bottom* “Like to like” saturated patchy connections map the same part of the surrounding cortical field onto homologous cell positions on the Möbius configuration within each macrocolumn, while at short range “like to like” saturated synaptic connections also form between homologous positions between local maps

Fig. 18.4 Simulated and real maps of orientation preference in V1, from [108]. *Top:* Simulation. Colours of the spectrum, from red to violet, represent average OP of V1 neurons for slow-moving visual lines of orientation $0 - \pi$. Adjacent macrocolumns, of diameter approx $300 \mu\text{m}$ are set within an hexagonal frame (the patch system) with OP forming colour wheels about OP singularities. Orientations and chiralities of the colour wheels are arranged to approach a minimum total of angular disparity from mirror reflection of OP between each macrocolumn and its neighbours. *Bottom:* Real OP. Visualized in the tree shrew by [13]. Superficial patchy connections are demarcated in black by a selective stain. Scale of macrocolumns is approximate to that of the simulation



a third dimension, to a 2:1 map formed by squaring a complex vector. In polar coordinates,

$$P(|\mathbf{R} - \mathbf{C}_j|, \vartheta) \mapsto p^{[2]}(|\mathbf{r} - \mathbf{C}_j|, \pm\vartheta + \varphi) \quad (18.15)$$

where \mathbf{C}_j is the origin of both P and $p^{[2]}$ for the j -th local map, and corresponds to the position of the OP singularity in that macrocolumn. ϑ is the polar angle of \mathbf{R} , chirality of the local map is indicated by $\pm\vartheta$, and φ is the orientation of the local map relative to the global map. $\vartheta + \varphi$ can be defined on the range $0 - 2\pi$ in both local and global maps, but is represented with apparent angle doubling in the local map, producing an apparent superposition of angles ϑ and $\vartheta + \varphi$. This describes the form of “contextual” connections [3, 53].

With eye-opening, let $O(P, t)$ be a visual image projected to V1 by the direct visual pathway. Laterally travelling waves of pulses and local field potentials

transmit that image to each local map with a point to point delay, $\frac{|\mathbf{R}-\mathbf{r}|}{v}$, where v now represents wave speed, so that

$$O(\mathbf{P}, t) \mapsto O \left(\mathbf{p}^{[2]}, t + \frac{|\mathbf{R} - \mathbf{r}|}{v} \right) \quad (18.16)$$

Suppose $O(\mathbf{P}, t)$ is a segment of the image of a visual line, travelling with uniform velocity, V_x , on the cortical surface, along an x -axis directed toward a macrocolumn with its singularity at \mathbf{C}_j , O has a component of its extension on the x -axis, O_x , and an orthogonal component of extension, on the y -axis, O_y . K_x is the dominant spatial frequency of O_x , and K_y is the dominant spatial frequency of O_y . Then the local map projection of O has a transformed spatial frequency in the x -axis but not in the y -axis—i.e.:

$$k_x \propto \frac{v}{v \pm V_x} K_x \quad (18.17)$$

$$k_y \propto K_y, \quad (18.18)$$

where k_x, k_y are the spatial frequencies in the local map projection of O , and the sign \pm in Eq. (18.17) depends on whether O is approaching or departing from \mathbf{C}_j . That is, O 's orientation in the global map is projected to the local map, with Doppler shift, producing an apparent difference in orientation, $\delta\vartheta$;

$$\delta\vartheta = \left| \tan^{-1} \frac{K_y}{K_x} - \tan^{-1} \frac{k_y}{k_x} \right| \quad (18.19)$$

Laterally transmitted contextual signals generally do not trigger cell firing, until the classic receptive field (cRF) is directly stimulated [3, 53] via the visual pathway. The cells that fire are those that reflect the supra-threshold summations of sub-threshold signals conveyed over the contextual, patchy, connections, and the direct pathway. The summation of contextual and direct cRF inputs will act as an impulse causing a transient breakdown of equilibrium, during which synapses that were in both saturated and sensitive state in equilibrium briefly generate substantial synaptic currents (see Fig. 18.5). Action potentials are triggered transiently in surrounding cells. Subsequently there is a restoration toward the equilibrium state on withdrawal of the stimulus. During the breakdown the mapping of activity from the global to the local map becomes

$$O(\mathbf{P}, t) \mapsto O \left(\mathbf{p}^2, t + \frac{|\mathbf{R} - \mathbf{r}|}{v} \right) \quad (18.20)$$

The change from Eq. (18.16) made by removal of the square brackets from $\mathbf{p}^{[2]}$ represents the breakdown's form, as itself a map from global to local scale, resembling a 2:1 complex-multiplication map, as initially described by Alexander et al. [2].

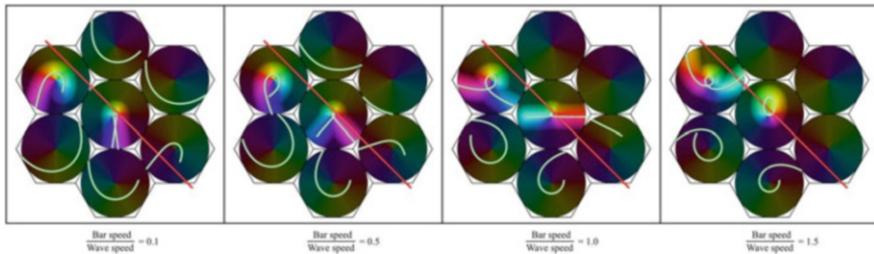


Fig. 18.5 Red line is the projection, via the direct visual pathway, of a line in the (monocular) visual field, oriented at 45° to the line of passage, and moving from left to right across a system of 7 macrocolumns in V1. Green lines represent the image of the red line, transmitted laterally with delay, by patchy (contextual) connections. Bright illumination against the cortical background represents the field of supra-threshold summation of direct pathway and contextual inputs, firing cells with a preferred orientation/velocity/length relation to the red line

18.2.5 Post-natal Effects of Learning, the Spatio-Temporal Filter Model, Dimension Reduction, and “Like To Like” Connections

Following eye opening, stimuli with regularly repeated spatial and temporal structure reach V1. Exposure to a repeated stimulus will lead to permanent synaptic consolidation of connections, in accordance with physiological versions of the Hebb rule, and the spatio-temporal learning rule [28–30, 72, 73, 96], overlaying any consolidated connections formed in the ante-natal, equilibrium condition. As remarked in the Introduction, Baker and Issa [6] have shown that all V1 response features can be described in terms of six variables—optimal values of orientation preference, spatial frequency preference, and temporal frequency preference, each associated with a Gaussian bandwidth of tuning of the cortical response to these features. These define three hypothetical filter processes. Stimulus variables in the present model have equivalents to those used in the spatio-temporal filter model. These are:

Spatio-temporal model	Present model
Object orientation	Orientation relative to the y -axis defined for Eqs. (18.17) and (18.18)
Object velocity	V_x
Object drift angle	$\tan^{-1}[K_y/K_x]$
Object spatial frequency	$K_x/\cos(\tan^{-1}[K_y/K_x])$
Object temporal frequency	$V_x K_x$

Repeated stimulation with a particular stimulus will therefore lead, under Hebbian learning, to maximization of the response to that stimulus, thus creating an apparent “tuning” of particular neurons to that particular combination of stimulus

features. Thus, the spatio-temporal model can be regarded as a consequence of the present model. Optimization by learning of the parameters for each of the three filters must be competitive between adjacent cells, providing the necessary condition for fitting response maps with continuity and completeness, by dimension-reduction methods [24, 25, 52]. Finally, the consolidation of saturated long-range patchy connections by Hebbian learning would result in mature “like to like” connections.

