



This is to certify that the

dissertation entitled

Bifurcations in Brain Dynamics

presented by

Eugene Izhikevich

has been accepted towards fulfillment of the requirements for

<u>Ph.D.</u> degree in <u>Mathematics</u>

Auc Hantt

(Frank C. Hoppensteadt) Major professor

Date April 18, 1996

MSU is an Affirmative Action/Equal Opportunity Institution



PLACE IN RETURN BOX to remove this checkout from your record. TO AVOID FINES return on or before date due.

DATE DUE	DATE DUE	DATE DUE
SEP 23 1997		
NEW 2 8 2005		

MSU is An Affirmative Action/Equal Opportunity Institution c/oirs/datadus.pm3-p.1

BIFURCATIONS IN BRAIN DYNAMICS

By

Eugene M. Izhikevich

A DISSERTATION

Submitted to Michigan State University in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Department of Mathematics

1996

ABSTRACT

BIFURCATIONS IN BRAIN DYNAMICS

BY

Eugene M. Izhikevich

Mathematical models of the brain are studied with the assumption that the connections between neurons are weak. This leads to weakly connected systems, which are called Weakly Connected Neural Networks (WCNNs). Local dynamics of the WCNNs is studied using bifurcation theory.

First it is proved that the WCNNs could have interesting local dynamics with possible applications to neurocomputers only near bifurcations. Then it is shown that near the bifurcations the WCNNs can be significantly simplified and reduced to canonical models.

Derivation and analysis of the canonical models for multiple (quasi-static) saddlenode, pitchfork and Andronov-Hopf bifurcations and multiple cusp singularities is presented. Mathematical analysis of the canonical models suggests a new neural network paradigm — non-hyperbolic neural networks. It also sheds some light on possible synaptic organizations of the brain. In particular, it reveals the relationship between synaptic architectures (anatomy) and dynamical properties (function) of networks of neural oscillators.

A part of this dissertation (Chapters 2 and 7) received the SIAM Student Paper Prize in applied mathematics for 1995.

DEDICATION

To April 29.

The day of birth of two people who have had a huge impact on my life for the last three years.

ACKNOWLEDGMENTS

This work could not be accomplished without guidance and wisdom of my Teacher Frank C. Hoppensteadt. It is a great fortune and pleasure to work with him. I am very grateful for this opportunity and for all kinds of investments he made in me and my research. It seems to me that for the last three years we have developed a relationship which goes beyond an academic level and I hope he feels the same way. I am proud to be called a student of Frank C. Hoppensteadt.

I would like to mention my former scientific supervisor, professor of Moscow State University, Georgii G. Malinetskii, who introduced me to the Dynamical System Theory.

I am especially grateful to Sheldon E. Newhouse for his contribution to my understanding of Dynamical System Theory and for his patience to my "active" style of learning.

I am thankful to the Society for Industrial and Applied Mathematics (SIAM) for the high appreciation of this work.

Finally, I would like to thank my wife Tatyana Izhikevich (Kazakova) for her understanding and support.

Contents

	List	of figures	x
1	Intr	roduction	1
	1.1	Overview	1
	1.2	Models in Mathematical Biology	6
	1.3	Neurobiological Background	9
	1.4	Neural Network Types	11
		1.4.1 Olfactory Bulb	13
		1.4.2 Networks of Excitable Elements	14
	1.5	Generic Equations	17
Ι	De	erivation of Canonical Models	20
2	We	akly Connected Neural Networks	21
	2.1	Hyperbolic Equilibrium	21
	2.2	Non-Hyperbolic Equilibrium	24
	2.3	The Center Manifold Reduction	25
	2.4	Canonical Models	30

		2.4.1 Multiple Saddle-Node Bifurcation	31
		2.4.2 Multiple Cusp Singularity	33
		2.4.3 Multiple Pitchfork Bifurcation	34
	2.5	Multiple Andronov-Hopf Bifurcation	34
		2.5.1 Equality of Frequencies and Attention	37
	2.6	Discussion	39
		2.6.1 Adaptation Condition and Psychology	41
3	Sing	gularly Perturbed WCNNs	44
	3.1	Basic Definitions	44
	3.2	Motivational Examples	46
	3.3	Reduction to Regular Perturbation Problem	50
	3.4	Center Manifold Reduction	51
	3.5	Canonical Models	52
		3.5.1 Multiple Quasi-Static Saddle-Node Bifurcations	55
		3.5.2 Multiple Quasi-Static Pitchfork Bifurcations	58
		3.5.3 Discussion of the case $\mu = \mathcal{O}(\varepsilon)$	60
	3.6	Multiple Quasi-Static Andronov-Hopf Bifurcations	61
	3.7	Conclusion	63
		3.7.1 Synaptic Organizations of the Brain	63
4	Wea	akly Connected Maps	65
	4.1	Hyperbolic Fixed Points	67

	4.2	Non-hyperbolic Fixed Points	68
		4.2.1 Multiple Saddle-Node Bifurcations	69
		4.2.2 Multiple Flip Bifurcations	70
	4.3	Connection With ODE	72
11	A	nalysis of Canonical Models	74
5	Mu	tiple Saddle-Node Bifurcation	75
	5.1	Saddle-Node on a Limit Cycle	75
	5.2	The VCON	78
	5.3	Preliminary Analysis	79
	5.4	Case $c \neq 0$	80
	5.5	Adaptation Condition Is Satisfied	82
6	Mu	tiple Andronov-Hopf Bifurcation	85
	6.1	Complex Synaptic Coefficients c_{ij}	85
	6.2	Oscillator Death and Self-ignition	88
	6.3	Synchronization and Convergence	91
7	Mul	tiple Cusp Singularity	94
	7.1	Extreme Values of Parameters	95
		7.1.1 Global behavior	95
		7.1.2 Strong Input From Receptors	97
		7.1.3 Extreme Psychological Condition	98

	7.2	Canonical Models as a GAS-Type NN	99
	7.3	Symmetric Synaptic Connections	101
	7.4	Hebbian Learning Rule For Synaptic	
		Matrix C	103
	7.5	Bifurcations for $r = 0$	105
		7.5.1 Stability of the Origin	106
		7.5.2 Stability of the Other Equilibria	108
	7.6	Bifurcations for $r \neq 0$ (two memorized images)	112
		7.6.1 The Reduction Lemma	112
		7.6.2 Recognition: Only One Image Is Presented	115
		7.6.3 Recognition: Two Images Are Presented	117
	7.7	Bistability of Perception	118
	7.8	Quasi-Static Variation of Parameter b	121
8	Qua	asi-Static Bifurcations	124
	8.1	Stability of the Equilibrium	124
	8.2	Dale's Principle and Synchronization	127
	8.3	Further Analysis of the Andronov-Hopf Bifurcation	130
	8.4	Proofs of Theorems 54 and 57	132
		8.4.1 Proof of Theorem 54	134
		8.4.2 Proof of Theorem 57	135

9 Non-Hyperbolic Neural Networks

9.1	Problem 1	139
9.2	Problems 2 and 3	142
10 Syna	ptic Organizations of the Brain	144
10.1	Neural Oscillators	147
	10.1.1 Multiple Andronov-Hopf Bifurcation	149
	10.1.2 Type A and B Neural Oscillators	151
10.2	Dale's Principle and Connectivity	153
10.3	Classification of Synaptic Organizations	156
10.4	Learning Dynamics	158
10.5	Memorization of Phase Information	161
10.6	Synaptic Organizations	166
11 Discu	ussion	173
11.1	Canonical Models and Normal Forms	173
11.2 \$	Synaptic Connections	174
11.3	Mathematical Conditions and Biology	175
11.4	Co-dimensions of the Models	176
11.5	List of Canonical Models	178
Bibliogr	aphy	179

List of Figures

1.1	Summary of bifurcations and canonical models for regularly perturbed	
	weakly connected systems	4
1.2	Summary of multiple quasi-static bifurcations and canonical models	
	for singularly perturbed weakly connected systems	5
1.3	Relationship between models in mathematical biology	8
1.4	A "typical" neuron. It receives signals from other neurons via synapses	
	and sends the output signal to the other neurons through its axon	10
1.5	a. Multiple Attractor (MA) neural network. b. Globally Asymptoti-	
	cally Stable (GAS) neural network.	12
1.6	Schematic representation of the Olfactory Bulb (OB) anatomy. Activ-	
	ity of the mitral cells is denoted by X_i and granule cells by Y_i	14
1.7	An excitable system which mimics a neuron dynamics. The neuron	
	can be near threshold (a), hyperpolarized (b) or depolarized (c). $\ \ .$	15
1.8	Basic principles of the non-hyperbolic neural network functioning. The	
	input is given as a parameter which perturbs the non-hyperbolic equi-	
	librium. Local bifurcations of the equilibrium affect global dynamics	
	of the network	16

2.1	Synaptic connections between neurons having different natural fre-	
	quencies are functionally insignificant. Therefore, the network can be	
	divided into subnetworks (pools) of oscillators having equal or ε -close	
	natural frequencies	38
3.1	Possible intersections of nullclines of the relaxation neuron (3.1)	47
3.2	An excitable system. There are initial conditions for which the system	
	(3.1) generates an action potential, or spike (dotted line)	48
3.3	Intersections of nullclines which do not correspond to an excitable sys-	
	tem. a. A relaxation oscillator with non-zero amplitude. b. A relax-	
	ation oscillator with zero amplitude	49
4.1	Thalamo-cortical interactions	66
4.2	Bifurcations of a mapping. a. Saddle-node bifurcation for $c_i p_i < 0$. b.	
	Saddle-node bifurcation for $c_i p_i > 0$. c. Flip bifurcation	69
5.1	Phase portraits for various c	76
5.2	Saddle-node bifurcation on a limit cycle	77
5.3	An intersection of nullclines in a relaxation system which exhibits	
	saddle-node bifurcation on a limit cycle	77
5.4	Dynamic behavior for various c_i	81
5.5	Co-existence of a local attractor and a global limit cycle	84
7.1	Global flow structures of the canonical models. a. System (7.1) for	
	$\sigma_1 = \sigma_2 = -1$ is bounded. b. System (7.1) for $\sigma_1 = +1$, $\sigma_2 = -1$ is	
	not bounded	97

Phase portrait of (7.2) working in the extreme psychological regime. 7.299 Phase portrait of the canonical model (7.13) of weakly connected neural 7.3network near multiple pitchfork bifurcation point for different values of the bifurcation parameter b. a. $b < -\beta_1$. b. $-\beta_1 < b < -\beta_2$. c. $-\beta_2 < b < -\beta_2 + (\beta_1 - \beta_2)/2$. d. $-\beta_2 + (\beta_1 - \beta_2)/2 < b$ 107 Bifurcation diagram. 7.4110 Every equilibrium point $\pm \sqrt{b + \beta_k} \xi^k$ becomes an attractor after the 7.5sequence of the pitchfork bifurcations. Every time b crosses $-\beta_k + \beta_k$ $(\beta_s - \beta_k)/2, \ s < k$, the ξ^s -direction becomes stable. 110 Phase portrait of the canonical model on the stable invariant plane 7.6 $\operatorname{span}(\xi^1,\xi^2)$. The first image is presented as an input onto the network. a. Input is weak, i.e. $|a| < a^*$. b. For $|a| = a^*$ there are fold bifurcations. c. For $|a| > a^*$ the canonical model is globally asymptot-116 Phase portrait of the canonical model on the stable invariant plane 7.7 $\operatorname{span}(\xi^1,\xi^2)$. The input is a mixture of two images ξ^1 and ξ^2 . a. Overall input is weak. b. Strong input and weak contrast. There is a co-existence of two attractors. 118 Bistability of perception. 7.8 120 7.9 Bifurcation diagrams for quasi-static variation of parameter b. a. The contrast c = 0. b. The first image is dominant. c. The second image is dominant. e d. Stroboscopic presentation of the phenomenon for 123

9.1	Depending upon the input $(r^1 \text{ or } r^2)$ the center manifold is tangent to	
	the corresponding memorized vectors $(v^1 \text{ or } v^2)$	141
10.1	Schematic representation of the neural oscillator. It consists of exci-	
	tatory (white) and inhibitory (shaded) populations of neurons. For	
	simplicity only one neuron from each population is pictured. White	
	arrows denote excitatory synaptic connections, black arrows denote	
	inhibitory synaptic connections	145
10.2	The neural oscillators (dotted boxes) are connected into a network.	
	The mitral cell makes contacts with other mitral cells and may have	
	contacts with other granule cells	145
10.3	A network of two neural oscillators. Open boxes are local populations	
	of excitatory neurons, shaded circles are local populations of inhibitory	
	neurons. The real numbers a_1, a_2, a_3, a_4 are entries of the Jacobian	
	matrix of <i>i</i> -th neural oscillator. The real numbers s_1, s_2, s_3, s_4 denote	
	the strength of synaptic connections	149
10.4	Differences in dynamic behavior of type A and B neural oscillators.	
	See text	152
10.5	Complex numbers $v_1, -v_2, v_3, -v_4$ as vectors on the complex plane.	
	For simplicity we depict $-v_2$ with zero real part. a. Type A neural	
	oscillator. b. Type B neural oscillator	154
10.6	Synaptic configurations that can exhibit the phenomenon described in	
	Theorem 60. Open boxes depict excitatory neurons and shaded circles	
	depict inhibitory neurons. A vertical pair of excitatory and inhibitory	
	neurons is one neural oscillator	155

- 10.7 Possible values of synaptic connections c_{ij} for different synaptic configurations S_{ij} satisfying Dale's principle. For synaptic configurations that are not explicitly depicted the possible values of c_{ij} may occupy all shaded areas. a. Type A neural oscillator. b. Type B neural oscillator157
- 10.8 Open boxes are excitatory neurons, shaded circles are inhibitory neurons. If there is an arrow between two neurons, then the synaptic contact is possible.

a. The synaptic organizations that cannot memorize phase information.

b. The synaptic organization that can either learn or unlearn phase information (but not both). If the network has more than two oscillators, then the Dale's principle will be violated during the learning.c. The synaptic organization that can learn phase information.

- 10.9 c_{ij} and c_{ji} must be inside the shaded area between $-v_4$ and $-v_2$. . . 169

Chapter 1

Introduction

To study the brain is a challenge not only for neurobiologists but also for mathematicians. It is important to understand how it works. It might help us to build a new generation of computers — neurocomputers. We would understand the nature of many brain diseases and could find ways of curing them. Another noble goal of studying of brain models is possible minimization of wet lab experiments that sacrifice animals.

1.1 Overview

In this work we develop bifurcation theory for the Weakly Connected Neural Networks (**WCNNs**) of the form

$$X_i = F_i(X_i, \lambda) + \varepsilon G_i(X_1, \dots, X_n, \lambda, \varepsilon), \quad \varepsilon \ll 1,$$
(1.1)

where X_i is the activity of the *i*-th neuron and λ is a bifurcation parameter.

Historically, the weakness of connections (condition $\varepsilon \ll 1$) arises as a technical auxiliary assumption needed for studying networks of oscillators. Although the assumption has strong neurophysiological justification, its importance has eluded researches so far. For our work it has paramount importance since it allows us to study (1.1) without even knowing the functions F_i and G_i that describe the real brain dynamics.

Indeed, since little is known about chemical and electro-physiological processes taking place inside real neurons, we do not have detailed information about the dynamical system (1.1). Nevertheless, we would like to know its dynamic behavior. Especially, we are interested in the regimes where (1.1) has interesting neurocomputational properties, i.e. when it can serve as a prototype for a new generation of computers – neurocomputers. We discuss these issues here.

In Chapter 2 we consider WCNN (1.1) at an equilibrium point and study its local dynamics using bifurcation theory. First we show (Section 2.1) that the WCNN must be near a multiple bifurcation point in order to have non-trivial dynamics. This is necessary (but not sufficient) condition for the WCNN to be interesting from the neuro-computational point of view. In biological terms this corresponds to the case when neurons are near thresholds and are susceptible to external perturbations.

Then, using the Center Manifold Theorem (Section 2.3), we show that near multiple bifurcations the WCNN (1.1) can be significantly simplified and reduced to canonical models. In Chapter 2 we derive canonical models for multiple saddlenode, pitchfork and Andronov-Hopf bifurcations and multiple cusp singularities (see flowchart in Figure 1.1).

In Chapter 3 we analyze local dynamics of singularly perturbed weakly connected systems of the form

$$\begin{cases} \mu X'_i = F_i(X_i, Y_i, \lambda) + \varepsilon P_i(X, Y, \lambda, \varepsilon) \\ Y'_i = G_i(X_i, Y_i, \lambda) + \varepsilon Q_i(X, Y, \lambda, \varepsilon) \end{cases} \quad i = 1, \dots, n, \quad \mu \ll 1, \quad \varepsilon \ll 1$$
(1.2)

Such systems describe dynamics of networks of relaxation oscillators or excitable systems or any other systems which have several time scales. We show (Section 3.3) that in many cases (1.2) can be reduced to (1.1). The cases when this reduction is

impossible correspond to quasi-static bifurcations in (1.2). In Chapter 3 we derive canonical models for multiple quasi-static saddle-node, pitchfork and Andronov-Hopf bifurcations (see flowchart in Figure 1.1).

In Chapter 4 we study local dynamics of weakly connected mappings of the form

$$X_i \mapsto F(X_i, \lambda) + \varepsilon G_i(X, \lambda, \varepsilon), \quad \varepsilon \ll 1.$$

We derive canonical models for multiple saddle-node and flip bifurcations. Then we reveal the relationship between these canonical models and the canonical models for weakly connected systems of ordinary differential equations.

One of the principal achievements of this work is derivation of canonical models for the WCNNs. The canonical models have few non-linear terms, nevertheless, they capture the qualitative behavior of all WCNNs of the form (1.1) or (1.2), including possibly the real brain. Thus, in order to study the neuro-computational properties of the brain it is reasonable to study canonical models first.

In Chapter 5 we continue to analyze multiple saddle-node bifurcation. In particular, we are interested in how much the canonical models can tell us about the dynamics of the original WCNN (1.1).

We study canonical models for multiple Andronov-Hopf bifurcation in Chapters 6 and 10. In particular, we find conditions under which the model can operate as a gradient (Hopfield-type) neural network (Section 6.3).

Since the pitchfork bifurcation is a particular case of cusp singularity we study both canonical models in Chapter 7. We show that they can operate as classical neural networks. We concentrate our efforts on the case when the models are taught to recognize only two images (Section 7.6). In this case behavior of the models is somehow close to behavior of the human brain, especially when the two images are presented simultaneously. We illustrate a psychological phenomenon – bistability of

$$\begin{array}{cccc} D_{x_i}f_i & \underline{\qquad \mbox{hyperbolic}} & Linear uncoupled \\ \dot{x}_i = D_{x_i}f_ix_i \\ \hline & \mbox{non-hyperbolic} \\ \hline & \mbox{eigenvalues} & \underline{\mbox{a pair of pure imaginary}} & Canonical model for multiple \\ Andronov-Hopf bifurcation \\ & z'_i = a_i z_i + b_i z_i |z_i|^2 + \sum c_{ij} z_j \\ \hline & \mbox{one} \\ \hline & zero \\ \hline & \mbox{Adaptation condition} & \underline{NO} \\ & \mbox{Adaptation condition} & \underline{NO} \\ & \mbox{f}_i(0,\lambda) + \varepsilon g_i(0,\lambda,\rho,0) = \mathcal{O}(\overline{\varepsilon^2}) & \mbox{i a f}_i(x_i,\lambda) + \varepsilon g_i(0,\lambda,\rho,\varepsilon) \\ \hline & \mbox{VES} \\ & f''_{xx} & \underbrace{\neq 0} \\ & \mbox{Canonical model for multiple} \\ & \mbox{gathered bifurcation} \\ & \mbox{a'_i = r_i + b_i x_i + x_i^2 + \sum c_{ij} x_j \\ \hline & \mbox{But} \\ & \mbox{f}_{xxx}^{'i} \neq 0 \\ \hline & \mbox{Canonical model for multiple} \\ & \mbox{pitchfork bifurcation} \\ & \mbox{a'_i = r_i + b_i x_i \pm x_i^3 + \sum c_{ij} x_j \\ & \mbox{avel cusp singularity} \\ & \mbox{a'_i = r_i + b_i x_i \pm x_i^3 + \sum c_{ij} x_j \\ & \mbox{avel cusp singularity} \\ & \mbox{a'_i = r_i + b_i x_i \pm x_i^3 + \sum c_{ij} x_j \\ & \mbox{avel cusp singularity} \\ & \mbox{avel cusp singularity} \\ & \mbox{a'_i = r_i + b_i x_i \pm x_i^3 + \sum c_{ij} x_j \\ & \mbox{avel cusp singularity} \\ & \mbox{avel cusp singular$$

Figure 1.1: Summary of bifurcations and canonical models for regularly perturbed weakly connected systems

 $\dot{x}_i = f_i(x_i, \lambda) + \varepsilon g_i(x, \lambda, \rho, \varepsilon)$

$$\begin{cases} \mu X_{i}^{t} = F_{i}(X_{i}, Y_{i}, \lambda) + \varepsilon P_{i}(X, Y, \lambda, \mu, \varepsilon) \\ Y_{i}^{t} = G_{i}(X_{i}, Y_{i}, \lambda) + \varepsilon Q_{i}(X, Y, \lambda, \mu, \varepsilon) \end{cases}$$

$$D_{x_{i}}F_{i} \qquad \text{hyperbolic} \qquad y_{i}^{t} = g_{i}(y_{i}, \lambda, \mu) + \varepsilon q_{i}(x, \lambda, \mu, \varepsilon)$$

$$\downarrow \text{non-hyperbolic} \qquad \qquad \text{Canonical model for multiple quasi-static} \\ \text{andronov-Hopf bifurcation} \\ \begin{cases} z_{i}^{t} = (a_{i} + A_{i}v_{i})z_{i} + b_{i}z_{i}|z_{i}|^{2} + \sum_{j=1}^{n} c_{ij}z_{j} \\ v_{i}^{t} = d_{i}(R_{i} + S_{i}|z_{i}|^{2} + T_{i}v_{i}) \end{cases}$$

$$\downarrow \text{cone} \\ zero \\ \downarrow (\varepsilon) \qquad = \mathcal{O}(\varepsilon) \qquad F_{y}^{t} \cdot G_{x}^{t} \rightarrow 0 \\ \downarrow < 0 \qquad \qquad \text{Linear, hyperbolic, uncoupled} \\ \begin{cases} \dot{x}_{i} = F_{y}^{t}y_{i} \\ \dot{y}_{i} = \mu_{1}G_{x}^{t}x_{i}, \end{cases}$$

$$\downarrow = \mathcal{O}(\varepsilon^{2}) \qquad \text{Canonical model for multiple quasi-static} \\ \text{Andronov-Hopf bifurcation} \\ z_{i}^{t} = a_{i}z_{i} + b_{i}z_{i}|z_{i}|^{2} + \sum_{j=1}^{n} c_{ij}z_{j} \\ \downarrow < 0 \qquad \qquad \text{Canonical model for multiple} \\ \begin{cases} \dot{x}_{i} = -y_{i} + x_{i}^{2} + \sum_{j=1}^{n} c_{ij}x_{j} \\ y_{i}^{t} = a_{i}(x_{i} - r_{i}) \\ y_{i}^{t} = a_{i}(x_{i} - r_{i}) \end{cases}$$

Canonical model for multiple quasi-static pitchfork bifurcation

Canonical model for multiple quasi-static cusp singularity

$$\begin{cases} x'_i = -y_i \pm x_i^3 + \sum_{j=1}^n c_{ij} x_j & \text{unfolding} \\ y'_i = a_i x_i & \end{cases} \begin{cases} x'_i = -y_i \pm x_i^3 + \sum_{j=1}^n c_{ij} x_j \\ y'_i = a_i (x_i - r_i) \end{cases}$$

Figure 1.2: Summary of multiple quasi-static bifurcations and canonical models for singularly perturbed weakly connected systems

perception using these models (Section 7.7).

Canonical models for multiple quasi-static saddle-node and pitchfork bifurcations are studied in Chapter 8. There we study how some biological constraints (such as Dale's principle) affect dynamics of the networks.

Studying the canonical models suggests a new neural network type – non-hyperbolic neural networks, which we discuss in Chapter 9. This new type utilizes the fact that local dynamics near a non-hyperbolic equilibrium can determine global dynamics of a system. This justifies our restriction to consider only local dynamics of the WCNNs.

The main result for our WCNN Theory is in Chapter 10. In it we analyze the relationship between synaptic organizations (anatomy) and dynamical properties (function) of the brain. In particular, we show that there are some synaptic organizations that have especially rich dynamic behavior. Comparison of our findings with neurophisiological data shows that these organizations (anatomical structures) are ubiquitous in the brain. We consider this not as a lucky coincidence, but as a sign that the theory of weakly connected neural networks developed in this dissertation is promising for study of the human brain¹.

1.2 Models in Mathematical Biology

Unlike in physics, there is no expectation that mathematical modeling in neurobiology could give exact quantitative results. Indeed, neurophysiological processes do not have known conservation laws or symmetry. We do not know and probably will never know all the nature laws that govern brain activity. Thus, we will possibly never be able to write an equation or a system of equations that describes completely a brain's

6

¹Chapters 2 and 7 received SIAM Student Paper Prize in applied mathematics for 1995

dynamics. Nevertheless, it would be interesting and helpful to have a model that somehow reflects and simulates the brain's computational abilities.

Most models in neuroscience can be divided into the following groups:

- Ordinary Language Models are used by biologists to explain how the human brain or some of its structures might work. These models are not mathematical and, hence, are suitably imprecise.
- Comprehensive Models are the result of an attempt to take into account all known neurophysiological facts and data. Usually they are too cumbersome and are not amenable to mathematical analysis. A typical example is Hodgkin's and Huxley's (1954) model.
- Empirical Models occur when one tries to construct a simple model reflecting one or more important neurophysiological observations. A typical example is Hopfield's (1982) network of McCulloch-Pitts (1943) neurons, based on the fact that neurons could be bistable elements. Although amenable to mathematical analysis, these models are far from reality.
- Canonical Models arise when one studies a dynamical system in critical parameter regimes. A typical example of a critical regime is a bifurcation, also referred to as being a phase transition. The major advantage of canonical models is that they describe qualitatively all dynamical systems including the real brain. Their drawback is that they are useful only when the brain operates near the critical regime. Several examples of canonical models are presented in this dissertation.

The division above is artificial. There are no exact boundaries between the model types. For example, the Hodgkin-Huxley model could be classified as an empirical



Figure 1.3: Relationship between models in mathematical biology

model because it reflects our knowledge of membrane properties in 1950's, which turned out to be far from complete. The canonical models are another example: They might be considered as empirical models too because they can be analyzed without resort to computers and/or because they illustrate some basic neurophysiological facts such as bistability. Each of the model types has its own advantages and drawbacks. Neither of them is better or worse than the others. For further discussion see chart in Figure 1.3.

Most biologists use the ordinary language models since they are precise where data are known and appropriately vague otherwise. These models do not require knowledge of mathematics. Typical example of such a model is given in the next section as an introduction to the neurophysiology.

Using comprehensive models could be a trap: The more neurophysiological facts are taken into consideration during the construction of the model, the more sophisticated and immense the model becomes. As a result, the model can become useless since it cannot be reasonably analyzed even with help of a computer. Moreover, as it usually turns out later, the model is not 'comprehensive' at all, i.e. it is only an illusion that the model does include all essential neurophysiological information.

Most mathematicians and physicists who study brain functions use empirical

models. Usually one has an idea borrowed from an ordinary language model and tries to construct a simple formula that illustrates the idea. The simpler the model the better. As a result, one has a system of ordinary differential or difference equations that might have some computational properties but could be irrelevant to the brain. We believe that if we completely understand the brain we would be able to explain how it works using a simple empirical model. Since we are far away from understanding the brain, invention of empirical models capable of performing something useful is more an art rather than a science. Some successful models are discussed in the review article by S. Grossberg (1988).

There are few examples of using the canonical models since there are only two of them known so far. First is the Voltage Controlled Oscillator Neuron (VCON, see Hoppensteadt 1986) model, which can arise when one considers a weakly connected network of excitable oscillators (see e.g. Ermentrout and Kopell 1986). The other canonical model occurs when one considers a weakly connected network of oscillators near Andronov-Hopf bifurcation. We discuss these models as well as derive and analyze other canonical models in subsequent chapters.

1.3 Neurobiological Background

It is believed that basic functional unit of the brain is a neuron. Neurons can be different in shape, size and function, but it is possible to describe typical neuron attributes.

A typical neuron consists of a cell body, a dendritic arborization and a long axon (see Figure 1.4). Due to nonlinear properties of membranes, neurons can generate action potentials, or voltage spikes — the basic mechanism of communication between neurons and brain structures.



Figure 1.4: A "typical" neuron. It receives signals from other neurons via synapses and sends the output signal to the other neurons through its axon.

A simplified description of neuron activity is the following: Its dendrites receive input signals from other neurons via synapses — the contacts between neurons; after spatio-temporal processing, which takes place in the dendrites and the cell body, a neuron might generate a response, the action potential, that propagates along the axon to other neurons and serves as an input onto their dendrites. Neurons can have synaptic contacts with many thousands of other neurons. How the neurons communicate is still a mystery. The signal could be encoded as the number of spikes per unit of time or as inter-spike intervals. It is also feasible that the signal can be encoded as phase differences since most of the neurons tend to generate the action potentials repeatedly.

Using various chemicals it is possible to suppress nonlinear membrane properties responsible for generation of the action potentials. In this case one action potential of a neuron induces very small changes in the membrane potential of another neuron (i.e. the excitatory post-synaptic potential (EPSP) is less than 1 percent of the amplitude of the spike). This observation lies at the foundation of a theory of weakly connected neural networks. Nevertheless, in normal conditions even small

10

perturbations of membrane conductance may evoke substantial changes in membrane potential because neurons are usually very close to thresholds and can easily generate action potentials. Thus, we can characterize a neural network of the human brain as being a highly nonlinear and extremely sensitive system. We will see in the next chapter that these are desirable properties for a brain to have.

1.4 Neural Network Types

It is believed that our memories correspond to (meta-stable) attractors in our brain's phase space, which is huge since the human brain has more than 10¹¹ neurons (Shepherd 1983). Convergence to an appropriate attractor is called recognition. There are many Neural Network (NN) models that can mimic recognition by association processes.

In general, NN is a network of interconnected simple elements, usually called neurons, that performs computational tasks, such as recognition by association, memorization, etc., on a given input (key) pattern. The pattern may be temporal or spatial or both. In the NN types discussed here, the input pattern does not depend upon time, i.e. input is quasi-static.

Most of the NNs dealing with static input patterns can be divided into two groups according to the way in which the pattern is presented:

- MA -type (Multiple Attractor NN.) The key pattern is given as an initial state of the network. And the network converges to one of many possible choices (Hopfield 1982, Grossberg 1988).
- GAS -type (Globally Asymptotically Stable NN.) The key pattern is given as a parameter which controls the location and shape of a unique attractor (Hirsch 1989).



Figure 1.5: a. Multiple Attractor (MA) neural network. b. Globally Asymptotically Stable (GAS) neural network.

In either case the NN must converge to an equilibrium. Each of MA-type network attractors corresponds to a prior memorized image. Each input pattern, considered as an initial state, lies in a domain of attraction of some of the attractors (see Figure 1.5a). A "Good" MA-type NN is one that can memorize many images; It does not have spurious (or false) memory, i.e. the attractors that do not correspond to any of the previously memorized images; All its attractors have "large" attraction domains. A typical example of the MA-type NN is the Hopfield network (Hopfield 1982), although it is not a "good" network since it has spurious memory.

The GAS-type NN has only one attractor that depends upon the input pattern as a parameter (see Figure 1.5b). Various inputs can place the attractor on various locations in the network's phase space. Learning in such a NN consists of adjusting the connections between the neurons so that the network can realize the mapping

$$\left\{\begin{array}{c} \text{Patterns} \\ \text{to be memorized.} \end{array}\right\} \rightarrow \left\{\begin{array}{c} \text{Prescribed locations} \\ \text{of the attractor.} \end{array}\right\}$$

A "good" NN of this type should memorize many such mappings so that if the key pattern to be recognized is near (in some metric) to a previously memorized pattern, then the resulting attractor is in the prescribed location.

All attractors usually considered in these NNs are equilibrium points, although

there are many attempts to understand the role of limit cycles and chaotic attractors in brain functioning (Baird 1986, Eckhorn et.al 1988, Gray 1994, Hoppensteadt 1989, 1986, Izhikevich and Malinetskii 1992,1993, Kazanovich and Borisyuk 1994, Skarda and Freeman 1987, Tsuda 1992).

We will show in subsequent chapters that the canonical models that we derive can work as both MA and GAS-type NNs.

Notice that the process of recognition in the NN types presented above is essentially a global phenomenon. But the canonical models describe only local dynamics of WCNN near some equilibrium. Studying the canonical models suggests a new NN type, which we discuss below.

1.4.1 Olfactory Bulb

Let us consider the well studied brain system comprising the Olfactory Bulb (OB). The neurons of OB receive signals from olfactory receptors, process information and send it to other parts of the brain. Anatomically the OB consists of neural oscillators (see Figure 1.6). The processing of information by the OB has been studied neurophysiologically (see, for example, Shepherd 1976,1983 and Skarda and Freeman 1987) and using mathematical models (see, for example, Baird 1986, Erdi et.al 1993, Izhikevich and Malinetskii 1993 and Li and Hopfield 1989).

It is believed that in the absence of the signals from olfactory receptors the OB's activity is chaotic and of low amplitude. W. Freeman called this background chaotic activity the dormant state. The signals from receptors make the chaotic activity coherent and periodic. Each inhaled odor excites its own pattern of spatial oscillations. From a mathematical point of view, each odor is represented by a limit cycle in the OB's phase space. To study the limit cycles requires global information about the



Figure 1.6: Schematic representation of the Olfactory Bulb (OB) anatomy. Activity of the mitral cells is denoted by X_i and granule cells by Y_i .

OB dynamics, but to study how the chaotic attractor corresponding to dormant state changes requires only information about local behavior near the attractor.

We hypothesize that it is possible to predict to which limit cycle the OB's activity will be attracted simply by studying the local bifurcations of the attractor. Thus, one could say that the future of the OB dynamics is determined by local events. This is the spirit of our non-hyperbolic NNs approach.

Unfortunately, it is difficult to study chaotic attractors. In this thesis we assume that when input from receptors is absent, the attractor corresponding to the dormant state is an equilibrium point. In this case one can think of the low-amplitude chaos as being noisy perturbations of an equilibrium. The limitation of this approach is that we neglect possible role of deterministic chaos in information processing (Izhikevich and Malinetskii 1992,1993).

1.4.2 Networks of Excitable Elements

The example of bifurcations in the olfactory bulb discussed above is vague. Below we present a precise mathematical illustration of the non-hyperbolic phenomenon that we study.



Figure 1.7: An excitable system which mimics a neuron dynamics. The neuron can be near threshold (a), hyperpolarized (b) or depolarized (c).

Consider an excitable system with nullclines as depicted in Figure 1.7a. Suppose the input to this system shifts one of the nullclines, say G = 0, to the right (excitation) or to the left (inhibition). Dynamical systems with such attributes mimic some important properties of real neuron dynamics and are studied in Chapter 3.

When the input is inhibitory (see Figure 1.7b), the equilibrium point is asymptotically stable. If the initial state of the system is close to the equilibrium, the dynamics converge to it. Neurobiologists might say that such neuron is hyperpolarized and silent.

When the input is excitatory (see Figure 1.7c), the equilibrium becomes unstable. Neurobiologists say that the neuron is depolarized and can generate action potentials or spikes. The spike is a global phenomenon in the sense that it is observable on the macro-level. But it is caused by a local event — temporary loss of stability by the equilibrium. This is possible because the equilibrium (the intersection of the nullclines) in the excitatory systems (Figure 1.7a) is nearly non-hyperbolic. One can say that excitatory systems can perform a trivial pattern recognition task: They can discriminate between excitatory and inhibitory inputs.

Now consider a network of such excitatory elements (neurons). Suppose the input is inhibitory and strong. Then all neurons are hyperpolarized. Activity of the



Figure 1.8: Basic principles of the non-hyperbolic neural network functioning. The input is given as a parameter which perturbs the non-hyperbolic equilibrium. Local bifurcations of the equilibrium affect global dynamics of the network.

network converges to some equilibrium. The network is silent. Next suppose the input is shifted quasi-statically towards excitation until the global equilibrium loses its stability. The dynamics of the network might produce some macroscopic changes. For example, some of the neurons can generate action potentials whereas the others remain silent. The active neurons can send their signals to other parts of the brain and trigger various behavioral responses, such as attack or escape. This active response of the neural network depends on how the equilibrium loses its stability, which in turn depends on the input and the connections between the neurons. Thus, global behavior of the network crucially depends on local processes taking place when the equilibrium becomes non-hyperbolic. We can summarize the above as the following:

NH -type (Non-Hyperbolic NN.) The input pattern is given as a parameter which perturbs a non-hyperbolic equilibrium (see Figure 1.8).

To summarize, we can say that even local information about the brain's dynamic behavior can be useful for understanding its neuro-computational properties.

In subsequent chapters we derive canonical models for general weakly connected neural networks. The canonical models describe brain's local activity, so we analyze them not only from MA and GAS-type NN point of view, but also from NH-type NN point of view.

1.5 Generic Equations

Before writing equations let us introduce some useful definitions. Among them there are two that cannot be strictly defined (like notion of the *set* in mathematics).

By local population of neurons we mean a set of strongly interconnected neurons that are close to each other, have approximately the same pattern of synaptic connections and dynamic behavior. For instance, it could comprise those neurons in a cortical column, in an olfactory bulb glomerulus, etc.

We can think of *activity* of the population as the number of action potentials per unit time, or as an amount of chemical neurotransmitter released by synaptic terminals, or any other physiological observable. Another point of view is that the activity of the population of neurons is a variable from a (possibly infinite-dimensional) Banach space that describes all electro-physiological and neuro-chemical properties of the neurons: Spatial distributions of membrane potentials and neurotransmitters; activity of receptors, ion channels and pumps, etc.

Of course, these are not the precise definitions required for mathematical modeling, but they give us some flexibility in interpretation of results. Indeed, any unfolded and explicit definition of them means a deliberate restriction of the set of phenomena that could be described by our mathematical modeling. Finally, the following abuse of notation is widely accepted in the neural network literature: We will call a population of neurons simply a neuron and the activity of the population the activity of the neuron.

Let M be a manifold and let $X_i \in M$ denote the activity of the *i*-th neuron.

