Margarita Kuzmina, Eduard Manykin, Evgeny Grichuk
Oscillatory Neural Networks

Margarita Kuzmina, Eduard Manykin,
Evgeny Grichuk

Oscillatory Neural Networks

In Problems of Parallel Information Processing

DE GRUYTER

Mathematics Subject Classification 2010
68T45, 68T05, 68T40, 68T42, 93C10, 93C15, 93C40, 34H20

Physics and Astronomy Classification Scheme 2010
05.45.Xt, 87.10.-e, 87.18.Sn, 87.85.dq, 87.85.St, 42.65.-k, 42.30.Tz

Authors
Margarita Kuzmina, Ph. D.
Russian Academy of Sciences
Keldysh Institute of Applied Mathematics
Miusskaya Sq. 4
125047 Moscow
Russia
mg.kuzmina@gmail.com

Prof. Dr. habil. Eduard Manykin
National Research Center
Kurchatov Institute
Kurchatov Sq. 1
123182 Moscow
Russia
manykin_ea@nrcki.ru

Evgeny Grichuk, Ph. D.
National Research Center
Kurchatov Institute
Kurchatov Sq. 1
123182 Moscow
Russia
grichuk_es@nrcki.ru

ISBN 978-3-11-026835-5
e-ISBN 978-3-11-026920-8

Library of Congress Cataloging-in-Publication Data
A CIP catalog record for this book has been applied for at the Library of Congress.

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.dnb.de.

Cover image: Thinkstock/iStock
Typesetting: le-tex publishing services GmbH, Leipzig
Printing and binding: Hubert & Co. GmbH & Co. KG, Göttingen
♾Printed on acid-free paper
Printed in Germany

www.degruyter.com

Preface

The last decade was marked by highly increased interest to the study of human brain functioning. Understanding of the human brain has been recently claimed as one of the greatest problems of 21st century science. An integrated brain understanding would offer great benefits for computational mathematics, medicine, and intellectual robotics dealing with creation of intelligent machines capable of thinking, learning, and demonstrating human reactions.

Intentions to create a complex large-scale multilevel model of the human brain, based on physiologically accurate imitations of brain structure performance, were declared repeatedly. In particular, in the frames of the famous interdisciplinary Blue Brain Project, started in 2005, a physiologically detailed rat brain neocortical column consisting of 1.2×10^6 neurons and 10^9 synapses was created in 2011. Another detailed large-scale model of a mammalian brain thalamocortical system, comprising 10^{11} spiking neurons and almost 10^{15} synapses, was created in 2008 at the Neuroscience Institute, San Diego, CA, USA. These achievements in computer modeling of the brain structures can naturally provide inspirations for new research in related areas, such as neuroinformatics, communication theory, neuroscience, medicine, and neurorobotics. Nowadays there is a variety of programs on creation of artificial brains for intelligent "android" robots emulating the human brain. The performance of created artificial brains should be tested (and compared with the human brain performance) in a representative set of tasks. The task of image recognition belongs to the simplest type of such tasks. The aspects of human consciousness should be traditionally expertized using the so-called Alan Turing test, initiated in the early 1950s. There are claims, based particularly on quick growth trends of computer power, that the whole brain emulation using conventional computers could be done by 2025.

Further extensive brain studies will definitely stimulate the development of new brain-inspired parallel adaptive computational algorithms. Neuromorphic methods of image and signal processing based on the brain visual cortex performance and its activity-dependent plasticity are of particular interest. The scope of problems concerning the brain memory also remains a challenging and interesting topic. Although it has been understood that the source of the brain memory consists in synaptic connections change, the character of memory data representation in the brain (memory structure) continues to remain a big puzzle.

The information processing in the brain is realized via spatiotemporal dynamics of patterns of excitation in neuronal ensembles arising by means of neuron interactions. Learning processes, equipped with the brain memory and self-organizing ability, modify the pattern structure and dynamics. A combination of analytical mathematical tools (such as approaches from statistical physics and information geometry) and the development of artificial neural networks (ANNs) models and associated learning algorithms constitute the theoretical background for the investigation of the

brain structures functioning. The design of oscillatory neural networks, a special kind of ANN models, provides a valuable contribution into the elucidation of the basic principles of the brain performance, because they allow us to study capabilities and limitations of synchronization-based information processing.