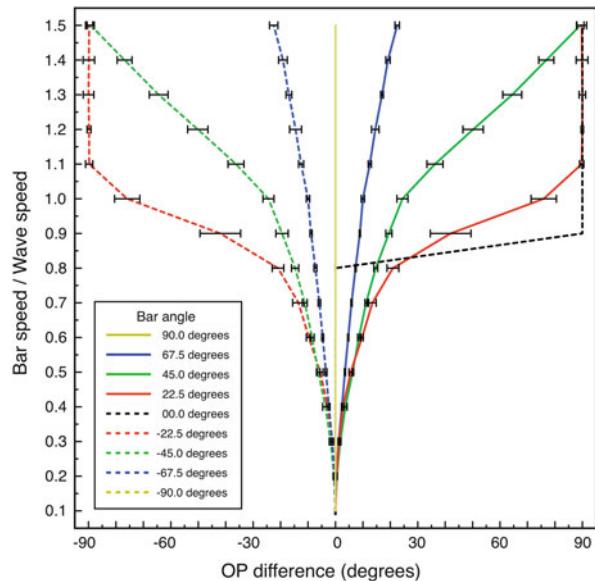
18.3 Simulations: A Critical Test

A critical test of our model, then, is whether we can reproduce in simulation the results of Basole et al. [8], without appeal to a priori feature-specific responses to orientation, spatial frequency, or temporal frequency, as in the spatio-temporal filter model—the band-width of tuning regarded as a post-natal effect, and not a primary explanation. Equation (18.20) was applied in simulations of an hexagonal array of seven adjacent macrocolumns. Results reported in Fig. 18.6 are for the central macrocolumn of the array of seven. Examples from the array are shown in Fig. 18.5, which shows the orthogonal transformation of apparent OP from the lowest to the highest bar speed for a moving line stimulus oriented at 45° to its line of passage. Again, details of the simulation and controls are given elsewhere [106, 109].

18.3.1 Effect of Object Velocity on Apparent Orientation Preference

A moving line in the visual field, relayed by the direct visual pathway to the cRF of each macrocolumn is represented as a red bar. In a single simulation the red bar travelled across the entire hexagonal array from left to right, with constant speed, direction and orientation. The orientation of the red bar to the line of passage is measured as *bar angle* from degrees, where the bar is oriented orthogonally to the direction of travel, to ±90°, where the bar is oriented in the direction of travel. The lag-transmitted image of the red bar, relayed as subthreshold activation to each macrocolumn via the superficial patch system, is shown in green, with illumination about the zone of subthreshold activation, to indicate that input to the cRF from the direct visual pathway and contextual signals caused triggering of action potentials. The average angle from the macrocolumn singularity to the centers of action potential generation (i.e., all points on the green line with illumination) was calculated at each time-step, and shown as a black arrow, thus indicating the part of the macrocolumn with a response preference (*apparent OP*) for the particular bar movement. A change in the sector of the macrocolumn that is maximally stimulated

Fig. 18.6 Change in apparent OP, and standard error of the estimate, as a function of bar speed to wave speed, for lines at different orientations to their directions of motion. Bar length 6 units



is equivalent to an equal change in the angle of approach of the bar needed to maintain stimulation of the same sector. The black arrow angle was averaged over a window beginning after the red bar had passed the center of the macrocolumn. Combinations of bar-length, orientation of the bar to the direction of movement, and bar speed, were then systematically varied in separate simulations. Their effects on OP, measured at the central local map of the hexagonal group, were obtained as *OP difference*, $\Delta\phi$ —a measure of the change in OP as a function of these variables. The *reference OP*, $\phi_0 \in [0, \pi]$, was the OP found at the lowest bar speed applied (bar speed/wave speed = 0.1) and the *apparent OP*, $\phi_1 \in [0, \pi]$, was the OP found at higher speeds.

Systematic results are shown in Fig. 18.6, which graphs OP difference versus bar speed/wave speed, for bar angles to $\pm 90^\circ$, calculated for a bar length of 6 units. Variation of bar length showed progressive lessening of the effect of velocity on OP for greater bar lengths.

For the case of bar-angle 0° (a line oriented orthogonally to its direction of passage, as in classical measurements of OP) no OP difference is seen until, as bar speed approaches wave speed, a 90° change in apparent OP takes place at a single increment in speed. This corresponds to transition to a “motion streak”, as object movement blurs resolution in the direction of motion. Increasing OP difference with bar speed at other bar angles is a more gradual development of the same effect—that is, mixing of responses to object speed and to object orientation.

These results match the findings of Basole et al. [8] and are consistent with effects of Doppler shifting of the image transferred from the global to the local map and further selected by the time of activation of the cRF.

To exclude alternative explanations, similar simulations were performed in which contextual (green bar) responses were constrained to occur only with a limited angular response within a macrocolumn. That is, a restricted response to the line, according only to its orientation was imposed, in analogy to conventional models of OP, but with conduction delays of “like to like” fibers included. Then, systematic variation of OP with bar velocity did not occur. A further comparison can be made to the predicted anatomical structure that would emerge if there were no competition for resources between synapses from the same neuron. In that case, OP maps would emerge with any given stimulus orientation represented twice about a singularity—which is not the case.

18.4 Conclusion

From our initial conjecture regarding the evolutionary path to encephalization, we have deduced a model of self-organization in V1 that explains otherwise disparate experimental data, and data which has presented paradoxes to standard explanations. The model’s properties also approach a biological optimum, achieving minimum metabolic cost per neuron, minimum total axonal length per connection, and efficient packing, minimizing transmission delays. In effect, the decline of stimulus cross-correlation with increasing distance in visual sensory space, and the corresponding decline of cortical pulse cross-correlation with increasing distance of cell separation, permits development of an internal reference frame for representation of visual events prior to direct visual experience—a *tabula rasa*—upon which subsequent learning can be etched. In the pre-vision state, synaptic couplings at equilibrium are highly orderly, thus offering high information storage capacity, as complex visual correlations become stored by subsequent Hebbian consolidation.

Beyond V1, we speculate that the model may be generally applicable throughout the neocortex. Cortical structure and dynamics, including patch connections, are similar throughout the cortex, and stimulus cross-correlations decline with distance in all sensory modalities—most obviously so for somatic sensation, but also with tone and position in the auditory system. The spatial distribution and intermixing of odour receptors (reviewed by Freeman [32]) implies an analogy even for olfaction. Similar ultra-small world representations might therefore form for all sensory cortices. Although outside primary sensory cortices a similar degree of orderliness of connections is not apparent, that does not exclude the applicability of the model elsewhere, because, as we have seen, the model may be applicable to V1 even in those species which lack strong anatomical ordering, and readily apparent orderliness is a geometrical consequence only for those cortical areas made up of cells with particularly long patch connections.