Suppose that the dynamics of each X_i can be described by the dynamical system

$$\dot{X}_i = F_i(X_i, \lambda), \quad \lambda \in \Lambda, \quad X_i \in M, \quad i = 1, \dots, n.$$
 (1.3)

where the Banach manifold Λ is a parameter space and n is the number of neurons. We will assume that the unknown functions F_i describing the real neuron activity are as smooth as it is necessary for our computations. Equations (1.3) considered as a system describe dynamics of uncoupled neural network because the dynamics of each neuron X_i depends only upon itself and does not depend upon the other neurons.

The basic assumption of the theory of weakly connected neural networks is that the contribution of activity of one neuron to activity of another one is very small, say, of order $\varepsilon \ll 1$.

Definition 1 A Weakly Connected Neural Network (WCNN) is a dynamical system of the form

$$\dot{X}_i = F_i(X_i, \lambda) + \varepsilon G_i(X_1, \dots, X_n, \lambda, \rho, \varepsilon), \qquad X_i \in M, \qquad i = 1, \dots, n,$$
(1.4)

where G_i describes synaptic connections between the neuron X_i and the other neurons X_1, \ldots, X_n and receptors $\rho \in \mathcal{R}$; \mathcal{R} is a Banach manifold and parameter $\varepsilon \ll 1$.

Note that the activity of *i*-th neuron X_i depends strongly only upon itself and the parameter λ . One can consider (1.4) as an ε -perturbation of the (uncoupled) system (1.3). We refer to (1.4) as being a regularly perturbed WCNN. For reasons that will become clear later, we might call (1.4) the equation that describes the *physiology* of WCNN and the state variable $X = (X_1, \ldots, X_n) \in M^n$ describes physiological state of the brain. In contrast, the parameter λ describes the *psychological* state of the brain. In the analysis below we assume that the network (1.4) works in a psychologically quasi-static regime, i.e. when λ is changed so slowly that it could be assumed to be constant. We will return to this issue in Section 2.6.1. Many processes in the human brain can be described by singularly perturbed, weakly connected dynamical systems of the form

$$\begin{cases} \mu X'_i = F_i(X_i, Y_i, \lambda) + \varepsilon P_i(X, Y, \lambda, \rho, \varepsilon) \\ Y'_i = G_i(X_i, Y_i, \lambda) + \varepsilon Q_i(X, Y, \lambda, \rho, \varepsilon) \end{cases} \quad i = 1, \dots, n.$$
(1.5)

Here each pair $(X_i, Y_i) \in M$ denotes activity of the *i*-th neuron. The small parameter μ denotes a ratio of time scales. Thus $\{X_i\}$ are "fast" and $\{Y_i\}$ are "slow" variables. Variables X_i and Y_i could also denote the activities of local populations of excitatory and inhibitory neurons, respectively. Then (1.5) could describe dynamics of weakly connected networks of relaxation neuron oscillators such as those depicted in Figure 1.6.

We study both dynamical systems here. Our strategy is the following: Suppose we know the behavior of each neuron and, hence, the behavior of the uncoupled systems. What new dynamic behavior can emerge as a result of introducing arbitrary small coupling ε between them?

The behavior of each neuron might be chaotic, and introducing coupling can destroy the chaos or make it more complicated. Even in the simpler case when the dynamics of each neuron is periodic, it is still possible to observe very complicated nonlinear phenomena such as chaos, phase locking or synchronization in the coupled system (Hoppensteadt 1993).

Below we study only the case when the dynamics of each neuron is convergent and the attractor is an equilibrium point. Hence, the dynamics of the network of uncoupled neurons is also convergent to a equilibrium point. We will study WCNN in some neighborhood of the equilibrium point and will show that under some conditions its dynamics is very rich.
Part I

Derivation of Canonical Models

Chapter 2

Weakly Connected Neural Networks

In this chapter we study local dynamics of a WCNN, described by the dynamical system

$$\dot{X}_i = F_i(X_i, \lambda) + \varepsilon G_i(X_1, \dots, X_n, \lambda, \rho, \varepsilon), \qquad i = 1, \dots, n,$$
(2.1)

near an equilibrium point. First we show that the dynamics is not interesting from neuro-computational point of view unless the equilibrium corresponds to a bifurcation point. In biological terms such neurons are said to be near a threshold. Using a Center Manifold reduction we prove the Fundamental Theorem of WCNN Theory, which says that neurons should be close to thresholds in order to participate in brain dynamics.

Then we show that a WCNN near a bifurcation point can be significantly simplified and reduced to a canonical model. We derive the canonical models for multiple saddlenode, pitchfork and Andronov-Hopf bifurcations and for multiple cusp singularities.

2.1 Hyperbolic Equilibrium

Since we are interested in local structure, we may assume that M is a Euclidean space \mathbb{R}^m for some m > 0. We may also assume that Λ and \mathcal{R} are some Banach spaces.

To avoid awkward formulas, we will use the following vector notation

$$X = (X_1, \dots, X_n)^{\mathsf{T}} \in M^n$$

$$F(X, \lambda) = (F_1(X_1, \lambda), \dots, F_n(X_n, \lambda))^{\mathsf{T}} : M^n \times \Lambda \to M^n$$

$$G(X, \lambda, \rho, \varepsilon) =$$

$$(G_1(X, \lambda, \rho, \varepsilon), \dots, G_n(X, \lambda, \rho, \varepsilon)) : M^n \times \Lambda \times \mathcal{R} \times \mathbb{R} \to M^n.$$

Then we can rewrite (2.1) more concisely as

$$\dot{X} = F(X,\lambda) + \varepsilon G(X,\lambda,\rho,\varepsilon).$$
(2.2)

Note that $F(X, \lambda)$ has diagonal structure. Without loss of generality we also may assume that when $\varepsilon = 0$ the equilibrium point is the origin X = 0 for $\lambda = 0$, i.e.

$$F(0,0)=0.$$

Let L_i be the Jacobian $m \times m$ -matrix of first partial derivatives of the function $F_i(X_i, 0)$ with respect to X_i at the origin

$$L_i = D_{X_i} F_i(0,0) = \left(\frac{\partial F_{ik}}{\partial X_{ij}}(0,0)\right)_{k,j=1,\dots,m}.$$

When L_i has no eigenvalues with zero real part, the equilibrium point is said to be *hyperbolic*.

It is easy to see that the Jacobian matrix $L = D_X F(0,0)$ of the right hand-side of (2.1) for $\varepsilon = 0$ has the form

$$L = \begin{pmatrix} L_1 & 0 & \cdots & 0 \\ 0 & L_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & L_n \end{pmatrix}.$$
 (2.3)

The main purpose of this section is to prove the following

Theorem 2 If the dynamics of each neuron is near a hyperbolic equilibrium, then the weakly connected network (2.2) of such neurons, the uncoupled network

$$\dot{X}_i = F_i(X_i, \lambda), \quad i = 1, \dots, n \tag{2.4}$$

and the linear system

$$\dot{X} = LX \tag{2.5}$$

have topologically equivalent local flow structures.

Hence, the local dynamics of (2.2) is the direct product of the single neuron dynamics and, therefore, the entire neural network is, in this sense, no more complex than a single neuron, which has locally linear dynamics. Therefore, a hyperbolic WCNN is not interesting as a brain model.

Proof. Since each L_i is hyperbolic, then so is L. In particular, it is nonsingular. The Implicit Function Theorem guarantees existence of a unique set of smooth function $\tilde{X} = h(\varepsilon, \lambda, \rho) : \mathbb{R} \times \Lambda \times \mathcal{R} \to \mathbb{R}^{mn}$ defined in a neighborhood of the origin such that $h(0, 0, \rho) = 0$ and

$$F(\tilde{X},\lambda) + \varepsilon G(\tilde{X},\rho,\varepsilon) = 0$$

for all sufficiently small ε , λ and bounded ρ . This means that near the equilibrium point X = 0 of the unperturbed (uncoupled) system (2.4) there is a unique equilibrium point \tilde{X} of the perturbed (coupled) system (2.2).

Note that we do not require that parameter $\rho \in \mathcal{R}$ be small. This is possible because we assume that $G(X, \lambda, \rho, \varepsilon)$ is a smooth function, which is multiplied by ε . Hence, for small ε it suffices to require only that ρ be bounded.

By continuity properties of the spectrum of linear operators no eigenvalues can cross the imaginary axis provided the parameters are sufficiently small. Hence, the equilibrium \tilde{X} is also hyperbolic with eigenspaces and invariant manifolds of the same dimensions as those of the unperturbed (uncoupled) system (2.4). In particular, if one of the equilibrium points is stable, then so is the other. Local topological equivalence of the flows of (2.4), (2.2) and the linear system (2.5) follows from Hartman-Grobman Theorem (Guckenheimer and Holmes 1983). \Box

Note that the above analysis is also valid when M is a Banach manifold. In that case, instead of Jacobian matrices we consider Fréchet derivatives, but along with hyperbolicity we must impose some technical conditions (for instance, the Fréchet derivatives must have bounded inverse), which are always met in finite dimensional case.

Corollary 3 The only equilibrium points that require further discussions are nonhyperbolic ones.

2.2 Non-Hyperbolic Equilibrium

We consider the unperturbed (uncoupled) system (2.4) and compare its behavior with the behavior of perturbed (coupled) system (2.2). We seek such changes that endow (2.2) with non-trivial neuro-computational abilities.

In the previous section it was shown that a necessary condition for a weakly connected neural network (2.2) to exhibit local nonlinear behavior is non-hyperbolicity. Next we consider (2.2) "near" a non-hyperbolic equilibrium, as explained later on.

Suppose that the origin X = 0 is a non-hyperbolic equilibrium point of (2.4) for $\lambda = 0$. This means that the Jacobian L has eigenvalues with zero real part. Due to the diagonal structure of L, this is possible only if one or more of the Jacobians L_i have eigenvalues with zero real part.

Depending upon the type and number of them there might be various cases:

- Some of the Jacobian matrices L_i have many eigenvalues with zero real part. Comprehensive analysis of this case is difficult and has not been done yet.
- Some of the Jacobian matrices have a pure imaginary pair of eigenvalues (i.e. the neurons are near an Andronov-Hopf bifurcation point). This case has attracted much attention in recent years due to its connection with synchronization phenomena (Ermentrout and Kopell 1990; Hoppensteadt and Izhikevich 1995; Kopell 1986). Comprehensive analysis of a system of two such neurons can be found in (Aronson et.al 1990). We study this case in Section 2.5.
- Some of the Jacobian matrices have only one simple zero eigenvalue, whereas the others have no eigenvalues with zero real part.

The last case appears to be the simplest one. Thus, the natural way to start studying weakly connected neural network dynamics is to explore this case. We concentrate our efforts on this case, although the theorem that we prove in the next section is applicable to any non-hyperbolic weakly connected system.

2.3 The Center Manifold Reduction

Without loss of generality we may reorder the system and so assume that only the first k equations in (2.1) have a non-hyperbolic Jacobian matrix L_i . We will study the most interesting case when all nonzero eigenvalues of L_i have negative real parts.

Let us represent the phase space $M \cong \mathbb{R}^m$ of *i*-th neuron (i = 1, ..., k) as a direct sum

$$M = E_i^c \oplus E_i^s,$$

where the center subspace E_i^c is spanned by the eigenvectors of L_i that correspond to eigenvalues with zero real parts and the stable subspace E_i^s is spanned by the other (generalized) eigenvectors. We will use the notation $E^c = E_1^c \times \cdots \times E_k^c$.

Theorem 4 Suppose that each of the first k Jacobian matrices

$$L_i = D_{X_i} F_i(0,0)$$

is non-hyperbolic. Then the WCNN (2.1) is locally governed by a dynamical system of the form

$$\dot{x}_i = f_i(x_i, \lambda) + \varepsilon g_i(x, \lambda, \rho, \varepsilon), \quad i = 1, \dots, k,$$
(2.6)

where $x_i \in E_i^c$ and

$$J_i = D_{x_i} f_i(0,0) = L_{i|_{E_i^c}}, \quad i = 1, \dots, k.$$

In particular, J_i have all eigenvalues with zero real parts.

More precisely, there is a function $Z : E^c \times \Lambda \times \mathcal{R} \times \mathbb{R} \to M^n$ such that any local solutions X(t) of (2.1) tend exponentially to

$$Z(x(t),\lambda,\rho,\varepsilon),$$

where $x(t) \in E^c$ is some solution of (2.6) determined by X.

Proof. Proof of the theorem uses the Center Manifold Theorem and the fact that the Center Manifold has a convenient weakly connected form. Our treatment of the Center Manifold Reduction is based on that of Iooss and Adelmeyer 1992, Section I.2.

Let $\pi_i^c: M \to E_i^c$ denote a projection operator such that

$$\ker \pi_i^c = E_i^s.$$

Let $x_i = \pi_i^c X_i \in E_i^c$. We use the notation $x = (x_1, \ldots, x_k)^{\mathsf{T}} \in E^c$. In order to apply the Center Manifold Theorem, we must consider the auxiliary system

$$\begin{cases} \dot{X}_i = F_i(X_i, \lambda) + \varepsilon G_i(X, \lambda, \rho, \varepsilon) \\ \dot{\lambda} = 0 \\ \dot{\rho} = 0 \\ \dot{\varepsilon} = 0 \end{cases} \qquad i = 1, \dots, n.$$

at the equilibrium point $(X, \lambda, \varepsilon) = (0, 0, 0)$. Its center subspace is

$$E^{c} \times \Lambda \times \mathcal{R} \times \mathbb{R} = \{(x, \lambda, \rho, \varepsilon) \mid x \in E^{c}; \lambda \in \Lambda; \rho \in \mathcal{R}; \varepsilon \in \mathbb{R}\}.$$

Applying the Center Manifold Theorem (Iooss and Adelmeyer 1992), we obtain the function

$$H = (H_1, \ldots, H_n) : E^c \times \Lambda \times \mathcal{R} \times \mathbb{R} \to E_1^s \times \cdots \times E_k^s \times M^{n-k}$$

with

$$H(0, 0, \rho, 0) = 0$$

and

$$D_x H(0,0,\rho,0) = 0, (2.7)$$

such that for λ and ε sufficiently small and for bounded $\rho \in \mathcal{R}$ the manifolds

$$\mathcal{M}(\lambda, \rho, \varepsilon) = \{x + H(x, \lambda, \rho, \varepsilon) \mid x \in E^c\}$$

are locally attractive and invariant with respect to (2.1). Furthermore, x on $\mathcal{M}(\lambda, \rho, \varepsilon)$ is governed by a dynamical system of the form

$$\dot{x}_i = \pi_i^c \left(F_i(x_i + H_i(x, \lambda, \rho, \varepsilon), \lambda) + \varepsilon G_i(x + H(x, \lambda, \rho, \varepsilon), \lambda, \rho, \varepsilon) \right)$$
(2.8)

for i = 1, ..., k. Since $\mathcal{M}(\lambda, \rho, \varepsilon)$ is locally attractive, all local solutions of (2.1) tend exponentially to those of (2.8). Therefore, the function Z mentioned in the theorem has the form

$$Z(x,\lambda,\rho,\varepsilon) = x + H(x,\lambda,\rho,\varepsilon)$$

for $x \in E^c$.

Let us show that H has a weakly connected form. For this, notice that $\mathcal{M}(\lambda, \rho, 0)$ is the Center Manifold for the *uncoupled* ($\varepsilon = 0$) system

$$X_i = F_i(X_i, \lambda)$$
 $i = 1, \ldots, n$

Hence, the function $H(x, \lambda, \rho, 0)$ has uncoupled form

$$H_i(x,\lambda,
ho,0)=\left\{egin{array}{cc} V_i(x_i,\lambda) & ext{for }i=1,\ldots,k,\ V_i(\lambda) & ext{for }i=k+1,\ldots,n, \end{array}
ight.$$

where V_i are some functions. Recall that we assume that all data are as smooth as it is necessary for our computations. Therefore, H depends smoothly on ε , and for $\varepsilon \neq 0$ we can rewrite it in the weakly connected form

$$H(x,\lambda,\rho,\varepsilon) = \begin{cases} V_i(x_i,\lambda) + \varepsilon W_i(x,\lambda,\rho,\varepsilon) & \text{for } i = 1,\ldots,k, \\ V_i(\lambda) + \varepsilon W_i(x,\lambda,\rho,\varepsilon) & \text{for } i = k+1,\ldots,n, \end{cases}$$

where W_i are some functions. The functions F_i in (2.8) can also be rewritten in the weakly connected form

$$F_i(x_i + V_i(x_i, \lambda) + \varepsilon W_i(x, \lambda, \rho, \varepsilon), \lambda) = F_i(x_i + V_i(x_i, \lambda), \lambda) + \varepsilon \hat{F}_i(x, \lambda, \rho, \varepsilon)$$

for some functions \hat{F}_i . If we denote for $i = 1, \ldots, k$

$$f_i(x_i,\lambda) = \pi_i^c F_i(x_i + V_i(x_i,\lambda),\lambda)$$

and

$$g_i(x,\lambda,\rho,\varepsilon) = \pi_i^c \left(\hat{F}_i(x,\lambda,\rho,\varepsilon) + G_i(x+H(x,\lambda,\rho,\varepsilon),\lambda,\rho,\varepsilon) \right)$$

then equations in (2.8) can be written as (2.6).

Now note that

$$J_{i} = D_{x_{i}} f_{i}(0,0) = D_{x_{i}} \pi_{i}^{c} F_{i}(x_{i} + V_{i}(x_{i},0),0)|_{x_{i}=0}$$

= $\pi_{i}^{c} D_{x_{i}} F_{i}(0,0) \cdot (E + D_{x_{i}} V_{i}(0,0)),$

where E is the unit matrix. From (2.7) it follows that $D_{x_i}V_i(0,0) = 0$. Recall that $D_{x_i}F_i(0,0) = L_i$. Hence

$$J_i = \pi_i^c L_i = L_i \pi_i^c = L_{i|_{E^c}}.$$

Remark 5 The theorem can be restated concisely as follows: The center manifold for a weakly connected system has weakly connected form.

Remark 6 A generalization of the theorem for the case when the Jacobian matrices may have eigenvalues with positive real parts is straightforward. The center manifold is not attractive in this case.

Most important cases when the Jacobian matrix L_i is non-hyperbolic correspond to bifurcation of the *i*-th neuron dynamics. Activity of such a neuron is sensitive to external influences, and biologists say that such a neuron is near threshold. The fact that only the first k neurons participate non-trivially in (2.6) motivates the following result:

Corollary 7 (The Fundamental Theorem of Weakly Connected Neural Network Theory.) In order to make a non-trivial contribution to brain dynamics, a neuron must be near threshold.

The Fundamental Theorem is not totally unexpected for neurobiologists. It corroborates their belief that in the human brain, which is undoubtly an extremely complex system, all neurons must be near thresholds or, otherwise, the brain could not be able to cope with its tasks.

2.4 Canonical Models

In this section we consider a weakly connected network of non-hyperbolic neurons described by a dynamical system of the form (2.6)

$$\dot{x}_i = f_i(x_i, \lambda) + \varepsilon g_i(x, \lambda, \rho, \varepsilon), \quad i = 1, \dots, n,$$

where $x_i \in \mathbb{R}$ is a scalar and

$$\frac{\partial}{\partial x_i} f_i(0,0) = 0 \tag{2.9}$$

for all i.

We study (2.6) for $\rho = \rho(\varepsilon)$ and $\lambda = \lambda(\varepsilon)$ such that $\rho(0) = \rho_0$ and

$$\lambda(\varepsilon) = 0 + \varepsilon \lambda_1 + \varepsilon^2 \lambda_2 + \mathcal{O}(\varepsilon^3), \quad \lambda_1, \lambda_2 \in \Lambda,$$
(2.10)

i.e. $\lambda(\varepsilon)$ is ε -close to the bifurcation value $\lambda = 0$. This is not a restriction on ρ and λ since we allow ρ_0, λ_1 and λ_2 to assume any values.

Definition 8 We say that λ satisfies the *adaptation condition* if

$$D_{\lambda}f_i(0,0)\lambda_1 + g_i(0,0,\rho_0,0) = 0$$
(2.11)

for all i.

The adaptation condition (2.11) is equivalent to the conditions

$$f_i(0,\lambda) + \varepsilon g_i(0,0,
ho,0) = \mathcal{O}(\varepsilon^2), \quad i = 1,\ldots,n,$$

which can be explained in ordinary language as: The internal parameter λ counterbalances (up to order ε) the steady state input from the entire network onto each neuron. Notice that the representation (2.10) is not necessary for the adaptation condition to be satisfied. Thus $\lambda(\varepsilon)$ can be any function (differentiable in ε), although it is easier to work with smooth functions of the form (2.10).

The adaptation condition can also be restated in terms of the original weakly connected system (2.1). In this case it suffices to demand that λ counterbalances the input from the network along with E^c direction.

2.4.1 Multiple Saddle-Node Bifurcation

A dynamical system

$$\dot{x}_i = f_i(x_i, \lambda), \quad x_i \in \mathbb{R}$$

is near saddle-node bifurcation point $x_i = 0$ for $\lambda = 0$ if

$$p_i = \frac{\frac{\partial}{\partial x_i} f_i(0,0) = 0}{\frac{\partial^2}{\partial x_i^2} f_i(0,0) \neq 0}$$
(2.12)

and

$$D_{\lambda}f_i(0,0) \neq 0,$$
 (2.13)

although in the theorem below we do not require condition (2.13).

If all equations in the uncoupled system (1.3) are near bifurcation points of the same type, then the bifurcation is called *multiple*.

Theorem 9 If the WCNN (2.6) is near a multiple saddle-node bifurcation and for the external input

$$\rho(\varepsilon) = \rho_0 + \varepsilon \rho_1 + \mathcal{O}(\varepsilon^2), \quad \rho_0, \rho_1 \in \mathcal{R}$$
(2.14)

the internal parameter λ satisfies the adaptation condition (2.11), then the invertible change of variables

$$x_i = \varepsilon p_i^{-1} \tilde{x}_i \tag{2.15}$$

and introduction of the "slow" time $\tau = \varepsilon t$ transforms (2.6) to

$$x'_{i} = r_{i} + b_{i}x_{i} + x_{i}^{2} + \sum_{j \neq i}^{n} c_{ij}x_{j} + \mathcal{O}(\varepsilon), \quad i = 1, \dots, n,$$
(2.16)

where $' = d/d\tau$ and we have erased ~. One can think of (2.16) as being the canonical model for multiple saddle-node bifurcations in WCNNs.

Proof. Using (2.10) and (2.14) we can write the initial part of the Taylor series of (2.6) in the form

$$\dot{x}_i = \varepsilon a_i + \varepsilon^2 d_i + \varepsilon \sum_{j=1}^n s_{ij} x_j + p_i x_i^2 + \text{h.o.t.}, \qquad (2.17)$$

where

$$a_{i} = D_{\lambda}f_{i} \cdot \lambda_{1} + g_{i} \in \mathbb{R}$$

$$d_{i} = D_{\lambda}f_{i} \cdot \lambda_{2} + (D_{\lambda}^{2}f_{i}) \cdot (\lambda_{1}, \lambda_{1}) + D_{\lambda}g_{i} \cdot \lambda_{1} + D_{\rho}g_{i} \cdot \rho_{1} + \frac{\partial}{\partial\varepsilon}g_{i},$$

$$s_{ij} = \frac{\partial g_{i}}{\partial x_{j}}$$

for $i \neq j$, and

$$s_{ii} = D_{\lambda} \frac{\partial f_i}{\partial x_i} \cdot \lambda_1 + \frac{\partial g_i}{\partial x_i},$$

where all derivatives are evaluated at $(x, \lambda, \rho, \varepsilon) = (0, 0, \rho_0, 0)$. The constants $p_i \neq 0$ were defined in (2.12).

Notice that the adaptation condition implies $a_i = 0$. If we denote

$$r_i = p_i d_i,$$

$$c_{ij} = p_i s_{ij} p_j^{-1},$$

$$b_i = c_{ii}$$

and use rescaling (2.15), then (2.17) transforms into (2.16). \Box

2.4.2 Multiple Cusp Singularity

A dynamical system

$$\dot{x}_i = f_i(x_i, \lambda), \quad x_i \in \mathbb{R}$$

is a near cusp singularity $x_i = 0$ for $\lambda = 0$ if

$$\frac{\partial}{\partial x_i} f_i(0,0) = \frac{\partial^2}{\partial x_i^2} f_i(0,0) = 0,$$

$$q_i = \frac{\partial^3}{\partial x_i^3} f_i(0,0) \neq 0$$
(2.18)

Theorem 10 If the WCNN (2.6) is near a multiple cusp singularity, if the external input is

$$\rho(\varepsilon) = \rho_0 + \sqrt{\varepsilon}\rho_{\frac{1}{2}} + \mathcal{O}(\varepsilon), \quad \rho_0, \rho_{\frac{1}{2}} \in \mathcal{R}$$
(2.19)

and if the internal parameter λ satisfies the adaptation condition (2.11), then the invertible change of variables

$$x_i = \sqrt{\varepsilon} |q_i|^{-\frac{1}{2}} \tilde{x}_i \tag{2.20}$$

and introduction of the "slow" time $\tau = \varepsilon t$ transforms (2.6) to

$$x'_{i} = r_{i} + b_{i}x_{i} + \sigma_{i}x_{i}^{3} + \sum_{j \neq i}^{n} c_{ij}x_{j} + \mathcal{O}(\sqrt{\varepsilon}), \quad \sigma_{i} = \pm 1, \quad i = 1, \dots, n, \quad (2.21)$$

where $' = d/d\tau$ and we have erased ~. One can think of (2.21) as being the canonical model for multiple cusp singularities in WCNNs.

Proof. The initial portion of the Taylor series of (2.6) has the form

$$\dot{x}_i = \varepsilon \sqrt{\varepsilon} d_i + \varepsilon \sum_{j=1}^n s_{ij} x_j + q_i x_i^3 + \text{h.o.t.}, \qquad (2.22)$$

where s_{ij} and q_i were defined earlier, and

$$d_i = D_
ho g_i(0,0,
ho_0,0) \cdot
ho_{rac{1}{2}}$$

If we denote

$$c_{ij} = |q_i|^{\frac{1}{2}} s_{ij} |q_i|^{-\frac{1}{2}}, b_i = c_{ii}, r_i = |q_i|^{\frac{1}{2}} d_i, \sigma_i = \text{sign } q_i$$

and rescale using (2.20), then (2.22) transforms into (2.21). \Box

2.4.3 Multiple Pitchfork Bifurcation

In Theorem 10 we demanded that the deviation of $\rho(\varepsilon)$ from ρ_0 be of order $\sqrt{\varepsilon}$ (Formula (2.19)). If we require that $\rho(\varepsilon)$ have the form as in (2.14), i.e. $\rho_{\frac{1}{2}} = 0$, then $r_i = 0$ in (2.21), and (2.21) has the form

$$x'_{i} = b_{i}x_{i} + \sigma_{i}x_{i}^{3} + \sum_{j \neq i}^{n} c_{ij}x_{j} + \mathcal{O}(\sqrt{\varepsilon}), \quad \sigma_{i} = \pm 1, \quad i = 1, \dots, n.$$
(2.23)

One can think of (2.23) as being the canonical model for a multiple *pitchfork* bifurcation in WCNNs because (2.23) is exactly what one receives if one considers (2.6) with functions f_i and g_i having Z_2 symmetry, i.e. invariant under the reflection $x_i \rightarrow -x_i$. In this case the adaptation condition (2.11) is satisfied automatically.

In canonical models (2.21) and (2.23) the choice $\sigma_i = +1$ ($\sigma_i = -1$) corresponds to a subcritical (respectively supercritical) bifurcations in the *i*-th neuron dynamics.

2.5 Multiple Andronov-Hopf Bifurcation

The assumption that connections between neurons are weak arose first as an auxiliary assumption for networks of oscillators near Andronov-Hopf bifurcations. In our work it is a primary assumption and networks of weakly connected oscillators are just one of many possible scenarious to be studied. We start discussion of this case below, and we continue it in Chapters 6 and 10. A dynamical system

$$\dot{x}_i = f_i(x_i, \lambda)$$

is near Andronov-Hopf bifurcation if the Jacobian matrix

$$L_i = D_{x_i} f_i(0,0)$$

has a simple pair of pure imaginary eigenvalues. Using the Center Manifold Reduction for WCNN (Theorem 4) we may assume without loss of generality that L_i is a 2×2 matrix.

Let $\pm i\Omega_i$ be the eigenvalues of L_i , where

$$\Omega_i = \sqrt{\det L_i}.$$

Let $v_i \in \mathbb{C}^2$ and $\bar{v}_i \in \mathbb{C}^2$ denote the (column) eigenvectors of L_i corresponding to eigenvalues $i\Omega_i$ and $-i\Omega_i$, respectively. Let w_i and \bar{w}_i be dual (row) vectors to v_i and \bar{v}_i . Let V_i be the matrix whose columns are v_i and \bar{v}_i , i.e.

$$V_i = (v_i, \bar{v}_i).$$

Notice that V_i^{-1} has w_i and \bar{w}_i as its rows.

Theorem 11 If the WCNN (2.1) is near a multiple Andronov-Hopf bifurcation, then there is an invertible change of variables

$$x_{i}(t) = \sqrt{\varepsilon} V_{i} \left(\begin{array}{c} e^{i\Omega_{i}t} z_{i}(\tau) \\ e^{-i\Omega_{i}t} \bar{z}_{i}(\tau) \end{array} \right) + \mathcal{O}(\varepsilon), \qquad (2.24)$$

where $\tau = \varepsilon t$ is a 'slow' time, which transforms the WCNN to

$$z'_{i} = b_{i}z_{i} + d_{i}z_{i}|z_{i}|^{2} + \sum_{\Omega_{i}=\Omega_{j}} c_{ij}z_{j} + \mathcal{O}(\sqrt{\varepsilon}), \qquad (2.25)$$

where $' = d/d\tau$, $b_i, d_i, z_i \in \mathbb{C}$, and the (synaptic) coefficients $c_{ij} \in \mathbb{C}$ are given by

$$c_{ij} = w_j \cdot D_{x_j} g_i \cdot v_i, \qquad (2.26)$$

System (2.25) is the canonical model for a WCNN near multiple a Andronov-Hopf bifurcation.

Proof. Consider the uncoupled ($\varepsilon = 0$) system

$$\dot{x}_i = f_i(x_i, 0).$$

The change of variables

$$x_i = V_i \left(\begin{array}{c} z_i \\ \bar{z}_i \end{array}\right) \tag{2.27}$$

transforms the system to

$$\dot{z}_i = \mathrm{i}\Omega_i z_i + h(z_i, \bar{z}_i), \qquad (2.28)$$

where h accounts for all nonlinear terms in z_i and \bar{z}_i . Next, we use the well-known fact from Normal Form Theory (Arnold 1984, Guckenheimer and Holmes 1983) that there is a near identity change of variables

$$\tilde{z}_i = z_i + p_i(z_i, \bar{z}_i), \qquad (2.29)$$

which transforms (2.28) to

$$\dot{z}_i = \mathrm{i}\Omega_i z_i + d_i z_i |z_i|^2 + \mathcal{O}(|z_i|^5)$$

for some $d_i \in \mathbb{C}$, where we have erased \sim .

Now suppose $\varepsilon \neq 0$ and (2.10) holds. Then we apply the composition of (2.27) and (2.29) to (2.6) and receive

$$\dot{z}_i = \mathrm{i}\Omega_i z_i + d_i z_i |z_i|^2 + \varepsilon \sum_{j=1}^n (c_{ij} z_j + e_{ij} \bar{z}_j) + \mathcal{O}(|z_i|^5, \varepsilon |z|^2, \varepsilon^2 |z|).$$

Introducing the 'slow' time $\tau = \varepsilon t$ and changing variables

$$z_i(t) = \sqrt{\varepsilon} e^{i\frac{\Omega_i}{\varepsilon}\tau} \tilde{z}_i(\tau)$$
(2.30)

transforms the system to

$$z_i' = d_i z_i |z_i|^2 + \sum_{j=1}^n \left(e^{i\frac{\Omega_j - \Omega_i}{\epsilon}\tau} c_{ij} z_j + e^{i\frac{-\Omega_j - \Omega_i}{\epsilon}\tau} e_{ij} \bar{z}_j \right) + \mathcal{O}(\sqrt{\epsilon}), \qquad (2.31)$$

where we have erased ~. After averaging all terms that have the factor $e^{i\frac{\delta}{\epsilon}\tau}$ for $\delta \neq 0$ vanish, and we obtain (2.25), where $b_i = c_{ii}$.

It is easy to see that (2.24) is a composition of (2.27), (2.29) and (2.30). \Box

Alternative proofs of this theorem can be found, for example, in Ermentrout and Kopell (1992) or in Hoppensteadt and Izhikevich (1995).

Notice the remarkable resemblance of (2.23) to (2.25). The latter has the same form as the former except that all variables are complex.

2.5.1 Equality of Frequencies and Attention

It is customary to call Ω_i the natural frequency of the *i*-th neuron. Since

$$\lambda(\varepsilon) = 0 + \varepsilon \lambda_1 + \mathcal{O}(\varepsilon^2)$$

each oscillator

$$\dot{x}_i = f_i(x_i,\lambda(arepsilon)) + arepsilon g_i(0,..x_i,..,0,\lambda(arepsilon),
ho,arepsilon)$$

for $\varepsilon \neq 0$ has natural frequency

$$\omega_i^{\star}(\varepsilon) = \Omega_i + \varepsilon \omega_i + \mathcal{O}(\varepsilon^2),$$

which is ε -close to Ω_i . Therefore, when we say that a pair of oscillators have equal frequencies, we actually mean ε -close frequencies.

One direct consequence of Theorem 11 is the following

Corollary 12 All neural oscillators can be divided into groups, or pools, according to their natural frequencies. Oscillators from different pools have different natural frequencies and interactions between them are negligible (see Figure 2.1).



Figure 2.1: Synaptic connections between neurons having different natural frequencies are functionally insignificant. Therefore, the network can be divided into subnetworks (pools) of oscillators having equal or ε -close natural frequencies

Proof. From (2.25) it follows that the *i*-th neuron dynamics depends on the *j*-th neuron dynamics only if $\Omega_i = \Omega_j$. Indeed, if $\Omega_j - \Omega_i = \omega_j^*(0) - \omega_i^*(0) \neq 0$, then the term $c_{ij}z_j$ in (2.31) vanishes after the averaging. \Box

Thus, oscillators from different pools work independently from each other even when they have nonzero synaptic contacts c_{ij} , i.e. one neuron can "feel" another one only when they have equal natural frequencies. It is reasonable to speculate that the brain has a mechanism to regulate the natural frequencies Ω_i of its neurons so that some of them can be entrained into different pools at different times simply by adjusting Ω_i . This might be related to such phenomena as attention and dominanta (Hoppensteadt 1991, Krukov 1991, Kazanovich and Borisyuk 1994).

Remark 13 There are interactions between neurons even when $\Omega_i \neq \Omega_j$, but these interactions have smaller order and are hidden in the term $\mathcal{O}(\sqrt{\varepsilon})$ in (2.25). They are noticeable only on time scales of order $\mathcal{O}(1/\sqrt{\varepsilon})$ (for "slow" time τ , or $\mathcal{O}(1/\varepsilon\sqrt{\varepsilon})$ for normal time t) and are negligible on shorter time scales.

For example, if all oscillators have different natural frequencies, then we must

38

study weakly connected dynamical systems of the form

$$z'_{i} = b_{i} z_{i} + d_{i} z_{i} |z_{i}|^{2} + \sqrt{\varepsilon} p_{i}(z_{1}, \dots, z_{n}, \sqrt{\varepsilon}, \tau), \quad z_{i} \in \mathbb{C}$$

$$(2.32)$$

for some functions p_i . If all Re $b_i < 0$, then the uncoupled ($\varepsilon = 0$) system has an asymptotically stable hyperbolic equilibrium $z_1 = \cdots = z_n = 0$ and dynamics of coupled ($\varepsilon \neq 0$) system is trivial. When Re $b_i > 0$ for at least two different *i*'s and corresponding Re $d_i < 0$, then (2.32) is a weakly connected network of limit-cycle oscillators. Such networks are studied elsewhere.

Suppose the whole network is divided into two or more subnetworks (pools) of oscillators each having equal frequencies. Then it is natural to study dynamics of one such pool first, and then study how these pools interact. In order to study the first problem we may assume that $\Omega_1 = \cdots = \Omega_n = \Omega$, i.e. the whole network is one such pool. We use this assumption in Chapters 6 and 10.

In order to study the interactions between the pools we have to consider a dynamical system of the form

$$X'_i = F_i(X_i) + \sqrt{\varepsilon}G_i(X_1, \dots, X_k, \sqrt{\varepsilon}, \tau), \quad i = 1, \dots, k,$$

where $X_i = (z_{i_1}, \ldots, z_{i_{m_i}})$ describes activity of the *i*-th pool and *k* is the number of pools. Such a system is weakly connected, it coincides with 2.1), and, hence, can be studied by the bifurcation methods developed in this work.

2.6 Discussion

Consider the canonical models. In the neural network literature it is customary to call the matrix $C = (c_{ij})$ the synaptic (or connection) matrix. It is believed that Cdescribes memorization of information by the brain. We will show that the same is true for the canonical models: They have interesting neuro-computational properties. In particular, they can perform pattern recognition and recall tasks by association a basic property of the human brain.

Parameters b_i are called bifurcation parameters because they depend on λ . Each parameter r_i depends on λ and ρ and has the meaning of rescaled external input on the *i*-th neuron.

It is convenient to think of $\rho_0 \in \mathcal{R}$ in the expression

$$\rho(\varepsilon) = \rho_0 + \varepsilon \rho_1 + \mathcal{O}(\varepsilon^2)$$

as a parameter representing an environment in which the WCNN (2.1) does something, whereas $\rho_1 \in \mathcal{R}$ denotes an input pattern from this environment. For example, ρ_0 may parameterize a degree of illumination and ρ_1 represents the shape of an object to be recognized.

Notice that the synaptic coefficients c_{ij} computed in the proof of Theorem 4 depend on the environment ρ_0 , but not on the input ρ_1 .

We can also ascribe meaning to each term in

$$\lambda(\varepsilon) = 0 + \varepsilon \lambda_1 + \mathcal{O}(\varepsilon^2), \qquad \lambda_1 \in \Lambda.$$

The first term (which we postulated to be 0, but it could have been any $\lambda_0 \in \Lambda$) means that (2.1) is near a non-hyperbolic equilibrium point, which corresponds to one of the multiple bifurcations considered above. We showed that this was a necessary condition for a WCNN to exhibit any non-linear behavior. The second term (which is $\varepsilon \lambda_1$) is of order ε . This explains the notion "near". So, the "near" means to be ε -close to the bifurcation value $\lambda_0 = 0$. The fact that coefficient $\lambda_1 \in \Lambda$ satisfies (2.11) means that the network (2.1) is adapted to the particular environment $\rho_0 \in \mathcal{R}$ in which it works.