The book is devoted to the description and analysis of various oscillatory neural network models. Neuromorphic dynamical synchronization-based methods of information processing can be developed based on oscillatory network models with controllable synchronization. Besides, recurrent oscillatory networks of associative memory can be designed. The oscillatory associative memory networks, closely related to analogous networks of complex-valued neurons, possess better memory characteristics compared with those inherent to recurrent associative memory networks of formal neurons, namely the higher memory storage capacity and poor extraneous memory.

The book consists of six chapters. The introductory material in the first chapter briefly presents the necessary data on the human brain and outlines the basic types of ANN models, the widely used classical and modern learning algorithms for ANN, and the well-known models of oscillatory neural networks. A short overview of modern branches of ANN theory, such as complicated small-world networks, multi-agent systems, and cognitive dynamical systems, is given, and brief information about biologically accurate large-scale computer models of the brain structures is also included into the first chapter.

The second chapter concerns oscillatory network versions of recurrent associative memory networks (so-called attractor oscillatory networks). The computation procedure executed by this type of networks is realized via network state relaxation into the vicinity of a stable attractor of network dynamics. The original result, obtained by the authors, related to the construction and analysis of associative memory networks of limit-cycle oscillators is presented. Chapter 2 also contains the most significant results of the macrodynamical approach for large-scale recurrent neural networks of associative memory, including the authors' original results. A short review of the most important results on complex-valued neural networks of associative memory, closely related to oscillatory associative memory networks, is also included in the chapter.

A series of oscillatory network models with synchronization-based performance are described in Chapter 3. In particular, the famous oscillatory network LEGION, the three-dimensional biologically plausible network model by Li imitating the brain visual cortex performance, and a three-dimensional oscillatory network model for image processing tasks developed by the authors are included. Besides, the oscillatory network models for the separation of mixed sound fluxes and for odor recognition, exploiting versions of dynamical binding principle, and two models including modeling of visual attention are presented in the chapter.

A neuromorphic oscillatory network model with controllable synchronization and self-organized performance, developed by the authors, for image processing tasks is described in Chapter 4. It is a spatially distributed network of limit-cycle oscillators localized at the nodes of a two-dimensional lattice isomorphic with an image pixel array.

Single oscillator dynamics is tunable by the brightness value of the image pixel corresponding to the oscillator. A number of network coupling principles were constructed for different image processing tasks. The model demonstrates a variety of capabilities, such as brightness segmentation of real multipixel gray-level images, color image segmentation, selective image segmentation, and the simplest version of visual scene analysis – successive selection of spatially separated fragments of a visual scene.

Chapter 5 contains information on the photon echo effect that interestingly adjoins to the dynamical binding phenomenon presumably exploited in the brain functioning. Photon echo is a nonlinear optical analog of the spin echo phenomenon that accompanies nuclear and electronic magnetic resonance. The effect consists in a coherent response of an optical resonant medium to ultrashort pulses of laser radiation. In contrast to nuclear and electronic magnetic resonance, photon echo admits a highly parallel type of information processing in spatial domain (inside a nonlinear crystal of one cubical centimeter volume about 10^8 radiation rays can be used simultaneously) and is characterized by high operation speed. The principles of optical computing based on photon echo are nowadays well developed, and experimental prototypes of various optoelectronic devices for parallel information processing, such as optical filters and splitters, logical elements, vector–matrix multipliers, and phase optical memory have been implemented.

Chapter 6 contains some recent authors' results related to an attempt to incorporate a model of stochastic oscillators into the problems of optical computations. The model of stochastic oscillators was used for simulation of the random electric field of quasimonochromatic polarized light beams. The single stochastic oscillator is a Ginzburg–Landau limit cycle oscillator with chaotically modulated limit cycle size and frequency. The feedforward network of stochastic oscillators was designed for modeling the evolution of the electric field of a composite light beam consisting of incoherent superposition of nonpolarized quasimonochromatic light beams. The oscillatory network method allows us to see the temporal dynamics of the electric field of the composite light beam in the setup, in which beam components undergo the external optical device actions of typical optical devices (such as polarizer and compensator) causing instantaneous transformation of the electrical field at discrete moments of time. The calculation of the correlation matrix of a composite light beam permits us to estimate the modifications of the correlation properties of the random electric field and observe the emergence of coherence in the composite system.