The principles of the model may also generalize to inter-areal interactions, during embryogenesis. Cortical areas project to and from other areas via cortico-cortical

connections, which, because their axons diverge and overlap at their terminations, project substantial parts of one area onto another, and are generally reciprocal between areas [14, 15]. We have argued above, that, because co-variance of activity declines with metric distance at both the scale of the patchy connections and within a macrocolumn, a homeotypic mapping between scales can emerge. By similar arguments, sets of macrocolumns at both the lower, V1, level and higher levels, could resonate with, and form preferential connections with, superimposed and overlapping groups at the other level, in accord with the developmental selection requirement to maximize joint synchrony. With the occurrence of eye-opening, Hebbian learning would then begin to overwrite the equilibrium resonance configuration between areas, in analogy to the process at intra-areal level—with the added property of associating concurrent patterns of activity in the V1 macrocolumns. A beginning on defining these reciprocal interrelations has been made elsewhere [106].

Consequently, we may come to an analysis of information flow in the brain's neural networks, in a new way. It has long been known that a macroscopic level, sensory inputs to, and motor outputs from, the cortex are arranged into topographic maps. The present model extends the topographic format to the millimetric scale, and implies that the raw material of cortical information flow is the interaction of spatially organized images. This differs from standard concepts of feature detection, which have dominated conceptions of cortical function since Hubel and Wiesel's famous observations of 1959 [48]. On the Möbius strip, spatial relationships of sensory representations maintain nearest-neighbour relations, and distances from singularities are associated with the distribution of conduction delay from surrounding cortex. Subsequent Hebbian-strengthened connections can bridge points with higher spatio-temporal correlations than accounted for by physical distance of separation in sensory space alone. When the same notion is extended to inter-areal connections, superposition of projections to higher cortical areas permits responses to ever more complex "features" combining stimulus aspects that are separated in visual space. At the ultimate level of expression at the motor cortex, the same organizational model is applicable in the reverse way to that of the sensory cortex—with pyramidal motor neurons substituted for direct visual pathway inputs. The resulting organization is one in which signal flows from sensory inputs to motor cortices could generate organized sensory-motor sequences in response to both externally generated inputs, and to autonomous, internally generated signals [32, 34, 105]. Cortico-cortical connections would permit extension to almost any level of hierarchical complexity—a modular property facilitating the evolution of encephalization.

Acknowledgements The material in this chapter was presented at the First Neural Field Conference, Reading University, UK, (2010), with support of JJW. Special acknowledgement is made of the courage and generosity of Adrienne Wright, in enabling this work, and its presentation on that occasion.

References

1. Albert, M.V., Schnabel, A., Field, D.J.: Innate visual learning through spontaneous activity patterns. *PLoS Comput. Biol.* **10** (2008). doi:10.1371/journal.pcbi.1000137
2. Alexander, D.M., Bourke, P.D., Sheridan, P., Konstandatos, O., Wright, J.J.: Intrinsic connections in tree shrew V1 imply a global to local mapping. *Vis. Res.* **44**, 857–876 (2004)
3. Angelucci, A., Bullier, J.: Reaching beyond the classical receptive field of V1 neurons; horizontal or feedback axons? *J. Physiol. (Paris)* **97**, 141–154 (2003)
4. Angelucci, A., Levitt, J.B., Lund, J.S.: Anatomical origins of the classic receptive field and modulatory surround field of single neurons in macaque visual cortical area V1. *Prog. Brain Res.* **136**, 373–388 (2002)
5. Bahrey, H.L.P., Moody, W.J.: Early development of voltage-gated ion currents and firing properties in neurons of the mouse cerebral cortex. *J. Neurophysiol.* **89**, 1761–1773 (2002)
6. Baker, T.I., Issa, N.P.: Cortical maps of separable tuning properties predict population responses of complex visual stimuli. *J. Neurophysiol.* **94**, 775–787 (2005)
7. Barber, M.J., Lichtman, J.W.: Activity-driven synapse elimination leads paradoxically to domination by inactive synapses. *J. Neurosci.* **19**, 9975–9985 (1999)
8. Basole, A., White, L.E., Fitzpatrick, D.: Mapping of multiple features in the population response of visual cortex. *Nature* **423**, 986–990 (2003)
9. Basole, A., Kreft-Kerekes, V., White, L.E., Fitzpatrick, D.: Cortical cartography revisited: a frequency perspective on the functional architecture of visual cortex. *Prog. Brain Res.* **154**, 121–134 (2006)
10. Blakemore, C., Cooper, G. F.: Development of brain depends on the visual environment. *Nature* **228**, 477–478 (1970)
11. Blakemore, C., Van Sluyters, R.C.: Innate and environmental factors in the development of the kitten's visual cortex. *J. Physiol. (Lond.)* **248**, 663–716 (1975)
12. Blaschke, A.J., Staley, K., Chun, J.: Widespread programmed cell death in proliferative and postmitotic regions of the fetal cerebral cortex. *Development* **122**, 1165–1174 (1996)
13. Bosking, W.H., Zhang, Y., Schofield, B., Fitzpatrick, D.: Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *J. Neurosci.* **17**(6), 2112–2127 (1997)
14. Boucsein, C., Nawrot, M., Schepel, P., Aertsen, A.: Beyond the cortical column: abundance and physiology of horizontal connections imply a strong role for inputs from the surround. *Front. Neurosci.* **5**, 32 (2011). doi:10.3389/fnins.2011.00032
15. Braitenberg, V., Schüiz, A.: Anatomy of the cortex: statistics and geometry. Springer, Berlin/New York (1991)
16. Bressler, S.L., Coppola, R., Nakamura R.: Episodic multiregional cortical coherence at multiple frequencies during visual task performance. *Nature* **366**, 153–156 (1993)
17. Buzás, P., Kovács, K., Ferecskó, A.S., Budd, J.M.L., Eysel, U.T., Kisvárdy Z.F.: Model-based analysis of excitatory lateral connections in the visual cortex. *J. Comp. Neurol.* **499**, 861–881 (2006)
18. Callaway, E.M., Katz, L.C.: Emergence and refinement of clustered horizontal connections in cat striate cortex. *J. Neurosci.* **10**, 1134–1153 (1990)
19. Carrera-Perpiñán, M.Á., Lister, R.J., Goodhill, G.J.: A computational model for development of multiple maps in primary visual cortex. *Cereb. Cortex* **15**, 1222–1233 (2005)
20. Chapman, C.L., Bourke, P.D., Wright, J.J.: Spatial eigenmodes and synchronous oscillation: coincidence detection in simulated cerebral cortex. *J. Math. Biol.* **45**, 57–78 (2002)
21. Cohen, R., Havlin, S.: Scale-free networks are ultra-small. *Phys. Rev. Lett.* **90**, 058701 (2003)
22. Downes, J.H., Hammond, M.W., Xydias, D., Spencer, M., Becerra, V.M., Warwick, K., Whalley, B.J., Nasuto, S.J.: Emergence of a small-world functional network in cultured neurons. *PLoS Comput. Biol.* **8**, e1002522 (2012)
23. Durack, J.C., Katz, L.C.: Development of horizontal projections in layer 2/3 of ferret visual cortex. *Cereb. Cortex* **6**, 178–183 (1996)