2.6.1 Adaptation Condition and Psychology

Suppose the adaptation condition (2.11) is violated. Consider, for example, the multiple saddle-node bifurcation (the other multiple bifurcations and singularities may be considered similarly). The rescaling

$$\begin{aligned} x_i &\to \varepsilon^{\frac{1}{2}} x_i \\ t &\to \varepsilon^{-\frac{1}{2}} t \end{aligned}$$

transforms (2.17) to

$$\dot{x}_i = a_i + p_i x_i^2 + \mathcal{O}(\sqrt{\varepsilon}), \quad i = 1, \dots, n$$
(2.33)

with $a_i \neq 0$ and $p_i \neq 0$.

The behavior of this system is predetermined by the environment ρ_0 and the internal parameter λ (a_i depends upon them). The dynamics of (2.33) is hyperbolic and, hence, locally linear. System (2.33) does not have memory and cannot perform the recognition task because it is insensitive to the input pattern ρ_1 . It can react only to ρ_0 and its reaction is locally trivial.

We see that the adaptation condition (2.11) is important. Without it, recognition is impossible. That is exactly what we expected from our experience. Indeed, when we come into a dark room from bright sunshine, we cannot see anything until we adapt to the new environment. The human brain is an extremely flexible system. It has many subsystems with different time scales that help it to adapt to any new environment. Their functions are not known completely yet. The fact that we did not make any assumptions about the Banach spaces \mathcal{R} and Λ gives us freedom in interpretations of our results. For example, we do not have to specify where exactly the adaptation takes place – in the cortex, in the retina, or in the pupil of the eye. It does take place. And there is a mechanism responsible for that. Let us formalize this observation. Recall that we called $X \in M^n$ and $\lambda \in \Lambda$ physiological and psychological variables, respectively. Thus, (2.1) describes physiology of the brain in contrast to the dynamical system

$$\dot{\lambda} = H(\lambda, X, \rho_0), \qquad \lambda \in \Lambda, \qquad X \in M^n, \qquad \rho_0 \in \mathcal{R}, \tag{2.34}$$

that describes *psychology* of the brain.

The division into physiology and psychology is convenient. In (2.1) we assumed that λ and ρ are constants. Hence, the physiological dynamics is a mere response of the brain to the input ρ from receptors. The response depends upon the internal state λ of the brain as a parameter. For different choices of λ the response could be (and frequently is) different, because dynamics of the canonical models depends upon λ , for example, through constants b_i and r_i .

What do we know about (2.34)? First of all, the characteristic time of (2.34) must be much slower than that of (2.1), so that we may assume λ to be constant in (2.1). We showed that in order to perform any interesting task, the system (2.1) must be near a non-hyperbolic equilibrium point, i.e. λ must be in an ε -neighborhood of the origin. After rescaling $\lambda \to \varepsilon \lambda$, $H \to \varepsilon H$, we may assume that $\lambda = \mathcal{O}(1)$ in (2.34), i.e. we identified it with λ_1 from (2.10).

We know how important for (2.1) the adaptation condition (2.11) is. Hence, it is natural to assume that in the (possibly infinite-dimensional) Banach space Λ the linear manifold

$$\Lambda(\rho_0) = \{ \lambda_1 \in \Lambda \mid D_{\lambda} f_i(0,0) \cdot \lambda_1 + g_i(0,0,\rho_0,0) = 0 \}$$

is a global attracting set.

So, λ -dynamics governed by (2.34) can be divided on two parts: Changes approaching to $\tilde{\Lambda}(\rho_0)$, which might be called *adaptation* dynamics, and dynamics along

 $ilde{\Lambda}(
ho_0)$

While λ is approaching $\tilde{\Lambda}(\rho_0)$ any recognition is impossible, and the neural network (2.1) can react only to the environment ρ_0 . After the adaptation is completed and λ is in an ε -neighborhood of $\tilde{\Lambda}(\rho_0)$, the adaptation condition (2.11) is satisfied and the neural network is able to perform a pattern recognition task. The internal parameter λ continues to stay ε -close to $\tilde{\Lambda}(\rho_0)$. By moving along $\tilde{\Lambda}(\rho_0)$, λ cannot destroy the ability of the neural network (2.1) to recognize something, but it can affect how (2.1) does it. The internal parameter λ has great impact on the way the brain perceives the world. The WCNN (2.1) together with (2.34) could have features that psychologists might call attention, emotions, feelings, etc. To the best of our knowledge, the WCNN (2.1) together with (2.34) has not been studied yet.

Chapter 3 Singularly Perturbed WCNNs

3.1 Basic Definitions

Following are some definitions that are used here.

Definition 14 A Relaxation Neuron (\mathbf{RN}) is a singularly perturbed dynamical system of the form

$$\begin{cases} \mu X' = F(X, Y, \lambda) \\ Y' = G(X, Y, \lambda), \end{cases}$$
(3.1)

where $' = d/d\tau$ and $0 < \mu \ll 1$ is a small (dimensionless) parameter reflecting a ratio of time scales; the vector $X \in \mathbb{R}^k$ denotes fast and $Y \in \mathbb{R}^m$ (relatively) slow variables; $\lambda \in \Lambda$ is a parameter. We assume that the functions $F : \mathbb{R}^k \times \mathbb{R}^m \times \Lambda \to \mathbb{R}^k$ and $G : \mathbb{R}^k \times \mathbb{R}^m \times \Lambda \to \mathbb{R}^m$ are as smooth as it is necessary for our computations.

Example 1 If X denotes activity of "fast" ion channels (e.g. Na⁺, Ca⁺⁺, etc) and Y denotes activity of "slow" ion channels (e.g. K⁺) in a model neuron, then (3.1) could describe a mechanism of generation of action potentials by a neuron. The typical examples of such dynamical systems are Hodgkin-Huxley equations (Hodgkin and Huxley 1954) and Fitzhugh-Nagumo equations (Fitzhugh 1969) but for $\mu \sim 1$.

Example 2 If X and Y denote the activities of local populations of excitatory and inhibitory neurons, respectively, then (3.1) could describe dynamics of relaxation

neuron oscillator. An example of such a neural oscillator is given by Wilson and Cowan (1973).

Introducing the "fast" time $t = \tau/\mu$, we can rewrite (3.1) in the form

$$\begin{cases} \dot{X} = F(X, Y, \lambda) \\ \dot{Y} = \mu G(X, Y, \lambda), \end{cases}$$
(3.2)

where dot denotes d/dt. When μ is small, we can assume that $\dot{Y} \approx 0$ and consider the reduced fast system

$$\dot{X} = F(X, Y, \lambda), \tag{3.3}$$

where Y and λ are treated as parameters.

Definition 15 We say that the singularly perturbed dynamical system (3.1) is at quasi-static bifurcation point (X^*, Y^*) for $\lambda = \lambda^*$ and $\mu \ll 1$ if the corresponding reduced fast system (3.3) is at a bifurcation point X^* when $(Y, \lambda) = (Y^*, \lambda^*)$.

In this article we study quasi-static saddle-node, pitchfork and Andronov-Hopf bifurcations.

Remark 16 The quasi-static bifurcations of (3.1) when X and Y are scalars could be mistaken for a Bogdanov-Takens bifurcations of (3.2) at the equilibrium point (X^*, Y^*) for $(\lambda, \mu) = (\lambda^*, 0)$. Suppose (3.3) is at saddle-node bifurcation. Then $F_X = 0$ and the Jacobian matrix of (3.2) for $\mu = 0$ is

$$\left(\begin{array}{cc} 0 & F_Y \\ 0 & 0 \end{array}\right),$$

which for $F_Y \neq 0$ corresponds to the Jacobian matrix for Bogdanov-Takens bifurcation. Nevertheless, there is a difference between them: In the Bogdanov-Takens bifurcation, perturbations of the equations for X' and Y' have the same order, whereas in quasi-static bifurcations the perturbations have essentially different orders of magnitude. They differ by factor $\mu \ll 1$. We will see later that local dynamics of (3.1) depends more on F than on G.

Next we consider networks of such neurons. Let $(X_i, Y_i) \in \mathbb{R}^k \times \mathbb{R}^m$ denote the activity of the *i*-th RN, i = 1, ..., n. We use the notations $X = (X_1, ..., X_n)^{\mathsf{T}} \in \mathbb{R}^{kn}$ and $Y = (Y_1, ..., Y_n)^{\mathsf{T}} \in \mathbb{R}^{mn}$.

Definition 17 A weakly connected network of relaxation neurons is a dynamical system of the form

$$\begin{cases} \mu X_i' = F_i(X_i, Y_i, \lambda) + \varepsilon P_i(X, Y, \lambda, \mu, \varepsilon) \\ Y_i' = G_i(X_i, Y_i, \lambda) + \varepsilon Q_i(X, Y, \lambda, \mu, \varepsilon) \end{cases} \quad i = 1, \dots, n \tag{3.4}$$

where ε is a small parameter, and the functions P_i and Q_i represent connections from the whole network to the *i*-th relaxation neuron (RN).

Note the diagonal structure of (3.4) when $\varepsilon = 0$ and there are no connections between RN's.

The weakly connected network (3.4) can be considered as an ε -perturbation of the uncoupled ($\varepsilon = 0$) system

$$\begin{cases} \mu X_i' = F_i(X_i, Y_i, \lambda) \\ Y_i' = G_i(X_i, Y_i, \lambda) \end{cases} \quad i = 1, \dots, n.$$
(3.5)

We assume that each RN in (3.5) has an equilibrium point (X_i^*, Y_i^*) for some common value $\lambda = \lambda^* \in \Lambda$ (recall that Λ is a multidimensional parameter space). Hence the system (3.5) has the equilibrium point $(X_1^*, \ldots, X_n^*, Y_1^*, \ldots, Y_n^*)$.

3.2 Motivational Examples

We study behavior of (3.4) in some neighborhood of the equilibrium point (X^*, Y^*) . In particular, we are interested in how the equilibrium loses its stability. As soon



Figure 3.1: Possible intersections of nullclines of the relaxation neuron (3.1).

as the phase point (X(t), Y(t)) leaves a $\mathcal{O}(1)$ -neighborhood of the equilibrium, we cannot say anything about subsequent behavior of (3.4). Thus, our analysis is *local*, not global. Still, this analysis is useful, much in the spirit of thresholds of epidemics, explosion modes in chemical kinetics, extinction of chain branched reactions, phase changes in physics, etc.

Example 3 Suppose X_i and Y_i in (3.5) are one-dimensional variables and the nullclines $F_i(X_i, Y_i, \lambda^*) = 0$ and $G_i(X_i, Y_i, \lambda^*) = 0$ intersect transversally as depicted in Figure 3.1a. Then, each RN has an asymptotically stable equilibrium (X_i^*, Y_i^*) , and hence, the uncoupled system (3.5) as a whole has an asymptotically stable equilibrium (X^*, Y^*) . The weakly connected system (3.4), which is an ε -perturbation of (3.5), also has an asymptotically stable equilibrium which is in some neighborhood of (X^*, Y^*) . Theorem 18, presented in Section 3.3, and the Fundamental Theorem of WCNN Theory (Section 2.3) ensure that (3.4) does not acquire any non-linear features that make its dynamics more interesting than that of (3.5).

Example 4 Suppose the nullclines $F_i(X_i, Y_i, \lambda^*) = 0$ and $G_i(X_i, Y_i, \lambda^*) = 0$ intersect non-transversally as it is depicted in Figure 3.1b. Obviously, under perturbations, the non-hyperbolic equilibrium (X_i^*, Y_i^*) may disappear or transform into a pair of equilibria. Thus, it is reasonable to expect that the weakly connected system (3.4)



Figure 3.2: An excitable system. There are initial conditions for which the system (3.1) generates an action potential, or spike (dotted line).

might have some non-linear properties that (3.5) does not have. This case can also be reduced to the one studied in the previous chapter.

These examples show that sometimes the existence of two time scales (slow t and fast $\tau = t/\mu$) is irrelevant when the network of RN's can be reduced to a network of non-relaxation neurons (see Section 3.3).

Example 5 (*Excitable Systems*). When the nullclines intersect as in Figure 3.2, the equilibrium point is globally attractive. Nevertheless, there are initial conditions, which are relatively close to the equilibrium, such that the dynamics of the RN lead to large changes in the state variables X and Y before the RN activity eventually returns to the equilibrium. This amplified response (dotted line in Figure 3.2) is called an action potential or spike. Such systems are called *excitable* (Alexander et.al. 1990). One can think of excitable systems as being systems that are near a threshold or phase transition.

This property can be observed when

$$\frac{\partial F_i}{\partial X_i}(X_i^\star, Y_i^\star, \lambda^\star) \approx 0, \qquad (3.6)$$



Figure 3.3: Intersections of nullclines which do not correspond to an excitable system. a. A relaxation oscillator with non-zero amplitude. b. A relaxation oscillator with zero amplitude.

i.e. when $(X_i^{\star}, Y_i^{\star})$ is near a point for which

$$\frac{\partial F_i}{\partial X_i} = 0.$$

If X_i is a vector, then condition (3.6) should be replaced by the condition that the Jacobian matrix

$$D_{X_i}F_i(X_i^\star,Y_i^\star,\lambda^\star)$$

has eigenvalues with zero (or close to zero) real parts. Such matrices are called *non-hyperbolic*.

If the closeness is compatible with the strength of connections ε , then weakly connected system (3.4) may have interesting non-linear properties. The ε -perturbations from the other neurons can force a RN to generate the action potential, or to remain silent.

The excitable system is one of the many relaxation systems under consideration. For example, the RN whose nullclines intersect as depicted in Figure 3.3 are not

49

excitable. Nevertheless, our theory is applicable to them too. In both cases condition (3.6) is satisfied and a network of such neurons can have interesting nonlinear properties. We start studying networks of such RN's in Section 3.4.

Below we show that if condition (3.6) is violated for all RN's, then the network of such RN's behaves similar to a network of non-relaxation neurons, which we studied in previous chapter.

3.3 Reduction to Regular Perturbation Problem

One can consider (3.4) as a singular perturbation of the unperturbed ($\mu = 0$) weakly connected system

$$\begin{cases} 0 = F_i(X_i, Y_i, \lambda) + \varepsilon P_i(X, Y, \lambda, \mu, \varepsilon) \\ Y'_i = G_i(X_i, Y_i, \lambda) + \varepsilon Q_i(X, Y, \lambda, \mu, \varepsilon) \end{cases} \quad i = 1, \dots, n.$$
(3.7)

The reduced system (3.7) is a quasi-static approximation to (3.4). Then the questions arise: When do (3.4) and (3.7) have similar dynamical properties; Can (3.7) be further simplified? Partial answers are contained in the following

Theorem 18 Suppose that (3.4) has an equilibrium point (X^*, Y^*, λ^*) for $\varepsilon = 0$. Suppose that each Jacobian matrix

$$D_{X_i} F_i(X_i^{\star}, Y_i^{\star}, \lambda^{\star}) \tag{3.8}$$

has all eigenvalues with negative real parts. Then for $(X, Y, \lambda, \mu, \varepsilon)$ sufficiently close to $(X^*, Y^*, \lambda^*, 0, 0)$ the singularly perturbed weakly connected system (3.4) is approximated by the regularly perturbed weakly connected system

$$Y_i' = g_i(Y_i, \lambda, \mu) + \varepsilon q_i(Y, \lambda, \mu, \varepsilon), \quad i = 1, \dots, n,$$
(3.9)

for some functions g_i and q_i . In particular, if the equilibrium point of (3.9) for small ε , μ and λ near λ^* is asymptotically (un)stable, then so is the equilibrium point of (3.4).

Proof of the theorem is based on Implicit Function arguments and singular perturbation techniques. We do not present the proof here because Theorem 18 is a corollary of Theorem 19, which we prove in the next section. There we also explain the notion *approximated*.

In the rest of this chapter we study (3.4) for the case when its reduction to (3.9) is impossible, i.e. when some (or all) of the Jacobian matrices (3.8) have eigenvalues with zero real parts. This occurs when the corresponding RN's are near thresholds and are sensitive to external perturbations.

3.4 Center Manifold Reduction

Consider (3.4) near the equilibrium point (X^*, Y^*, λ^*) for $\varepsilon = 0$. Without loss of generality we may assume that $(X^*, Y^*, \lambda^*) = (0, 0, 0)$.

Theorem 19 Suppose that each of the first n_1 Jacobian matrices

$$L_i = D_{X_i} F_i(0, 0, 0)$$

has some eigenvalues with zero real parts and all other eigenvalues with negative real parts. Let E_i^c $(i = 1, ..., n_1)$ be the eigensubspace spanned by the (generalized) eigenvectors corresponding to the eigenvalues with zero real parts, and the other $n - n_1$ Jacobian matrices have all eigenvalues with negative real parts. Then (3.4) is locally approximated by the lower dimensional dynamical system of the form

$$\begin{cases} \mu x_i' = f_i(x_i, Y_i, \lambda, \mu) + \varepsilon p_i(x, Y, \lambda, \mu, \varepsilon), & i = 1, \dots, n_1, \\ Y_i' = g_i(x_i, Y_i, \lambda, \mu) + \varepsilon q_i(x, Y, \lambda, \mu, \varepsilon), & i = 1, \dots, n_1, \\ Y_j' = g_j(Y_j, \lambda, \mu) + \varepsilon q_j(x, Y, \lambda, \mu, \varepsilon), & j = n_1 + 1, \dots, n, \end{cases}$$
(3.10)

where $x_i \in E_i^c$ and

$$J_i = D_{x_i} f_i(0,0,0,0) = L_{i|_{E_i^c}}, \quad i = 1, \ldots, n_1.$$

In particular, J_i have all eigenvalues with zero real parts.

More precisely, there is a function $Z : E^c \times \mathbb{R}^{mn} \times \Lambda \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^{kn} \times \mathbb{R}^{mn}$ such that all local solutions (X(t), Y(t)) of (3.4) tend exponentially to

$$Z(x(t), Y(t), \lambda, \mu, \varepsilon),$$

where (x(t), Y(t)) is some solution of (3.10).

Proof. If we rewrite (3.4) as

$$\begin{cases} X_i = F_i(X_i, Y_i, \lambda) + \varepsilon P_i(X, Y) \\ \dot{Y}_i = \mu \left(G_i(X_i, Y_i, \lambda) + \varepsilon Q_i(X, Y) \right) \end{cases} \quad i = 1, \dots, n.$$

where $\cdot = d/dt$ and $t = \tau/\mu$ is the "fast" time, then the result follows directly from Theorem 2.3. \Box

Corollary 20 Theorem 18 follows from Theorem 19 when $n_1 = 0$. In this case dim $E^c = 0$ and (3.10) has the same form as (3.9).

Remark 21 Without loss of generality we assume in the following that all Jacobian matrices L_i , $i = 1, ..., n_1$, have all eigenvalues with zero real parts. This means that we consider the weakly connected system (3.4) to be restricted to its center manifold.

3.5 Canonical Models

In this section we study the local dynamics of the singularly perturbed weakly connected system (3.4)

$$\begin{cases} \mu X_i' = F_i(X_i, Y_i, \lambda, \mu) + \varepsilon P_i(X, Y, \lambda, \mu, \varepsilon) \\ Y_i' = G_i(X_i, Y_i, \lambda, \mu) + \varepsilon Q_i(X, Y, \lambda, \mu, \varepsilon) \end{cases} \quad i = 1, \dots, n \tag{3.11}$$

for the case when each Jacobian matrix

$$L_i = D_{X_i} F_i(0, 0, 0, 0)$$

has one simple zero eigenvalue. By Theorem 19 and Remark 21 we may assume that each $X_i \in \mathbb{R}$, i.e. it is a one dimensional variable and

$$L_i = \frac{\partial}{\partial X_i} F_i(0,0,0,0) = 0, \qquad (3.12)$$

i.e. we restricted (3.4) to its center manifold.

Fix Y = 0 and $\varepsilon = 0$ and consider the reduced fast system

$$\mu X'_i = F_i(X_i, 0, \lambda, \mu) \tag{3.13}$$

at the equilibrium point $(X, \lambda) = (0, 0)$. When (3.12) holds and a bifurcation occurs in (3.13), then we say that the RN

$$\begin{cases} \mu X_i' = F(X_i, Y_i, \lambda, \mu) \\ Y_i' = G(X_i, Y_i, \lambda, \mu), \end{cases}$$
(3.14)

is at quasi-static bifurcation.

We study here the case when all equations in (3.14) undergo bifurcations of the same type simultaneously. Such bifurcations are called multiple. We derive canonical models for the simplest and most interesting cases; namely multiple quasi-static saddle-node, pitchfork and Andronov-Hopf (see Section 3.6) bifurcations.

Our analysis of the weakly connected system (3.4) is local in the sense that we study its dynamics in an ε -neighborhood of the equilibrium point $(X, Y, \lambda, \mu, \varepsilon) =$ (0, 0, 0, 0, 0). In particular, we study the dependence of solutions of (3.4) on parameters $\lambda, \mu, \varepsilon$ by supposing that

$$\lambda(\varepsilon) = 0 + \varepsilon \lambda_1 + \mathcal{O}(\varepsilon^2), \qquad \lambda_1 \in \Lambda \mu(\varepsilon) = 0 + \varepsilon \mu_1 + \varepsilon^2 \mu_2 + \mathcal{O}(\varepsilon^3), \quad \mu_1, \mu_2 \in \mathbb{R}.$$
(3.15)

The analysis below does not depend crucially on λ_1 . In contrast to this, the values of μ_1 and μ_2 are important. We derive canonical models when $\mu_1 = 0$. The case $\mu_1 \neq 0$ is discussed in Section 3.5.3.

We also assume that the following transversality condition

$$D_{Y_i}F_i(0,0,0,0) \cdot D_{X_i}G_i(0,0,0,0) \neq 0, \quad i = 1, \dots, n,$$
 (3.16)

is satisfied for all *i*. If this condition is satisfied, we say that the relationship between X_i and Y_i is non-degenerate.

Before proceeding to derivations of the canonical models we prove the following result:

Lemma 22 There is a mapping $\tilde{Y} : \mathbb{R} \to \mathbb{R}^{mn}$ such that the initial portion of the Taylor expansion of (3.4) at $(X, Y) = (0, \tilde{Y}(\varepsilon))$ has the form

$$\begin{cases} \dot{X}_i = A_i Y_i + d_i X_i^2 + e_i X_i^3 + \varepsilon \sum_{j=1}^n C_{ij} X_j + \mathcal{O}(XY, \varepsilon Y, Y^2, X^3, \varepsilon X^2, \varepsilon^2 X) \\ \dot{Y}_i = \mu(\varepsilon R_i + B_i X_i) + \mu \mathcal{O}(Y, X^2, \varepsilon X, \varepsilon^2), \end{cases}$$
(3.17)

for i = 1, ..., n. In particular, the equations for \dot{X}_i do not have terms $\varepsilon, \varepsilon^2, ...$

Proof. From the transversality condition (3.16) it follows that

$$A_i = D_{Y_i} F_i(0, 0, 0, 0) \neq 0, \quad i = 1, \dots, n.$$
(3.18)

Applying the Implicit Function Theorem to the system of algebraic equations

$$0 = F_i(0, Y_i, \lambda(\varepsilon), \mu(\varepsilon)) + \varepsilon P_i(0, Y, \lambda(\varepsilon), \mu(\varepsilon), \varepsilon), \quad i = 1, \dots, n$$

gives a unique smooth function

$$Y(arepsilon) = Y(\lambda(arepsilon),\mu(arepsilon),arepsilon)$$

such that

$$F_i(0, \tilde{Y}_i(\varepsilon), \lambda(\varepsilon), \mu(\varepsilon)) + \varepsilon P_i(0, \tilde{Y}(\varepsilon), \lambda(\varepsilon), \mu(\varepsilon), \varepsilon) \equiv 0, \quad i = 1, \dots, n$$
(3.19)

for all sufficiently small ε . Let $(X, Y) \in \mathbb{R}^n \times \mathbb{R}^{mn}$ be local coordinates at $(0, \tilde{Y}(\varepsilon))$. Equation (3.19) guarantees that the Taylor expansions of equations for X'_i in (3.4) at $(0, \tilde{Y}(\varepsilon))$ do not have terms of order $\varepsilon, \varepsilon^2, \ldots$ not multiplied by state variables. Then, the initial portion of the Taylor expansion of (3.4) is defined in (3.17), where $A_i : \mathbb{R}^m \to \mathbb{R}$ are defined in (3.18), and

$$B_i = D_{X_i} G_i(0, 0, 0, 0) \in \mathbb{R}^m,$$

$$C_{ij} = \frac{\partial}{\partial X_j} P_i(0, 0, 0, 0, 0) \in \mathbb{R}$$
(3.20)

for $i \neq j$, and

$$C_{ii} = rac{\partial}{\partial X_i} \left(D_{Y_i} F_i(0,0,0,0) rac{d \check{Y}_i(0)}{d \epsilon} + D_\lambda F_i(0,0,0,0) \lambda_1 + rac{\partial}{\partial \mu} F_i(0,0,0,0) \mu_1 + P_i(0,0,0,0,0)
ight) \in \mathbb{R}$$

for i = j. Also

$$d_i = rac{1}{2} rac{\partial^2}{\partial X_i^2} F_i(0,0,0,0) \in \mathbb{R},$$

 $e_i = rac{1}{6} rac{\partial^3}{\partial X_i^3} F_i(0,0,0,0) \in \mathbb{R}$

and

$$R_{i} = D_{Y_{i}}G_{i}(0,0,0,0)\frac{d\tilde{Y}_{i}(0)}{d\varepsilon} + D_{\lambda}G_{i}(0,0,0,0)\lambda_{1} + \frac{\partial}{\partial\mu}G_{i}(0,0,0,0)\mu_{1} + Q_{i}(0,0,0,0,0)$$

We study here the case

$$\mu(\varepsilon) = \varepsilon^2 \mu_2 + \mathcal{O}(\varepsilon^3)$$

i.e. $\mu_1 = 0$. The case $\mu_1 \neq 0$ presents no problem and is discussed in Section 3.5.3.

3.5.1 Multiple Quasi-Static Saddle-Node Bifurcations

Each equation in the reduced fast system (3.13) is at saddle-node bifurcation if

$$\frac{\partial}{\partial X_i}F_i(0,0,0,0)=0,$$
$$D_{\lambda}F_{i}(0,0,0,0) \neq 0 \tag{3.21}$$

and

$$d_{i} = \frac{1}{2} \frac{\partial^{2}}{\partial X_{i}^{2}} F_{i}(0, 0, 0, 0) \neq 0$$
(3.22)

for all i. The theorem presented below is valid even when (3.21) is violated.

Theorem 23 Suppose that the singularly perturbed weakly connected system (3.4) is at a multiple quasi-static saddle-node bifurcation point, that the transversality condition (3.16) is satisfied, and that

$$\mu(\varepsilon) = \varepsilon^2 \mu_2 + \mathcal{O}(\varepsilon^3), \quad \mu_2 \neq 0.$$

Then the change of variables

$$\begin{aligned} x_i &= \varepsilon^{-1} d_i X_i &\in \mathbb{R} \\ y_i &= -\varepsilon^{-2} d_i A_i Y_i &\in \mathbb{R} \end{aligned} \qquad i = 1, \dots, n$$
 (3.23)

reduces (3.4) locally to

$$\begin{cases} x'_i = -y_i + x_i^2 + \sum_{j=1}^n c_{ij} x_j + \mathcal{O}(\varepsilon) \\ y'_i = a_i (x_i - r_i) + \mathcal{O}(\varepsilon) \end{cases} \quad i = 1, \dots, n, \qquad (3.24)$$

where $x_i, y_i \in \mathbb{R}$ are scalar variables and $a_i \neq 0$. We refer to (3.24) as being the canonical model for the multiple quasi-static saddle-node bifurcation in a weakly connected network of RN's.

Proof. If we use (3.23) and rescale the time $\tau = \varepsilon t$, then (3.17) transforms to (3.24), where

$$a_{i} = -\mu_{2}A_{i}B_{i}$$

$$r_{i} = d_{i}A_{i}R_{i}(A_{i}B_{i})^{-1}$$

$$c_{ij} = d_{i}C_{ij}d_{j}^{-1}$$
(3.25)

for all *i* and *j*. Note that $A_iB_i \neq 0$ due to transversality condition (3.16). Hence $a_i \neq 0$ and r_i is finite. \Box

Remark 24 Note that while $Y_i \in \mathbb{R}^m$, the canonical variable $y_i \in \mathbb{R}$ is a scalar. Therefore, the change of variables (3.23) cannot be invertible if m > 1 and the canonical model describes the dynamics of (3.17) projected into a linear subspace. Thus, we might lose some dynamics taking place along the kernel of the projection.

Remark 25 Using the translations

$$\begin{array}{l} \tilde{x}_i = x_i - r_i \\ \tilde{y}_i = y_i - r_i^2 - \sum_{j=1}^n c_{ij} r_j \end{array} \quad i = 1, \dots, n$$

we can rewrite (3.24) in the form (where we erase \sim)

$$\begin{cases} x'_{i} = -y_{i} + (2r_{i} + c_{ii})x_{i} + x_{i}^{2} + \sum_{j \neq i}^{n} c_{ij}x_{j} + \mathcal{O}(\varepsilon) \\ y'_{i} = a_{i}x_{i} + \mathcal{O}(\varepsilon) \end{cases} \quad i = 1, \dots, n, \quad (3.26)$$

which is sometimes more convenient. For, example, it is easy to see that (3.26) has a unique equilibrium (x, y) = (0, 0) up to terms of order $\mathcal{O}(\varepsilon)$, which is asymptotically stable when $r_i \to -\infty$.

Remark 26 With the obvious change of variables, the canonical model (3.24) can be written as

$$y_i'' - c_{ii}y_i' - (y_i')^2 + a_iy_i = \sum_{j \neq i}^n c_{ij}y_j' + \mathcal{O}(\varepsilon) \quad i = 1, \dots, n.$$
(3.27)

If all $a_i > 0$, then the left-hand side of each equation in (3.27) is the classical model for studying singular Andronov-Hopf bifurcations (Baer and Erneux 1986,1992; Eckhaus 1983).

Remark 27 We see that in the canonical model (3.24) only connections between "fast" variables X_i are significant. The other three types of connections have order ε or ε^2 and hence are negligible. Indeed, (3.24) can be rewritten in the form

$$\begin{cases} x'_i = -y_i + x_i^2 + \sum_{j=1}^n c_{ij} x_j + \varepsilon \sum_{j=1}^n d_{ij} y_j + \mathcal{O}(\varepsilon) \\ y'_i = a_i (x_i - r_i) + \varepsilon \sum_{j=1}^n e_{ij} x_j + \varepsilon^2 \sum_{j=1}^n f_{ij} y_j + \mathcal{O}(\varepsilon) \end{cases} \quad i = 1, \dots, n,$$

for some $d_{ij}, e_{ij}, f_{ij} \in \mathbb{R}$. We discuss possible applications of this fact in Section 3.7.1.

3.5.2 Multiple Quasi-Static Pitchfork Bifurcations

Suppose that the conditions

$$\frac{\partial}{\partial X_i}F_i(0,0,0,0) = \frac{\partial^2}{\partial X_i^2}F_i(0,0,0,0) = 0$$

and

$$e_i = \frac{1}{6} \frac{\partial^3}{\partial X_i^3} F_i(0, 0, 0, 0) \neq 0$$
(3.28)

are satisfied for all i. Then the uncoupled system (3.13), which describes fast dynamics of (3.4), is near multiple cusp singularity. Using Lemma 22 one can see that the reduced fast weakly connected system

$$\mu X_i' = F_i(X_i, Y_i, \lambda, \mu) + \varepsilon P_i(X, Y, \lambda, \mu, \varepsilon)$$
(3.29)

has a family of equilibrium points $(X, Y, \lambda, \mu, \varepsilon) = (0, \tilde{Y}(\varepsilon), \lambda(\varepsilon), \mu(\varepsilon), \varepsilon)$ for all sufficiently small ε . If we additionally demand that

$$\frac{\partial^2}{\partial \lambda \partial X} F_i(0,0,0,0) \neq 0 \tag{3.30}$$

for all *i*, then (3.29) at $(X, Y) = (0, \tilde{Y}(\varepsilon))$ has a multiple *pitchfork bifurcation*. The result presented below is valid even when (3.30) is violated.

Theorem 28 Suppose that singularly perturbed weakly connected system (3.4) is at a multiple quasi-static pitchfork bifurcation point, transversality condition (3.16) is satisfied, and

$$\mu(\varepsilon) = \varepsilon^2 \mu_2 + \mathcal{O}(\varepsilon^3), \quad \mu_2 \neq 0,$$

then the change of variables

$$\begin{aligned} x_i &= \sqrt{|e_i|\varepsilon^{-1}}X_i &\in \mathbb{R} \\ y_i &= -\sqrt{|e_i|\varepsilon^{-3}}A_iY_i &\in \mathbb{R} \end{aligned} \qquad i = 1, \dots, n \end{aligned}$$
(3.31)

reduces (3.4) to

$$\begin{cases} x'_i = -y_i + \sigma_i x_i^3 + \sum_{j=1}^n c_{ij} x_j + \mathcal{O}(\sqrt{\varepsilon}) \\ y'_i = a_i x_i + \mathcal{O}(\sqrt{\varepsilon}) \end{cases} \quad i = 1, \dots, n, \qquad (3.32)$$

where $a_i \neq 0$ and $\sigma_i = sign e_i = \pm 1$ (e_i was defined in (3.28)). One can think of (3.32) as being the canonical model for the multiple quasi-static pitchfork bifurcation in weakly connected networks of RN's when in addition (3.30) is satisfied.

Proof. If we use (3.31) and rescale time $\tau = \varepsilon t$, then (3.17) transforms to (3.32), where

$$a_i = -\mu_2 A_i B_i$$

$$c_{ij} = \sqrt{|e_i|} C_{ij} \sqrt{|e_j|^{-1}}$$

for all i and j. Note that $A_iB_i \neq 0$ due to (3.16). Hence $a_i \neq 0$. \Box

Remark 29 With an obvious change of variables the canonical model (3.32) can be rewritten as

$$y_i'' - c_{ii}y_i' - \sigma_i(y_i')^3 + a_iy_i = \sum_{j \neq i}^n c_{ij}y_j' + \mathcal{O}(\sqrt{\varepsilon}) \quad i = 1, \dots, n.$$
(3.33)

If all $a_i > 0$ and $\sigma_i = -1$, then the left-hand side of each equation in (3.33) is a van der Pol's oscillator in Lienard representation.

Notice that the oscillators are connected through the derivatives y'_j ("fast" variables), not y_j ("slow" variables) as is usually assumed (Grasman 1987, Section 3.1).

Remark 30 An unfolding of the cusp singularity should include the terms x_i^2 in the equations for x'_i . This is equivalent to introducing constant terms in the equations for \dot{y}_i :

$$\begin{cases} x'_i = -y_i + \sigma_i x_i^3 + \sum_{j=1}^n c_{ij} x_j + \mathcal{O}(\sqrt{\varepsilon}) \\ y'_i = a_i (x_i - r_i) + \mathcal{O}(\sqrt{\varepsilon}) \end{cases} \quad i = 1, \dots, n.$$

When $\sigma_i = -1$ and $a_i > 0$, this is a network of connected Bonhoeffer-Van der Poloscillators.

Remark 31 If

$$\mu(\varepsilon) = 0 + \mathcal{O}(\varepsilon^k), \quad k > 2,$$

then (3.4) can also be reduced to the canonical models (3.24) and (3.32), respectively. But in these cases

$$a_i = \mathcal{O}(\varepsilon^{k-2}), \quad i = 1, \dots, n,$$

i.e. each oscillator in the canonical models is a relaxation oscillator.

3.5.3 Discussion of the case $\mu = \mathcal{O}(\varepsilon)$

Next, we discuss the case

$$\mu(\varepsilon) = 0 + \varepsilon \mu_1 + \mathcal{O}(\varepsilon^2), \quad \mu_1 > 0$$

with the additional assumption that in the weakly connected system (3.4) each variable Y_i is one-dimensional. According to Lemma 22, system (3.4) can be expanded in the Taylor series shown in (3.17). Rescaling

$$\begin{aligned} x_i &= \varepsilon^{-\frac{1}{2}} X_i \\ y_i &= \varepsilon^{-1} Y_i, \quad i = 1, \dots, n \\ \tau &\leftarrow \varepsilon^{\frac{1}{2}} \tau \end{aligned}$$

transforms (3.17) to

$$\begin{cases} x_{i}' = A_{i}y_{i} + d_{i}x_{i}^{2} + \sqrt{\varepsilon}(e_{i}x_{i}^{3} + f_{i}x_{i}y_{i} + \sum_{j=1}^{n} C_{ij}x_{j}) + \mathcal{O}(\varepsilon) \\ y_{i}' = \mu_{1}B_{i}x_{i} + \sqrt{\varepsilon}(\mu_{1}R_{i} + g_{i}y_{i} + h_{i}x_{i}^{2}) + \mathcal{O}(\varepsilon), \end{cases} \qquad i = 1, \dots, n$$
(3.34)

 $A_i B_i > 0$

which is a perturbation of the uncoupled system

$$\begin{cases} x_i' = A_i y_i + d_i x_i^2 \\ y_i' = \mu_1 B_i x_i, \end{cases} \quad i = 1, \dots, n$$
(3.35)

near the equilibrium point (x, y) = (0, 0).

Theorem 32 If

(3.36)

for all i, then the weakly connected system (3.34), the uncoupled system (3.35) and the linear uncoupled system

$$\frac{d}{d\tau} \begin{pmatrix} x_i \\ y_i \end{pmatrix} = \begin{pmatrix} 0 & A_i \\ \mu_1 B_i & 0 \end{pmatrix} \begin{pmatrix} x_i \\ y_i \end{pmatrix}, \quad i = 1, \dots, n$$

are topologically conjugate.

Proof. Inequality (3.36) and $\mu_1 > 0$ imply that $(x_i, y_i) = (0, 0)$ is a hyperbolic saddle point for this system. The result follows from Implicit Function and Hartman-Grobman Theorems. \Box

Remark 33 When $\mu_1 A_i B_i > 0$, the weakly connected system (3.4) does not have interesting local nonlinear neuro-computational properties since it is essentially linear and uncoupled in that case.

If all $A_i B_i < 0$, then each Jacobian matrix

$$\left(\begin{array}{cc} 0 & A_i \\ \mu_1 B_i & 0 \end{array}\right)$$

has a pair of pure imaginary eigenvalues. In this case the WCNN (3.34) is at a multiple Andronov-Hopf bifurcation point. We studied this case in Section 2.5.

Finally, we note that it is not clear which facts derived in this subsection remain valid when we drop the assumption that each Y_i is a scalar.

3.6 Multiple Quasi-Static Andronov-Hopf Bifurcations

In this section we study the case when in the singularly perturbed weakly connected system (3.4) each Jacobian matrix

$$L_{i} = D_{X_{i}} F_{i}(0, 0, 0, 0) = \begin{pmatrix} a_{i1} & a_{i2} \\ a_{i3} & a_{i4} \end{pmatrix}$$

has a pair of purely imaginary non-zero eigenvalues $\pm i\Omega_i$. This corresponds to a multiple quasi-static non-singular Andronov-Hopf bifurcation of (3.4).

Theorem 34 Suppose that (3.4) is at multiple quasi-static Andronov-Hopf bifurcation point, then the invertible change of variables

$$X_{i} = \sqrt{\varepsilon} \begin{pmatrix} 1 & 1\\ \frac{a_{i4} + i\Omega_{i}}{a_{i2}} & \frac{a_{i4} - i\Omega_{i}}{a_{i2}} \end{pmatrix} \begin{pmatrix} e^{i\Omega_{i}\tau} z_{i}\\ e^{-i\Omega_{i}\tau} \bar{z}_{i} \end{pmatrix} + \mathcal{O}(\varepsilon) \quad i = 1, \dots, n,$$

$$Y_{i} = \varepsilon v_{i}$$

reduces (3.4) to

$$\begin{cases} z'_{i} = (a_{i} + A_{i}v_{i})z_{i} + b_{i}z_{i}|z_{i}|^{2} + \sum_{\Omega_{j} \neq i}^{n} c_{ij}z_{j} + \mathcal{O}(\sqrt{\varepsilon}) \\ v'_{i} = d_{i}(R_{i} + S_{i}|z_{i}|^{2} + T_{i}v_{i} + \mathcal{O}(\sqrt{\varepsilon})) \end{cases} \qquad i = 1, \dots, n, \quad (3.37)$$

where $z_i \in \mathbb{C}$, $v_i \in \mathbb{R}^m$ and $a_i, b_i, c_{ij} \in \mathbb{C}$; $A_i : \mathbb{R}^m \to \mathbb{C}$; $R_i, S_i \in \mathbb{R}^m$; $T_i : \mathbb{R}^m \to \mathbb{R}^m$ and $d_i = \mu/\varepsilon$.

We do not present a proof here. Derivation of (3.37) coincides (after obvious modifications) with the derivation of the canonical model for multiple Andronov-Hopf bifurcations in weakly connected networks of non-relaxation neurons and can be found in Section 2.5.

Notice again that in the canonical model (3.37) (as well as in (2.25)) only those RN's interact that have equal natural frequencies Ω_i . If $\Omega_i \neq \Omega_j$ for some *i* and *j*, then the *i*-th and *j*-th RN's do not "feel" each other even when $c_{ij} \neq 0$ and $c_{ji} \neq 0$. The *i*-th RN can turn on and off its connections with other RN's simply by adjusting its natural frequency Ω_i . This feature might be related to such a phenomenon as attention.

3.7 Conclusion

We have analyzed local dynamics of the singularly perturbed weakly connected system (3.4) at an equilibrium point and showed that in many cases it is governed locally by well-studied systems like regularly perturbed weakly connected systems (3.9) (Theorem 18). In other interesting cases the dynamics of (3.4) are governed by the canonical models (3.24), (3.32) or (3.37). To the best of our knowledge these models have not been analyzed yet. We think they have interesting computational properties with possible applications to neurocomputers. We present some basic analysis of (3.24) and (3.32) in Chapter 8.

3.7.1 Synaptic Organizations of the Brain

We saw (Remark 27) that in the canonical models only connections between "fast" variables X_i are significant and the other three types of connections ("fast" \rightarrow "slow", "slow" \rightarrow "fast", "slow" \rightarrow "slow") are negligible. Even this simple fact has some important biological implications.

Suppose (3.1) describes a mechanism of generation of action potentials by a neuron (as in Example 1). Then the significance of "fast" \rightarrow "fast" connections means that the synaptic transmission between neurons is triggered by "fast" ion channels (in fact, by Na⁺ and Ca⁺⁺; The former are responsible for depolarization, the latter open synaptic vesicles). Since the synaptic transmission mechanism is well studied now (Shepherd 1983), this observation does not carry much new information.

Suppose (3.1) describes the activity of the relaxation neural oscillator as in Example 2. Then "fast" \rightarrow "fast" connections are synaptic connections between excitatory neurons. Thus, one can conclude that information is transmitted from one part of the brain to another one through excitatory neurons, while the inhibitory neurons

serve only local purposes. And indeed, copious neurophisiological data (Rakic 1976; Shepherd 1976) suggest that excitatory neurons usually have long axons capable of forming distant synaptic contacts, while inhibitory neurons are local-circuit neurons having short axons (or without axons at all). They provide reciprocal inhibition. This division into relay- and inter-neurons has puzzled biologists for decades (see Rakic 1976). Our analysis shows that even if the inhibitory neurons had long axons, their impact onto other parts of the brain would be negligible. Hence, the long-axon inhibitory neurons are functionally insignificant.

The analysis above attempts to describe why a brain might have the anatomical structure it does. But we are still far away from satisfactory explanations of this problem. We continue our study of possible synaptic organizations of the brain in Chapter 10.

Chapter 4 Weakly Connected Maps

In previous chapters we studied dynamics of weakly connected networks governed by a system of ordinary differential equations (ODE). It is also feasible to consider weakly connected networks of difference equations, or mappings, of the form

$$X_i \mapsto F_i(X_i, \lambda) + \varepsilon G_i(X, \lambda, \rho, \varepsilon), \quad i = 1, ..., n, \quad \varepsilon \ll 1,$$
(4.1)

where variables $X_i \in \mathbb{R}^m$, parameters $\lambda \in \Lambda$, $\rho \in \mathcal{R}$ and functions F_i and G_i have the same meaning as in previous chapters. Difference equations of the form (4.1) arise when one studies Poincare maps of flows.

A weakly connected mapping (4.1) can also arise as a time-T map of the flows. Indeed, consider a periodically forced weakly connected network of the form

$$\dot{X}_i = \tilde{F}_i(X_i, \lambda, P_i(t)) + \varepsilon \tilde{G}_i(X, \lambda, \rho, Q(t), \varepsilon),$$

where $P_i(t)$ and Q(t) are *T*-periodic functions. Such systems can describe thalamocortical interactions (see Figure 4.1). In this case $P_i(t)$ and Q(t) denote inputs from the thalamus to the *i*-th cortical column X_i . Knowing X(0) one can find X(T), which is some function of X(0), λ , ρ and ε . It is easy to check that the function has a weakly connected form and can be written as (4.1).



Figure 4.1: Thalamo-cortical interactions

Recall that our strategy is to compare dynamic behavior of the uncoupled ($\varepsilon = 0$) system

$$X_i \mapsto F_i(X_i, \lambda), \quad i = 1, ..., n$$

$$(4.2)$$

and coupled ($\varepsilon \neq 0$) system (4.1). Obviously, the uncoupled system (4.2) is not interesting as a model of the brain. We are looking for such regimes and parameter values which endow the coupled system (4.1) with "interesting" neuro-computational properties.

As in previous chapters, we study dynamics of (4.2) near a fixed point $X^* = (X_1^*, ..., X_n^*)^{\mathsf{T}}$ for some $\lambda^* \in \Lambda$. Thus, we have

$$X_i^{\star} = F_i(X_i^{\star}, \lambda^{\star})$$

for all *i*. Without loss of generality we may assume $X^* = 0$ for $\lambda^* = 0$.

Notice that the time-T map having a fixed point corresponds to a continuous time dynamical system having a limit cycle.

Analysis of weakly connected mappings is parallel to analysis of weakly connected systems of ODE, which we performed in Chapter 2. As one can expect, local dynamics near a fixed point is not interesting when the point is hyperbolic (Section 4.1). Therefore, the only points that deserve our attention are non-hyperbolic. In this chapter we study dynamics of (4.1) in some neighborhood of a non-hyperbolic fixed point corresponding to saddle-node and flip bifurcations (Sections 4.2.1 and 4.2.2). First we derive canonical models and then we reveal the relationship between them and the canonical models for continuous time weakly connected neural networks (Section 4.3).

4.1 Hyperbolic Fixed Points

A fixed point $X_i^* = 0$ is said to be *hyperbolic* if the Jacobian matrix

$$L_i = D_{X_i} F_i(0,0)$$

does not have eigenvalues of unit modulus. Notice that the Jacobian matrix for the uncoupled system (4.2) near the fixed point $X^* = (X_1^*, ..., X_n^*)^{\mathsf{T}}$ has the form

$$L = \begin{pmatrix} L_1 & 0 & \cdots & 0 \\ 0 & L_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & L_n \end{pmatrix}.$$

and the fixed point X^* is hyperbolic if and only if each X_i^* is hyperbolic.

Theorem 35 If the dynamics of each neuron is near a hyperbolic fixed point, then the weakly connected network (4.1) of such neurons, the uncoupled network (4.2) and the linear mapping

$$X \mapsto LX$$

have topologically equivalent local orbit structures.

In this case the local behavior of the weakly connected system (4.1) is essentially linear and uncoupled.

Proof of the theorem uses the Implicit Function and Hartman-Grobman Theorems and coincides (with the obvious modifications) with the proof of the analogous theorem for ODE (Theorem 2 in Chapter 2).

4.2 Non-hyperbolic Fixed Points

It follows from the previous section that the only fixed points requiring further discussion are non-hyperbolic ones. An important case when such fixed points occur corresponds to bifurcations in dynamics of (4.2). In this section we study the simplest bifurcations of those: saddle-node and flip types. For each, the Jacobian matrix L_i has only one eigenvalue on the unit circle: +1 for saddle-node and -1 for the flip bifurcation. We assume that the other eigenvalues are inside the unit circle.

The Center Manifold Theorem for the maps (analogous to the one for flows used in Section 2.3) guarantees that there is a locally attractive invariant manifold in \mathbb{R}^{mn} on which the dynamics of (4.1) is governed by the lower-dimensional system of the form

$$x_i \mapsto f_i(x_i, \lambda) + \varepsilon g_i(x, \lambda, \rho, \varepsilon),$$
 (4.3)

where each $x_i \in \mathbb{R}$ is a scalar and

$$f_i(0,0)=0, \quad rac{\partial}{\partial x_i}f_i(0,0)=\pm 1.$$

As before, we assume that

$$\lambda = \lambda(\varepsilon) = 0 + \varepsilon \lambda_1 + \mathcal{O}(\varepsilon^2)$$

and

$$\rho = \rho(\varepsilon) = \rho_0 + \varepsilon \rho_1 + \mathcal{O}(\varepsilon^2)$$

for some $\lambda_1 \in \Lambda$ and $\rho_0, \rho_1 \in \mathcal{R}$.



Figure 4.2: Bifurcations of a mapping. a. Saddle-node bifurcation for $c_i p_i < 0$. b. Saddle-node bifurcation for $c_i p_i > 0$. c. Flip bifurcation.

4.2.1 Multiple Saddle-Node Bifurcations

Consider the weakly connected system (4.3) near a multiple saddle-node bifurcation point $x^* = 0$ for $\lambda^* = 0$. The initial portion of the Taylor series of the right-hand side of (4.3) at the origin for small ε has the form

$$x_i + p_i x_i^2 + \varepsilon (c_i + \sum_{j=1}^n c_{ij} x_j) + \text{h.o.t.},$$

where

$$p_i = \frac{1}{2} \frac{\partial^2}{\partial x_i^2} f_i(0,0) \neq 0,$$

$$c_{ij}=rac{\partial}{\partial x_j}g_i(0,0,
ho_0,0)$$

and

$$c_i = D_\lambda f_i(0,0)\lambda_1 + g_i(0,0,\rho_0,0).$$

Recall that we called $c_i = 0$ the *adaptation condition* for the *i*-th neuron dynamics.

Suppose $c_i \neq 0$ for all *i*. Then the rescaling $x = \sqrt{\varepsilon} \tilde{x}$ transforms the weakly connected mapping to

$$\tilde{x}_i \mapsto \tilde{x}_i + \sqrt{\varepsilon}(c_i + p_i \tilde{x}_i^2) + \mathcal{O}(\varepsilon)$$

which is essentially uncoupled (although non-linear). Indeed, the *i*-th neuron's dynamics is determined by the constants c_i and p_i . Depending upon the sign of their product the dynamics can have two qualitatively different local phase portraits:

- If c_ip_i < 0, then there is a pair of fixed points. One of them is stable and the other one is not (see Figure 4.2a).
- If $c_i p_i > 0$, then there are no fixed points. For any initial conditions the dynamics eventually leaves some neighborhood of the origin (see Figure 4.2b).

Local dynamics of the *i*-th neuron is not affected by the dynamics of the other neurons, provided the other neuron activities are bounded. We see that the neuron activity depends only on the environment ρ_0 and internal state λ_1 (since c_i depends upon them) and does not depend on the other neuron activities.

If the adaptation condition $(c_i = 0)$ is satisfied for all *i*, then after the rescaling $x = \varepsilon \tilde{x}/p_i$ the weakly connected system (4.3) transforms to

$$\tilde{x}_i \mapsto \tilde{x}_i + \varepsilon \left(r_i + \tilde{x}_i^2 + \sum_{j=1}^n \tilde{c}_{ij} \tilde{x}_j \right) + \mathcal{O}(\varepsilon^2),$$
(4.4)

for $\tilde{c}_{ij} = c_{ij}p_j/p_i$ and some constants $r_i \in \mathbb{R}$. We call (4.4) the canonical model for a multiple saddle-node bifurcation in weakly connected maps.

4.2.2 Multiple Flip Bifurcations

A weakly connected mapping (4.3) is near flip a bifurcation point $x^* = 0$ for $\lambda^* = 0$ if

$$f_i(0,0) = 0, \quad \frac{\partial}{\partial x_i} f_i(0,0) = -1$$

(see Figure 4.2c). The flip bifurcation does not have an analogue for one-dimensional flows. Thus, it is natural to expect some new qualitative features. The most prominent feature of the multiple flip bifurcation is that the adaptation condition is not necessary for the local dynamics of the weakly connected mappings to exhibit "interesting" properties.

Theorem 36 If the weakly connected mapping (4.3) is near a multiple flip bifurcation point, then there is an invertible change of variables which transforms (4.3) to

$$x_i \mapsto x_i + \sqrt{\varepsilon} \left(r_i x_i \pm x_i^3 + \sum_{j=1}^n c_{ij} x_j \right) + \mathcal{O}(\varepsilon).$$
 (4.5)

We call (4.5) the canonical model for multiple flip bifurcations in weakly connected mappings.

Proof. Let us denote the k-th iteration of the mapping (4.3) by x^k . Thus, we have

$$x_i^{k+1} = f_i(x_i^k, \lambda) + \varepsilon g_i(x^k, \lambda, \rho, \varepsilon).$$

The initial portion of the Taylor series of the right-hand side at the origin is given by

$$-x_{i}^{k} + p_{i}(x_{i}^{k})^{2} + q_{i}(x_{i}^{k})^{3} + \varepsilon(c_{i} + \sum_{j=1}^{n} c_{ij}x_{j}^{k}) + \text{h.o.t.}$$

Let us find how x^{k+2} depends on x^k . For this we have to evaluate the composition

$$\begin{aligned} x_i^{k+2} &= f_i(f_i(x_i^k,\lambda) + \varepsilon g_i(x^k,\lambda,\rho,\varepsilon),\lambda) + \varepsilon g_i(f(x^k,\lambda) + \varepsilon g(x^k,\lambda,\rho,\varepsilon),\lambda,\rho,\varepsilon) \\ &= x_i^k - 2\varepsilon \left(c_i p_i x_i^k + \sum_{j=1}^n c_{ij} x_j^k \right) - 2(q_i + p_i^2)(x_i^k)^3 + \text{h.o.t.} \end{aligned}$$

Now let $x_i = \sqrt{\varepsilon} \tilde{x}_i / \sqrt{2|q_i + p_i^2|}$, $\tilde{r}_i = -2c_i p_i$ and $c_{ij} = -\frac{1}{2} \tilde{c}_{ij} \sqrt{|q_j + p_j^2|/|q_i + p_i^2|}$, then we obtain

$$\tilde{x}_i^{k+2} = \tilde{x}_i^k + \sqrt{\varepsilon} \left(\tilde{r}_i \tilde{x}_i^k + \sigma_i (\tilde{x}_i^k)^3 + \sum_{j=1}^n \tilde{c}_{ij} \tilde{x}_j^k \right) + \mathcal{O}(\varepsilon),$$

where

$$\sigma_i = -\operatorname{sign}(q_i + p_i^2) = \pm 1.$$

Finally notice that this result does not depend on the adaptation condition $c_i = 0$. \Box

4.3 Connection With ODE

There is an intimate relationship between the canonical models for weakly connected maps and ODE. The former have a form

$$x_i^{k+1} = x_i^k + \tau h_i(x^k) + \mathcal{O}(\tau^2), \ \ \tau \ll 1,$$

for some functions h_i . If one considers the iteration x^k as a value of a smooth function at the moment $k\tau$, then he can rewrite the equation above as

$$x_i((k+1)\tau) = x_i(k\tau) + \tau h_i(x(k\tau)) + \mathcal{O}(\tau^2)$$

or

$$\frac{x_i(t+\tau)-x_i(t)}{\tau}=h_i(x(t))+\mathcal{O}(\tau),$$

where $t = k\tau$. The equation above is Euler's discretization of the ODE

$$\frac{dx_i}{dt} = h_i(x).$$

In this sense, we see that the canonical model for multiple saddle-node bifurcation

$$x_i \mapsto x_i + \varepsilon \left(r_i + x_i^2 + \sum_{j=1}^n c_{ij} x_j \right) + \mathcal{O}(\varepsilon^2)$$

is the discretization of that for ODE

$$\frac{dx_i}{dt} = r_i + x_i^2 + \sum_{j=1}^n c_{ij} x_j + \mathcal{O}(\varepsilon).$$

Similarly, the canonical model for a multiple flip bifurcation

$$x_i \mapsto x_i + \sqrt{\varepsilon} \left(r_i x_i \pm x_i^3 + \sum_{j=1}^n c_{ij} x_j \right) + \mathcal{O}(\varepsilon)$$

is the discretization of

$$\frac{dx_i}{dt} = r_i x_i \pm x_i^3 + \sum_{j=1}^n c_{ij} x_j + \mathcal{O}(\sqrt{\varepsilon}),$$

This relationship between canonical models allows us to concentrate our efforts on studying the canonical models for multiple saddle-node and pitchfork bifurcations in ODE, which we perform in subsequent chapters.

Part II

Analysis of Canonical Models

Chapter 5 Multiple Saddle-Node Bifurcation

In this chapter we continue to study the weakly connected neural network

$$\dot{x}_i = f_i(x_i, \lambda) + \varepsilon g_i(x, \lambda, \rho, \varepsilon)$$
(5.1)

near a multiple saddle-node bifurcation. In particular, we explore an important case where such a bifurcation occurs.

5.1 Saddle-Node on a Limit Cycle

Consider a dynamical system

$$\dot{x} = f(x, \lambda), \tag{5.2}$$

where $x \in [-1, 1]/(\{-1\} = \{1\})$ has a meaning of phase variable and the function f is periodic in x. Suppose (5.2) is near a saddle-node bifurcation point x = 0 for $\lambda = 0$, i.e.

$$f(0,0) = 0$$
 and $\frac{\partial}{\partial x}f(0,0) = 0$

but

$$p = \frac{1}{2} \frac{\partial^2}{\partial x^2} f(0,0) \neq 0.$$



Figure 5.1: Phase portraits for various c

Also suppose that f(x,0) > 0 for all $x \neq 0$. This implies p > 0. Consider the leading terms of the Taylor series of f for small x and $\lambda = \varepsilon \lambda_1 + \mathcal{O}(\varepsilon^2)$

$$f(x, \lambda) = \varepsilon c + px^2 + \text{h.o.t.},$$

where

$$c = D_{\lambda} f(0,0) \lambda_1. \tag{5.3}$$

Posible phase portraits of such system for various c are depicted in Figure 5.1. The saddle-node bifurcation occurs when c = 0.

We see that the *local* behavior near the non-hyperbolic equilibrium x = 0 has global ramifications: For c < 0 there are two equilibria – a stable node (filled circle) and a saddle (open circle). For any initial condition not at the saddle the dynamics approaches the node. When we consider saddle-node bifurcations, the saddle and the node are close to each other. This means that there could be some perturbations which move x(t) to the right hand side, beyond the saddle. In this case the activity makes one rotation (an excursion, an action potential or spike) and eventually approaches the node. Such a dynamical system is the typical (and possibly the simplest) example of an *excitatory* system. It resembles the excitatory properties of real neurons.

For c = 0 the saddle and node coalesce, and for c > 0 there is no equilibrium, in which case the dynamics of (5.2) is periodic. Such behavior can be observed in general dynamical systems when there is an asymptotically stable limit cycle with a



Figure 5.2: Saddle-node bifurcation on a limit cycle



Figure 5.3: An intersection of nullclines in a relaxation system which exhibits saddlenode bifurcation on a limit cycle

non-hyperbolic equilibrium on it. In two-dimensional case there is also an equilibrium inside the limit cycle (see Figure 5.2). Saddle-node bifurcation on a limit cycle can also arise in relaxation systems with the nullclines intersected as in Figure 5.3

Let us return to the dynamical system (5.2). As we see, local bifurcations of a nonhyperbolic equilibrium x = 0 can have a global effect. In order to find out whether (5.2) has periodic activity or not it is suffices to determine the sign of c defined in (5.3). If c < 0, then there are two equilibria

$$\pm \sqrt{\frac{\varepsilon|c|}{p}} + \mathcal{O}(\varepsilon)$$

and the dynamics converges. If c > 0, then there are no equilibria (local or global) and dynamics is periodic. The remarkable feature of the saddle-node bifurcation on a limit cycle is that even though the dynamics becomes periodic, x(t) spends most time in a neighborhood of the origin; that is, where the non-hyperbolic equilibrium used to be. We prove this fact in Lemma 37 below and use it throughout this chapter. In particular, we can determine how various neurons interact simply by analyzing local behavior near a saddle-node bifurcation.

5.2 The VCON

There is an invertible change of variables $x = A(\phi, \varepsilon)$ (Ermentrout and Kopell 1986) which transforms the dynamical system (5.2) to a simpler system

$$\phi' = (1 - \cos \phi)p + (1 + \cos \phi)c + \mathcal{O}(\sqrt{\varepsilon}), \tag{5.4}$$

which is an example of a Voltage Controlled Oscillator Neuron (VCON) (Hoppensteadt 1983). In the VCON, $' = d/d\tau$ where $\tau = \sqrt{\varepsilon}t$ is a "slow" time variable.

Notice that VCON is at saddle-node bifurcation when c = 0. For c < 0 it has two equilibria and for c > 0 it has none. The phase portrait of a VCON coincides with the one depicted in Figure 5.1. One can think of VCON as the canonical model for the saddle-node bifurcation on a limit cycle.

The Ermentrout-Kopell's change of variables $x = A(\phi, \varepsilon)$ is defined as follows:

- If $|x| \leq \sqrt[4]{\varepsilon}$, then $A(\phi, \varepsilon) = \sqrt{\varepsilon} \tan \frac{\phi}{2}$.
- If |x| > √ε, then A(φ,ε) is extended to a one-to-one smooth map satisfying some technical conditions.

VCON models such as (5.4) have attracted attention in part due to the fact that their dynamics is described in terms of phase variables. Thus, the hard mathematical problem of converting physical variables to phase variables is eliminated by choosing the phase-like variables in the first place. VCON models are also appealing to electrical engineers since they can be constructed as electrical circuits. There are many phenomena which can be studied using VCON models (Hoppensteadt 1986). The most intriguing of them is the problem of synchronization and attention. We study this problem in Section 2.5. In this chapter we use a VCON model to illustrate various dynamic behaviors.

5.3 Preliminary Analysis

In this section we are interested in behavior of (5.2) for c > 0. As we have already said, its dynamics is periodic, but what is the period? Since the period of oscillations in VCON (5.4) for $c \neq 0$ in terms of the "slow" time is $\mathcal{O}(1)$, the period of oscillations in (5.2) in terms of the normal time $t = \tau/\sqrt{\varepsilon}$ is $\mathcal{O}(\frac{1}{\sqrt{\varepsilon}})$. Actually it is $\frac{\pi}{\sqrt{\varepsilon cp}}$ but we do not need this expression.

Next, how much time does x(t) spend away from some small neighborhood of the origin, say away from $[-\sqrt[4]{\varepsilon}, +\sqrt[4]{\varepsilon}]$? To determine this it suffices to consider the Ermentrout-Kopell transformation

$$x = \sqrt{\varepsilon} \tan \frac{\phi}{2}.$$

Notice that its inverse is the transformation

$$\phi = 2 \arctan rac{x}{\sqrt{arepsilon}},$$

maps $\left[-\sqrt[4]{\varepsilon}, +\sqrt[4]{\varepsilon}\right]$ one-to-one onto $\left[-\pi + 2\sqrt[4]{\varepsilon}, \pi - 2\sqrt[4]{\varepsilon}\right]$. From (5.4) it follows that

$$\phi' \approx 2c$$

when $\phi \approx \pi$ or $-\pi$. Thus, it takes

$$\frac{2\sqrt[4]{\varepsilon}}{c} + \mathcal{O}(\sqrt{\varepsilon})$$

units of slow time τ for ϕ to cover the distance from $\pi - 2\sqrt[4]{\varepsilon}$ to $\pi + 2\sqrt[4]{\varepsilon}$. In the terms of normal time $t = \tau/\sqrt{\varepsilon}$ it takes $\frac{2}{c\sqrt[4]{\varepsilon}}$. Now recall that $\phi \in (\pi - 2\sqrt[4]{\varepsilon}, \pi + 2\sqrt[4]{\varepsilon})$ corresponds to $|x| > \sqrt[4]{\varepsilon}$. Thus, we proved the following result:

Lemma 37 During one oscillation (action potential) x(t) spends

$$\mathcal{O}(\frac{1}{\sqrt[4]{\varepsilon}})$$

away from $\sqrt[4]{\varepsilon}$ -neighborhood of the origin and

$$\mathcal{O}(\frac{1}{\sqrt{\varepsilon}})$$

in the $\sqrt[4]{\varepsilon}$ -neighborhood of the origin.

Since $1/\sqrt[4]{\varepsilon}$ is less than $1/\sqrt{\varepsilon}$ one can conclude that the activity x(t) spends most of the time in the neighborhood of the origin.

5.4 Case $c \neq 0$

Now consider the weakly connected neural network (5.1). Suppose that in the uncoupled ($\varepsilon = 0$) system

$$\dot{x}_i = f_i(x_i, \lambda), \quad i = 1, ..., n$$
 (5.5)

each equation describes dynamics having a saddle-node bifurcation on a limit cycle. Consider the parameter c_i defined by

$$c_i = D_\lambda f_i(0,0)\lambda_1 + g_i(0,0,\rho_0,0).$$

Theorem 38 • If $c_i > 0$, then the *i*-th neuron oscillates with the period $\mathcal{O}(\frac{1}{\sqrt{\epsilon}})$. The activities of the other neurons do not change the qualitative behavior.

• If $c_i < 0$, then the *i*-th neuron is silent. More precisely, its activity $x_i(t)$ stays in $\sqrt{\varepsilon}$ -neighborhood of the origin and the activities of the other neurons do not affect it (see Figure 5.4).



Figure 5.4: Dynamic behavior for various c_i

Proof. Suppose all the variables x_i are small, say $|x_i| \leq \sqrt[4]{\varepsilon}$. Then the dynamics of the *i*-th neuron is governed by

$$\dot{x}_i = \varepsilon c_i + p_i x_i^2 + \varepsilon \sum c_{ij} x_j + \text{h.o.t.}$$
 (5.6)

In this case we say that the network is in Stage I. If at least one of x_j is large, say $|x_j| > \sqrt[4]{\varepsilon}$, then the *i*-th neuron dynamics is governed by the system

$$\dot{x}_i = \varepsilon c_i + p_i x_i^2 + \varepsilon \hat{g}_i(x) + \text{h.o.t.}$$
(5.7)

for some function \hat{g}_i . We call this Stage II.

Since the number of neurons n is finite, the network spends $\mathcal{O}(\frac{1}{\sqrt[4]{\varepsilon}})$ in Stage II and, hence, the rest of the time (namely $\mathcal{O}(\frac{1}{\sqrt{\varepsilon}})$) in Stage I.

Suppose $c_i > 0$, i.e. the *i*-th neuron in the uncoupled network (5.5) oscillates. Can weak connections between the neurons stop the oscillations or distort substantially the period? The answer is NO. Indeed, the oscillations can be stopped only when x_i is small enough. Suppose the network is in Stage II. Its dynamics is governed by (5.7). The value of the function \hat{g}_i could be negative and greater than c_i in absolute value. In this case

$$\dot{x}_i = -\mathcal{O}(\varepsilon)$$

but it could happen only during a time period of total length $\mathcal{O}(\frac{1}{\sqrt{\epsilon}})$. Hence $x_i(t)$ could be pulled back for no more than $\mathcal{O}(\epsilon \cdot \frac{1}{\sqrt{\epsilon}}) = \mathcal{O}(\epsilon^{3/4})$, but the dynamics in Stage I pushes $x_i(t)$ forward more than $\mathcal{O}(\epsilon \cdot \frac{1}{\sqrt{\epsilon}}) = \mathcal{O}(\sqrt{\epsilon})$. Distance $\mathcal{O}(\epsilon^{3/4})$ is negligible in comparison with $\mathcal{O}(\sqrt{\epsilon})$ for $\epsilon \to 0$.

The case $c_i < 0$ can be analyzed using the same arguments with obvious modifications. \Box

It is not correct to assume that the network is uncoupled for $c_i \neq 0$. Indeed, if say $c_1 > 0$ and $c_2 > 0$, then the first and second neurons oscillate and one can observe such phenomena as synchronization or entrainment. Unfortunately, to study these phenomena one needs to consider the network on a very large time scale.

If, say, $c_3 < 0$, then the system still can be wobbling in a neighborhood of $-\sqrt{\varepsilon c_i/p_i}$ due to the influences from the other neurons, but the oscillations have very small amplitude and, again, are negligible for an observer.

The theorem above covers the case $c_i \neq 0$. What if $c_i = 0$ for some, but not all *i*? In this case it is possible to show that activity can be convergent or periodic. For example, consider a network of VCONs of the form

$$\begin{cases} \dot{\phi}_1 = 1 - \cos \phi_1 + \varepsilon c_{12} \sin^2 \phi_2 \\ \dot{\phi}_2 = 1 - \cos \phi_2 + \varepsilon c_2 \end{cases}$$

where $c_2 > 0$, and, hence, ϕ_2 oscillates. It is easy to see that in the case $c_{12} < 0$ the variable ϕ_1 stays in a neighborhood of the origin, but in the case $c_{12} > 0$ it oscillates.

5.5 Adaptation Condition Is Satisfied

Now suppose $c_i = 0$ for all *i*. Recall that in this case we say that the adaptation condition is satisfied. In Chapter 2 we showed that the local dynamics of (5.2) is

governed by the canonical model

$$x'_{i} = r_{i} + b_{i}x_{i} + x_{i}^{2} + \sum_{j=1}^{n} c_{ij}x_{j}.$$
(5.8)

How much can we tell about global dynamics of (5.2) studying this canonical model? Obviously, all attractors of the canonical model are also local attractors of the original weakly connected neural network and vice-versa, since (5.8) is obtained from (5.2) (essentially) by rescaling.

Can the weakly connected neural network have an attractor which the canonical model does not have? The answer is YES, but as one expects, the attractor cannot be local. Following is an example of a weakly connected neural network with non-trivial global behavior which the canonical model cannot have.

Consider a weakly connected network of VCONs of the form

$$\begin{cases} \dot{\phi}_1 = 1 - \cos\phi_1 - \varepsilon^2 (1 + \cos\phi_1) + \varepsilon V(\phi_2) \\ \dot{\phi}_2 = 1 - \cos\phi_2 - \varepsilon^2 (1 + \cos\phi_2) + \varepsilon V(\phi_1) \end{cases}$$
(5.9)

where the 2π -periodic function $V(\phi)$ is given on $[-\pi, \pi]$ by the formula

$$V(\phi) = \begin{cases} 0, & \text{if } |\phi| < \pi/2\\ a, & \text{otherwise} \end{cases}$$

for some constant a > 0. It is easy to check that such system is near a multiple saddle-node bifurcation for $\phi_1 = \phi_2 = 0$. The canonical model for this system is

$$x'_i = -1 + x_i^2, \quad i = 1, 2.$$

It is uncoupled because V'(0) = 0. Obviously, the only attractor is a stable node x = (-1, -1). Nevertheless, it is possible to show that (5.9) has a family of global limit cycles for all sufficiently large a. One such limit cycle is depicted in Figure 5.5.

Thus, the analysis of the canonical model does not necessarily reveal global dynamics of the original weakly connected neural network.



Figure 5.5: Co-existence of a local attractor and a global limit cycle

Chapter 6

Multiple Andronov-Hopf Bifurcation

Recall that a WCNN near a multiple Andronov-Hopf bifurcation point is governed (see Theorem 11) by the dynamical system of the form

$$z'_{i} = b_{i}z_{i} + d_{i}z_{i}|z_{i}|^{2} + \sum_{j \neq i}^{n} c_{ij}z_{j}, \qquad (6.1)$$

where $b_i, c_{ij}, d_i, z_i \in \mathbb{C}$. In this chapter we study general properties of this canonical model.

6.1 Complex Synaptic Coefficients c_{ij}

It is sometimes convenient to rewrite (6.1) in polar coordinates: Let $z_i = r_i e^{i\phi_i}$, $b_i = \rho_i + i\omega_i$, $d_i = \alpha_i + i\beta_i$ and $c_{ij} = |c_{ij}|e^{i\psi_{ij}}$, then (6.1) is equivalent to the system

$$\begin{cases} r_{i}' = \rho_{i}r_{i} + \alpha_{i}r_{i}^{3} \\ + \sum_{j \neq i}^{n} |c_{ij}|r_{j}\cos(\phi_{j} + \psi_{ij} - \phi_{i}) \\ \phi_{i}' = \omega_{i} + \beta_{i}r_{i}^{2} \\ + \frac{1}{r_{i}}\sum_{j \neq i}^{n} |c_{ij}|r_{j}\sin(\phi_{j} + \psi_{ij} - \phi_{i}). \end{cases}$$
(6.2)

We see that $|c_{ij}|$ represents the strength of synaptic connections between the *j*-th and *i*-th neurons while $\psi_{ij} = \text{Arg } c_{ij}$ encodes phase information of the synaptic connections, which we call the network's *natural phase differences*.

Indeed, consider a network consisting of two identical neurons having synaptic connections only in one direction, for example, from z_2 to z_1 . Suppose also that $\operatorname{Im} d_i \equiv \beta_i = 0$, i.e. the frequency does not depend on the amplitude. Such a network is governed by the dynamical system

$$\begin{cases} r_1' = \rho r_1 + \alpha r_1^3 + |c_{12}| r_2 \cos(\phi_2 + \psi_{12} - \phi_1) \\ \phi_1' = \omega + \frac{1}{r_1} |c_{12}| r_2 \sin(\phi_2 + \psi_{12} - \phi_1) \\ r_2' = \rho r_2 + \alpha r_2^3 \\ \phi_2' = \omega. \end{cases}$$

The unique stable solution of this system satisfies

$$\phi_2(\tau) + \psi_{12} - \phi_1(\tau) \equiv 0 \mod 2\pi,$$

i.e. the oscillators have constant phase difference ψ_{12} . This motivates our definition of the natural phase difference since it occurs "naturally" in a pair of neurons connected in one direction.

Notice that when $\psi_{12} < 0$, an observer sees that one of the oscillators (in this case z_1) oscillates with some time delay. Obviously, it would be wrong to prescribe this to spike propagation or synaptic transmission delays. As we will see later, the coefficient ψ_{12} may take on many values depending on the synaptic organization of the network, but not the speed of transmission of spikes through axons or dendrites.

If the neurons are connected in both direction (i.e. $c_{12} \neq 0$ and $c_{21} \neq 0$), then the phase difference between them generically differs from $\operatorname{Arg} c_{12}$ or $\operatorname{Arg} c_{21}$. Nevertheless, even for a network of n interconnected neurons we can prove the following result

Lemma 39 If $c_{ij} \neq 0$, then there are values of the parameters ρ_1, \ldots, ρ_n such that

$$\phi_i(\tau) - \phi_j(\tau) \mod 2\pi \to \psi_{ij},\tag{6.3}$$

i.e. the i-th and j-th oscillators have constant phase difference ψ_{ij} .

Proof. Fix j and let $\rho_j = 1, \rho_i = -\frac{1}{\delta}$ for $i \neq j$. After rescaling $(z_i \rightarrow \delta z_i \text{ for } i \neq j)$ the system (6.1) transforms to

$$\begin{cases} z_j' = (1 + \mathrm{i}\omega_j)z_j + d_j z_j |z_j|^2 + \mathcal{O}(\delta), \\ z_i' = \frac{1}{\delta}(c_{ij}z_j - z_i) + \mathcal{O}(1), \quad i \neq j. \end{cases}$$

Applying singular perturbation methods (see Hoppensteadt (1993)) to the second equation we see that

$$z_i(\tau) = c_{ij} z_j(\tau) + \mathcal{O}(\delta, e^{-\frac{\tau}{\delta}}).$$

After a short initial transient period (e.g., $\tau > \mathcal{O}(|\delta \log \delta)|)$ this equation gives (6.3) for $\delta \to 0$. \Box

It follows from this proof that the choice of parameters corresponds to the case when the amplitude of j-th neuron is much bigger than that of the other neurons. One can consider the j-th oscillator as the leading one that synchronizes the whole network. A similar phenomenon was studied by Kazanovich and Borisyuk (1994).

The representation (6.2) of the canonical model is interesting since one can obtain valuable information about behavior of coupled oscillators simply by looking at (6.2), without any further mathematical analysis.

Indeed, it is easy to see that the impact of one oscillator on the amplitude of another one is maximal when their phases are synchronized so that (6.3) holds, since

$$\cos(\phi_j + \psi_{ij} - \phi_i) = \cos(\psi_{ij} - \psi_{ij}) = 1$$

reaches its maximal value then. If the two oscillators are completely out of phase, i.e. if

$$\phi_i(\tau) - \phi_j(\tau) \mod 2\pi = \psi_{ij} \pm \frac{\pi}{2},$$

then $\cos(\phi_j + \psi_{ij} - \phi_i) = 0$ and the influence of *j*-th neuron on the amplitude r_i of *i*-th neuron is negligible even when $|c_{ij}|$ is very large!