24. Durbin, R., Mitchison, G.: A dimension reduction framework for understanding cortical maps. *Nature* **343**, 644–647 (1990)
25. Durbin, R., Willshaw, D.J.: An analogue approach to the travelling salesman problem using an elastic net method. *Nature* **326**, 689–691 (1987)
26. Eckhorn, R., Bauer, R., Jordon, W., Brosch, M., Kruse, W., Monk, M., Reitboeck, H.J.: Coherent oscillations: a mechanism of feature linking in the visual cortex? *Biol. Cybern.* **60**, 121–130 (1988)
27. Eckhorn, R., Reitboeck, H.J., Arndt, M., Dicke, P.: Feature linking via synchronization among distributed assemblies: simulations of results from cat visual cortex. *Neural Comput.* **2**, 293–307 (1990)
28. Elliot, T.: Stability against fluctuations: scaling, bifurcations, and spontaneous symmetry breaking in stochastic models of synaptic plasticity. *Neural Comput.* **23**, 674–734 (2011)
29. Elliott, T., Shadbolt, N.R.: Multiplicative synaptic normalization and a nonlinear Hebb rule underlie a neurotrophic model of competitive synaptic plasticity. *Neural Comput.* **14**, 1311–1322 (2002)
30. Enoki, R., Hu, Y-L., Hamilton, D., Fine, A.: Expression of long-term plasticity at individual synapses in hippocampus is graded, bi-directional, and mainly pre-synaptic: optic quantal analysis. *Neuron* **62**, 242–253 (2009)
31. Erwin, E., Miller, K.D.: Correlation-based development of ocularly-matched orientation maps and ocular dominance maps: determination of required input activity structures. *J. Neurosci.* **18**, 9870–9895 (1998)
32. Freeman, W.J.: *Mass Action in the Nervous System*. Academic, New York (1975)
33. Freeman, W.J.: Predictions on neocortical dynamics derived from studies in paleocortex. In: *Induced Rhythms of the Brain*. Birkhäuser, Boston (1991)
34. Freeman, W.J., Quiroga, R.Q.: *Imaging Brain Function*. Springer, New York/Heidelberg/Dordrecht/London (2013)
35. Gilbert, C.D., Wiesel, T.N.: Morphology and intracortical projections of functionally characteristic neurons in cat visual cortex. *Nature* **280**, 120–125 (1979)
36. Gilbert, C.D., Wiesel, T.N.: Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. *J. Neurosci.* **9**, 2432–2442 (1989)
37. Gray, C.M., König, P., Engel, A.K., Singer, W.: Oscillatory responses in cat visual cortex exhibit intercolumnar synchronisation which reflects global stimulus properties. *Nature* **388**, 334–337 (1989)
38. Grossberg, S., Olson, S.J.: Rules for the cortical map of ocular dominance and orientation columns. *Neural Netw.* **7**, 883–894 (1994)
39. Harris, A.E., Ermentrout, G.B., Small, S.L.: A model of ocular column development by competition for trophic factor. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 9944–9949 (1997)
40. Hashimoto, K., Tsujita, M., Miyazaki, T., Kitamura, K., Yamazaki, M., Shin, H-S., Watanabe, M., Sakimura, K., Kano, M.: Postsynaptic P/Q-type Ca^{2+} channel in Purkinje cell mediates synaptic competition and elimination in developing cerebellum. *PNAS* **108**, 9987–9992 (2011)
41. Hassenstaub, A., Shu, Y., Haider, B., Krauschaar, U., Duque, A., McCormick, D.A.: Inhibitory postsynaptic potentials carry synchronized frequency information in active cortical networks. *Neuron* **47**, 423–435 (2005)
42. Hebb, D.: *The Organization of Behavior*. Wiley, New York (1949)
43. Heck, N., Golbs, A., Riedemann, T., Sun, J-J., Lessmann, V., Luhmann, H.J.: Activity dependent regulation of neuronal apoptosis in neonatal mouse cerebral cortex. *Cereb. Cortex* **18**, 1335–1349 (2008)
44. Higginbotham, H., Yokota, Y., Anton, E.S.: Strategies for analyzing neuronal progenitor development and neuronal migration in the developing cerebral cortex. *Cereb. Cortex* **21**, 1465–1474 (2011)
45. Hiramoto, M., Cline, H.: Convergence of multisensory inputs in the *xenopus* tadpole tectum. *Dev. Neurobiol.* **69**, 959–971 (2009)

46. Hirsch, J.A., Gilbert, C.D.: Synaptic physiology of horizontal connections in the cat's visual cortex. *J. Neurosci.* **11**, 1800–1809 (1991)
47. Horton, C.H., Adams, D.L.: The cortical column: a structure without a function. *Phil. Trans. R. Soc. B.* **360**, 837–862 (2005)
48. Hubel, D.H., Wiesel, T.N.: Receptive fields of single neurones in the cat's striate cortex. *J. Physiol.* **148**, 574–591 (1959)
49. Issa, P., Rosenberg, A., Husson, T.R.: Models and measurements of functional maps in V1. *J. Neurophysiol.* **99**, 2745–2754 (2008)
50. James, W.: *Psychology (Briefer Course)*. Holt, New York (1890)
51. Kathuri, N., Lichtman, J.W.: The role of neuronal identity in synaptic competition. *Nature* **424**, 430 (2003). doi:10.1038/nature01836
52. Kohonen, T.: Self-organized formation of topologically correct feature maps. *Biol. Cybern.* **43**, 59–69 (1982)
53. Li, W., Their, P., Wehrhahn, C.: Contextual influence on orientation discrimination of humans and responses of neurons in V1 of alert monkeys. *J. Neurophysiol.* **83**, 941–954 (2000)
54. Linsker, R.: From basic network principles to neural architecture: emergence of spatial opponent cells. *Proc. Natl. Acad. Sci. U.S.A.* **83**, 7508–7512 (1986)
55. Linsker, R.: From basic network principles to neural architecture: emergence of orientation selective cells. *Proc. Natl. Acad. Sci. U.S.A.* **83**, 8390–8394 (1986)
56. Linsker, R.: From basic network principles to neural architecture: emergence of orientation columns. *Proc. Natl. Acad. Sci. U.S.A.* **83**, 8779–8783 (1986)
57. MacLean, P.D.: A triune concept of the brain and behavior. In: Boag, T.J., Campbell, D. (eds.) *The Hincks Memorial Lectures*, pp. 6–66. University of Toronto Press, Toronto (1973)
58. Marks, G.A., Shaffery, J.P., Okensberg, A., Speciale, S.G., Roffwarg, H.P.: A functional role for REM sleep in brain maturation. *Behav. Brain Res.* **69**, 1–11 (1995)
59. McGuire, B.A., Gilbert, C.D., Rivlin, P.K., Wiesel, T.N.: Targets of horizontal connections in macaque primary visual cortex. *J. Comp. Neurol.* **305**, 370–392 (1991)
60. Miller, K.D.: A model for the development of simple cell receptive fields and the ordered arrangement of orientation columns through the activity dependent competition between ON- and OFF-center inputs. *J. Neurosci.* **14**, 409–441 (1994)
61. Miller, K.D.: Synaptic economics: competition and cooperation in correlation-based synaptic plasticity. *Neuron* **17**, 371–374 (1996)
62. Mirmiran, M.: The function of fetal/neonatal rapid eye movement sleep. *Behav. Brain Res.* **69**, 13–22 (1995)
63. Mitchison, G., Crick, F.: Long axons within the striate cortex: their distribution, orientation, and patterns of connection. *Brain Pharmacol.* **79**, 3661–3665 (1982)
64. Miyashita, M., Tanaka, S.: A mathematical model for the self-organization of orientation columns in visual cortex. *NeuroReport* **3**, 69–72 (1992)
65. Model, P.G., Bornstein, M.B., Crain, S.M., Pappas, G.D.: An electron microscopic study of the development of synapses in cultured fetal mouse cerebrum continuously exposed to xylocaine. *J. Cell Biol.* **49**, 362–371 (1971)
66. Montague, P.R.: The resource consumption principle: attention and memory in volumes of neural tissue. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 3691–3623 (1996)
67. Muir, D.R., Douglas, R.J.: From neural arbours to daisies. *Cereb. Cortex* **21**, 1118–1133 (2011)
68. Muir, D.R., Da Costa, N.M.A., Girardin, C.C., Naaman, S., Omer, D.B., Ruesch, E., Grinvald, A., Douglas, R.J.: Embedding of cortical representations by the superficial patch system. *Cereb. Cortex* **21**, 2244–2260 (2011)
69. Newman, J.D., Harris, J.C.: The scientific contributions of Paul D: MacLean. *J. Nerv. Ment. Dis.* **197**, 3–5 (2009)
70. Obermayer, K., Ritter, H., Schulten, K.: A principle for the formation of the spatial structure of cortical feature maps. *Proc. Natl. Acad. Sci. U.S.A.* **87**, 8345–8349 (1990)
71. Obermayer, K., Ritter, H., Schulten, K.: A model for the development of the spatial structure of retinotopic maps and orientation columns. *IEICE Trans. Fundam.* **E75A**, 537–545 (1992)

72. O'Connor, D.H., Wittenberg, G.M., Wang, S.S.-H.: Dissection of bidirectional synaptic plasticity into saturable unidirectional processes. *J. Neurophysiol.* **94**, 1565–1573 (2005)
73. O'Connor, D.H., Wittenberg, G.M., Wang, S.S.-H.: Graded bidirectional synaptic plasticity is composed of switch-like unitary events. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 9679–9684 (2005)
74. Okamoto, H., Ichikawa, K.: A model for molecular mechanisms of synaptic competition for a finite resource. *Biosystems* **55**, 65–71 (2000)
75. Paik, S.-B., Ringach, D.L.: Retinal origin of orientation maps in visual cortex. *Nat. Neurosci.* **14**, 919–925 (2011)
76. Papez, J.W.: A proposed mechanism of emotion. *Arch. Neurol. Psychiatry* **38**, 725–743 (1937)
77. Price, D.J.: The postnatal development of clustered intrinsic connections in area 18 of the visual cortex in kittens. *Dev. Brain Res.* **24**, 31–38 (1986)
78. Rakic, P.: Specification of cerebral cortical areas. *Science* **241**, 170–176 (1988)
79. Ringach, D.L.: On the origin of the functional architecture of the cortex. *PLoS One* **2** e251 (2007)
80. Robinson, P.A., Rennie, C.J., Wright, J.J.: Synchronous oscillations in the cerebral cortex. *Phys. Rev. E* **57**, 4578–4588 (1998)
81. Rockland, K.S., Lund, J.S.: Intrinsic laminar lattice connections in primate visual cortex. *J. Comp. Neurol.* **216**, 303–318 (1983)
82. Ruthazer, E.S., Stryker, M.P.: The role of activity in the development of long-range horizontal connections in area 17 of the ferret. *J. Neurosci.* **16**, 7253–7269 (1996)
83. Schillen, T.B., König, P.: Binding by temporal structure in multiple feature domains of an oscillatory neural network. *Biol. Cybern.* **70**, 397–405 (1994)
84. Scholl, D.A.: *The Organization of the Cerebral Cortex*. Wiley, New York (1956)
85. Sherk, H., Stryker, M.P.: Quantitative study of orientation selectivity in visually inexperienced kittens. *J. Neurophysiol.* **39**, 63–70 (1976)
86. Shi, Y., Kirwan, P., Smith, J., Robinson, H.P.C., Livesey, F.J.: Human cerebral cortex development from pluripotent stem cells to functional cortical synapses. *Nat. Neurosci.* **15**, 477–486 (2012)
87. Singer, W.: Neuronal synchrony: a versatile code for the definition of relations? *Neuron* **24**, 49–65 (1999)
88. Sperry, R.W.: Problems Outstanding in the Evolution of Brain Function. James Arthur Lecture on the Evolution of the Human Brain. The American Museum of Natural History, New York (1964)
89. Steriade, M.: Corticothalamic resonance, states of vigilance and mentation. *Neuroscience* **101**, 243–276 (2000)
90. Swindale, N.V.: A model for the formation of orientation columns. *Proc. R. Soc. B.* **215**, 211–230 (1982)
91. Swindale, N.V.: A model for the coordinated development of columnar systems in primate striate cortex. *Biol. Cybern.* **66**, 217–230 (1992)
92. Swindale, N.V.: The development of topography in the visual cortex: a review of models. *Netw.: Comput. Neural Syst.* **7**, 161–247 (1996)
93. Tanaka, S.: Theory of self-organization of cortical maps: mathematical framework. *Neural Netw.* **3**, 625–640 (1990)
94. Thomaïdou, D., Mione, M.C., Cavanagh, J.F.R., Parnavelas, J.G.: Apoptosis and its relation to the cell cycle in the developing cerebral cortex. *J. Neurosci.* **17**, 1075–1085 (1997)
95. Traub, R.D., Whittington, M.A., Stanford, I.M., Jefferys, J.G.R.: A mechanism for generation of long-range synchronous fast oscillations in the cortex. *Nature* **383**, 621–624 (1996)
96. Tsukada, M., Fukushima, Y.: A context dependent mechanism in hippocampal CA1 networks. *Bull. Math. Biol.* **73**, 417–435 (2010). doi:10.1007/s11538-010-9566-8
97. van Ooyen, A.: Competition in the development of nerve connections: a review of models. *Netw.: Comput. Neural Syst.* **12**, R1–R47 (2001)
98. van Ooyen, A., Willshaw, D.J.: Competition for neurotrophic factor in the development of nerve connections. *Proc. R. Soc. Lond. B.* **266**, 883–892 (1999)