It is also easy to see that the larger is an oscillator amplitude r_j , the larger its impact on the other oscillators. Conversely, if *i*-th neuron has very small amplitude r_i , then it is susceptible to the influences of the other oscillators because of the term

$$\frac{1}{r_i}\sum_{j\neq i}^n |c_{ij}| r_j \sin(\phi_j + \psi_{ij} - \phi_i),$$

which can grow as $r_i \to 0$.

We next study some local properties of (6.1), in particular, the stability of the origin $z_1 = \cdots = z_n = 0$.

6.2 Oscillator Death and Self-ignition

Note that the canonical model (6.1) always has an equilibrium point $z_1 = \cdots = z_n = 0$ for any choice of parameters. If all ρ_i are negative numbers with sufficiently large absolute values, then the equilibrium point is stable.

In this section we study how the equilibrium can lose its stability. Using the canonical model we also illustrate two well-known phenomena: Oscillator death (or quenching, or Bar-Eli effect) and coupling-induced spontaneous activity (or self-ignition). We consider (6.1) for the most interesting case when $\text{Re } d_i < 0$ for all *i*. This corresponds to supercritical Andronov-Hopf bifurcation, i.e. to a birth of stable limit cycle.

We start from the observation that in the canonical model each oscillator is governed by a dynamical system of the form

$$z' = (\rho + \mathrm{i}\omega)z + dz|z|^2. \tag{6.4}$$

Obviously, if $\rho \leq 0$, then the equilibrium z = 0 is stable. As ρ increases through $\rho = 0$ (6.4) undergoes Andronov-Hopf bifurcation. As a result, the equilibrium losses its stability and (6.4) has stable limit cycle of radius $\mathcal{O}(\sqrt{\rho})$ for $\rho > 0$.

We can characterize the qualitative differences in dynamic behavior of (6.4) for $\rho \leq 0$ and $\rho > 0$ as follows:

- When $\rho \leq 0$, dynamical system (6.4) describes *intrinsically passive* element incapable of sustaining periodic activity.
- When ρ > 0, dynamical system (6.4) describes intrinsically active oscillator, or pacemaker.

In the uncoupled network (C = 0)

$$z_i' = (\rho + i\omega)z_i + d_i z_i |z_i|^2, \quad i = 1, \dots, n$$
 (6.5)

of such oscillators the equilibrium point $z_1 = \cdots = z_n = 0$ is stable for $\rho \le 0$ and unstable for $\rho > 0$

We ask what happens when we consider the canonical model (6.1) with nonzero matrix $C = (c_{ij})$? A partial answer is given in the following result.

Lemma 40 Let α denote the largest real part of all eigenvalues of connection matrix $C = (c_{ij})$. Consider the network of identical oscillators governed by

$$z_i' = (\rho + i\omega)z_i + d_i z_i |z_i|^2 + \sum_{j=1}^n c_{ij} z_j, \ i = 1, \dots, n.$$
(6.6)

The equilibrium point $z_1 = \cdots = z_n = 0$ is stable if

 $\rho < -\alpha$.

It is unstable if

$$\rho > -\alpha$$
.

Proof. The full system has the form

$$\begin{cases} z_i' = (\rho + i\omega)z_i + d_i z_i |z_i|^2 + \sum_{j=1}^n c_{ij} z_j, \\ \bar{z}'_i = (\rho - i\omega)\bar{z}_i + \bar{d}_i \bar{z}_i |z_i|^2 + \sum_{j=1}^n \bar{c}_{ij} \bar{z}_j, \end{cases} i = 1, \dots, n.$$
(6.7)

It is easy to see that the origin $z_1 = \overline{z}_1 = \cdots = z_n = \overline{z}_n = 0$ is always an equilibrium point of (6.7). The $(2n) \times (2n)$ Jacobian matrix J at the origin has the form

$$J = \begin{pmatrix} (\rho + i\omega)E + C & 0\\ 0 & (\rho - i\omega)E + \bar{C} \end{pmatrix},$$

where 0 and E are the $n \times n$ zero and identity matrices, respectively.

Suppose $\lambda_1, \ldots, \lambda_n$ are eigenvalues of C counted with their multiplicity and suppose that v_1, \ldots, v_n are corresponding (generalized) eigenvectors. Direct computation shows that J has 2n eigenvalues

$$\rho + i\omega + \lambda_1, \quad \rho - i\omega + \bar{\lambda}_1, \dots, \ \rho + i\omega + \lambda_n, \quad \rho - i\omega + \bar{\lambda}_n$$

and 2n corresponding eigenvectors

$$\left(\begin{array}{c}v_1\\0\end{array}\right), \left(\begin{array}{c}0\\\bar{v}_1\end{array}\right), \ldots, \left(\begin{array}{c}v_n\\0\end{array}\right), \left(\begin{array}{c}0\\\bar{v}_n\end{array}\right),$$

where 0 denotes a vector of zeros. Stability of (6.7) is determined by the eigenvalues of J with maximal real part. These eigenvalues are of the form

$$\rho + i\omega + \lambda, \quad \rho - i\omega + \overline{\lambda},$$

where λ are eigenvalues of C. Let $\alpha = \text{Re } \lambda$, then the origin is stable if

$$\rho + \alpha < 0,$$

and unstable if

$$\rho + \alpha > 0.$$

In the rest of this section we use Lemma 40 to illustrate two interesting effects.

• If $\alpha < 0$, then the network (6.6) is stable even when

$$0 < \rho < -\alpha$$
.

That is, even though each oscillator is a pacemaker, the coupled system may approach z = 0. This effect, which can be called *oscillator death*, was studied numerically by Bar-Eli (1985) and analytically for general systems by Aronson et al. (1990).

• If $\alpha > 0$, then the network (6.6) can exhibit spontaneous activity even when

$$-\alpha < \rho < 0,$$

i.e. when each oscillator is intrinsically passive, coupling can induce synchronous activity in the network. This effect, which can be called *self-ignition* (Rapaport 1952), is discussed in details by Kowalski et al. (1992).

Remark 41 If C has only one eigenvalue with maximal real part, then (6.6) undergoes an Andronov-Hopf bifurcation as ρ increases through $\rho = -\alpha$.

Note that in this case the coordinates of the limit cycle depend upon the matrix C; more precisely, upon the eigenvector that corresponds to the "leading" eigenvalue of C. Thus, to understand the dynamics of the canonical model (6.6), one should understand possible structures of the connection matrix $C = (c_{ij})$. We will do this in Chapter 10.

6.3 Synchronization and Convergence

In this section we reveal the conditions under which the canonical model (6.1) can operate as an MA-type NN (see Section 1.4).

First, we assume that all d_i are real and negative. Without loss of generality we may take $d_i = -1$. Thus, we study dynamical system of the form

$$z'_{i} = (\rho_{i} + \omega_{i})z_{i} - z_{i}|z_{i}|^{2} + \sum_{j \neq i}^{n} c_{ij}z_{j}.$$
(6.8)
We take advantage of the fact that the system is invariant under rotations to prove the following theorem.

Theorem 42 (COHEN-GROSSBERG CONVERGENCE THEOREM FOR OSCILLATORY NEURAL NETWORKS) If in the canonical model (6.8) all neurons have equal center frequencies $\omega_1 = \cdots = \omega_n = \omega$ and the matrix of synaptic connections $C = (c_{ij})$ is self-adjoint, i.e.

$$c_{ij}=\bar{c}_{ji},$$

then the neural network dynamics converges to a limit cycle. On the limit cycle all neurons have constant phase shifts, which corresponds to synchronization of the network activity.

Proof. In the rotating coordinate system $u_i = e^{-i\omega\tau} z_i(\tau)$ (6.8) becomes

$$u_i' = \rho_i u_i - u_i |u_i|^2 + \sum_{j=1}^n c_{ij} u_j, \quad i = 1, \dots, n.$$
(6.9)

Note that the mapping $U:\mathbb{C}^{2n}\to\mathbb{R}$ given by

$$U(u_1, ..., u_n, \bar{u}_1, ..., \bar{u}_n) = -\sum_{i=1}^n \left(\rho_i |u_i|^2 - \frac{1}{2} |u_i|^4 + \sum_{j=1}^n c_{ij} \bar{u}_i u_j \right)$$

is a global Liapunov function for (6.9). Indeed, it is continuous, bounded below (because it behaves like $\frac{1}{2}|u|^4$ for large u), satisfies

$$u_i' = -\frac{\partial U}{\partial \bar{u}_i}, \quad \bar{u}_i' = -\frac{\partial U}{\partial u_i},$$

and, hence,

$$\frac{dU}{d\tau} = \sum_{i=1}^{n} \left(\frac{\partial U}{\partial u_i} u_i' + \frac{\partial U}{\partial \bar{u}_i} \bar{u}_i' \right) = -2 \sum_{i=1}^{n} |u_i'|^2 \le 0.$$

Notice that $\frac{dU}{d\tau} = 0$ precisely when $u_1' = \cdots = u_n' = 0$, i.e. at the equilibrium point of (6.9). Let $u^* \in \mathbb{C}^n$ be such a point. Then, while the solution u of (6.9) converges

to u^* , the solution of (6.8) converges to the limit cycle $z = e^{i\omega\tau}u^*$. Obviously, any pair of oscillators have constant phase difference on this limit cycle. \Box

It should be noted that dynamics of (6.8) can converge to different limit cycles depending upon the initial conditions and the choice of the parameters ρ_1, \ldots, ρ_n . For fixed parameters there could be many such limit cycles corresponding to different memorized images (Baird 1986, Li and Hopfield 1989).

The theorem states that the canonical model (6.8) is an MA-type NN model (Hopfield 1982, Grossberg 1988), but instead of equilibrium points, the network dynamics converge to limit cycles, as was postulated by Baird (1986). Whether this new feature renders the canonical model any advantages over the classical Hopfield model or not is still an open question.

We continue our studying of the canonical model in Chapter 10.

Chapter 7 Multiple Cusp Singularity

We analyze the canonical model (2.21) for $\varepsilon \to 0$. Thus, we study the system

$$x'_{i} = r_{i} + b_{i}x_{i} + \sigma_{i}x_{i}^{3} + \sum_{j=1}^{n} c_{ij}x_{j}, \qquad (7.1)$$

where x_i, r_i, b_i, c_{ij} are real variables and $\sigma_i = \pm 1$.

Before proceeding further, we explain the meaning of the data in (7.1). Each x_i depends on X_i and is a scalar that describes in some sense activity of the *i*-th neuron, and the vector $x = (x_1, \ldots, x_n)^{\top} \in \mathbb{R}^n$ describes a physiological state of the network; The parameter $r_i \in \mathbb{R}$ is an external input from receptors to the *i*-th neuron. It depends on λ and ρ . Each $b_i \in \mathbb{R}$ is an internal parameter, which also depends upon λ and ρ . The vector $(b_1, \ldots, b_n)^{\top} \in \mathbb{R}^n$ is a multidimensional bifurcation parameter; $C = (c_{ij}) \in \mathbb{R}^{n \times n}$ is a matrix of synaptic connections between neurons. There is strong neurobiological evidence that synapses are responsible for associative memorization and recall (Shepherd 1983). We will see this realized in the canonical model (7.1) but from a rigorous mathematical point of view.

It should be stressed that (7.1) is an interesting dynamical system in itself without any connection to the WCNN theory. It exhibits useful behavior from a computational point of view and deserves to be studied per se.

The simplicity of (7.1) (in comparison with (2.1)) is misleading. It is very difficult

to study it for an arbitrary choice of the parameters. We will study it using bifurcation theory. Nevertheless, we can answer interesting questions about (7.1) only by making some assumptions about the parameters $\{r_i\}$, $\{b_i\}$ and C. In Section 7.1.2 we assume that the input from receptors is very strong, i.e. all r_i have large absolute values. In Section 7.1.3 we consider extreme choices for b_i . Section 7.3 is devoted to the study of (7.1) under the classical assumption that the matrix of synaptic connections Cis symmetric. This restriction arises naturally when one considers Hebbian learning rules, which are discussed in Section 7.4.

In Section 7.5 we study bifurcations in the canonical model (7.1) when the external input $r_i = 0$ for all *i*. In this case (7.1) coincides with the canonical model for multiple pitchfork bifurcations in the WCNNs. Section 7.6 is devoted to studying the canonical model when only one or two images are memorized. In Sections 7.7 and 7.8 we illustrate the phenomena of bistability of perception and decision making in analogy with problems in psychology.

We start studying (7.1) by asking the following question: What is the behavior of its solutions far from the origin x = 0, i.e. outside some ball $B_0(R) \subset \mathbb{R}^n$ with large radius R?

7.1 Extreme Values of Parameters

7.1.1 Global behavior

Let $B_0(R) \subset \mathbb{R}^n$ denote a ball with center at the origin and radius R. By the term "global behavior" we mean a flow structure of a dynamical system outside the ball $B_0(R)$ for sufficiently large R.

A dynamical system is bounded if there is R > 0 such that $B_0(R)$ attracts all trajectories, i.e. for any initial condition x(0) there exists t_0 such that $x(t) \in B_0(R)$ for all $t \ge t_0$. Obviously, all attractors of such a system lie inside $B_0(R)$.

Theorem 43 A necessary and sufficient condition for (7.1) to be bounded is that $\sigma_1 = \cdots = \sigma_n = -1$, i.e. (7.1) must be

$$x'_{i} = r_{i} + b_{i}x_{i} - x_{i}^{3} + \sum_{j=1}^{n} c_{ij}x_{j}.$$
(7.2)

This corresponds to a multiple supercritical cusp singularity.

Proof. We are interested in the flow structure of (7.1) outside some ball $B_0(R)$ with sufficiently large radius R. Let $\varepsilon = R^{-2}$ be a small parameter. After the rescaling $x_i \to \sqrt{\varepsilon} x_i, t \to \varepsilon^{-1} t$, (7.1) can be rewritten as

$$x'_i = \sigma_i x_i^3 + \varepsilon (b_i x_i + \sum_{j=1}^n c_{ij} x_j + \sqrt{\varepsilon} r_i),$$

which must be studied outside the unit ball $B_0(1)$.

Note that this is an ε -perturbation of the uncoupled system

$$x'_i = \sigma_i x_i^3, \qquad 1 \le i \le n. \tag{7.3}$$

Obviously, the unit ball $B_0(1)$ attracts all trajectories of (7.1) if and only if all $\sigma_i < 0$. Any ε -perturbation of (7.3) has the same property provided ε is small enough. \Box

So, the flow structure of (7.1) outside a ball with sufficiently large radius looks like that of (7.3), as depicted in Figure 7.1a for the case $\sigma_1 = \sigma_2 = -1$ and in Figure 7.1b when σ_1 , is positive.

To be bounded is a desirable property in applications. Any initial condition x(0)lies in a domain of attraction of some attractor that lies somewhere inside $B_0(R)$. Hence, for any x(0) we have at least a hope to find the asymptotic dynamics. From now on we will consider only (7.2) as the canonical model of the WCNN near a multiple cusp singularity point, i.e. we study the supercritical cusp singularity.



Figure 7.1: Global flow structures of the canonical models. a. System (7.1) for $\sigma_1 = \sigma_2 = -1$ is bounded. b. System (7.1) for $\sigma_1 = +1$, $\sigma_2 = -1$ is not bounded.

7.1.2 Strong Input From Receptors

What is the behavior of the canonical model (7.2) when the external input from receptors is very strong, i.e. when the parameter

$$R = \min |r_i|$$

is very large?

Let $\varepsilon = R^{-\frac{2}{3}}$ and rescale variables by setting: $x_i = \varepsilon^{\frac{1}{2}} \tilde{x}_i$, $r_i = \varepsilon^{\frac{3}{2}} \tilde{r}_i$ and $\tau = \varepsilon t$. Then (7.2) can be rewritten as (after dropping ~)

$$x'_{i} = r_{i} - x_{i}^{3} + \varepsilon (b_{i}x_{i} + \sum_{j=1}^{n} c_{ij}x_{j}).$$
(7.4)

System (7.4) is ε -perturbation of the uncoupled system

$$x'_{i} = r_{i} - x_{i}^{3}. (7.5)$$

It is obvious that (7.5) has only one equilibrium point $x = (r_1^{\frac{1}{3}}, \ldots, r_n^{\frac{1}{3}})^{\mathsf{T}}$ for any external input r_1, \ldots, r_n . The Jacobian at that point is

$$L = \begin{pmatrix} -3r_1^{\frac{4}{3}} & 0 & \cdots & 0\\ 0 & -3r_2^{\frac{2}{3}} & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & -3r_n^{\frac{2}{3}} \end{pmatrix}.$$
 (7.6)

Note that according to the rescaling all $|r_i| \ge 1$, hence all diagonal elements in L are negative. Thus, the equilibrium point is a stable node.

All of the equilibrium points considered above are hyperbolic. Any ε -perturbations of (7.5) do not change the qualitative picture provided ε is small enough. So, the phase portrait of (7.2) for strong external inputs is qualitatively the same as that of (7.5).

Thus, when the input from receptors is strong (in comparison with the synaptic connections c_{ij} or the internal parameters b_i), then the network dynamics approach the unique equilibrium. At this equilibrium each neuron is either depolarized (excited) or hyperpolarized (inhibited) depending on the sign of the input r_i .

7.1.3 Extreme Psychological Condition

It is convenient to assume that the set of parameters $\{b_i\}$ describes the psychological state of the WCNN because it affects the way that the network reacts to external inputs. One might speculate that when

$$\beta = \min_i |b_i|$$

is very large, the network is working in an "extreme psychological condition".

We use the same method of analysis of (7.2) as in previous sections. Let $\varepsilon = \beta^{-1}$ be a small parameter. By rescaling $(x_i = \sqrt{\varepsilon} \tilde{x}_i, b_i = \varepsilon \tilde{b}_i, \tau = \varepsilon t)$ we can rewrite (7.2) as

$$x'_{i} = b_{i}x_{i} - x_{i}^{3} + \varepsilon (\sum_{j=1}^{n} c_{ij}x_{j} + \sqrt{\varepsilon}r_{i}).$$

$$(7.7)$$

The weakly connected system (7.7) is an ε -perturbation of the uncoupled system

$$x_i' = b_i x_i - x_i^3. (7.8)$$

In (7.8) each equation has either one or three equilibrium points depending upon



Figure 7.2: Phase portrait of (7.2) working in the extreme psychological regime. a. All $b_i > 0$. b. $b_1 > 0$, $b_2 < 0$.

the sign of b_i . The points are $x_i = 0$ for any b_i and in addition $x_i = \pm \sqrt{b_i}$ for $b_i > 0$. The Jacobian of (7.8) is

$$\begin{pmatrix} b_1 - 3x_1^2 & 0 & \cdots & 0 \\ 0 & b_2 - 3x_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_n - 3x_n^2 \end{pmatrix}.$$

Obviously, the equilibrium point $x = (x_1, \ldots, x_n)$ of (7.8) is a stable node if

$$x_i = \begin{cases} 0 & \text{if } b_i < 0\\ \pm \sqrt{b_i} & \text{if } b_i > 0. \end{cases}$$

It is a saddle if some (but not all) of these conditions on x_i are violated, and it is an unstable node if all of the conditions are violated. Note that the presence of saddles and the unstable node is possible only if there exist at least one positive b_i . If all $b_i < 0$, then there is only one equilibrium point $x = (0, ..., 0)^{\mathsf{T}}$, which is the stable node (see Figure 7.2).

7.2 Canonical Models as a GAS-Type NN

Systems that have only one asymptotically stable equilibrium point and do not have any other attractors are called *globally asymptotically stable* systems (Hirsch 1989). Such systems are good candidates for GAS-type NNs (see Section 1.4). We have already seen that (7.2) is globally asymptotically stable when the input r from receptors is strong. (The values of the other parameters, viz. b and C, are irrelevant). This is also true when (7.2) is considered for negative b_i with large absolute values.

In the previous section we have assumed that all $|b_i|$ are large. The remarkable fact is that for (7.2) to be globally asymptotically stable it is suffices to require that $b = (b_1, \ldots, b_n)^{\mathsf{T}}$ take only intermediate values comparable to those of entries of C. A more accurate statement is the following:

Theorem 44 The dynamical system (7.2)

$$x'_{i} = r_{i} + b_{i}x_{i} - x_{i}^{3} + \sum_{j=1}^{n} c_{ij}x_{j}$$

is globally asymptotically stable if

$$b_i < -\left(c_{ii} + \frac{1}{2}\sum_{j \neq i} |c_{ij} + c_{ji}|\right)$$
(7.9)

for all $1 \leq i \leq n$.

Proof. We will prove it using Hirsch's theorem (Hirsch 1989). Let L be the Jacobian of (7.2) at a point $x = (x_1, \ldots, x_n)^{\mathsf{T}}$. Hirsch's theorem claims that if there is a constant $-\delta < 0$ such that

$$\langle L\xi,\xi\rangle \leq -\delta\langle\xi,\xi\rangle$$

for all $\xi = (\xi_1, \dots, \xi_n)^{\mathsf{T}} \in \mathbb{R}^n$ then (7.2) is globally asymptotically stable. Here $\langle \xi, \eta \rangle$ denote the inner (dot) product of vectors ξ and η .

It is easy to check that

$$\langle L\xi,\xi\rangle = \sum_{i=1}^{n} (b_i - 3x_i^2)\xi_i^2 + \sum_{i,j=1}^{n} c_{ij}\xi_i\xi_j$$

$$=\sum_{i=1}^{n} (b_{i} + c_{ii} - 3x_{i}^{2})\xi_{i}^{2} + \frac{1}{2}\sum_{\substack{i,j=1\\i\neq j}}^{n} (c_{ij} + c_{ji})\xi_{i}\xi_{j}$$

$$\leq \sum_{i=1}^{n} (b_{i} + c_{ii})\xi_{i}^{2} + \frac{1}{4}\sum_{\substack{i,j=1\\i\neq j}}^{n} |c_{ij} + c_{ji}| (\xi_{i}^{2} + \xi_{j}^{2})$$

$$= \sum_{i=1}^{n} \left(b_{i} + c_{ii} + \frac{1}{2}\sum_{j\neq i} |c_{ij} + c_{ji}| \right) \xi_{i}^{2} < -\delta \sum_{i=1}^{n} \xi_{i}^{2}$$

where $-\delta = \max_i (b_i + c_{ii} + \frac{1}{2} \sum_{j \neq i} |c_{ij} + c_{ji}|)$. We used here the inequality

$$\xi_i \xi_j \leq \frac{1}{2} (\xi_i^2 + \xi_j^2).$$

The inequality (7.9) guarantees that $-\delta < 0$ and, hence, all the conditions of Hirsch's theorem are satisfied. This completes the proof. \Box

Note that the inequality (7.9) is much more appealing than the requirement that the absolute values of b_i be arbitrary large. Thus, even for "reasonable" values of the internal parameters b_i , the dynamics of (7.2) is globally asymptotically stable.

Another remarkable fact is that the external input $r \in \mathbb{R}^n$ does not come into the condition (7.9). What it does affect is the location of the unique attractor of the network. Therefore, the canonical model (7.2) for b_i satisfying (7.9) can work as a GAS-type NN.

7.3 Symmetric Synaptic Connections

In our analysis of (7.2) above we have assumed that some of its parameters take their extreme values. The dynamics of (7.2) always converges in these cases, i.e. its attractors are stable equilibrium points. Next we consider (7.2) without the requirement that any of its parameters take extreme, possibly implausible values.

Theorem 45 (Cohen-Grossberg) If the matrix of synaptic connections C is symmetric, then the neural network

$$x'_{i} = r_{i} + b_{i}x_{i} - x_{i}^{3} + \sum_{j=1}^{n} c_{ij}x_{j}$$

is a gradient system.

The symmetry of C is a strong requirement, but it arises naturally if one considers the Hebbian learning rules, which will be discussed in the next section. Such systems are widely studied in the neural network literature (see review by S. Grossberg 1988).

Proof. To prove the theorem it suffices to present a function $U: \mathbb{R}^n \to \mathbb{R}$ satisfying

$$x_i' = -\frac{\partial U}{\partial x_i}.$$

It is easy to check that

$$U(x) = -\sum_{i=1}^{n} (r_i x_i + \frac{1}{2} b_i x_i^2 - \frac{1}{4} x_i^4) - \frac{1}{2} \sum_{i,j=1}^{n} c_{ij} x_i x_j$$

is such a function for (7.2). \Box

Note that far away from the origin U(x) behaves like $\frac{1}{4} \sum_{i=1}^{n} x_i^4$, hence U(x) is bounded below.

Being a gradient system imposes many restrictions on the possible dynamics of (7.2). For example, its dynamics cannot be oscillatory or chaotic. Nevertheless, this property is considered to be very useful from computational point of view, which we discuss later.

To what type of NNs does the canonical model (7.2) belong? It is clear from previous sections that for one choice of the parameters (7.2) has many attractors and, hence, is a candidate for MA-type NN, whereas for other choices of the parameters it is globally asymptotically stable and, hence, is the GAS-type NN. We will show in the next section that (7.2) stands somewhere between MA and GAS-types and, hence, will be considered as a new NN type.

We are interested in the basic principles of the human brain functioning. Hence, we will study only the qualitative behavior of the canonical model and neglect quantitative features. The main tools in the analysis below come from bifurcation theory. Unfortunately, comprehensive analysis of the model for arbitrary r_i , b_i and C is formidable unless we impose some additional restrictions onto the parameter spaces. Among them there are two we discuss now.

First of all, we will study the canonical models when

$$b_1=\cdots=b_n=b.$$

Thus, instead of n bifurcation parameters b_1, \ldots, b_n we have only one $b \in \mathbb{R}$.

The second assumption concerns the matrix of synaptic connections C, which is responsible for learning in the NNs.

7.4 Hebbian Learning Rule For Synaptic Matrix C

Little is known about the processes of learning in a human brain. One basic approach to studying learning and recognition in artificial NNs is to assume that the synaptic matrix is constructed according to Hebb's learning rule (Hebb 1949)

$$c_{ij} = \frac{1}{n} \sum_{s=1}^{m} \beta_s \xi_i^s \xi_j^s, \qquad 1 \le i, j \le n,$$
(7.10)

where $\xi^s = (\xi_1^s, \dots, \xi_n^s)^{\mathsf{T}} \in \mathbb{R}^n$, $s = 1, \dots, m$ are key patterns to be memorized; The constants β_s measure "strength" or "quality" of the memory about the patterns ξ^s ; The number of memorized patterns, m, cannot exceed n. The Hebbian learning rule

(7.10) can be rewritten in the more convenient way

$$C = \frac{1}{n} \sum_{s=1}^{m} \beta_s \xi^s (\xi^s)^{\mathsf{T}}, \qquad (7.11)$$

where $^{\top}$ means transpose.

It is easy to see that the synaptic matrix C constructed according to (7.11) is symmetric. It is also true that any symmetric matrix C can be represented as (7.11) for some, possibly non-unique, choice of the orthogonal vectors ξ^1, \ldots, ξ^m . Thus, we have proved the following

Proposition 46 The matrix of synaptic connections C is symmetric if and only if there is a set of orthogonal patterns ξ^1, \ldots, ξ^m such that C is constructed according to the Hebbian learning rule (7.11).

Note that the orthogonal vectors ξ^1, \ldots, ξ^m are eigenvectors of C. If we normalize them such that

$$|\xi^s|^2 = \langle \xi^s, \xi^s \rangle = \sum_{i=1}^n (\xi^s_i)^2 = n,$$

then the constants β_s are eigenvalues of C and (7.11) is the spectral decomposition of C. We assume that $\beta_1 \geq \ldots \geq \beta_m > 0$. If m < n, then there is an n - mdimensional eigenspace ker $C \subset \mathbb{R}^n$ corresponding to the zero eigenvalue. We denote this eigenvalue by $\beta_{m+1} = 0$.

To summarize we can say that the Hebbian learning rule for orthogonal patterns gives a way of constructing the matrix of synaptic connections such that each pattern is an eigenvector of the matrix corresponding to positive eigenvalue.

In our analysis below we will assume that the learning rule is Hebbian. This assumption imposes significant restrictions on possible dynamic behavior of the canonical model. For example, it follows from the Section 7.3 that (7.2) is a gradient dynamical system. In order to make all computations without resorting to computer simulations we also assume that $|\xi_i^s| = 1$ for all *i*, i.e.

$$\xi^s = (\pm 1, \dots, \pm 1)^{\mathsf{T}}, \qquad 1 \le s \le m.$$

For these purposes we introduce the set

$$\Xi^n = \{\xi \in \mathbb{R}^n, \ \xi = (\pm 1, \dots, \pm 1)^\top\} \subset \mathbb{R}^n.$$

$$(7.12)$$

We will also need for our analysis an orthogonal basis for \mathbb{R}^n that contains vectors only from Ξ^n . This basis always exists if $n = 2^k$ for some integer k > 0. The assumption that $\xi^s \in \Xi^n$ might look artificial, but it is very important in neurocomputer applications and in digital circuit design.

Recall that we are interested in qualitative behavior of the canonical models. All attractors that we will study below are hyperbolic. Hence, if we perturb the parameters r_1, \ldots, r_n , b_1, \ldots, b_n and C, i.e. if we violate the assumptions made above, the qualitative behavior will be the same provided the perturbations are not very large.

7.5 Bifurcations for r = 0

We start the bifurcation analysis of the canonical model (7.2) for the special case when there are no receptor inputs, i.e. when $r_1 = \cdots = r_n = 0$. Thus, we are interested in qualitative behavior of the dynamical system

$$x'_{i} = bx_{i} - x_{i}^{3} + \sum_{j=1}^{n} c_{ij}x_{j} \quad 1 \le i \le n.$$
(7.13)

It is easy to check that the canonical model (7.13) describes a WCNN with Z_2 symmetry $x \to -x$ near a *multiple pitchfork* bifurcation point. Symmetry means that for b > 0 each neuron is essentially a bistable element with two states: excitation $(x_i = \sqrt{b})$ and inhibition $(x_i = -\sqrt{b})$.

7.5.1 Stability of the Origin

Note that (7.13) always has a equilibrium point $x_1 = \cdots = x_n = 0$. It follows from Section 7.1.3, that the origin is the only equilibrium point, which is a stable node, for b << -1, or for b satisfying (7.9). We also know that for $b \gg 1$ the canonical model has many attractors (see Figure 7.2). What we do not know is the behavior of the model for intermediate values of b.

Thus, we have the following questions: What happens while b is increasing? How does the origin lose its stability? How many and of what type are the new equilibrium points? What is the relationship between them and the synaptic matrix $C = (c_{ij})$? These and other questions are studied in this section.

Let L be the Jacobian of the right-hand side of (7.13) at the origin. It is easy to see that

$$L = bE + C,$$

where E is the unit matrix. Let β_1, \ldots, β_m be the (distinct) eigenvalues of C ordered such that

$$\operatorname{Re}\beta_1 \geq \cdots \geq \operatorname{Re}\beta_m$$

counted with multiplicity. Obviously, L has m eigenvalues

$$\lambda_s = b + \beta_s, \qquad 1 \le s \le m$$

with the same eigenvectors as those of C. The matrix L has all eigenvalues with negative real parts, and, hence, the origin is a stable equilibrium point for (7.13) if and only if $b < -\text{Re}\beta_1$ (see Figure 7.3a).



Figure 7.3: Phase portrait of the canonical model (7.13) of weakly connected neural network near multiple pitchfork bifurcation point for different values of the bifurcation parameter b. a. $b < -\beta_1$. b. $-\beta_1 < b < -\beta_2$. c. $-\beta_2 < b < -\beta_2 + (\beta_1 - \beta_2)/2$. d. $-\beta_2 + (\beta_1 - \beta_2)/2 < b$.

If β_1 is a real eigenvalue with multiplicity one, then (7.13) undergoes a pitchfork bifurcation when b crosses $-\beta_1$. For b slightly larger than $-\beta_1$ the origin is a saddle surrounded by two sinks (see Figure 7.3b) and those are the only equilibrium points for (7.13).

If (β_1, β_2) is a pair of complex conjugate eigenvalues with multiplicity one, then we can observe the Andronov-Hopf bifurcation for $b = -\text{Re}\beta_1$.

For $b > -\beta_1$ it is possible to observe the birth of a pair of saddles or an unstable limit cycle every time b crosses $-\text{Re}\beta_s$, where β_s is an eigenvalue with multiplicity one.

For the eigenvalues with multiplicity more than one bifurcations could be more complicated. Nevertheless, we will consider some of them later.

Recall that each neuron is bistable only for b > 0. For negative b there is only one stable state $x_i = 0$ and, hence, it is "passive". But when the neurons are connected they acquire a new property: bistability for $-\beta_1 < b < 0$. This is the property that each neuron alone cannot have. Thus a network of "passive" elements can exhibit "active" properties. This is called self-ignition and has already been studied in Section 6.2 for oscillatory neural networks. We encounter this phenomena frequently in our analysis of brain function.

7.5.2 Stability of the Other Equilibria

It is noteworthy that we have not restricted the synaptic matrix C yet. All the bifurcations discussed above take place for any C. In return for this generality, we cannot trace the new equilibrium points and study their stability. Fortunately, we can do it if we assume that the synaptic matrix C is constructed according to the Hebbian learning rule

$$C = \frac{1}{n} \sum_{s=1}^{m} \beta_s \xi^s (\xi^s)^{\mathsf{T}}, \qquad \beta_1 \ge \cdots \ge \beta_m > 0$$

and that the memorized images $\xi^1, \ldots, \xi^m \in \Xi^n$ are orthogonal. For simplicity we assume that all β_s are different. At the end of this section we will discuss the case $\beta_1 = \cdots = \beta_m$. Let $x_i = y_s \xi_i^s$ for i = 1, ..., n. Then

$$x'_{i} = y'_{s}\xi^{s}_{i} = by_{s}\xi^{s}_{i} - y^{3}_{s}(\xi^{s}_{i})^{3} + \beta_{s}y_{s}\xi^{s}_{i}.$$

After (dot) multiplication by ξ_i^s , we have

$$y'_{s} = (b + \beta_{s})y_{s} - y_{s}^{3}.$$
(7.14)

This equation has only one equilibrium point $y_s = 0$ for $b < -\beta_s$ and for $b > -\beta_s$ there are three points $y_s = 0$, $y_s = \pm \sqrt{b + \beta_s}$. Hence the original system (7.13) has two new equilibrium points $x = \pm \sqrt{b + \beta_s} \xi^s$ after b crosses $-\beta_s$.

Note that the pair of new equilibrium points lies on the line spanned by the memorized pattern ξ^s . Every attractor lying on or near $span(\xi^s)$ is called an attractor corresponding to the pattern ξ^s . When the network activity x(t) approaches such an attractor, we say that the NN has recognized the memorized image ξ^s .

To our surprise only the pair $x = \pm \sqrt{b + \beta} \xi^1$ is a pair of stable nodes, whereas

the others $x = \pm \sqrt{b + \beta_s} \xi^s$, $s \ge 2$ are pairs of saddles, at least when b is near $-\beta_s$ (see Figure 7.3c).

Let us study the stability of $x = \pm \sqrt{b + \beta_k} \xi^k$ for some $1 \le k \le m$. The matrix of linearization L at this point is

$$L = (b - 3(b + \beta_k))E + C$$

It has eigenvalues

$$\lambda_s = -2(b+\beta_k) + \beta_s - \beta_k, \qquad 1 \le s \le m+1.$$

where $\beta_{m+1} = 0$ corresponds to ker C. Note that $\lambda_1 \ge \cdots \ge \lambda_m > \lambda_{m+1}$. The maximum eigenvalue $\lambda_1 = -2(b + \beta_k) + \beta_1 - \beta_k$ is always negative only for k = 1. For $k \ge 2$ the inequality $\lambda_1 < 0$ gives us the condition

$$b > -\beta_k + \frac{\beta_1 - \beta_k}{2}.\tag{7.15}$$

One could say that in the life of the equilibrium point $\sqrt{b + \beta_k} \xi^k$ there are two major events: Birth (when $b = -\beta_k$) and maturation ($b = -\beta_m + (\beta_1 - \beta_m)/2$) when the point becomes a stable node. For k = 2 see Figure 7.3d and 7.4.

It is easy to see that when $b = -\beta_k$ the eigenvalues of L are

$$\lambda_1 \geq \cdots \geq \lambda_{k-1} \geq 0 = \lambda_k \geq \lambda_{k+1} \geq \cdots \geq \lambda_m$$

So, $\sqrt{b+\beta_k} \xi^k$ is the saddle such that k-1 directions corresponding to ξ^1, \ldots, ξ^{k-1} are unstable (see Figure 7.5a). Every time *b* crosses $-\beta_k + (\beta_s - \beta_k)/2$, s < k there is a pitchfork bifurcation. As a result, the direction corresponding to ξ^s becomes stable and there appears a new pair of saddles lying in span (ξ^s, ξ^k) (see Figure 7.5b,c and d).

To summarize, we can say that for $b < -\beta_1$ the only equilibrium point is the origin, but for $b > -\beta_m + (\beta_1 - \beta_m)/2$ there are *m* pairs of stable nodes corresponding to the memorized images ξ^1, \ldots, ξ^m and many saddles lying in between these nodes.



Figure 7.4: Bifurcation diagram.



Figure 7.5: Every equilibrium point $\pm \sqrt{b + \beta_k} \xi^k$ becomes an attractor after the sequence of the pitchfork bifurcations. Every time b crosses $-\beta_k + (\beta_s - \beta_k)/2$, s < k, the ξ^s -direction becomes stable.

Recall that we referred to the attractors that do not correspond to any of the memorized images as being spurious memory. Is there any spurious memory in (7.13)? The answer is *yes.* Fortunately, it happens for large *b*. Indeed, when b > 0 all eigenvalues of the matrix of linearization at the origin are positive, not only $\lambda_1, \ldots, \lambda_m$. The new unstable directions correspond to ker *C* (of course, if m < n). It is easy to check that for $0 < b < \frac{\beta_1}{2}$ all the equilibrium points lying in ker *C* are saddles (except the origin, which is an unstable node), whereas for $b > \frac{\beta_1}{2}$ there are 2(n - m) stable nodes among them.

In order to avoid spurious memory, one should keep the bifurcation parameter below the critical value $\frac{\beta_1}{2}$. Actually, it is more reasonable to demand that b be negative. By this means we guarantee that nothing interesting is going on in the directions orthogonal to the all memorized images. But we must be cautious because not all equilibrium points corresponding to memorized patterns are stable for b < 0. Indeed, the stability condition (7.15) for b < 0 can be satisfied only if

$$\beta_k > \frac{\beta_1}{3}.$$

Thus, all memorized images are stable nodes and successful recognition is possible if the weight β_m of the weakest image is greater than one third of that of the strongest one.

Obviously, we do not have this kind of problem when $\beta_1 = \cdots = \beta_m = \beta > 0$. For $b < -\beta$ the NN is globally asymptotically stable. For $b = -\beta$ there is a multiple pitchfork bifurcation with the birth of 2m stable nodes corresponding to the memorized images. For $-\beta < b < 0$ these nodes are the only attractors¹ and behavior of the NN (7.13) is very simple in the directions orthogonal to the span (ξ^1, \ldots, ξ^m) . Thus, the system (7.13) can work as a typical MA-type NN. If the initial condition

¹Actually, the correct statement is that these are the only attractors that bifurcated from the origin. Whether there are other attractors or not is still an open question.

x(0) is an input from receptors, then the activity x(t) of (7.13) approaches the closest attractor, which corresponds to one of the previously memorized images. It is believed that this simple procedure is a basis for construction of new generations of computers — neurocomputers.

Nevertheless, our opinion is that this is too far from the basic principles of how a real brain functions (despite the fact that we know almost nothing about these principles). In the next section we explore another, more realistic, approach.

7.6 Bifurcations for $r \neq 0$ (two memorized images)

As we have already mentioned, comprehensive analysis of the canonical model

$$x'_{i} = r_{i} + bx_{i} - x_{i}^{3} + \sum_{j=1}^{n} c_{ij}x_{j}$$
(7.16)

is formidable. Even for a symmetric synaptic matrix $C = (c_{ij})$ it is difficult although there is a Liapunov function in that case. Hence, every piece of information about (7.16) obtained by analytical tools is precious.

The next step in studying (7.16) is to assume that the number of memorized images $m \leq 2$. The key result in this direction is the Reduction Lemma that enables us to reduce the number of independent variables to 2.

7.6.1 The Reduction Lemma

Suppose that $\xi^1, \ldots, \xi^n \in \Xi^n$ form an orthogonal basis for \mathbb{R}^n and that ξ^1 and ξ^2 coincide with memorized images, where Ξ is defined in (7.12). Let

$$y_s = \frac{1}{n} \langle x, \xi^s \rangle = \frac{1}{n} \sum_{i=1}^n x_i \xi_i^s$$

be the projection of $x \in \mathbb{R}^n$ onto $\frac{1}{n}\xi^s$. Obviously, x can be represented as the sum

$$x = \sum_{s=1}^{n} y_s \xi^s.$$
 (7.17)

A similar decomposition is possible for any input $r \in \mathbb{R}^n$: Let

$$a_s = \frac{1}{n} \langle r, \xi^s \rangle,$$

then

$$r = \sum_{s=1}^{n} a_s \xi^s.$$
 (7.18)

Let us prove the following

Lemma 47 (Reduction Lemma) If

$$C = \frac{1}{n} (\beta_1 \xi^1 (\xi^1)^{\mathsf{T}} + \beta_2 \xi^2 (\xi^2)^{\mathsf{T}})$$

for orthogonal $\xi^1, \xi^2 \in \Xi^n$, b < 0 and $a_3 = \cdots = a_n = 0$, then the plane $y_3 = \cdots = y_n = 0$ is a stable invariant manifold for (7.16) and dynamics on the manifold are governed by the system

$$\begin{cases} y_1' = a_1 + (b + \beta_1)y_1 - 3y_1y_2^2 - y_1^3 \\ y_2' = a_2 + (b + \beta_2)y_2 - 3y_2y_1^2 - y_2^3. \end{cases}$$
(7.19)

The condition $a_3 = \cdots = a_n = 0$ means that the input from receptors r is in the span (ξ^1, ξ^2) . For example, if we studied the olfactory system, then with this restriction only two odors are inhaled and recognizable.

Proof. Substituting (7.17) and (7.18) into (7.16) gives

$$\sum_{s=1}^{n} y'_{s} \xi^{s}_{i} = \sum_{s=1}^{n} a_{s} \xi^{s}_{i} + b \sum_{s=1}^{n} y_{s} \xi^{s}_{i} - \sum_{s=1}^{n} \sum_{p=1}^{n} \sum_{q=1}^{n} y_{s} y_{p} y_{q} \xi^{s}_{i} \xi^{p}_{i} \xi^{q}_{i} + \sum_{s=1}^{n} y_{s} \beta_{s} \xi^{s}_{i},$$

where $\beta_3 = \cdots = \beta_n = 0$. Projecting both sides onto $\frac{1}{n}\xi^k$ gives

$$y'_{k} = a_{k} + (b + \beta_{k})y_{k} - \sum_{s=1}^{n} \sum_{p=1}^{n} \sum_{q=1}^{n} y_{s}y_{p}y_{q}(\frac{1}{n}\sum_{i=1}^{n} \xi_{i}^{k}\xi_{i}^{s}\xi_{i}^{p}\xi_{i}^{q}), \qquad 1 \le k \le n.$$
(7.20)

Note that

$$\frac{1}{n}\sum_{i=1}^{n}\xi_{i}^{k}\xi_{i}^{s}\xi_{i}^{p}\xi_{i}^{q} = \begin{cases} 1, & \text{if } k = s, p = q \text{ or } k = p, s = q \text{ or } k = q, s = p \\ d_{kspq}, & \text{if all indices are different} \\ 0, & \text{otherwise,} \end{cases}$$

where $d_{kspq} \in \mathbb{R}$ are some constants. We used the assumption that $\xi_i^s = \pm 1$ for any s and *i*.

If the number of the memorized images m were greater than 2, then all equations in (7.20) would contain the constants d_{kspq} . It is possible to eliminate them if we consider (7.20) on the plane $y_3 = \cdots = y_n = 0$ for $m \leq 2$. Indeed, the product $y_s y_p y_q$ is always zero unless $1 \leq s, p, q \leq 2$. The inequality guarantees that at least two indices coincide. Hence, the sum $\frac{1}{n} \sum_{i=1}^{n} \xi_i^k \xi_i^s \xi_i^p \xi_i^q$ is either 1 or 0. It is 1 when all the indices are equal (this gives y_k^3) or when k is equal to only one of the three indices s, p, q (there are three such possibilities and, hence, the term $3y_k y_{3-k}^2$). Thus, the system (7.20) on the plane can be rewritten as (7.19).

We still must show that the plane is a stable invariant manifold.

From the Lemma's conditions we know that $a_3 = \cdots = a_n = 0$ and $\beta_3 = \cdots = \beta_n = 0$. Let's fix y_1 and y_2 and consider them as parameters. Keeping only linear terms, we can rewrite (7.20) for $k \ge 3$ as

$$y'_k = by_k - 3y_k(y_1^2 + y_2^2) + \text{higher order terms}, \quad 3 \le k \le n.$$

The plane is invariant because $y'_3 = \ldots = y'_n = 0$ on it. It is stable because $b - 3(y_1^2 + y_2^2) < b < 0$, which follows from the Lemma's condition that b < 0. \Box

It is still an open question whether the invariant plane is globally asymptotically stable or not. Our conjecture is that for b < 0 it is true, but we do not need this for our analysis below. If $\beta_1 = \beta_2 = \beta$, then it is easy to check that (7.19) can be rewritten as

$$\begin{cases} u' = s + (b + \beta)u - u^{3} \\ v' = c + (b + \beta)v - v^{3}, \end{cases}$$
(7.21)

where $u = y_1 + y_2$, $v = y_1 - y_2$, $s = a_1 + a_2$ and $c = a_1 - a_2$. The advantage of (7.21) is that it is uncoupled and each equation can be studied independently.

7.6.2 Recognition: Only One Image Is Presented

Without loss of generality we may assume that it is ξ^1 , i.e.

$$r_i = a\xi_i^1.$$

Assuming that all the conditions of the Reduction Lemma are satisfied and that $\beta_1 = \beta_2 = \beta$ we can rewrite (7.16) as (7.21). Note that s = c = a. Thus, the dynamics on the (u, v) plane is the direct product of two identical equations

$$z' = a + (b + \beta)z - z^3.$$

If $b + \beta < 0$, then there is only one equilibrium point for any a. The dynamics on the (u, v) plane is qualitatively the same as that of the canonical model of WCNN near multiple pitchfork bifurcation point (7.13) for $b + \beta_1 < 0$ which is depicted in Figure 7.3a.

Suppose $b + \beta > 0$. There are three equilibrium points when $|a| < a^*$, where

$$a^{\star} = 2\left(\frac{b+\beta}{3}\right)^{\frac{3}{2}}$$

Hence, (7.21) has nine equilibrium points. Again, there is not any qualitative distinction between the phase portrait of (7.13) depicted in Figure 7.3d and that of (7.21), which we depict in Figure 7.6a for a > 0. We see that $|a| < a^*$ is too weak to produce any qualitative changes in the dynamics of the canonical model (7.16) in comparison with (7.13). Nevertheless, it is easy to see that the domain of attraction of the



Figure 7.6: Phase portrait of the canonical model on the stable invariant plane span (ξ^1, ξ^2) . The first image is presented as an input onto the network. a. Input is weak, i.e. $|a| < a^*$. b. For $|a| = a^*$ there are fold bifurcations. c. For $|a| > a^*$ the canonical model is globally asymptotically stable.

equilibrium point corresponding to the presented image ξ^1 is much bigger than the attraction domains of the other equilibrium points. By the term attraction domain *size* we mean here the distance from the attractor to the closest saddle. We use this definition in order to be able to compare domains that have infinite volumes.

When the parameter a crosses $\pm a^*$ one can observe two fold (saddle-node) bifurcations and one co-dimension-2 bifurcation (see Figure 7.6b). All of them take place simultaneously due to the fact that (7.21) is a direct product of two identical equations. We consider these bifurcations elsewhere when we study the canonical model for WCNNs near multiple fold bifurcation point.

If the input $r = a\xi^1$ is sufficiently strong (i.e. if $|a| > a^*$), then there is only one equilibrium point, which is a stable node (see Figure 7.6c). The equilibrium point is globally asymptotically stable in this case.

We see that the canonical model (7.16) can work as GAS-type NN when the input strength a is strong enough, viz.

$$|a| > 2\left(\frac{b+\beta}{3}\right)^{\frac{3}{2}}.$$

We performed all the analysis above for the case of one presented and two memorized images.

7.6.3 Recognition: Two Images Are Presented

Without loss of generality we may assume in this case that

$$r = a_1 \xi^1 + a_2 \xi^2$$

for $a_1, a_2 > 0$. If $\beta_1 = \beta_2 = \beta$ and all the conditions of the Reduction Lemma are satisfied, then the canonical model (7.16) can be reduced to the two-dimensional system (7.21)

$$\begin{cases} u' = s + (b + \beta)u - u^3\\ v' = c + (b + \beta)v - v^3. \end{cases}$$

We cannot reduce (7.21) to a one-dimensional system because in general $s \neq c$. The constant $s = a_1 + a_2$ has obvious meaning of overall *strength* of the input from receptors, whereas $c = a_1 - a_2$ is the *contrast* of the input. When c > 0 (c < 0) we say that ξ^1 (ξ^2) is dominant.

In order to determine the qualitative behavior of (7.21) we have to compare s and c with the bifurcation value a^* . When both s and c are less than a^* , the qualitative phase portrait of (7.21) depicted in Figure 7.7a coincides with that of (7.13) depicted in Figure 7.3d provided $b + \beta > 0$.

Very interesting behavior arises when the overall input from receptors s is strong enough, i.e. when $s > a^*$. Then, (7.21) generically has either one or three equilibrium points (see Figure 7.7b). Its behavior is determined by the equation

$$v' = c + (b + \beta)v - v^3, \quad v \in \mathbb{R}.$$
 (7.22)

Obviously, the dynamics of (7.22) depends crucially not only upon which image is dominant but also upon how dominant it is. If $|c| < a^*$, then there is a co-existence



Figure 7.7: Phase portrait of the canonical model on the stable invariant plane span (ξ^1, ξ^2) . The input is a mixture of two images ξ^1 and ξ^2 . a. Overall input is weak. b. Strong input and weak contrast. There is a co-existence of two attractors.

between these two images (see Figure 7.7b). Both equilibrium points are stable. If $|c| > a^*$ then only one image survives, viz. the dominant image.

One possibility to explain the co-existence of two attractors corresponding to two different images is that the NN cannot distinguish between them when the contrast |c| is small. One could say that the two-attractor state corresponds to the "I do not know" answer. We prefer another explanation suggested by the psychological experiment described in the next section.

7.7 Bistability of Perception

In the previous sections we showed that if the conditions of the Reduction Lemma are satisfied and the overall input from receptors is strong $(s > a^*)$, then the canonical model behaves qualitatively like the equation (7.22)

$$v' = c + bv - v^3, \qquad v \in \mathbb{R},$$

where $c = a_1 - a_2$ is the contrast between two images $a_1\xi^1$ and $a_2\xi^2$ and b is a real parameter (we incorporated β into b, so b can be positive or negative).

We have already mentioned that if the contrast is weak $(|c| < a^*)$ then (7.22) has two attractors corresponding to the previously memorized images ξ^1 and ξ^2 .

First of all, note that the co-existence of two attractors contradicts the GAS-type NN paradigm which requires that the NN have only one attractor. We must accept the fact that the brain is a very complicated system having many attractors. Its dynamic behavior depends not only upon the input r, the synaptic memory C and the psychological state b but also upon a short-term past activity (which sometimes is called a short-term memory (Grossberg 1988)). In our case this is the initial condition x(0). Obviously, which attractor will be selected by the NN depends upon the initial state. Simultaneous existence of several attractors for the input that is a mixture of images suggests the following hypothesis: The NN perceives the ambiguous input according to the network's past short-term activity x(0).

The behavior of the *artificial* NN (7.16) is similar to the behavior of the real human brain in the following psychological experiment (Attneave 1971): The fourth figure from the left in the top row depicted in Figure 7.8 was shown to be perceived with equal probability as the face of a man or the body of a girl. If the figure is included in a sequence, then its perception depends upon the direction in which the sequence is viewed.

This phenomena was studied from catastrophe theory point of view (Poston and Stewart 1978, Stewart and Peregoy 1983) and it was shown that there is a onedimensional section of a cusp catastrophe in the human perception of the figures.

The remarkable fact is that the WCNN approximated by (7.22) also exhibits the cusp catastrophe. Suppose ξ^1 and ξ^2 represent girl's body and man's face images, respectively. If we fix b > 0 and vary the image contrast $c = a_1 - a_2$, then the artificial NN also has the same bistable perception of the presented images $a_1\xi^1$ and $a_2\xi^2$ (see



Figure 7.8: Bistability of perception.

bottom row in Figure 7.8).

What we have not explained yet is the switching of our attention (say, from girl's body to man's face and back) while we observe an ambiguous picture. These oscillations in our perception cannot be explained by the catastrophe theory. We can tackle this problem by embedding the WCNN (2.1) into the Λ -space, i.e. by allowing the internal parameter λ to vary (see Section 2.6.1). This idea was used by Ditzinger and Haken (1989).

As we can see, the canonical model (7.16) can work as the MA and GAS-type NNs simultaneously. Indeed, its dynamics crucially depends upon the input $r \in \mathbb{R}^n$ from receptors. If the input is strong enough and there is no ambiguity, then (7.16) has only one attractor and, hence, works as the GAS-type NN. If the input is weak or ambiguous, then (7.16) can have many attractors and, hence, can work as the MA-type neural network.

We think that the real brain might use similar principles. Consider, for example, the olfactory system (Baird 1986, Erdi et.al 1993, Li and Hopfield 1989, Skarda and Freeman 1987). It is believed that each inhaled odor has its own attractor – a stable limit cycle. The analysis of the canonical model (7.16) suggests that when an animal inhales a mixture of the odors, the appropriate limit cycles become stable so that

120

there is a one-to-one correspondence between the inhaled odors and the attractors.

Similar results were obtained by studying another NN (Izhikevich and Malinetskii 1993), but the attractors there were chaotic.

7.8 Quasi-Static Variation of Parameter b

In the two preceding sections we studied the behavior of the canonical model (7.16) for fixed b. We varied the contrast c and saw that there were two attractors when the contrast was weak. The NN recognized one of the two presented images according to the initial conditions, not to dominance of one of them over the other.

Of course, the attraction domain of the stronger image was bigger than that of the weaker one, but the network could not determine which image was dominant. One possibility to do it is to collect statistics over many trials for random initial conditions.

There is another possibility of determining which image is dominant. We have to fix the contrast c and vary the bifurcation parameter b very slowly so that we can neglect the transient processes and assume that x(t) is arbitrarily close to attractor. Such variation of the parameter is called *quasi-static*.

Recall that for b sufficiently small the canonical model (7.16) is globally asymptotically stable, i.e. it has only one attractor a stable node. For large b, system (7.16) has many attractors.

Suppose we start from small b. Then for any initial condition x(0) the activity x(t) approaches the unique attractor, and after some transient process x(t) is in a small neighborhood of the attractor. Let us increase b quasi-statically. The activity x(t) remains in the small neighborhood provided the attractor is hyperbolic.

Suppose the input is a combination of two previously memorized images ξ^1 and ξ^2 . Suppose also that all conditions of the Reduction Lemma are satisfied, $\beta_1 = \beta_2 = \beta$ and $s = a_1 + a_2$ are large enough. Then the qualitative behavior of the canonical model (7.16) is governed by the dynamical system (7.22).

$$v' = c + bv - v^3, \qquad v \in \mathbb{R},$$

Suppose c = 0, i.e. the input images have equal strength. As we expected, (7.22) has only one attractor v = 0 for small b. Here *small* means $b < -\beta$. If we increase b quasi-statically, the activity v(t) is always in a neighborhood of the origin provided $b < -\beta$.

When $b = -\beta$, the NN must choose one of the stable branches of the pitchfork bifurcation diagram depicted in Figure 7.9a. The nonzero contrast c is a perturbation (or an imperfection, see Golubitsky and Shaeffer (1979)) of the pitchfork bifurcation (see Figure 7.9b and c). No matter how small the contrast c is, the NN correctly chooses the corresponding branch provided the quasi-static increasing of b is slow enough. The case when the first image ξ^1 is dominant is depicted in Figure 7.9b. The stroboscopic presentation of the phenomenon is depicted in Figure 7.9d.

One can speculate that when the internal parameter b crosses a bifurcation value, the NN "thinks". Choosing one of the stable branches could be called the "decision making". Prolonged staying near the unstable branch could be called the "I don't know" state. Thus, we speculate that some types of non-hyperbolic behavior exhibited by the canonical model near bifurcation points are intimately connected to such vague psychological processes as recognition and thinking.



Figure 7.9: Bifurcation diagrams for quasi-static variation of parameter b. a. The contrast c = 0. b. The first image is dominant. c. The second image is dominant. e d. Stroboscopic presentation of the phenomenon for $b < -\beta$, $b = -\beta$ and $b > -\beta$, respectively.

Chapter 8 Quasi-Static Bifurcations

In this chapter we analyze the canonical models (3.24) and (3.32) for singularly perturbed WCNNs in the special case

$$a_1=\cdots=a_n=a>0$$

An appropriate change of coordinates and taking the limit as $\varepsilon \to 0$ transform the canonical models to

$$\begin{cases} x'_{i} = -y_{i} + r_{i}x_{i} + x_{i}^{2} + \sum_{j=1}^{n} c_{ij}x_{j} \\ y'_{i} = x_{i} \end{cases} \quad i = 1, \dots, n$$
(8.1)

and

$$\begin{cases} x'_{i} = -y_{i} + r_{i}x_{i} \pm x_{i}^{3} + \sum_{j=1}^{n} c_{ij}x_{j} \\ y'_{i} = x_{i} \end{cases} \quad i = 1, \dots, n,$$
(8.2)

respectively. Here $r_i \in \mathbb{R}$ describes input to the *i*-th relaxation neuron. We use the notation $X = (x_1, \ldots, x_n)^{\mathsf{T}} \in \mathbb{R}^n$; $Y = (y_1, \ldots, y_n)^{\mathsf{T}} \in \mathbb{R}^n$ and $C = (c_{ij}) \in \mathbb{R}^{n \times n}$.

8.1 Stability of the Equilibrium

Note that the canonical models (8.1) and (8.2) always have a unique equilibrium point, namely, the origin $(X, Y) = (0, 0) \in \mathbb{R}^n \times \mathbb{R}^n$. In this section we study the stability of the origin, as determined by the Jacobian matrix

$$L = \begin{pmatrix} R+C & -E \\ E & 0 \end{pmatrix}, \tag{8.3}$$

where E is the unit $n \times n$ -matrix, $C = (c_{ij})$, and

$$R = \begin{pmatrix} r_1 & 0 & \cdots & 0 \\ 0 & r_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_n \end{pmatrix}.$$
 (8.4)

A matrix is called *hyperbolic* if all its eigenvalues have non-zero real parts. It is *stable* if all its eigenvalues have negative real parts. An eigenvalue with the largest (most positive) real part is the *leading* (dominant) eigenvalue.

Theorem 48 The Jacobian matrix $L \in \mathbb{R}^{2n \times 2n}$ defined in (8.3) has the following properties:

(a) L is non-singular.

(b) L is stable if and only if R + C is stable.

(c) L is hyperbolic if and only if R + C is hyperbolic.

(d) If R+C has a zero eigenvalue, then L has a pair of pure imaginary eigenvalues $\pm i$.

(e) If R + C has a pair of purely imaginary eigenvalues, then L has two pairs of pure imaginary eigenvalues.

Proof. Suppose $Lv = \mu v$ for some $\mu \in \mathbb{C}$ and a non-zero vector $v \in \mathbb{C}^{2n}$. We use notation

$$v=\left(\begin{array}{c}v_1\\v_2\end{array}\right),$$

where $v_1, v_2 \in \mathbb{C}^n$. Then

$$(L - \mu E)v = \begin{pmatrix} R + C - \mu E & -E \\ E & -\mu E \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$
$$= \begin{pmatrix} (R + C - \mu E)v_1 - v_2 \\ v_1 - \mu v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Thus, we have

$$(R + C - \mu E)v_1 - v_2 = 0 v_1 - \mu v_2 = 0 .$$
 (8.5)

(a) If $\mu = 0$, then the second equation implies $v_1 = 0$. The first equation implies $v_2 = 0$, therefore v = 0. This means that L cannot have zero eigenvalue, hence it is always non-singular.

(b) From part (a) we know that $\mu \neq 0$. Then $v_2 = \mu^{-1}v_1$ and the eigenvector v of L has the form

$$v = \left(\begin{array}{c} v_1\\ \mu^{-1}v_1 \end{array}\right).$$

The first equation in (8.5) gives

$$(R + C - (\mu + \mu^{-1})E)v_1 = 0$$

for some non-zero $v_1 \in \mathbb{R}^n$. Hence the matrix

$$R + C - (\mu + \mu^{-1})E \tag{8.6}$$

is singular. Its eigenvalues are

$$\lambda_i - (\mu + \mu^{-1}), \quad i = 1, \dots, k,$$

where $\lambda_1, \ldots, \lambda_k$ are those of R+C. Since it is singular, at least one of the eigenvalues of (8.6) should be zero. Hence μ is a solution of one of the equations

$$\lambda_i = \mu + \mu^{-1}, \quad i = 1, \dots, k.$$
 (8.7)

It is easy to check that if $\operatorname{Re} \lambda_i < 0$, then $\operatorname{Re} \mu < 0$ and vice-versa.

(c) Condition (8.7) reveals the relationship between eigenvalues of L and R + C. If L is non-hyperbolic, then from part(a) it follows that μ is pure imaginary. Hence λ is pure imaginary or zero. Conversely, if λ is zero or pure imaginary, then μ is pure imaginary.

(d) and (e) follow from (8.7). \Box
Corollary 49 The equilibrium point loses its stability through a (possibly multiple) Andronov-Hopf bifurcation.

We study the Andronov-Hopf bifurcation in Section 8.3. In particular, we are interested in when the bifurcation is subcritical or supercritical and how this depends on the input.

Next, we analyze some neurophysiological consequences of Theorem 48.

8.2 Dale's Principle and Synchronization

One application of (3.4) is in modeling weakly connected networks of relaxation neural oscillators. In this case x_i and y_i denote rescaled activities of local populations of excitatory and inhibitory neurons, respectively. Each coefficient c_{ij} describes the strength of synaptic connections from x_j to x_i .

We say that x_j is depolarized when $x_j > 0$. Notice that if $c_{ij} > 0$ ($c_{ij} < 0$), then depolarization of x_j facilitates (impedes) that of x_i . We call such synapses excitatory (inhibitory). Copious neurophysiological data suggest that excitatory neurons have only excitatory synapses. This observation is usually referred to as *Dale's Principle* (Dale 1935; Shepherd 1983). In our case it implies that $c_{ij} \ge 0$ for all *i* and *j* (where $c_{ij} = 0$ corresponds to absence of a synapse from x_j to x_i).

In this section we show that Dale's principle imposes some restriction on local dynamics of the canonical models when the origin loses stability. In particular, we prove that the neural oscillators can synchronize.

Theorem 50 Suppose the synaptic matrix $C = (c_{ij})$ satisfies Dale's principle $(c_{ij} \ge 0$ for all i and j). Then generically

- The equilibrium of (8.1) and (8.2) loses stability via an Andronov-Hopf bifurcation.
- The network's local activity is in-phase synchronized, i.e. any two neural oscillators have nearly zero phase difference.

Proof. Let

$$\rho = \min_{1 \le i \le n} (r_i + c_{ii})$$

and consider matrix A defined by

$$A = R + C - \rho E.$$

Dale's principle ensures that A has non-negative entries. The Perron-Frobenius Theorem (Gantmacher 1959) applied to A shows that the leading eigenvalue λ of A is real and non-negative and that the corresponding eigenvector u has only non-negative entries. Typically, λ has multiplicity one and u has positive entries. The leading eigenvalue of R + C is $\lambda + \rho$, which is real and also has multiplicity one. Theorem 48 guarantees that when the equilibrium loses stability, the Jacobian matrix L has only one pair of pure imaginary eigenvalues $\pm i$. Thus, the multiple Andronov-Hopf bifurcation is not typical in this sense.

From the proof of Theorem 48 it follows that the corresponding eigenvectors of L have the form

$$\left(\begin{array}{c} u\\ \mp\imath u\end{array}\right),$$

where u was defined above. Local dynamics near the equilibrium is described by

$$\begin{pmatrix} X(t) \\ Y(t) \end{pmatrix} = \begin{pmatrix} u \\ -\iota u \end{pmatrix} z(t) + \begin{pmatrix} u \\ +\iota u \end{pmatrix} \bar{z}(t) + \text{ higher-order terms,}$$
(8.8)

where $z(t) \in \mathbb{C}$ is small. The activity of each neural oscillator has the form

$$\begin{pmatrix} x_i(t) \\ y_i(t) \end{pmatrix} = 2u_i \begin{pmatrix} \operatorname{Re} z(t) \\ \operatorname{Im} z(t) \end{pmatrix} + \text{ higher-order terms}$$

We can express the activity of the *i*-th oscillator through that of the *j*-th oscillator by

$$\begin{pmatrix} x_i(t) \\ y_i(t) \end{pmatrix} = \frac{u_i}{u_j} \begin{pmatrix} x_j(t) \\ y_j(t) \end{pmatrix} + \text{ higher-order terms,}$$
(8.9)

where $u_i/u_j > 0$ because u has positive entries. Hence the *i*-th and *j*-th oscillators have zero phase difference (up to some order). \Box

If the Andronov-Hopf bifurcation is supercritical, i.e. there is a birth of a stable limit cycle, then in-phase synchronization is asymptotic. More precisely, all local solutions have the form (8.8), where z(t) is small and periodic and the higher order terms remain sufficiently small as $t \to \infty$.

If the Andronov-Hopf bifurcation is subcritical (i.e. there is a death of an unstable limit cycle), then z(t) in (8.8) grows as $t \to \infty$ and in-phase synchronization is only *local*. The higher-order terms in (8.8) can grow with time and after a while they can be significant.

Remark 51 If Dale's principle is not satisfied, then:

- Multiple Andronov-Hopf bifurcation with exactly two pairs of pure imaginary eigenvalues is also generic. This follows from Theorem 48, part (e).
- Either in-phase or anti-phase synchronization is possible, i.e. the phase difference between any two oscillators could be nearly 0 or π . This follows from (8.9) because u_i/u_j is a scalar and could be positive or negative.

Remark 52 There have been many studies of existence and stability of in-phase and anti-phase solutions in linearly coupled relaxation oscillators. See, for example Belair and Holmes (1984), Storti and Rand (1986), Somers and Kopell (1993), Kopell and Somers (1995) and Mirollo and Strogatz (1990). Our results complement and extend those of these authors since they perform global analysis of two coupled oscillators, while we perform local analysis of n coupled oscillators.

Remark 53 An important difference between weakly connected networks of relaxation and non-relaxation oscillators (Hoppensteadt and Izhikevich 1995b) is that in the former the phase differences are usually either 0 or π , but in the latter they may assume arbitrary values.

8.3 Further Analysis of the Andronov-Hopf Bifurcation

In this section we study the Andronov-Hopf bifurcation in canonical models when R + C has a simple zero eigenvalue. Our major goal is to determine when it is subcritical or supercritical and how this depends on the matrix C and the inputs r_1, \ldots, r_n .

We begin with an analysis of the canonical model (8.2)

$$\begin{cases} x'_i = -y_i + r_i x_i + \sigma_i x_i^3 + \sum_{j=1}^n c_{ij} x_j \\ y'_i = x_i \end{cases} \quad \sigma_i = \pm 1, \quad i = 1, \dots, n$$

because it is simpler than that of (8.1).

Let $v_1 = (v_{11}, \ldots, v_{1n})^{\top} \in \mathbb{R}^n$ be the normalized eigenvector of R + C corresponding to the zero eigenvalue. Let $w_1 = (w_{11}, \ldots, w_{1n}) \in \mathbb{R}^n$ be dual to v_1 , i.e.

$$w_1v_1 = \sum_{i=1}^n w_{1i}v_{1i} = 1$$

and w_1 is orthogonal to the other (generalized) eigenvectors of R + C.

Theorem 54 If the parameter a defined by

$$a \equiv \frac{3}{4} \sum_{i=1}^{n} \sigma_i w_{1i} v_{1i}^3 \tag{8.10}$$

is positive (negative), then the Andronov-Hopf bifurcation in (8.2) is subcritical (supercritical).

Proof of the theorem is given in Section 8.4.

In neural network studies it is frequently assumed that synapses are modified according to Hebbian learning rule. This implies that the synaptic matrix C is symmetric. It is also reasonable to consider a network of approximately similar oscillators. They can have different quantitative features, but their qualitative behavior should be comparable. These two observations motivate the following result

Corollary 55 Suppose that

- 1. The synaptic matrix $C = (c_{ij})$ is symmetric, and
- 2. All oscillators have the same type, i.e.

$$\sigma_1=\cdots=\sigma_n=\sigma.$$

If $\sigma = +1$ ($\sigma = -1$), then the Andronov-Hopf bifurcation in (8.2) is always subcritical (supercritical).

Proof. If C is symmetric, then so is R + C. Since every symmetric matrix has orthogonal eigenvectors, we have $v_1 = w_1^{\mathsf{T}}$. Therefore (8.10) can be rewritten as

$$a = \frac{3}{4}\sigma \sum_{i=1}^{n} v_{1i}^4$$

and its sign is determined by σ . \Box

Remark 56 We can relax assumption 1 in the corollary simply by requiring that v_1 be orthogonal to the other (generalized) eigenvectors.

Analysis of the canonical model (8.1)

$$\begin{cases} x'_{i} = -y_{i} + r_{i}x_{i} + x_{i}^{2} + \sum_{j=1}^{n} c_{ij}x_{j} \\ y'_{i} = x_{i} \end{cases} \quad i = 1, \dots, n$$

is more complicated than that of (8.2). Since we are interested mostly in the case when the synaptic matrix C is symmetric, we assume at the very start that it is.

Theorem 57 The Andronov-Hopf bifurcation in (8.1) for a symmetric matrix $C = (c_{ij})$ is always subcritical.

Proof of the theorem is given in Section 8.4.

8.4 Proofs of Theorems 54 and 57

The canonical models (8.1) and (8.2) can be written concisely in the form

$$Z' = F(Z), \tag{8.11}$$

where $Z = (X, Y) \in \mathbb{R}^{2n}$, $F : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ and F(0) = 0. The Jacobian matrix L = DF at the equilibrium is given by (8.3). From Theorem 48 it follows that the equilibrium Z = 0 loses stability via an Andronov-Hopf bifurcation when R + C has one simple zero eigenvalue and the other eigenvalues lie in the left half-plane. Let v_1 be the eigenvector of R + C corresponding to the zero eigenvalue. Then L has a pair of pure imaginary eigenvalues $\pm i$ with the corresponding eigenvectors

$$\left(\begin{array}{c} v_1\\ \mp \imath v_1\end{array}\right). \tag{8.12}$$

To determine the type of bifurcation that occurs we restrict (8.11) to the center manifold, which is tangent to the center subspace:

$$E^{c} = \operatorname{span}\left\{ \left(\begin{array}{c} v_{1} \\ 0 \end{array} \right), \left(\begin{array}{c} 0 \\ v_{1} \end{array} \right) \right\} = \left\{ \left(\begin{array}{c} v_{1} \\ 0 \end{array} \right) x + \left(\begin{array}{c} 0 \\ v_{1} \end{array} \right) y \mid x, y \in \mathbb{R} \right\}, \quad (8.13)$$

where x and y can be treated as coordinates on E^c . On the manifold (8.11) has the normal form

$$\begin{cases} x' = -y + f(x, y) \\ y' = x + g(x, y), \end{cases}$$
(8.14)

where f and g denote the non-linear terms in x and y. Then the Andronov-Hopf bifurcation is subcritical (supercritical) if the parameter

$$a = \frac{1}{16} (f_{xxx} + f_{xyy} + g_{xxy} + g_{yyy} + f_{xy}(f_{xx} + f_{yy}) - g_{xy}(g_{xx} + g_{yy}) - f_{xx}g_{xx} + f_{yy}g_{yy})$$
(8.15)

is positive (negative). For derivation of (8.15) see Guckenheimer and Holmes (1983). Note that during the center manifold reduction it suffices to compute f and g only up to third order terms in x and y.

Our treatment of the center manifold reduction is based on that of Iooss and Adelmeyer (1992). To perform the reduction we must introduce some objects: Let $E^s \subset \mathbb{R}^{2n}$ be the stable subspace spanned by the (generalized) eigenvectors of Lcorresponding to the eigenvalues of L having negative real parts. Thus, we have the splitting

$$\mathbb{R}^{2n} = E^c \oplus E^s.$$

Let $\pi_c : \mathbb{R}^{2n} \to E^c$ and $\pi_s : \mathbb{R}^{2n} \to E^s$ be projectors such that

ker
$$\pi_c = E^s$$
 and ker $\pi_s = E^c$.

If w_1 is a dual vector to v_1 , then π_c is given by

$$\pi_c = \begin{pmatrix} v_1 w_1 \\ v_1 w_1 \end{pmatrix}, \tag{8.16}$$

where v_1w_1 denotes the $n \times n$ -matrix defined by the tensor product $v_1 \otimes w_1$. Note also that π_c and π_s commute with L. The Center Manifold Theorem ensures that there is a mapping $\Psi: E^c \to E^s$ with

$$\Psi(0)=0$$
 and $D\Psi(0)=0$

such that the manifold ${\boldsymbol{\mathcal{M}}}$ defined by

$$\mathcal{M} = \{ v + \Psi(v) \mid v \in E^c \}$$

is invariant and locally attractive. The reduced system has the form

$$v' = \pi_c F(v + \Psi(v)) \tag{8.17}$$

and (8.14) is just (8.17) written in local coordinates on E^c .

The initial portion of Taylor's expansion of the function $\Psi(v)$, which defines the center manifold, can be determined from the equation

$$D\Psi(v)v' = \pi_s F(v + \Psi(v)), \qquad (8.18)$$

where v' is defined in (8.17).

We do not need (8.18) for proving Theorem 54 because the canonical model (8.2) does not have quadratic terms in x and y. But we use (8.18) in proof of Theorem 57.