99. von der Malsburg, C.: Self organization of orientation sensitive cells in the striate cortex. *Kybernetik* **14**, 85–100 (1973)
100. Whittington, M.A., Faulkner, H.J., Doherty, H.C., Traub, R.D.: Neuronal fast oscillations as a target site for psychoactive drugs. *Pharmacol. Ther.* **86**, 171–190 (2000)
101. Wiesel, T.N., Hubel, D.H.: Ordered arrangement of orientation columns in monkeys lacking visual experience. *J. Comp. Neurol.* **158**, 307–318 (1974)
102. Willshaw, D.J., von der Malsburg, C.: How patterned neural connections can be set up by self-organization. *Proc. R. Soc. B.* **194**, 431–435 (1976)
103. Witten, T.A., Sander, L.M.: Diffusion-limited aggregation, a kinetic critical phenomenon. *Phys. Rev. Lett.* **47**, 1400–1403 (1981)
104. Wright, J.J.: Generation and control of cortical gamma: findings from simulation at two scales. *Neural Netw.* **22**, 373–384 (2009)
105. Wright, J.J.: Attractor dynamics and thermodynamic analogies in the cerebral cortex: synchronous oscillation, the background EEG, and the regulation of attention. *Bull. Math. Biol.* **73**(2), 436–457 (2010). doi:10.1007/s11538-010-9562-z
106. Wright, J.J., Bourke, P.D.: On the dynamics of cortical development: synchrony and synaptic self-organization. *Front. Comput. Neurosci.* **7**, 4 (2013)
107. Wright, J.J., Bourke, P.D., Chapman, C.L.: Synchronous oscillation in the cerebral cortex and object coherence: simulation of basic electrophysiological findings. *Biol. Cybern.* **83**, 341–353 (2000)
108. Wright, J.J., Alexander, D.M., Bourke, P.D.: Contribution of lateral interactions in V1 to organization of response properties. *Vis. Res.* **46**, 2703–2720 (2006)
109. Wright, J.J., Bourke, P.D.: On the dynamics of cortical development: synchrony and synaptic self-organization. *Front. Comput. Neurosci.* **7**, 4 (2013). doi:10.3389/fncom.2013.00004
110. Yakovlev, P.I.: Motility, behaviour and the brain; stereodynamic organization and neural co-ordinates of behavior. *J. Nerv. Ment. Dis.* **107**, 313–335 (1948)
111. Yu, Y-C., He, S., Chen, S., Fu, Y., Brown, K.N., Yao, X-H., Ma, J., Gao, K.P., Sosinsky, G.E., Huang, K., Shi, S-H.: Preferential electrical coupling regulates neocortical lineage-dependent microcircuit assembly. *Nature* **486**, 113–118 (2012)

Index

- Action selection, 343, 344, 346, 350, 352, 362
Adaptation, 119–122, 124, 130, 131, 133, 134, 140, 169, 187–190, 192, 200, 205, 207, 350, 468
 α -function, 4, 5
Amari equation, 28, 289, 290, 301, 314
Amplitude equations, 1, 14, 17, 18, 120, 126, 129, 207
Anesthetic, general, 394
Anesthetics, 384, 388
Aplysia, 379
Apoptosis, 457, 459, 460, 462, 464, 468
Arnoldi iterations, 196
Asynchronous states, 241
Autoregressive moving average (ARMA) model, 387, 389
Axonal conduction velocity, 371
Axonal propagation, 436
- Background-relative-luminance contrast (BRLC), 327
Backpropagation, 33
Banach space, 310
Benzodiazepine, 383, 384
Bernoulli process, 287
Bessel function, 161, 189, 191
Beta buzz, 384, 385
Bifurcation, 224, 375–379
 analysis, 376
 Hopf, 25, 91, 119, 121, 122, 129, 133, 134, 139, 141–143, 192, 194, 224
 metabifurcation analysis, 377
 numerical, 187, 188, 208
 saddle-node, 24, 139, 141, 189, 192, 197, 408
 Shil'nikov saddle node, 376
- Turing-Hopf, 17, 409
Bispectral Index, 386
Bistability, 169
Bogdanov-Takens point, 376
BOLD (blood oxygen level dependent), 425, 427
Boundary conditions, 8, 111, 121, 137, 156, 168, 176, 178, 196, 217, 219, 240, 245, 258, 263, 264, 301
BRAIN (Brain Research through Advancing Innovative Neurotechnologies) Initiative, vi
Brain Activity Map, vi
Brain wave equation, 19
Breathers, 119–121, 134, 143, 149, 176, 183, 184, 187, 192, 193, 198, 199, 202
Brownian motion, 82, 91, 251
Bumps, 2, 3, 21, 22, 24, 26, 97, 100, 101, 103, 105–109, 119–121, 124, 127, 130–134, 136, 138–149, 154, 163, 166, 167, 171, 183, 187, 188, 213, 223, 225, 226, 328, 344, 347, 349, 350, 357, 360, 362
 stimulus-induced, 121
 stimulus-locked, 121, 146
 traveling, 97, 105, 119–121, 130–134, 136, 138, 139, 141, 143–149
- Cable equation, 5, 37
Canonical microcircuit model, 433
Center manifold, 187, 208
Chapman-Kolmogorov equation, 238
Cognition, vi, 3, 319, 320, 330, 331, 337, 342, 361–363, 367
 cognitive functions, 341, 342, 350
Coherence, 213