8.4.1 Proof of Theorem 54

From (8.13) we see that

$$v' = \begin{pmatrix} v_1 \\ 0 \end{pmatrix} x' + \begin{pmatrix} 0 \\ v_1 \end{pmatrix} y'.$$
(8.19)

Since π_c and L commute, $\pi_c \Psi = 0$ and $\pi_c v = v$ for $v \in E^c$, we have

$$\pi_c L(v + \Psi(v)) = L\pi_c(v + \Psi(v)) = Lv.$$
(8.20)

Therefore the right-hand side of (8.17) is

$$L\begin{pmatrix} v_1\\ 0 \end{pmatrix} x + L\begin{pmatrix} 0\\ v_1 \end{pmatrix} y + \pi_c \begin{pmatrix} (\sigma_1(v_{11}x)^3, \dots, \sigma_n(v_{1n}x)^3) + \text{h.o.t.} \\ 0 \end{pmatrix}$$
$$= \begin{pmatrix} 0\\ v_1 \end{pmatrix} x + \begin{pmatrix} -v_1\\ 0 \end{pmatrix} y + \pi_c \begin{pmatrix} v_1x^3 \sum_{i=1}^n w_{1i}\sigma_iv_{1i}^3 + \text{h.o.t.} \\ 0 \end{pmatrix}.$$

Multiplying by $(w_1, 0)$ and $(0, w_1)$ gives

$$\begin{cases} x' = -y + x^3 \sum_{i=1}^n \sigma_i w_{1i} v_{1i}^3 + \text{h.o.t.} \\ y' = x. \end{cases}$$

The parameter a defined in (8.15) is

$$a=\frac{3}{4}\sum_{i=1}^n\sigma_iw_{1i}v_{1i}^3$$

8.4.2 Proof of Theorem 57

If the connection matrix C is symmetric, then so is R + C. Therefore, R + C has n orthogonal eigenvectors v_1, \ldots, v_n forming a basis for \mathbb{R}^n . As before, the vector v_1 corresponds to the zero eigenvalue of R + C and v_2, \ldots, v_n correspond to the other eigenvalues $\lambda_2, \ldots, \lambda_n$, which are negative. Using the proof of Theorem 48 we can define the stable subspace of L by

$$E^{s} = \operatorname{span}\left\{ \left(\begin{array}{c} v_{2} \\ 0 \end{array} \right), \left(\begin{array}{c} 0 \\ v_{2} \end{array} \right), \dots, \left(\begin{array}{c} v_{n} \\ 0 \end{array} \right), \left(\begin{array}{c} 0 \\ v_{n} \end{array} \right) \right\}$$

and

$$\pi_{s} = \left(\begin{array}{c} \sum_{k=2}^{n} v_{k} v_{k}^{\mathsf{T}} \\ \sum_{k=2}^{n} v_{k} v_{k}^{\mathsf{T}} \end{array} \right).$$

To determine the parameter a we must find the quadratic terms of $\Psi(v)$. Let

$$\Psi\left(\left(\begin{array}{c}v_1\\0\end{array}\right)x+\left(\begin{array}{c}0\\v_1\end{array}\right)y\right)=x^2\left(\begin{array}{c}p_1\\q_1\end{array}\right)+xy\left(\begin{array}{c}p_2\\q_2\end{array}\right)+y^2\left(\begin{array}{c}p_3\\q_3\end{array}\right)+\text{h.o.t.},$$

where

$$p_i = \sum_{k=2}^n p_{ik} v_k$$
 and $q_i = \sum_{k=2}^n q_{ik} v_k$

for i = 1, 2, 3. Since

$$\begin{cases} x' = -y + \text{h.o.t.} \\ y' = x + \text{h.o.t.}, \end{cases}$$

the left-hand side of (8.18) is

$$D\Psi(v)v' = xy \left(\begin{array}{c} 2p_3 - 2p_1\\ 2q_3 - 2q_1 \end{array}\right) + (x^2 - y^2) \left(\begin{array}{c} p_2\\ q_2 \end{array}\right) + \text{h.o.t.}$$
(8.21)

Since π_s commutes with L, and $\pi_s v = 0$, we have

$$\pi_{s}L(v+\Psi(v)) = L\pi_{s}(v+\Psi(v)) = L\Psi(v) = x^{2} \begin{pmatrix} \sum_{k=2}^{n} \lambda_{k}v_{k}p_{1k} - q_{1} \\ p_{1} \end{pmatrix} + xy \begin{pmatrix} \sum_{k=2}^{n} \lambda_{k}v_{k}p_{2k} - q_{2} \\ p_{2} \end{pmatrix} + y^{2} \begin{pmatrix} \sum_{k=2}^{n} \lambda_{k}v_{k}p_{3k} - q_{3} \\ p_{3} \end{pmatrix} + \text{h.o.t.}$$
(8.22)

Thus, the right-hand side of (8.18) is

$$L\Psi(v) + \pi_s \left(\begin{array}{c} (v_{1i}x)_i^2 \\ 0 \end{array} \right) = L\Psi(v) + x^2 \left(\begin{array}{c} \sum_{k=2}^n v_k \sum_{i=1}^n v_{ki} v_{1i}^2 \\ 0 \end{array} \right)$$
(8.23)

Combining like-terms in (8.21),(8.22) and (8.23) and considering projections on each v_k , we obtain the system

$$\begin{pmatrix} p_{2k} \\ q_{2k} \end{pmatrix} = \begin{pmatrix} \lambda_k p_{1k} - q_{1k} \\ p_{1k} \end{pmatrix} + \begin{pmatrix} \sum_{i=1}^n v_{ki} v_{1i}^2 \\ 0 \end{pmatrix}$$
$$2 \begin{pmatrix} p_{3k} \\ q_{3k} \end{pmatrix} - 2 \begin{pmatrix} p_{1k} \\ q_{1k} \end{pmatrix} = \begin{pmatrix} \lambda_k p_{2k} - q_{2k} \\ p_{2k} \end{pmatrix}$$
$$- \begin{pmatrix} p_{2k} \\ q_{2k} \end{pmatrix} = \begin{pmatrix} \lambda_k p_{3k} - q_{3k} \\ p_{3k} \end{pmatrix},$$

which is be solved for $p_{ik}, q_{ik}, i = 1, 2, 3; k = 2, ..., n$. Below we will use only the values of p_{1k} and p_{3k} , which are given by

$$p_{1k} = -p_{3k} = -\frac{2\lambda_k}{4\lambda_k^2 + 9} \sum_{i=1}^n v_{ki} v_{1i}^2.$$
(8.24)

Now let us determine the reduced system (8.17). Its left-hand side is given by (8.19). Using (8.20), we can find its right-hand side: it is

$$L\begin{pmatrix} v_{1} \\ 0 \end{pmatrix} x + L\begin{pmatrix} 0 \\ v_{1} \end{pmatrix} y + \pi_{c} \begin{pmatrix} (v_{1i}x + \sum_{k=2}^{n} v_{ki}(x^{2}p_{1k} + xyp_{2k} + y^{2}p_{3k}) + \text{h.o.t.})_{i}^{2} \\ 0 \end{pmatrix} = = \begin{pmatrix} 0 \\ v_{1} \end{pmatrix} x + \begin{pmatrix} -v_{1} \\ 0 \end{pmatrix} y + \begin{pmatrix} v_{1} \sum_{i=1}^{n} v_{1i}(v_{1i}^{2}x^{2} + 2v_{1i}x \sum_{k=2}^{n} v_{ki}(x^{2}p_{1k} + xyp_{2k} + y^{2}p_{3k})) + \text{h.o.t.} \\ 0 \end{pmatrix}.$$

Multiplying by $(v_1^{\mathsf{T}}, 0)$ and $(0, v_1^{\mathsf{T}})$ gives

$$\begin{cases} x' = -y + x^2 \sum_{i=1}^n v_{1i}^3 + x^3 \sum_{i=1}^n \sum_{k=2}^n 2v_{1i}^2 v_{ki} p_{1k} + \\ xy^2 \sum_{i=1}^n \sum_{k=2}^n 2v_{1i}^2 v_{ki} p_{3k} + \text{h.o.t.} \\ y' = x. \end{cases}$$

It follows from (8.15) that

$$a = \frac{1}{4} \sum_{i=1}^{n} v_{1i}^{2} \sum_{k=2}^{n} v_{ki} (3p_{1k} + p_{3k}).$$

Using (8.24) we see that

$$a = -\sum_{k=2}^{n} \left(\sum_{i=1}^{n} v_{ki} v_{1i}^{2}\right)^{2} \frac{\lambda_{k}}{4\lambda_{k}^{2} + 9} > 0$$

since each $\lambda_k < 0$.

Chapter 9 Non-Hyperbolic Neural Networks

Let us return to the question how the canonical models might perform pattern recognition tasks. First we outline the main idea, then we present rigorous mathematical considerations.

Consider the canonical models for multiple subcritical pitchfork bifurcation

$$x'_{i} = r_{i}x_{i} + x_{i}^{3} + \sum_{j=1}^{n} c_{ij}x_{j}, \qquad (9.1)$$

for multiple subcritical Andronov-Hopf bifurcation

$$z'_{i} = r_{i}z_{i} + d_{i}z_{i}|z_{i}|^{2} + \sum_{j=1}^{n} c_{ij}z_{j}, \quad \text{Re} \, d_{i} > 0,$$
(9.2)

for multiple quasi-static saddle-node bifurcation

$$\begin{cases} x'_{i} = -y_{i} + r_{i}x_{i} + x_{i}^{2} + \sum_{j=1}^{n} c_{ij}x_{j} \\ y'_{i} = x_{i} \end{cases},$$
(9.3)

and pitchfork bifurcation

$$\begin{cases} x'_{i} = -y_{i} + r_{i}x_{i} + x_{i}^{3} + \sum_{j=1}^{n} c_{ij}x_{j} \\ y'_{i} = x_{i} \end{cases}$$
(9.4)

Notice that these models have an equilibrium – the origin, for any choice of parameters. Stability of the origin is determined by the Jacobian matrix R + C, where

$$R = \begin{pmatrix} r_1 & 0 & \cdots & 0 \\ 0 & r_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_n \end{pmatrix},$$

and $C = (c_{ij})$ is the synaptic matrix. Both R and C are complex-valued for (9.2) and real valued for the other canonical models. The origin is stable if all eigenvalues of R + C have negative real parts and is unstable otherwise.

The origin loses its stability via subcritical pitchfork bifurcation (for (9.1)) or Andronov-Hopf bifurcation (for the other models). In some neighborhood of the bifurcation point the direction along the eigenvector corresponding to leading eigenvalue of R + C becomes unstable and the activity vector moves along this direction. After a while it leaves a small neighborhood of the origin and an observer notices some macroscopic changes in dynamics of the canonical models (see Figure 1.8). Thus the local event – loss of stability by the origin, produces a global effect. This is the keyidea of the non-hyperbolic NN approach, which we described in Section 1.4. Below we explain in detail the idea outlined above.

9.1 Problem 1

Given an input vector $r^k = (r_1^k, \ldots, r_n^k) \in \mathbb{R}^n$ we can construct diagonal matrix R^k by

$$R^{k} = \begin{pmatrix} r_{1}^{k} & 0 & \cdots & 0 \\ 0 & r_{2}^{k} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{n}^{k} \end{pmatrix}$$

We use the following notation: $\lambda_k \in \mathbb{R}$ denotes the leading eigenvalue of the matrix $R^k + C$, i.e. the eigenvalue with the largest real part. The vector $u^k \in \mathbb{R}^n$ denotes an eigenvector of $R^k + C$ corresponding to λ_k , if the leading eigenvalue is unique and simple, i.e. it has multiplicity one.

Suppose we are given a set of input vectors $\{r^1, \ldots, r^m\} \subset \mathbb{R}^n$ and a set of key patterns $\{v^1, \ldots, v^m\} \subset \mathbb{R}^n$ to be memorized. Consider the following problem

PROBLEM 1. Find a matrix $C \in \mathbb{R}^{n \times n}$ such that for all matrices $\mathbb{R}^k + C$, k = 1, ..., m the leading eigenvalues $\lambda_k = 0$ are simple and $u^k = v^k$.

Suppose that given $r^1, ..., r^m$ and $v^1, ..., v^m$ there is such a matrix C. Then the canonical models can perform the pattern recognition tasks in the sense described next:

We say that the k-th input r^k from external receptors is given if the parameters r_1, \ldots, r_n in the canonical models are given by

$$r_i = r_i^k + \rho, \quad i = 1, \dots, n, \tag{9.5}$$

where $\rho \in \mathbb{R}$ is a scalar bifurcation parameter. Then, for $\rho < 0$ the equilibrium point of the canonical models (the origin) is stable; for $\rho = 0$ there is a bifurcation; and for $\rho > 0$ the equilibrium is unstable.

For the canonical model (9.1) the equilibrium loses stability through subcritical pitchfork bifurcation (see Section 7.5). For small positive ρ the canonical model dynamics approach the center manifold, which is tangent to the center subspace $E \in \mathbb{R}^n$ defined by

$$E = \operatorname{span}\left\{u^k\right\},\,$$

where u^k is the eigenvector of $R^k + C$ corresponding to the leading eigenvalue λ_k . According to Problem 1, the vector u^k coincides with the memorized pattern v^k . Thus, when the k-th input pattern is given, the activity of the canonical models is close to the linear subspace E which is determined by the memorized pattern v^k . A rough sketch of local dynamics is depicted in Figure 9.1

In the canonical models (9.2), (9.3) and (9.4) the origin loses stability through Andronov-Hopf bifurcations (see Remark 41 and Section 8.3). The center manifold



Figure 9.1: Depending upon the input $(r^1 \text{ or } r^2)$ the center manifold is tangent to the corresponding memorized vectors $(v^1 \text{ or } v^2)$

is tangent to the center subspace defined by

$$E = \operatorname{span}\left\{ \begin{pmatrix} u^{k} \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \bar{u}^{k} \end{pmatrix} \right\}.$$
(9.6)

If the Andronov-Hopf bifurcation is supercritical for all inputs r^k , then the new-born stable limit cycles lie close to the corresponding center subspaces defined by (9.6). Thus, for each input r^k there is an attractor – stable limit cycle, which has location prescribed by the memorized pattern v^k . This was observed in the olfactory bulb (see Section 1.4.1). The only difference is that the limit cycles in the canonical models have small radii. Finally, we note that the supercritical Andronov-Hopf bifurcations can be observed in the canonical model (8.2) for $\sigma_i = -1$, $i = 1, \ldots, n$. Therefore, (8.2) can function like a GAS-type neural network, but there are two steady states: an unstable equilibrium (the origin) and a stable limit cycle.

If the Andronov-Hopf bifurcations is subcritical for all inputs r^k , then the canonical models are like non-hyperbolic neural networks. Indeed, after the bifurcation, which is local, the dynamics leave some neighborhood of the equilibrium point along the direction determined by one of the memorized pattern v^1, \ldots, v^m .

For example, suppose the network recognizes v^1 , and that $|v_1^1| \gg |v_2^1|$. Then

the first neural oscillator oscillates with the amplitude much larger than that of the second one. If the Andronov-Hopf bifurcation is supercritical, then the attractor has the same property. If the bifurcation is subcritical, then these oscillations are observed locally and might persist globally. In both cases, an experimenter on the olfactory bulb discovers that there is a spatial pattern of oscillations: various sites of the olfactory bulb oscillate with the same frequency but with different amplitudes (Skarda and Freeman 1987). In that case the results predicted by our analysis of the canonical models agree with the neurophysiological experiments.

9.2 Problems 2 and 3

To the best of our knowledge, Problem 1 is still unresolved. We do not know any general method that allows one to construct such synaptic matrix C, although for some special cases, such as m = 1, the construction is trivial. Below we present alternative problems that can be easier to resolve.

PROBLEM 2. Find matrix a C such that each λ_k is simple and real and $u^k = v^k$ for all k.

Since we do not require that $\lambda_k = 0$, the bifurcations occur for $\rho = -\lambda_k$, where ρ is defined in (9.5).

PROBLEM 3. Find matrix a C such that the leading eigenvalues λ_k of $\mathbb{R}^k + C, k = 1, ..., m$, are real and simple and the corresponding eigenvectors u^k are pairwise orthogonal.

Note that we do not require that $u^k = v^k$ here. The requirement that all u^k be orthogonal means that the response of the neural network on various inputs is as different as possible, even when the inputs to be memorized are similar (like a cat and a dog).

Chapter 10

Synaptic Organizations of the Brain

Neurophysiological studies of various brain structures (see Rakic 1976 and Shepherd 1976) show that there is a pattern of local synaptic circuitry in many parts of the brain: Local populations of excitatory and inhibitory neurons have extensive and strong synaptic connections between each other so that action potentials generated by the former excite the latter, which in turn, reciprocally inhibit the former (see Figure 10.1). They can be motoneurons and Renshaw interneurons in spinal cord; Mitral and granule cells in olfactory bulb; Pyramidal cells and thalamic interneurons in corticothalamic system, etc. Such pairs of interacting excitatory and inhibitory populations of neurons can also be found in cerebellum, hippocampus, olfactory cortex and neocortex (Shepherd 1976). This is one of the basic mechanisms for the generation of periodic activity in the brain. Such a pair is called a *neural oscillator*.

The neural oscillators within one brain structure can be connected into a network because the excitatory (and sometimes inhibitory) neurons can have synaptic contacts with other, distant, neurons. For example, in the olfactory bulb the mitral cells have contacts with other mitral cells (see Figure 10.2), whereas the granule cells apparently do not make any distant contacts, they do not even have axons. Their only purpose



Figure 10.1: Schematic representation of the neural oscillator. It consists of excitatory (white) and inhibitory (shaded) populations of neurons. For simplicity only one neuron from each population is pictured. White arrows denote excitatory synaptic connections, black arrows denote inhibitory synaptic connections



Olfactory Bulb

Figure 10.2: The neural oscillators (dotted boxes) are connected into a network. The mitral cell makes contacts with other mitral cells and may have contacts with other granule cells

is to provide a reciprocal dendro-dendritic inhibition for the mitral cells. Sometimes inhibitory neurons can also have long axons; for example, the periglomerular cells in the olfactory bulb.

Though on the local level all neural oscillators appear to be similar, the type of connections between them may differ. In this case we say that the networks of such oscillators have different *synaptic organization*. For instance, in Figure 10.2 the contacts between mitral cells and distant granule cells (dashed line) might or might not exist. These cases correspond to various synaptic organizations and, hence, to various dynamical properties of the network. The notion of synaptic organization is closely related to the notion of anatomy of the brain. Thus, in this chapter we study relationships between anatomy and functions of the brain. For example, we show that some synaptic organizations allow the network to memorize time delays, or phase deviation information, whereas the others do not allow such a possibility.

For the sake of clarity we always depict only two neural oscillators and the synaptic connections only in one direction, as in Figure 10.2. It is implicitly assumed that the network consists of many neural oscillators and the synaptic connections of the same type exist between any two oscillators and in all directions.

We have already used Dale's principle (Dale 1935; Shepherd 1983) in Chapter 8. Recall that it says: The excitatory neurons may have only excitatory synaptic connections with other neurons and inhibitory neurons may have only inhibitory synaptic connections. This principle imposes some restrictions on possible synaptic organization of the networks and we study how it affects the dynamical properties of the networks.

There is neurophysiological data about the importance of oscillations and chaos in the brain (Eckhorn et al. 1988, Gray 1994 and Skarda and Freeman 1987), and there have been many studies of the role of oscillations in the processing of information by the brain. Most of them devoted to study of synchronization phenomenon (Aronson et al. 1990, Hoppensteadt 1989, Kazanovich and Borisyuk 1994, von der Malsburg and Buhmann 1992 and Schuster and Wagner 1990). Pattern memorization and recognition by oscillatory neural networks have been studied, for example, by Baird (1986), Erdi et al. (1993) and Li and Hopfield (1989). To the best of our knowledge, there have been no attempts to connect dynamical properties, such as the possibility to learn a pattern, with the synaptic organization of a network. We think that such attempts could reveal the relationship between function and structure of the real brain.

10.1 Neural Oscillators

A neural oscillator is described by a dynamical system of the form

$$\begin{cases} \dot{x} = f(x, y, \lambda) \\ \dot{y} = g(x, y, \lambda), \end{cases}$$

where $x, y \in \mathbb{R}$ are the activity of excitatory and inhibitory neurons, respectively. A typical example of the neural oscillator is Wilson and Cowan model (1972,1973).

In this chapter we assume that the activities x and y are one-dimensional variables. This is a technical assumption made to simplify our computations and to allow us to use Dale's principle.

We call the oscillator *neural* to emphasize its connections with neuroscience. Thus, using Dale's principle we can gain some information about f and g, viz.

$$rac{\partial g}{\partial x} \geq 0 \quad ext{ and } \quad rac{\partial f}{\partial y} \leq 0,$$

for all x, y and λ . The first inequality means that x excites y. The second inequality means that y inhibits x.

Let $(x_i, y_i) \in \mathbb{R}^2$ denote the activity of *i*-th neural oscillator for i = 1, ..., n.

A weakly connected network of neural oscillators is a dynamical system of the form

$$\begin{cases} \dot{x}_i = f_i(x_i, y_i, \lambda) + \varepsilon p_i(x_1, y_1, \dots, x_n, y_n, \varepsilon) \\ \dot{y}_i = g_i(x_i, y_i, \lambda) + \varepsilon q_i(x_1, y_1, \dots, x_n, y_n, \varepsilon) \end{cases}$$
(10.1)

satisfying for all x, y, λ and $i \neq j$ Dale's principle, namely

$$\frac{\partial p_i}{\partial x_j} \ge 0, \quad \frac{\partial q_i}{\partial x_j} \ge 0, \quad \frac{\partial p_i}{\partial y_j} \le 0 \text{ and } \frac{\partial q_i}{\partial y_j} \le 0,$$
 (10.2)

where $p_i, q_i : \mathbb{R}^{2n+1} \to \mathbb{R}$ are functions that represent synaptic connections from the whole network onto the *i*-th neural oscillator. One can think of (10.1) with (10.2) as being a generalization of Wilson's-Cowan's model of an oscillatory neural network (Wilson and Cowan 1972,1973).

Remark 58 We do not demand that the connections within an oscillator are weak. For our analysis we need weakness of connections between neural oscillators.

In order to apply the WCNN theory that we developed in Chapter 2 we have to impose the additional assumption that each neural oscillator is not a pacemaker, i.e. that each equation in the uncoupled ($\varepsilon = 0$) system

$$\begin{cases} \dot{x}_i = f_i(x_i, y_i, \lambda) \\ \dot{y}_i = g_i(x_i, y_i, \lambda) \end{cases} \quad i = 1, \dots, n,$$
(10.3)

has a stable equilibrium point, say (0,0) for $\lambda = 0$. So we have

$$f_i(0,0,0) = 0$$

$$g_i(0,0,0) = 0$$

Throughout this chapter we denote the Jacobian matrix for i-th neural oscillator by

$$L_i = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}_i = \frac{\partial(f_i, g_i)}{\partial(x_i, y_i)} = \begin{pmatrix} \frac{\partial f_i}{\partial x_i} & \frac{\partial f_i}{\partial y_i} \\ \frac{\partial g_i}{\partial x_i} & \frac{\partial g_i}{\partial y_i} \end{pmatrix},$$



Figure 10.3: A network of two neural oscillators. Open boxes are local populations of excitatory neurons, shaded circles are local populations of inhibitory neurons. The real numbers a_1, a_2, a_3, a_4 are entries of the Jacobian matrix of *i*-th neural oscillator. The real numbers s_1, s_2, s_3, s_4 denote the strength of synaptic connections

where all derivatives are evaluated at the equilibrium point (0,0). The matrices L_1, \ldots, L_n could differ because we do not assume that the neural oscillators are identical. Sometimes we denote entries of the Jacobian matrices by $a_{i1}, a_{i2}, a_{i3}, a_{i4}$ to stress that they belong to L_i .

The matrices

$$S_{ij} = \begin{pmatrix} s_1 & s_2 \\ s_3 & s_4 \end{pmatrix}_{ij} = \frac{\partial(p_i, q_i)}{\partial(x_j, y_j)} = \begin{pmatrix} \frac{\partial p_i}{\partial x_j} & \frac{\partial p_i}{\partial y_j} \\ \frac{\partial q_i}{\partial x_j} & \frac{\partial q_i}{\partial y_j} \end{pmatrix}$$
(10.4)

denote rescaled synaptic connections from j-th to i-th neural oscillator (see Figure 10.3) because the actual synaptic connections have order ε and look like εS_{ij} . Note that according to Dale's principle (10.2) the synaptic matrices have signature

$$\begin{pmatrix} + & - \\ + & - \end{pmatrix}$$

10.1.1 Multiple Andronov-Hopf Bifurcation

Recall that the Fundamental Theorem of WCNN Theory requires that (10.1) be near a multiple bifurcation, or otherwise its dynamics is not interesting from neurocomputational point of view. It is reasonable to require that the bifurcation be Andronov-Hopf.

$$\operatorname{tr} L_i = a_1 + a_4 = 0$$
 and $\operatorname{det} L_i = a_1 a_4 - a_2 a_3 > 0$.

Recall that we associate with each neural oscillator its natural frequency

$$\Omega_i = \sqrt{\det L_i} = \sqrt{a_{i1}a_{i4} - a_{i2}a_{i3}}$$

As usual, we assume that

$$\lambda = \lambda(\varepsilon) = 0 + \varepsilon \lambda_1 + \mathcal{O}(\varepsilon^2) \tag{10.5}$$

for some $\lambda_1 \in \Lambda$. Thus, all neural oscillators are ε -close to the Andronov-Hopf bifurcation and their natural frequencies are

$$\Omega_i + \varepsilon \omega_i$$

for some $\omega_i \in \mathbb{R}$. Recall that according to Corollary 12 we consider networks of oscillators having equal natural frequencies, i.e. $\Omega_i = \Omega_j = \Omega$.

Recall that the WCNN (10.1) near multiple Andronov-Hopf bifurcation is governed by the canonical model (see Theorem 11)

$$z_i' = b_i z_i + d_i z_i |z_i|^2 + \sum_{j \neq i}^n c_{ij} z_j$$

We restrict our attention to the case when all coefficients d_i are real and negative. This implies that the frequency of oscillators does not depend upon their amplitudes. The condition $d_i < 0$ implies that the Andronov-Hopf bifurcation for each oscillator is supercritical. This corresponds to a birth of a stable limit cycle. Without loss of generality we can take $d_i = -1$. Thus, we study the dynamical system

$$z_i' = (\rho_i + i\omega_i)z_i - z_i|z_i|^2 + \sum_{j \neq i}^n c_{ij}z_j$$
(10.6)

for i = 1, ..., n.

Synaptic Coefficients

In this chapter we need the exact relationship between complex-valued synaptic coefficients c_{ij} and actual synaptic connections S_{ij} . This is important for neurobiological interpretations of the theory developed below. It will allow us to interpret all results obtained by studying (10.6) in terms of the original WCNN (10.1), i.e. in terms of excitatory and inhibitory populations of neurons and interactions between them.

Lemma 59 The relationship between c_{ij} and S_{ij} is given by

$$c_{ij} = \frac{1}{2} \left(1 + \frac{ia_4}{\Omega} \quad , \quad -\frac{ia_2}{\Omega} \right)_i S_{ij} \left(\begin{array}{c} 1 \\ \frac{a_4 + i\Omega}{a_2} \end{array} \right)_j.$$
(10.7)

Proof. The eigenvectors of L_i are

$$v_i = \begin{pmatrix} 1\\ rac{a_4 + \mathrm{i}\Omega}{a_2} \end{pmatrix}_i$$
 and $\bar{v}_i = \begin{pmatrix} 1\\ rac{a_4 - \mathrm{i}\Omega}{a_2} \end{pmatrix}_i$

The dual vectors are

$$w_i = \frac{1}{2} \left(1 + i \frac{a_4}{\Omega}, -i \frac{a_2}{\Omega} \right)_i$$

and

$$\bar{w}_i = \frac{1}{2} \left(1 - \mathrm{i} \frac{a_4}{\Omega}, \ \mathrm{i} \frac{a_2}{\Omega} \right)_i.$$

The result follows from expressions (2.26) and (10.4). \Box

10.1.2 Type A and B Neural Oscillators

Recall that according to Dale's principle $a_2 < 0$ and $a_3 > 0$. Hence, each Jacobian matrix L generically either has signature

$$\begin{pmatrix} + & - \\ + & - \end{pmatrix}$$
 or $\begin{pmatrix} - & - \\ + & + \end{pmatrix}$.

In first case we say that the neural oscillator is of type A, in the second case type B.



Figure 10.4: Differences in dynamic behavior of type A and B neural oscillators. See text

Consider the differences between these types. Suppose the neural oscillator is at the equilibrium point (0,0) and we apply a short impulse to its excitatory neuron so that its new activity becomes $0 + \delta x$, where δx is small. Since the equilibrium (0,0)is stable and the Andronov-Hopf bifurcation is non-degenerate, the activity vector (x(t), y(t)) approaches (0,0).

If the neural oscillator is of type A, the relaxation to (0,0) has the following form: The excitatory neuron increases its activity x further increasing the inhibitory neuron activity y. After a while, y reciprocally inhibits x and both neuron activities decrease (see Figure 10.4a).

If the neural oscillator is of type B, then there is no such initial amplified response of excitatory neuron activity, but there is an increasing of inhibitory neuron activity (see Figure 10.4b). We see that in type A neural oscillators it is possible for excitatory and inhibitory neurons to reach the peak of activity approximately simultaneously, whereas for type B neural oscillators it is not.

The differences between type A and B neural oscillators were not essential for our mathematical analysis so far. Nevertheless, they are crucial in the next section and in Section 10.6 where we show that type A neural oscillators have some interesting properties: In some sense they are "smarter".

10.2 Dale's Principle and Connectivity

There are some neurophysiological implications of formula (10.7). We first observe that if the WCNN (10.1) is uncoupled then so is the canonical model (10.6). Indeed, if $p_i = q_i = 0$ for all *i*, then

$$S_{ij} = \begin{pmatrix} \frac{\partial p_i}{\partial x_j} & \frac{\partial p_i}{\partial y_j}\\ \frac{\partial q_i}{\partial x_j} & \frac{\partial q_i}{\partial y_j} \end{pmatrix} = \begin{pmatrix} 0 & 0\\ 0 & 0 \end{pmatrix}$$

and, hence, from (10.7) $c_{ij} = 0$ for all i and j.

It turns out that the converse is not true. There could be nonzero connection functions p_i and q_i such that the synaptic coefficient $c_{ij} = 0$. This means that although two neural oscillators can be physically connected $(S_{ij} \neq 0)$, the synaptic connections between them are not effective because the canonical model is uncoupled $(c_{ij} = 0)$. The existence of such a phenomenon follows from the next theorem.

Theorem 60 If the *i*-th and the *j*-th neural oscillators are of type A, then there are nonzero synaptic configurations S_{ij} between the *i*-th and *j*-th neural oscillators such that $c_{ij} = 0$. Such configurations can be found that satisfy Dale's principle. If the neural oscillators are of type B, then such synaptic configurations always violate Dale's principle.

Proof. Consider formula (7.10) for c_{ij} . After rearrangement we see that it is equivalent to

$$c_{ij} = v_1 s_1 + v_2 s_2 + v_3 s_3 + v_4 s_4, (10.8)$$

where

$$v_1 = \frac{1}{2} + \mathrm{i}\frac{a_{i4}}{2\Omega},$$

$$v_2 = \frac{a_{j4} - a_{i4}}{2a_{j2}} + \mathrm{i}\frac{a_{j4}a_{i4} + \Omega^2}{2a_{j2}\Omega},$$



Figure 10.5: Complex numbers $v_1, -v_2, v_3, -v_4$ as vectors on the complex plane. For simplicity we depict $-v_2$ with zero real part. a. Type A neural oscillator. b. Type B neural oscillator

$$v_3 = -\mathrm{i}rac{a_{i2}}{2\Omega},$$
 $v_4 = rac{a_{i2}}{2a_{i2}} - \mathrm{i}rac{a_{j4}a_{i2}}{2a_{j2}\Omega}$

are four complex numbers. Consider v_1, v_2, v_3 and v_4 as vectors in \mathbb{R}^2 . Obviously, they are linearly dependent. Hence the equation

$$0 = v_1 s_1 + v_2 s_2 + v_3 s_3 + v_4 s_4, \tag{10.9}$$

has nontrivial solutions.

In order to satisfy Dale's principle s_1 and s_3 must be non-negative and s_2 and s_4 non-positive. Thus we must find solutions for

$$0 = v_1 s_1 + (-v_2)|s_2| + v_3 s_3 + (-v_4)|s_4|$$
(10.10)

with non-negative coefficients s_1 , $|s_2|$, s_3 and $|s_4|$.

If the neural oscillators are of type B, then all of the complex numbers $v_1, -v_2, v_3$ and $-v_4$ have positive imaginary parts (see Figure 10.5b). All their non-trivial linear combinations with non-negative coefficients also have positive imaginary parts. Thus equation (10.9) and Dale's principle cannot be satisfied simultaneously.

If the neural oscillators are of type A, then $v_1, -v_2, v_3, -v_4$ have imaginary parts with different signs (see Figure 10.5a), and it is easy to see that (10.10) can be satisfied. Indeed, one can take a linear combination of v_1 and $-v_4$ so that the imaginary part



Figure 10.6: Synaptic configurations that can exhibit the phenomenon described in Theorem 60. Open boxes depict excitatory neurons and shaded circles depict inhibitory neurons. A vertical pair of excitatory and inhibitory neurons is one neural oscillator

(projection on vertical axis) of $v_1s_1 + (-v_4)|s_4|$ is zero and add v_3 or $-v_2$ or both to kill the real part (projection on horizontal axis). Thus, equation (10.9) and Dale's principle can be satisfied simultaneously. \Box

Theorem 60 could be a revelation for neuroscientists. Indeed, it is probably difficult to accept that the existence of synaptic connections between neurons from two different neural oscillators does not necessarily imply that the oscillators interact. We have already seen this for oscillators that have different natural frequencies Ω (see Corollary 12). In Theorem 60 the neural oscillators could be identical; they can even act as if they were synchronized, but the synaptic connections between them are not responsible for that. If we somehow suppress one of them, the other one will not "feel" it. Its amplitude and phase will be unchanged. It should be noted that the effect described above is a synergetic phenomenon, and can be observed on the level of interacting neural oscillators, but possibly not on the level of interacting neurons.

A direct consequence of the preceding proof is the following

Corollary 61 The phenomenon described in Theorem 60 can be observed only in networks with the synaptic configurations depicted in Figure 10.6.

Proof. From Figure 10.5 it follows that (10.10) has non-trivial solutions with nonnegative coefficients only if $s_1 \neq 0$, $s_4 \neq 0$ and either $s_2 \neq 0$ or $s_3 \neq 0$. The result follows from Figure 10.3. \Box In all of the synaptic configurations the inhibitory neurons should be long-axon neurons capable of forming long-distance synaptic contacts (i.e. contacts from one neural oscillator to another one), which is rare in the brain. All of the brain structures we studied are not of the type depicted in Figure 10.6.

10.3 Classification of Synaptic Organizations

According to Lemma 39 the values $\psi_{ij} = \text{Arg } c_{ij}$ decode phase information. Let us determine possible values of the natural phase differences ψ_{ij} for various synaptic configurations S_{ij} satisfying Dale's principle. It is easy to do this using (10.8) and the vectors from Figure 10.5. Indeed, if for example, S_{ij} has the form

$$S_{ij} = \left(\begin{array}{cc} s_1 & 0\\ s_3 & 0 \end{array}\right),$$

where s_1 and s_3 are some positive numbers, then

$$c_{ij} = v_1 s_1 + v_3 s_3,$$

and, hence, possible values of c_{ij} are linear combinations of v_1 and v_3 with positive coefficients. Values of c_{ij} may be anywhere between lines spanned through v_1 and v_3 in Figure 10.5a or b. Thus, ψ_{ij} is an angle between Arg v_1 and $\pi/2$ as shown in Figure 10.7.

Complete classification of c_{ij} for all S_{ij} is given in Figure 10.7. We see that type A neural oscillators are able to reproduce the entire range of natural phase differences if they are suitably connected, while type B oscillators cannot. Indeed, all possible values $\psi_{ij} = \text{Arg } c_{ij}$ for type B oscillators are between Arg v_1 and Arg $(-v_4)$. Type B neural oscillators cannot even have natural phase difference $\psi_{ij} = 0$, whereas the type A oscillators can.



Figure 10.7: Possible values of synaptic connections c_{ij} for different synaptic configurations S_{ij} satisfying Dale's principle. For synaptic configurations that are not explicitly depicted the possible values of c_{ij} may occupy all shaded areas. a. Type A neural oscillator. b. Type B neural oscillator

Using the classification in Figure 10.7 we can solve a number of problems. Knowing the phase shift between two neural oscillators we can find possible synaptic configurations that can produce the shift; Knowing changes in synapses we can find changes in phase shifts and vice versa, etc.

We will use this classification below when we analyze possible synaptic organizations from the point of view of memorization of phase information.

156

10.4 Learning Dynamics

Much is not known about learning in the human brain, but our major hypotheses about the learning dynamics appear to be consistent with observations. We assume that

- Learning results from modifying synaptic connections between neurons (Hebb 1949).
- Learning is local, i.e. the modification depends upon activities of pre- and postsynaptic neurons and does not depend upon activities of the other neurons.
- The modification of synapses is slow compared with characteristic times of neuron dynamics.
- If either pre- or post-synaptic neurons or both are silent, then no synaptic changes take place except for exponential decay, which corresponds to forget-ting.

These assumptions in terms of the WCNN (10.1) have the following implications: The first hypothesis states that learning is described by modification of the matrices S_{ij} defined in (10.4). Recall that the actual synaptic connections have order ε . We denote them by W_{ij} . Thus

$$W_{ij} = \begin{pmatrix} w_1 & w_2 \\ w_3 & w_4 \end{pmatrix}_{ij} = \varepsilon S_{ij}.$$
 (10.11)

The second hypothesis says that for fixed i and j the entries of W_{ij} are modified according to equations of the form

$$\begin{cases} w_1' = h_1(w_1, x_i, x_j), \\ w_2' = h_2(w_2, x_i, y_j), \\ w_3' = h_3(w_3, y_i, x_j), \\ w_4' = h_4(w_4, y_i, y_j), \end{cases}$$
(10.12)

where (x_i, y_i) are local coordinates at the origin. We introduce the "slow" time $\tau = \varepsilon t$ to account for the third hypothesis. We say that a neural oscillator is silent if its activity is at an equilibrium point, i.e. it does not oscillate. Then the fourth hypothesis says that

$$\begin{aligned} h(w,0,y) &= h(w,x,0) = h(w,0,0) = \tilde{h}(w) \\ &= -\gamma w + \delta w^2 + \dots, \end{aligned}$$

for all x and y, so that h has the form

$$h(w, x, y) = -\gamma w + \theta xy + \delta_1 wx + \delta_2 wy + \delta w^2 + \dots$$
(10.13)

It follows from (10.11) that a synaptic coefficient w is of order ε . From (2.24) we know that the activities x and y are of order $\sqrt{\varepsilon}$. After rescaling by $w \to \varepsilon s$, $x \to \sqrt{\varepsilon}x$, $y \to \sqrt{\varepsilon}y$, we obtain the learning rule

$$s' = -\gamma s + \theta x y + \mathcal{O}(\sqrt{\varepsilon}), \qquad (10.14)$$

which we refer to as the Hebbian synaptic modification rule. Note that although we consider general functions h, after the rescaling only two constants γ and θ are significant to leading order. They are the rate of memory fading and the rate of synaptic plasticity, respectively.

We assume that the fading rate γ is positive and the same for all synapses. The plasticity rates can differ for different synapses. To distinguish them we write θ_{ijk} for i, j = 1, ..., n; k = 1, 2, 3, 4.

In order to understand how learning influences the dynamics of the canonical model (10.6) we must calculate the changes in c_{ij} .

Lemma 62 If all the conditions listed above are satisfied, then

$$c_{ij}' = -\gamma c_{ij} + k_{ij2} z_i \bar{z}_j + k_{ij3} \bar{z}_i z_j, \qquad (10.15)$$

where

$$k_{ij2} = \frac{1}{2} \left(\theta_{ij3} + \theta_{ij1} + \sigma_j (\theta_{ij4} + \theta_{ij2}) + \frac{ia_{i1}}{\Omega} (\theta_{ij3} - \theta_{ij1} + \sigma_j (\theta_{ij4} - \theta_{ij2})) \right)$$

$$k_{ij3} = \frac{1}{2} \left(1 + \frac{ia_{i4}}{\Omega} \right) \left[\theta_{ij1} - \theta_{ij3} + \left(\frac{a_{j4} + i\Omega}{a_{j2}} \right)^2 (\theta_{ij2} - \theta_{ij4}) \right]$$

$$\sigma_j = -\frac{a_{j3}}{a_{j2}} > 0.$$
(10.16)

Proof. Using (10.7) we see that

$$c_{ij}' = \frac{1}{2} \left(1 + \frac{\mathrm{i}a_4}{\Omega} \ , \ -\frac{\mathrm{i}a_2}{\Omega} \right)_i \left(\begin{array}{c} s_1' & s_2' \\ s_3' & s_4' \end{array} \right)_{ij} \left(\begin{array}{c} 1 \\ \frac{a_4 + \mathrm{i}\Omega}{a_2} \end{array} \right)_j.$$

From (10.11) we have

$$s_{ijk}' = \frac{1}{\varepsilon} w_{ijk}', \quad k = 1, 2, 3, 4$$

for all i and j. So (10.12) and (10.13) imply that

$$\begin{cases}
c_{ij}' = -\gamma c_{ij} + \frac{1}{2\varepsilon} \times \\
\left\{ \left(1 + \frac{ia_4}{\Omega}, -\frac{ia_2}{\Omega} \right)_i \left(\begin{array}{c} \theta_1 x_i x_j & \theta_2 x_i y_j \\ \theta_3 y_i x_j & \theta_4 y_i y_j \end{array} \right)_{ij} \left(\begin{array}{c} 1 \\ \frac{a_4 + i\Omega}{a_2} \end{array} \right)_j + \text{h.o.t.} \right\}.
\end{cases} (10.17)$$

Using (2.24) for x_i and y_i we can rewrite (10.17) in terms of z_i and z_j as

$$c_{ij}' = -\gamma c_{ij} + k_{ij1} e^{\frac{i2\Omega}{\epsilon}\tau} z_i z_j + k_{ij2} z_i \bar{z}_j + k_{ij3} \bar{z}_i z_j + k_{ij4} e^{\frac{i-2\Omega}{\epsilon}\tau} \bar{z}_i \bar{z}_j + \mathcal{O}(\sqrt{\epsilon}), \qquad (10.18)$$

where $k_{ij1}, k_{ij2}, k_{ij3}$ and k_{ij4} are some coefficients. We are not interested in k_{ij1} and k_{ij4} because after averaging all terms containing $e^{\frac{i m\Omega}{\epsilon}\tau}$ with $m \neq 0$ disappear, and we will have

$$c_{ij}' = -\gamma c_{ij} + k_{ij2} z_i \bar{z}_j + k_{ij3} \bar{z}_i z_j + \mathcal{O}(\sqrt{\varepsilon}).$$

Taking the limit $\varepsilon \to 0$ gives (10.15). It is easy to check that k_{ij2} and k_{ij3} are given as shown in (10.16). \Box

Remark 63 The fourth hypothesis about the learning dynamics is redundant. Indeed, if we drop it, then the function h(w, x, y) defined by (10.13) acquires linear terms in x and y. They add linear terms of the form $e^{i(\Omega/\epsilon)\tau}z$ and $e^{-i(\Omega/\epsilon)\tau}\bar{z}$ to (10.18), which eventually vanish after the averaging.

Note that we assumed little about the actual learning dynamics. Nevertheless, the family of possible learning rules (10.15) that satisfy our assumptions is apparently narrow. In the next section we show that to be "useful" the learning rule (10.15) must satisfy the additional conditions: Im $k_{ij2} = 0$ and $k_{ij3} = 0$. Using this and (10.16) we can determine what restrictions must be imposed on the plasticity rates $\theta_{ij1}, \ldots, \theta_{ij4}$, and, hence, onto the possible organization of the network so that it can memorize phase information, which we discuss next.

10.5 Memorization of Phase Information

We develop here the concept of memorization of *phase differences*. By this we understand the following: If during a learning period neuron A excites neuron B such that B generates an action potential with time delay τ , then changes occur so that whenever A generates an action potential then so does B with the same time delay τ .

Since in the real brain neurons tend to generate the action potentials repeatedly, instead of the time delay we will be interested in phase difference between dynamics of the neurons A and B. So, if during a learning period, two neural oscillators generate action potentials with some phase difference, then after the learning is completed, they can reproduce the same phase difference.

Whether memorization of phase differences is important or not is a neurophysiological question. We suppose here that it is important. Then, we would like to understand what conditions must be imposed on a network's architecture to ensure it can memorize phase differences.

The memorization of phase information in terms of the canonical model (10.6) means the following: Suppose during a learning period the oscillator activities $z_i(\tau)$ are given so that the phase differences $\operatorname{Arg} z_i \bar{z}_j$ are kept fixed. We call the pattern of the phase differences the *image* to be memorized. Suppose also that the synaptic coefficients c_{ij} are allowed to evolve according to the learning rule (10.15). Then we say that the canonical model *memorized* the image if there is an attractor in the z-space such that when the activity $z(\tau)$ is on the attractor, the phase differences between the oscillators coincide with those to be learned.

Theorem 64 Suppose the neural oscillators have equal center frequencies $\omega_1 = \cdots = \omega_n = \omega$. Consider the weakly connected network of such oscillators governed by

$$z_i' = (\rho_i + i\omega)z_i - z_i|z_i|^2 + \sum_{j=1}^n c_{ij}z_j, \quad i = 1, \dots, n,$$

together with the learning rule (10.15). The network can memorize phase differences of at least one image if and only if

$$k_{ij2} > 0 \quad and \quad k_{ij3} = 0,$$
 (10.19)

i.e. the learning rule (10.15) has the form

$$c_{ij}' = -\gamma c_{ij} + k_{ij} z_i \bar{z}_j, \quad i \neq j, \tag{10.20}$$

where k_{ij} , i, j = 1, ..., n, are positive real numbers.

Proof. Let us introduce the new rotating coordinate system $e^{i\omega\tau}z_i(\tau)$. In the new coordinates the canonical model becomes

$$z_i' = \rho_i z_i - z_i |z_i|^2 + \sum_{j=1}^n c_{ij} z_j, \quad i = 1, \dots, n.$$
 (10.21)

First, we prove that (10.19) is a sufficient condition. Our goal is to show that after learning is completed, the dynamical system (10.21) has an attractor such that the phase differences on the attractor coincide with those of the memorized pattern.

Let (10.21) be in the learning mode such that the phase differences $\phi_i(\tau) - \phi_j(\tau)$ (mod 2π) =Arg z_i -Arg z_j (mod 2π) =Arg $z_i \bar{z}_j$ are kept fixed. Then, according to the learning rule (10.20) the coefficients c_{ij} approach $\frac{k_{ij}}{\gamma} z_i \bar{z}_j$ and, hence, ψ_{ij} approaches Arg $z_i \bar{z}_j$, where $c_{ij} = |c_{ij}| e^{i\psi_{ij}}$. Note that ψ_{ij} satisfies

$$\psi_{ij} = -\psi_{ji} \quad \text{and} \quad \psi_{ij} = \psi_{ik} + \psi_{kj} \tag{10.22}$$

for any i, j and k. We must show that after learning is completed the neural network can reproduce a pattern of activity having the memorized phase differences ψ_{ij} .

We assumed that during learning all activities $z_i \neq 0$ so that the phases ϕ_i of the oscillators are well-defined. It is easy to see that after learning is complete we have $c_{ij} \neq 0$ for $i \neq j$.

Consider (10.21) in polar coordinates

$$\begin{cases} r_i' = \rho_i r_i - r_i^3 + \sum_{j=1}^n |c_{ij}| r_j \cos(\phi_j + \psi_{ij} - \phi_i) \\ \phi_i' = \frac{1}{r_i} \sum_{j=1}^n |c_{ij}| r_j \sin(\phi_j + \psi_{ij} - \phi_i). \end{cases}$$
(10.23)

Let us show that the radial components, determined by

$$r_i' = \rho_i r_i - r_i^3 + \sum_{j=1}^n |c_{ij}| r_j \cos(\phi_j + \psi_{ij} - \phi_i), \qquad i = 1, \dots, n$$
(10.24)

are bounded. Indeed, let $B(0, R) \subset \mathbb{R}^n$ be a ball at the origin with arbitrarily large radius R > 0. Consider the flow of (10.24) outside the ball. After the rescaling $r_i \to Rr_i, \ \tau \to R^{-2}\tau$, the system (10.24) becomes

$$r_i' = -r_i^3 + \mathcal{O}(R^{-2}), \qquad i = 1, \dots, n,$$

which is an R^{-2} -perturbation of

$$r_i' = -r_i^3, \qquad i = 1, \dots, n.$$
 (10.25)
For any initial conditions the activity vector of (10.25) is inside a unit ball B(0,1) after some finite transient. Any perturbations of (10.25) has the same property. Therefore, after the finite transition interval the activity vector of (10.24) is inside B(0, R) for any initial conditions and any values of ϕ_1, \ldots, ϕ_n . Hence, all attractors of (10.23) lie inside the cylinder $B(0, R) \times \mathbb{R}^n \subset \mathbb{R}^{2n}$.

Fix index k. It is easy to check that the hyperplane

$$\phi_i = \phi_k + \psi_{ik}, \quad i \neq k \tag{10.26}$$

is a global invariant manifold. Indeed, using (10.22) we have

$$\phi_j + \psi_{ij} - \phi_i = \phi_k + \psi_{jk} + \psi_{ij} - \phi_k - \psi_{ik} = \psi_{ij} - \psi_{ij} = 0$$
(10.27)

for all i and j, and, hence,

$$\phi_i'=0$$

From (10.27) we obtain the same invariant manifold for any other choice of k.

In order to study the stability of the manifold consider the auxiliary system

$$\phi_i' = \frac{1}{r_i} \sum_{j=1}^n |c_{ij}| r_j \sin(\phi_j + \psi_{ij} - \phi_i), \quad i \neq k,$$
(10.28)

where r_1, \ldots, r_n and ϕ_k are fixed. Since all attractors of (10.23) are inside the cylinder, we may assume that $r_i < R$ for all *i*. The Jacobian matrix of (10.28), say $J = (J_{ij})_{i,j \neq k}$, is diagonal-dominant because

$$J_{ij} = \begin{cases} \frac{r_j}{r_i} |c_{ij}| & i \neq j, \\ -\frac{1}{r_i} \sum_{m=1}^n |c_{im}| r_m & i = j \end{cases}$$

and

$$J_{ii} + \sum_{j \neq i}^{n} J_{ij} = -\frac{r_k}{r_i} |c_{ik}| < -r_k \mu_k < 0,$$

where

$$\mu_k = \frac{1}{R} \min_{i \neq k} |c_{ik}|.$$

This means that all eigenvalues of J have negative real parts; Hence, (10.26) is an asymptotically stable equilibrium point for (10.28). Therefore, in the original 2n-dimensional system (10.23) the flow is directed everywhere toward the invariant manifold, at least for $(r_1, \ldots, r_n) \neq (0, \ldots, 0)$. Hence, the manifold contains an attractor of (10.23). Moreover, it is possible to prove that the complement of its domain of attraction has measure zero, i.e. this is the only attractor for (10.21).

Note that on the manifold the phase differences satisfy

$$\phi_i - \phi_j = \psi_{ij}.$$

Thus (10.19) is a sufficient condition memorization and recall of phase differences.

It should be stressed that the oscillators have constant phase shifts on the manifold even when the attractor is not an equilibrium point. For example, if the attractor were chaotic, then one could observe an interesting phenomenon: The oscillator's amplitudes r_1, \ldots, r_n have chaotic activity whereas their phases ϕ_1, \ldots, ϕ_n have constant differences ψ_{ij} . Thus, the synchronization does not necessarily mean that the entire network's activity is on the limit cycle.

Next, we show that conditions (10.19) are necessary. Since the pattern of activity to be memorized and the values of ρ_1, \ldots, ρ_n are not specified, it is assumed that the network can learn and reproduce phase shifts of any activity pattern $z^* = (z_1^*, \ldots, z_n^*)^{\mathsf{T}}$ for any choice of ρ_1, \ldots, ρ_n .

The phase difference between *i*-th and *j*-th oscillators during the learning period is Arg z_i^* -Arg $z_j^* \pmod{2\pi} = \operatorname{Arg} z_i^* \overline{z}_j^*$. Hence, the same value must be reproduced after the learning is completed. From Lemma 39 it follows that the network can always reproduce the phase shifts $\phi_i - \phi_j \pmod{2\pi} = \psi_{ij} = \operatorname{Arg} c_{ij}$. Therefore the equality $\psi_{ij} = \operatorname{Arg} z_i^* \overline{z}_j^*$ must be satisfied. This is possible, for any z^* , only if (10.19) holds. Hence (10.19) is necessary. \Box Note the similarity of (10.20) and the Hebbian rule (10.14). The only difference is that in (10.20) the variables c and z are complex-valued.

Let us rewrite (10.20) in polar coordinates: If $c_{ij} = |c_{ij}|e^{i\psi_{ij}}$ and $z_i = r_i e^{i\phi_i}$, then

$$\begin{cases} |c_{ij}|' = -\gamma |c_{ij}| + k_{ij}r_ir_j\cos(\phi_i - \phi_j - \psi_{ij}) \\ \psi_{ij}' = \frac{1}{|c_{ij}|}k_{ij}r_ir_j\sin(\phi_i - \phi_j - \psi_{ij}). \end{cases}$$

From the second equation it is clear that

$$\psi_{ij} \to \phi_i - \phi_j \pmod{2\pi}$$

as we expected on the basis of Lemma 39. Notice that if $\psi_{ij} = \phi_i - \phi_j$, then $\cos(\phi_i - \phi_j - \psi_{ij}) = 1$ and first equation coincide with the Hebbian learning rule.

Since we know how k_{ij2} and k_{ij3} depend upon the original WCNN we can restate the results of Theorem 64 in terms of (10.1). Almost all results discussed in the next section are straightforward consequences of the following

Corollary 65 A weakly connected network of neural oscillators can memorize phase differences if and only if the plasticity rates satisfy

$$\theta_{ij1} = \theta_{ij3}, \qquad \theta_{ij2} = \theta_{ij4} \tag{10.29}$$

and

$$k_{ij} = (\theta_{ij1} + \sigma_j \theta_{ij2}) > 0, \tag{10.30}$$

where $\sigma_j > 0$ is defined in (10.16).

The proof follows from application of the condition (10.19) in Theorem 64 to the representation (10.16) in Lemma 62.

10.6 Synaptic Organizations

In this section we apply Corollary 10.30 to various synaptic organizations.

Corollary 66 The rate of synaptic plasticity is locally determined by the pre-synaptic neurons.

Proof. Indeed, the constants θ_{ij1} and θ_{ij3} determine the rate of synaptic plasticity from the same pre-synaptic neuron x_j onto different post-synaptic neurons x_i and y_i , which belong to the same neural oscillator. It follows from condition (10.29) that the plasticity rates must coincide. Similarly the same result is true for θ_{ij2} and $\theta_{ij4} \square$

It would be incorrect to think that the values of the actual synaptic connections S_{ij} depend exclusively upon pre-synaptic neurons. The Corollary merely claims that the pre-synaptic neurons are the only neurons that regulate the rates of plasticity, the "speed" of modification, but not the modification itself.

This corollary is consistent with neurobiology. It is known (Shepherd 1983) that many chemicals pass through the axon of a neuron. Modifications and growth of the axon terminals depend crucially upon these chemicals and, hence, upon the neuron. So, it should be expected that the rate of modification, which is connected with learning, also depends crucially upon the pre-synaptic neuron.

Let us study possible synaptic organizations of the network from the point of view of learning of phase differences. Suppose that $\theta_{ij1} = 0$ for some $i \neq j$. That is, there is no modification of synapses between *j*-th and *i*-th excitatory neurons except for fading (atrophy). So, even if a synapse s_{ij1} between x_j and x_i existed at the beginning, it would atrophy with time. Thus, without loss of generality we may assume that $\theta_{ij1} = 0$ means that formation and growth of synapses from x_j onto x_i is impossible. The same consideration can be applied to the $\theta_{ij2}, \theta_{ij3}$ and θ_{ij4} . In figure 10.8 we draw arrows from one neuron to another only when the corresponding plasticity rate is nonzero, i.e. only if synaptic contact between the two neurons is possible. Different choices of the arrows correspond to different synaptic organizations of the neural



Figure 10.8: Open boxes are excitatory neurons, shaded circles are inhibitory neurons. If there is an arrow between two neurons, then the synaptic contact is possible.

a. The synaptic organizations that cannot memorize phase information.

b. The synaptic organization that can either learn or unlearn phase information (but not both). If the network has more than two oscillators, then the Dale's principle will be violated during the learning.

c. The synaptic organization that can learn phase information.

d. The synaptic organization that can both learn and unlearn phase information.

network.

Corollary 67 The synaptic organizations depicted in Figure 10.8a cannot memorize phase information.

Proof. According to the condition (10.29) if one of the plasticity rates is zero, then so should be the other one corresponding to it. Thus, the arrows must be in pairs, i.e. if a neuron has synaptic contacts with some neural oscillator then it must have access to both excitatory and inhibitory neurons of the neural oscillator. Obviously, none of the architectures on Figure 10.8a satisfies this condition. \Box

167



Figure 10.9: c_{ij} and c_{ji} must be inside the shaded area between $-v_4$ and $-v_2$

Corollary 68 The synaptic organization depicted in Figure 10.8b can memorize phase information only if the network consists of two type A neural oscillators and the phase difference to be memorized is close to π .

Proof. From the proof of Theorem 64 it follows that Arg $c_{ij} = -\text{Arg } c_{ji}$, i.e. c_{ij} and c_{ji} must lie on two symmetric rays from the origin (see Figure 10.9). In order to satisfy Dale's principle both beams must be inside the shaded area between $-v_4$ and $-v_2$ (see classification in our previous paper). This is impossible if the neural oscillators are of type B. If they are of type A, then the phase difference to be memorized should be sufficiently close to π (within (Arg v_4)-neighborhood).

Suppose the network consists of more than two oscillators. Then the phase difference between first and second and between second and third neural oscillators should be close to π . Hence the phase difference between first and third oscillators is close to 0. Thus c_{13} and c_{31} violate Dale's principle. \Box

Since networks seldom have only two elements we can conclude that this synaptic organization is not much better than those depicted in Figure 10.8a.

We see that the only candidates for the synaptic organizations that can memorize phase differences are those depicted in Figure 10.8c and d. To discuss them we use the condition (10.30).

What are the signs of θ_{ij1} and θ_{ij2} ? We show that $\theta_{ij1} \ge 0$ as it was postulated by Hebb (1949). Indeed, if two excitatory neurons generate action potentials simultaneously, then the case $\theta_{ij1} \ge 0$ corresponds to increasing of strength of excitatory synaptic connections between them. For θ_{ij2} we have a slightly different situation. The pre-synaptic neuron y_j is inhibitory. It is not clear what changes take place in the synapses from y_j to x_i (or y_i) if they fire simultaneously. We consider both cases: $\theta_{ij2} \ge 0$, which corresponds to decreasing of strength of the inhibitory synapse $s_{ij2} < 0$, and $\theta_{ij2} \le 0$, which corresponds to increasing of the strength of s_{ij2} .

The case $\theta_{ij2} \ge 0$ is straightforward. Indeed, since $\sigma_j > 0$ we have $\theta_{ij1} + \sigma_j \theta_{ij2} > 0$ and both synaptic organizations depicted on Figure 10.8c and d can memorize phase differences.

The case $\theta_{ij2} \leq 0$ requires more attention. Obviously, the synaptic organization in Figure 10.8c can memorize phase differences because $\theta_{ij2} = 0$ for it, and hence condition (10.30) is satisfied.

The synaptic organization depicted in Figure 10.8d needs special discussion. In this case k_{ij} can be positive or negative depending upon the relative values of θ_{ij1} and θ_{ij2} (see Figure 10.10).

Note that when $k_{ij} < 0$ the network memorizes not the presented image but its inverse (photographic negative). Sometimes it is convenient to think of this being that the network unlearns the image. So, the synaptic organization on Figure 10.8d is able not only learn but also unlearn information simply by adjusting the plasticity rates θ_{ij1} and θ_{ij2} .

The special choice of the plasticity constants such that

$$k_{ij} = (\theta_{ij1} + \sigma_j \theta_{ij2}) = 0$$

is interesting in the following sense: Since the plasticity constants are not zero, there are undoubtly some changes in synaptic coefficients s_{ijk} between the neurons.



Figure 10.10: For different choices of the plasticity rates θ the oscillatory neural network can learn (k > 0), unlearn (k < 0) or passively forget (k = 0) phase information. Here $k = \theta_1 + \sigma \theta_2$. Negative value of θ_2 corresponds to increasing of strength of the inhibitory synapses. Positive value – to decreasing of the strength.

Nevertheless, the network as a whole does not learn anything because for $k_{ij} = 0$

$$c_{ij}' = -\gamma c_{ij}$$

as it follows from (10.20). Moreover, the network forgets (loses) information because $c_{ij} \rightarrow 0$ as τ increases. Hence, the full-connected synaptic organization can exhibit a broad dynamic repertoire.

By applying arguments similar to those used in the proof of Corollary 68, we can see that in both synaptic organizations the neural oscillators must be of type A, or else Dale's principle will be violated. Moreover, in the architecture on Figure 10.8c the phase difference to be memorized should be between $-\text{Arg } v_1$ and $\text{Arg } v_1$. Thus we have demonstrated the following result:

Corollary 69 The only synaptic organizations that can memorize phase information are those depicted on Figure 10.8c and d for type A neural oscillators. In both cases the excitatory neurons are long-axon neurons capable of forming synaptic connections with distant neurons and the inhibitory neurons might have long (case c) or short (case d) axons.

Note that in the synaptic organization depicted in Figure 10.8c only excitatory neurons may have long axons. The inhibitory neurons can make synaptic contacts

only between themselves and nearby excitatory neurons within the same neural oscillator (see Figure 10.3).

Copious neurophysiological data (Shepherd 1983, Rakic 1976) suggest that excitatory neurons usually have long axons and inhibitory neurons are local-circuit interneurons. It is believed that the inter-neurons process information locally whereas the long-axon neurons transmit it to other regions of brain (Rakic 1976).

In the model that we studied above there is no local processing of information. Each neural oscillator works in the very primitive regime – oscillation. Nevertheless, even this simple neural network suggests that it is very important to see this natural division into local circuit inter-neurons and long-axon relay neurons.

Chapter 11

Discussion

11.1 Canonical Models and Normal Forms

One of the most important results presented in this work is the reduction of general WCNN to canonical models. The canonical models have few non-linear terms, nevertheless, they capture the qualitative behavior of the original WCNN. Another method of simplification of dynamical systems is known as the Normal Form Theory (Arnold 1982, Guckenheimer and Holmes 1983).

Recall that the Normal Form Theory considers a dynamical system

$$\dot{x} = f(x)$$

at an equilibrium point. Let L be the Jacobian matrix at the equilibrium and let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of L. Then there is a near identity change of variables

$$y = x + h(x)$$

that transforms the original system to a new system of the form

$$\dot{y} = g(y)$$

with the following properties: The nonlinear function g has the smallest number of non-linear terms, i.e. it has only the resonant terms corresponding to integer-valued

relationships between the eigenvalues of L, namely, the relationships of the form

$$\lambda_s = \sum_{k=1}^n m_k \lambda_k,$$

where each m_k is a non-negative integer and $\sum m_k \ge 2$. The relationship above is called a *resonance*. Thus, by the change of variables y = x + h(x) it is possible to "kill" all non-linear terms except the resonant ones. The resulting system is called *the normal form*. Thus the question: Are the canonical models that we derived in this thesis the normal forms for WCNNs? The answer is NO.

Consider a WCNN

$$\dot{x}_i = f_i(x_i) + \varepsilon g_i(x), \quad x_i \in \mathbb{R}, \quad i = 1, \dots, n$$
(11.1)

near, say, multiple saddle-node bifurcation. Since $df_i/dx_i = 0$ for all *i*, the Jacobian matrix for (11.1) is a zero $n \times n$ -matrix. Since it has *n* zero eigenvalues, all terms in $f = (f_1, \ldots, f_n)$ are resonant. Therefore, the normal form for (11.1) coincides with (11.1), whereas the canonical model for multiple saddle-node bifurcation generically differs from (11.1). We see that direct application of the Normal Form Theory to WCNNs is useless. Nevertheless, it must be acknowledged that the Normal Form Theory helped us when we studied multiple Andronov-Hopf bifurcations in Section 2.5.

11.2 Synaptic Connections

Our use of the word "synaptic" is an abuse of language, which has become customary in the neural network literature. There is no reason for the connections between neurons to be *exclusively* synaptic. One neuron can affect another using other than direct synapse interactions. For instance, a neuron can cause secretion of neuropeptides and hormones that reach other neurons by passive diffusion or advection. Such connections can play significant roles in processing information by the brain though they are not synaptic. Nevertheless, they can be taken into account by the functions G_i in the WCNN (2.1) and, hence, are accounted for in the synaptic coefficients c_{ij} .

11.3 Mathematical Conditions and Biology

When we analyze the WCNNs, we impose some mathematical conditions, such as the adaptation condition

$$D_{\lambda}f_{i}\lambda_{1} + g_{i}(0,0,\rho_{0},0) = 0, \qquad (11.2)$$

or such as $\Omega_i = \Omega_j$, etc. There are many of them in Chapters 2 and 3.

Caution should be used in application of these conditions to biological objects. Biologists might say that we do not know (and probably will never know) all the underlying laws which govern the biological systems. Hence we might not know the exact values of the variables needed for checking the mathematical conditions. Moreover, due to constant perturbations from the outside world, the values of biological variables fluctuate. Thus, it could be unlikely that an exact condition, say (11.2), is satisfied for any reasonable period of time.

Mathematician might add that for generic f_i and g_i (11.2) holds only for λ_1 from a set of measure zero. Therefore, exact conditions such as (11.2) could be meaningless for biological systems.

Recall that all WCNNs have a small parameter ε . In our analysis ε operates as a basic unit of measure. All the conditions that we impose on the WCNN dynamics must be satisfied up to $\mathcal{O}(\varepsilon)$: The network should be in some ε -neighborhood of a multiple bifurcation point; Natural frequencies Ω_i should be ε -close to each other; The adaptation condition (11.2) should be satisfied up to order ε , etc. Thus, all conditions that we impose on WCNNs are not "exact". By allowing fluctuations of order ε we make the conditions meaningful for biological systems.

11.4 Co-dimensions of the Models

Let us count how many conditions should be satisfied so that WCNN (2.1) is governed, for example, by the canonical model (2.16). The number of conditions is called the *co-dimension*.

First, each neuron must be at a saddle-node bifurcation, there n conditions. Second, the adaptation condition (2.11) gives another n conditions. Thus, the codimension of (2.16) as a canonical model of (2.1) is 2n. Similar calculations show that the co-dimension of (2.21) is 3n. Thus, there is a natural question: Which of the models is more generic?

Suppose we have adapted networks of 100 and 10 neurons near multiple saddlenode bifurcation and cusp singularity, respectively. From the pure mathematical point of view the second network is more generic because there are only 30 conditions to be satisfied (instead of 200 for the first network).

From the common sense of a neurobiologist the answer is not so obvious. It is well known that despite the quantitative differences, the neurons are qualitatively similar. If there is an electro-physiological mechanism that forces a neuron to be near threshold and to be adapted, then the same mechanism should be present in the other 99 neurons in the 100-neuron network. In other words, if two conditions are satisfied for one neuron, then it is physiologically plausible that approximately the same conditions are satisfied for the other neurons. Thus, there are only two nonsimilar conditions imposed on the co-dimension-2n network, whereas there are three conditions in co-dimension-3n networks. We see that there is an apparent discrepancy between the mathematical and neural network notions of co-dimension. Another aspect of co-dimension is that when we have a dynamical system and a set of bifurcation parameters, it could be difficult (even numerically) to find a bifurcation point of high co-dimension. But if we know that there are interesting neurocomputational properties near some multiple bifurcation with high co-dimension, then it is an easy task to construct a dynamical system which is close to the multiple bifurcation.

11.5 List of Canonical Models

Below is the list of canonical models derived in this thesis. They are presented according to the order of appearance in the text.

Co-di- men- sion	Canonical model	Conditions	Reference
2 <i>n</i>	$x_i' = r_i + b_i x_i + x_i^2 + \sum c_{ij} x_j$	Saddle-node bifurcation, adaptation condition	Theorem 9
3n	$x'_i = r_i + b_i x_i \pm x_i^3 + \sum c_{ij} x_j$	Cusp singularity, adaptation condition	Theorem 10
2n – 1	$z'_i = b_i z_i + d_i z_i z_i ^2 + \sum c_{ij} z_j$	Andronov-Hopf bifurcation or quasi-static pitchfork bifurcation, equality of frequencies	Theorem 11 and Section 3.5.3
2 <i>n</i>	$\begin{array}{l} x_i' = -y_i + r_i x_i + x_i^2 + \sum c_{ij} x_j \\ y_i' = a_i x_i \end{array}$	Quasi-static saddle-node bifurcation, $\mu = \mathcal{O}(\epsilon^2)$	Theorem 23
3 <i>n</i>	$egin{aligned} x_i' &= -y_i + r_i x_i \pm x_i^3 + \sum c_{ij} x_j \ y_i' &= a_i x_i \end{aligned}$	Quasi-static pitchfork bifurcation, $\mu = \mathcal{O}(\epsilon^2)$	Theorem 28
2 <i>n</i>	$z'_{i} = A_{i}v_{i}z_{i} + b_{i}z_{i} z_{i} ^{2} + \sum_{i}c_{ij}z_{j}$ $v'_{i} = d_{i}(R_{i} + S_{i} z_{i} ^{2} + T_{i}v_{i})$	Quasi-static Andronov-Hopf bifurcation, equality of frequencies	Theorem 34
2 <i>n</i>	$x_i \mapsto x_i + \varepsilon \left(r_i + x_i^2 + \sum c_{ij} x_j\right)$	Saddle-node bifurcation, adaptation condition	Section 4.2.1
n	$x_i \mapsto x_i + \varepsilon \left(r_i x_i \pm x_i^3 + \sum c_{ij} x_j \right)$	Flip bifurcation,	Theorem 36

Derivation and analysis of the canonical models is considered to be the major result of the thesis.

BIBLIOGRAPHY

BIBLIOGRAPHY

J.C. Alexander, E.J. Doedel and H.G. Othmer (1990), On the Resonance Structure in a Forced Excitable System, SIAM J.Appl.Math., 50:1373-1418

V.I. Arnold (1982), Geometrical Methods in the Theory of Ordinary Differential Equations, Springer-Verlag, NY, (russian original, Moscow 1977)

D.G. Aronson, G.B. Ermentrout and N. Kopell (1990), Amplitude Response of Coupled Oscillators, Physica D 41:403-449

F. Attneave (1971), Multistability in Perception, Scientific American 225:63-71

S.M. Baer and T. Erneux (1992), Singular Hopf Bifurcation to Relaxation Oscillations II, SIAM J.Appl.Math., 52:1651-1664

S.M. Baer and T. Erneux (1986), Singular Hopf Bifurcation to Relaxation Oscillations, SIAM J.Appl.Math., 46:721-739

B. Baird (1986), Nonlinear dynamics of pattern formation and pattern recognition in the rabbit olfactory bulb, Physica D 22:150-175

K. Bar-Eli (1985), On the stability of coupled chemical oscillators, Physica D 14:242-252

J. Belair and P.Holmes (1984), On linearly coupled relaxation oscillations, Quarterly of Appl.Math 42:193-219

H.H. Dale (1935), Pharmacology and the nerve endings, Proc Roy Soc Med 28:319-332

T. Ditzinger and G. Haken (1989), Oscillations in the Perception of Ambiguous Patterns, Biological Cybernetics 61:279-287

W. Eckhaus (1983), Relaxation Oscillations Including a Standard Chase of French Ducks, Lecture Notes in Math.,985:432-449 R. Eckhorn, R. Bauer, W. Jordan, M. Brosch, W. Kruse, M. Munk, H.J. Reitboeck (1988), Coherent Oscillations: A mechanism of feature link in the visual cortex? Biol Cybern 60:121-130

P. Erdi, T. Grobler, G. Barna and K.Kaski (1993), Dynamics of the Olfactory bulb: bifurcations, learning, and memory, Biological Cybernetics 69:57-66

B. Ermentrout and N. Kopell (1994), Learning of phase lags in coupled neural oscillators, Neural Computation 6:225-241

G.B. Ermentrout and N. Kopell (1991), Multiple pulse interactions and averaging in systems of coupled neural oscillators, J.Math. Biol. 29:195-217

G.B. Ermentrout and N. Kopell (1990), Oscillator Death in Systems of Coupled Neural Oscillators, SIAM J. Appl. Math. 50:125-146

R. Fitzhugh (1969), Mathematical Models of Excitation and Propagation in Nerve, in Biological Engineering, H.P.Schwan, ed., McGraw-Hill, New York, NY pp.1-85

P. Frankel and T. Kiemel (1993), Relative phase behavior of two slowly coupled oscillators, SIAM J.Appl.Math 53:1436-1446

Z. Gajic, D. Petkovski and X. Shen (1990), Singularly Perturbed and Weakly Coupled Linear Control Systems, Lecture Notes in Control and Information Sciences, Springer-Verlag

Gantmacher (1959), Applications of the Theory of Matrices, Interscience, New York

J. Grasman (1987), Asymptotic Methods for Relaxation Oscillations and Applications, Springer-Verlag, New York.

C.M. Gray (1994), Synchronous oscillations in neuronal systems: mechanism and functions, Journal of Computational Neurosci 1:11-38

C.M. Gray, P. Konig, A.K. Engel and W. Singer (1989), Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties, Nature 338:334-337

M. Golubitsky and D. Shaeffer (1979), A theory for imperfect bifurcation via singularity theory, Communications on Pure and Applied Mathematics, 32:21-98

S. Grossberg (1988), Nonlinear neural networks: Principles, mechanisms, and architectures, Neural Networks 1:17-61

J. Guckenheimer and D. Holmes (1983), Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields, Springer, New York

D.O. Hebb (1949), The Organization of behavior. Wiley, New York

M.W. Hirsch (1989), Convergent activation dynamics in continuous time networks, Neural Networks 2:331-349

A.L. Hodgkin and A.F. Huxley (1954), A quantitative description of membrane current and application to conduction and excitation in nerve, J.Physiol., 117:500-544

J.J. Hopfield (1982), Neural networks and physical systems with emergent collective computational abilities, Proc Natl Acad Sci USA 79:2554-2558

F.C. Hoppensteadt and E.M. Izhikevich (1995a), Canonical Models for Bifurcations from Equilibrium in Weakly Connected Neural Networks, World Congress on Neural Networks, Washington DC, extended abstract, I:80-83

F.C. Hoppensteadt and E.M. Izhikevich (1995b), Synaptic Organizations and Dynamical Properties of Weakly Connected Neural Oscillators: I. Analysis of Canonical Model. Biological Cybernetics, accepted

F.C. Hoppensteadt (1993), Analysis and simulations of chaotic systems, Springer-Verlag, New York

F.C. Hoppensteadt (1991), The searchlight hypothesis, J Math Biol 29:689-691

F.C. Hoppensteadt (1989), Intermittent chaos, self-organization, and learning from synchronous synaptic activity in model neuron networks, Proc Natl Acad Sci USA 86:2991-2995

F.C. Hoppensteadt (1986), An Introduction to the Mathematics of Neurons, Cambridge Univ.Press, Cambridge, U.K.

G. Iooss and M. Adelmeyer (1992), *Topics in Bifurcation Theory*, Advanced series in nonlinear dynamics, vol.3, World Scientific.

E.M. Izhikevich and G.G. Malinetskii (1993), A neural network with chaotic behavior, preprint #17, Inst.Appl.Mathem., Russia Academy of Sciences (in russian)

E.M. Izhikevich and G.G. Malinetskii (1992), A possible role of chaos in neurosystems, Dokl.Akad.Nauk 326:626-632 (translated in Sov.Phys.Docl.(1993) 37(10) October 1992:492-495)

Ya.B. Kazanovich and R.M. Borisyuk (1994), Synchronization in a neural network of phase oscillators with the central element, Biol Cybern 71:177-185

N. Kopell and D.Somers (1995), Anti-phase solutions in relaxation oscillators coupled through excitatory interactions, J.Math.Biol. 33:261-280

N. Kopell (1986), Coupled oscillators and locomotion by fish. In: Levin S. (Ed) Nonlinear Oscillations in Biology and Chemistry. Lecture Notes in Biomathematics, Springer-Verlag.

J.M. Kowalski, G.L. Albert, B.K. Rhoades and G.W. Gross (1992), Neuronal networks with spontaneous, correlated bursting activity: Theory and Simulations, Neural Networks 5:805-822

V.I. Kryukov (1991), An attention model based on principle of dominanta, In: Holden AV, Kryukov VI (eds) Neurocomputers and attention. I. Neurobiology, synchronization and chaos, Manchester University Press, Manchester, 319-352

Z. Li and J.J. Hopfield (1989), Modeling the olfactory bulb and its neural oscillatory processings, Biological Cybernetics 61:379-392

C. von der Malsburg C and J. Buhmann (1992), Sensory segmentation with coupled neural oscillators, Biol Cybern 67:233-242

R.E. Mirollo and S.H. Strogatz (1990), Synchronization of pulse-coupled biological oscillators, SIAM J.Appl.Math. 50:1645-1662

J.C. Neu (1979), Coupled chemical oscillators, SIAM Journal of Applied Mathematics 37:307-315

G.M. Peponides and P.V. Kokotovic (1983), Weak connections, time scales, and aggregation of nonlinear systems, IEEE Trans. on Systems, Man, and Cybernetics, SMC-13:527-532

T. Poston and I. Stewart (1978), Nonlinear modeling of multistable perception, Behavioral Science 23:318-334

P. Rakic (1976), Local Circuit Neurons, MIT Press, Cambridge, Mass

A. Rapaport (1952), Ignition phenomenon in random nets, Bulletin of Mathematical Biophysics 14:35-44

G.M. Shepherd (1983), Neurobiology, Oxford University Press, New York

G.M. Shepherd (1976), Models of LCN Function in the Olfactory Bulb, In: Rakic P (Ed.) Local Circuit Neurons, MIT Press, Cambridge, Mass

C.A. Skarda and W.J. Freeman (1987), How brain makes chaos in order to make sense of the world, Behav Brain Sci 10:161-195

D. Somers and N. Kopell (1993), Rapid synchronization through fast threshold modulation, Biological Cybernetics 68:393-407

D.W. Storti and R.H. Rand (1986), Dynamics of two strongly coupled relaxation oscillators, SIAM J.Appl.Math. 46:56-67

I.N. Stewart and P.L. Peregoy (1983), Catastrophe theory modeling in psychology, Psychological Bulletin, 94:336-362

I. Tsuda (1992), Dynamic Link of Memory – chaotic memory map in nonequilibrium neural networks, Neural Networks 5:313-326

H.R. Wilson and J.D. Cowan (1973), A Mathematical Theory of the Functional Dynamics of Cortical and Thalamic Nervous Tissue, Kybernetik 13:55-80

H.R. Wilson and J.D. Cowan (1972), Excitatory and inhibitory interaction in localized populations of model neurons, Biophys J 12:1-24

M. Zak (1989), Weakly connected neural nets, Appl Math Lett 3:131-135