- Compartment model, 272, 287, 292
 Conditional density, 448
 Condition-of-satisfaction, 303, 316
 Conductance-based model, 242, 439
 Connection, 3, 77, 98, 103, 106, 115, 213–216,
 219, 222, 223, 226, 230, 232, 241,
 265, 284, 370, 417–419, 421, 422,
 424, 427, 439, 458, 461, 463, 465,
 466, 476
 patchy, 461, 466
 stochastic, 214, 225
 Connectivity, vi, 2, 3, 8, 28, 32, 80, 129, 176,
 188, 213, 215, 219, 233, 270, 361,
 369, 372, 375, 410, 417–421, 424,
 425, 427, 429, 436, 437, 440–442,
 445–447, 465
 global, 417, 418
 heterogeneous, 213
 homogeneous, 418, 419
 local, 422
 Connectome, vi, 418, 421, 425
 Context-free grammar, 307
 Convolution, 8, 19, 101, 120, 155, 156, 177,
 178, 184, 214, 215, 289, 301, 302,
 436, 437, 442
 Cortex, 3, 6, 48, 63, 105, 115, 117, 129, 153,
 182, 214, 236, 270, 320, 324, 345,
 367, 368, 370, 382, 383, 394, 396,
 398, 400, 403, 406, 407, 409–412,
 421, 425, 436, 439, 440, 442, 451,
 457–461, 464, 477
 auditory, 442, 450
 visual, vi, 105, 117, 129, 182, 214, 320,
 439, 451, 457, 459
 Criticality, 81, 84, 91, 92, 122, 202
 Cylinder set, 307
- Decision making, 341–343, 350, 351, 361, 362
 Delay, 58, 176, 178, 180, 183, 184, 332, 381,
 422, 424, 477
 Denotational semantics, 303
 Diffusion, 5, 61, 79, 80, 175, 200, 203, 208,
 236, 239, 251–254, 258, 269, 272,
 275, 276, 282, 289, 290, 292, 394,
 399, 402, 406, 407, 409, 411, 412,
 414
 diffusion tensor, 290, 421, 424
 Dipole, 6, 270, 272, 289, 292, 293, 436
 Discretization, 32, 36, 37, 178, 180, 250
 Dispersion curves, 402
 Distribution of population activation (DPA),
 321–324
 DNF model, 343, 344, 362
 Drifting, 120, 194, 202
 Dynamic causal modeling (DCM), 3, 436, 452
 Dynamic cognitive modeling (DCM), 34, 303
 Dynamic field automata, 300, 304, 305
 Dynamic field theory (DFT), 299, 319, 320,
 324, 327, 330, 331, 336, 337, 341
 Dynamic neural fields (DNF), 301, 302, 324,
 341, 343, 347, 361
- Eigenfunction, 133, 142, 163, 165, 201, 203
 Eigenmode, 14, 134
 Eigenvalue, 22–24, 79, 91, 124, 133, 137, 142,
 163, 196, 200, 203, 209, 399, 407,
 414
 Eigenvector, 122
 Electroencephalogram (EEG), 1, 3, 6, 19, 21,
 48, 54, 82, 91, 153, 269–271, 292,
 293, 367–370, 375, 380, 382–384,
 386, 388, 401, 408, 410, 421, 428,
 435, 436, 444
 Embryogenesis, 458, 459, 476
 Ephaptic interaction, 270, 293, 368
 Evans function, 22, 23, 138, 147–149, 166,
 219, 249
 Event-related brain potentials (ERP), 35
 evoked potentials, 419
 Extracellular space, 269, 270, 272, 275, 287,
 292
- Feature detection, 477
 Feature space, 299, 301
 Feedback, 7, 58, 119, 167, 188, 224, 270, 330,
 369, 381, 386, 468
 Fick's law, 289
 Field potential
 dendritic (DFP), 270, 271, 275, 278, 280,
 282, 285, 290, 292, 293
 local (LFP), 75, 270–272, 285, 293, 382,
 433–436, 444, 451, 452
 Firing rate function, 2, 18, 54, 79, 130, 137,
 146, 166, 187, 190, 208, 214, 236,
 256, 438
 Fixed point, 70, 79, 84, 88, 91, 216, 256, 263,
 265, 303, 328
 Fokker-Planck equation, 60, 61, 70, 86, 239,
 240, 250, 264, 265, 292
 Fourier transform, 10, 11, 19, 38, 79, 155, 157,
 160, 161, 171, 177, 179, 181, 184,
 195, 214, 228, 444, 445
 Fredholm alternative, 15, 201
 Free energy, 434, 448, 450, 451
 Frobenius-Perron equation, 299, 311, 312, 315

- Front(s), 2, 23, 28, 81, 120, 146, 154, 213, 214, 216, 217, 219, 223, 225, 230, 232, 233, 235, 247, 249–251, 253, 255–258, 421
traveling, 120, 145, 248, 249, 258
- Front speed, 217
- Functional Magnetic Resonance Imaging (fMRI), vi, 367, 384, 407, 421, 428, 436
- γ -amino butyric acid (GABA), 377, 383–385, 413, 424
- Gödel numbers, 309, 315
- GABA_A receptor, 384–386
- Gap junction, 368
- Generalized shift, 307–309
- Geometric visual hallucinations, 3, 127
- Ginzburg-Landau equations, 17
- Glauber states, 73
- Glia cells, 285, 293
- Goal inference, 343, 345, 350, 354, 359
- Green's function, 4, 5, 19, 76, 156, 261, 278, 281, 419
- Hamilton equations, 235, 263, 264
- Hamiltonian, 51, 53, 73, 158, 262, 263, 265
- Hankel transform, 161
- Head direction system, 153
- Heaviside function, 3, 57, 97, 98, 117, 130, 154, 158, 183, 190, 194, 244, 348
- Hebb rule, 33, 363, 457, 473
- Heteroclinic, 34, 263, 265, 303
- Hilbert space, 31
- Hodgkin-Huxley equations, 22
- Homoclinic, 24, 158, 192, 375, 379
- Human Brain Project, vi
- Human-robot interaction, 342, 343
- Ill-posed, 29, 30, 34
- Inhibition, 2, 12, 36, 58, 88, 101, 120, 127, 187, 214, 223, 303, 325, 328, 330, 335, 336, 350, 381, 382, 394, 396, 439
- Instability, 10, 11, 13, 14, 117, 133, 134, 139, 141, 147, 166, 187, 188, 194, 200, 203, 208, 325, 333, 404, 405, 407, 409, 414, 425, 469
detection, 325, 328
drift, 131, 134, 142, 189, 202, 208
memory, 337
selection, 330
- Turing, 9, 11, 12, 404, 409
- Integrate-and-fire (IF) neuron, 237, 238
- Interface, 26–28, 167, 184, 187, 188, 190, 198, 208, 337, 429
- Interneurons, 284, 330, 381, 418, 437
- Inverse problems, 1, 3, 29
- Involution, 158
- Ion channels, 236, 274
- Isoflurane, 373, 410
- Jacobian, 87, 192, 414
- Jansen-Rit model, 433
- Joint action, 341, 343, 344, 351, 362
- Kernel, 17, 28, 30, 32, 34, 36, 122, 134, 135, 177, 178, 181, 188, 196, 283, 301, 303, 314, 325, 350, 420, 437, 441, 442, 445
- Kirchhoff's laws, 276
- Koopman approach, 311
- Kramers-Moyal expansion, 60, 249
- Lagrangian, 68, 74
- Langevin equation, 87, 239, 243, 249, 265
- Laplace
assumption, 448
Laplacian, 196, 420, 429
transform, 10, 38, 82, 257
Variational, 448
- Leaky integrate-and-fire (LIF) model, 242, 272, 282, 284, 287, 293
- Leaky integrator rate (LIR) model, 288
- Liley model, 375, 377, 379, 383
- Long range connections, 214, 230, 231
- Long wavelength approximation, 160
- Loss of consciousness (LOC), 407, 410, 411
- Lotka-Volterra equations, 50, 51, 53
- Lyapunov exponent, 182, 379, 407
- Macrostate, 311
- Magnetoencephalogram (MEG), 3, 368, 375, 421, 428, 433–435, 442, 451, 452
- Marginal velocity distribution, 374
- Master equation, 60, 61, 64, 65, 68, 69, 72, 74, 85, 86, 91, 236, 245, 249, 258, 265
- Maxwell
construction, 395, 396
equations, 6, 272, 291–293
- McCulloch-Pitts model, 48
- Mean-field, 3, 72, 77, 79, 84, 87, 293, 369

- Metabifurcation, 376, 377
 Mexican hat, 2, 10, 24, 122, 140, 325
 MindScope Project, vi
 Mirror system, 345
 Möbius strip, 469, 477
 Moment generating function, 67, 68
 Moore-Penrose pseudoinverse, 34
 Multistability, 378
- Nernst-Planck equation, 272, 290–292
 Neurodynamics, 342
 Neuroinformatics, 421, 427
 Neurosteroid, 384
 Neurotransmitter, 4, 274, 287, 372, 377, 381, 411
 N-methyl-d-aspartate (NMDA), 373, 424, 425
 Noise, 3, 53, 87, 236, 241, 244–246, 249, 255, 258, 284, 328, 375, 381, 394, 401, 413, 425, 448, 450, 463
 channel, 236, 448
 observation, 447, 448
 Nonlinear dynamical automata, 300
 Non-synaptic plasticity, 368
 Normal form, 127, 129, 143
 Novikov's theorem, 249
 Numerical continuation, 171, 195, 197
- Observables, vi, vii, 54, 75, 270, 320
 Ohm's law, 290
 Operator, 4, 8, 22, 23, 29–35, 61, 62, 65, 73, 108, 148, 156, 160, 198, 200, 201, 203, 209, 251, 253, 303, 420
 annihilation, 63
 bosonic, 64
 compact, 30, 31
 creation, 63
 self-adjoint, 15, 31, 39, 200, 251
 time-ordering, 278, 281
 Orientation preference (OP), 461, 462, 469, 473–476
 Ornstein-Uhlenbeck process, 239, 255, 284
 Oscillations, 47, 91, 92, 159, 225, 285, 389, 393, 407, 409, 464
 alpha band, 380, 382–384
 gamma band, 382
 synchronous, 460, 464
 Oscillator algebra, 67
- Parse, 308
 Path integral, 65, 67, 68, 72, 74, 79, 235, 245, 255, 258, 259, 262, 265
- Patterns, vi, 2, 7, 10, 12, 17, 18, 24, 34, 97, 123, 126, 127, 129, 153, 168, 199, 221, 226, 232, 320, 324, 328, 344, 348, 350, 362, 368, 378, 398, 409, 410, 424, 425, 468, 477
 spatiotemporal, 121, 407
 Pauli spin matrices, 65
 Peaks, 82, 91, 123, 322, 324, 325, 328, 330, 336, 435, 442, 443
 self-stabilized, 332, 337
 self-sustained, 327, 337
 Perception, 3, 322, 332, 333, 342, 343, 363, 379
 Perceptron, 35, 54
 Persistent fluctuations, 213, 221, 222, 228, 232, 233
 Perturbation expansion, 15
 Phase transition, 69, 80, 92, 394, 398, 408, 411
 Population, 2, 6, 7, 55, 60, 62, 77, 78, 86, 88–91, 102, 176, 188, 214, 215, 236–238, 241–243, 245, 271, 285, 289, 321, 324, 344, 346–349, 352, 357, 360, 362, 371, 377, 382, 418, 419, 425, 435, 438, 446
 coding, 321, 322, 324
 model, 7, 79, 188
 Potential, 5, 8, 98, 237, 238, 243, 270, 274, 275, 279, 325, 370, 411, 463
 action, 4, 7, 56, 287, 460, 474
 extracellular, 6
 postsynaptic, 370, 372, 384, 394, 399, 409, 438
 resting, 4, 5, 275, 276, 410, 413, 463
 Power law, 89, 230, 255, 457, 465
 Power spectra, 272
 Prediction, 84, 92, 308, 336, 341, 342, 362, 383, 414, 447
 Propofol, 385–387, 394, 410, 413
 Pulses, 21, 34, 81, 139, 184, 213, 214, 219, 223, 370, 371, 460, 464, 471
 traveling, 119
 Pyramidal cells, 6, 270, 272, 274, 292, 437, 440, 449
- Quantum field theory, 61, 62, 72
- Recovery of consciousness (ROC), 408
 Regularization, 30–32
 Tikhonov, 32, 34
 Renormalization, 72, 77, 92, 255
 Representation, 18, 35, 50, 64, 66, 73, 77, 115, 189, 259, 260, 302, 303, 309, 315,

- 320, 332, 345, 347, 354, 357, 398, 423, 462, 476
occupation number, 63
spectral, 31
- Resting state (RSN), 9, 10, 81, 348, 375, 425
network, 425
- Retinotopic map, 105
- Reversible, 153, 158, 159, 384
- Rings, 111, 120, 121, 123, 153, 163–165, 171, 180, 199
- Saccades, 321, 322, 327, 328, 330
- Scattering, 208, 398
- Short-term memory, 153
- Sigmoid function, 97, 247, 248, 371, 421, 463
- Simulation, vi, 53, 105, 123, 134, 181, 207, 247, 293, 345, 346, 354, 357, 360, 362, 363, 403, 405–407, 414, 418, 427, 474
numerical, 182, 195, 198
- Singular perturbation theory, 84, 120
- Singular vector, 31
- Sloshers, 119, 143
- Small world, 460, 465
- Snakes, 159, 164
- Solitons, 205
- Spike frequency adaptation, 120, 167, 169, 188, 189, 207
- Spike rate, 183, 274, 320, 325
- Spots, 18, 120, 128, 187–189, 191, 192, 194, 197, 200, 202, 203, 205–207, 327
- Stability analysis, 97, 120, 134, 207, 305, 398
- Stochastic neural field models, 3
- Stuart-Landau equation, 16, 205
- Sub-threshold solutions, 337
- Superior colliculus, 322, 328
- Symbolic dynamics, 307
- Symbologram, 309, 310, 315
- Symmetry, 13, 53, 129, 135, 142, 160, 163, 165, 190, 193, 201, 270, 300, 398, 421, 463, 469
- Synapse, 2, 4, 6, 25, 78, 237, 274, 275, 330, 368, 382, 385, 411, 458, 468
axo-myelenic, 368
excitatory, 270, 272, 276, 283, 292
inhibitory, 106, 270, 274–276, 292
synaptic depression, 3, 169, 188
- Syntactic language processing, 35, 303, 307
- System-size expansion, 61, 69, 72, 85, 87, 92
- Thalamus, 368, 410
thalamic input, 285, 382, 440
- Thermodynamic limit, 236, 241
- The Virtual Brain (TVB), 417, 418, 427, 429
- Threshold crossing, 135, 237, 248
- Time constant, 106, 107, 122, 176, 244, 282, 301, 314, 377
- Transfer function, 120, 176, 183, 387, 418, 444, 445
- Tuning curves, 320–322
- Turing machine, 300, 305, 307, 308, 315
- van der Waals gas, 397, 398, 403
- Vector symbolic architectures, 302
- Visual hallucinations, 153, 213
- Warburg impedance, 292
- Waves
spiral, 105, 120, 167–169, 171
traveling, 125, 129, 135, 147, 149, 237, 248, 258, 362, 409, 421, 461
- Wilson-Cowan equations, 2, 57, 58, 61, 72, 78, 87, 93
- Winnerless competition, 303
- Wizard hat, 157, 192, 205
- Working memory, 187, 232, 327, 330–333, 335, 336, 341, 350, 362
- XOR, 35, 36