We hope that the book will be interesting and helpful for researchers (both young and proficient) who intend to deal with the problems in such areas as modeling dynamical ANNs and creation of parallel adaptive neuromorphic algorithms.

Acknowledgments

One of the authors (M. K.) is very grateful to Prof. M. V. Maslennikov (KIAM RAS, Moscow, Russia) for deep interest to the author's works on ANNs, attention, and permanent support. The author is very thankful for hospitality and interesting and helpful discussions to Prof. S. Amari (RIKEN BSI, Japan), Prof. S. Shinomoto (Kyoto University, Japan), Prof. Y. Kuramoto (Kyoto University, Japan), Prof. J. Hertz (Nordita, Denmark), Prof. Ph. de Forcrand (ETH Zurich, Switzerland), and Prof. M. Hasler (EPFL Lausanne, Switzerland). It is a great pleasure for the author to thank Prof. A. A. Frolov (Institute of Higher Nervious Activity and Neurophysiology RAS, Moscow, Russia), Prof. W. L. Dunin-Barkowsky (Institute of System Research, RAS, Moscow, Russia), Prof. N. Farhat (Penn Engineering, USA), Prof. A. Cichocki (RIKEN BSI, Japan), Prof. L. van Hemmen (TU Munich, Germany), Prof. V. G. Redko (Institute of System Research RAS, Moscow, Russia), Prof. A. I. Samarin (Neurocybernetics Research Institute, Rostov-on-Don, Russia), Dr. Ya. B. Kazanovich (IMPB RAS, Pushino, Moscow Region, Russia), Prof. R. M. Borisyuk (IMPB RAS, Pushino, Moscow Region, Russia; University of Plymouth, UK), Dr. A. A. Ezhov (TRINITI, Troitsk, Russia), and Prof. V. G. Yakhno (Applied Physics Institute RAS, Nizhny Novgorod, Russia) for many interesting and stimulating discussions.

Contents

1 Introduction

1.1 Brain: structure and function principles

The brain is a complex, multilevel, hierarchically organized, adaptive system composed of a range of interconnected modules. The brain modules are multilayered densely connected neural networks with complicated interconnection structure, containing a great variety of feedback loops. The total number of brain neurons is about 10^{10}–10^{11}, and the total number of interneuronal connections is about 10^{15}. There are thousands of anatomically various biological neurons (see Figures 1.1 and 1.2). However, general morphology allows us to classify all these cells as neurons (the cells of star-like shape). Neurons are characterized by a central cell, or soma, that may be of different shapes. Neurons can be viewed as having an input and output poles. The input pole consists of extensively branching tree-like extensions of the soma membrane known as dendrites. The output pole, called the axon, arises as a single structure from the soma. The length of the majority of dendrites is not greater than 1 mm, and the length of axons varies from parts of millimeter to meters. Neurons are capable of generating electrochemical pulses (spikes) propagating through axons and activating other neurons. The information transmission is realized through synapses (chemical transmission). Thus, dendrites can be considered as receptive zones, axons act as transmission lines, and synapses are elementary structural and functional units that mediate interactions between neurons.

A brain neuron may have about 10^4 different inputs and may send its output up to 2×10^5 other neurons. As a result of massive interconnections, the brain is an exclusively efficient informational structure. Although neurons are typically about five to six orders of magnitude slower than silicon logic gates (about 10^{-3} s per operation), the brain far exceeds modern computers in their capabilities and performance. In ad-

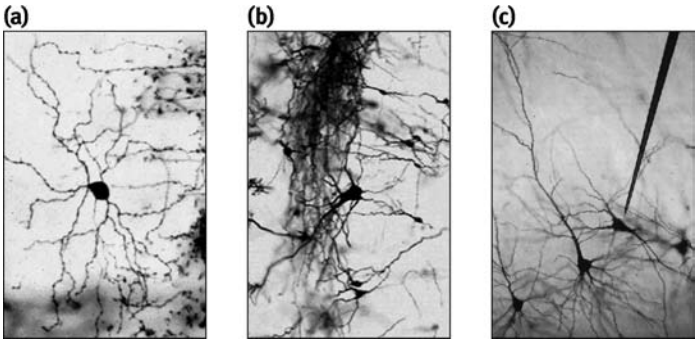

Fig. 1.1. Real neurons of mammal brains: (a) neuron of cat auditory cortex; (b) neurons of urchin olfactory bulb; (c) neurons of human visual cortex [49].

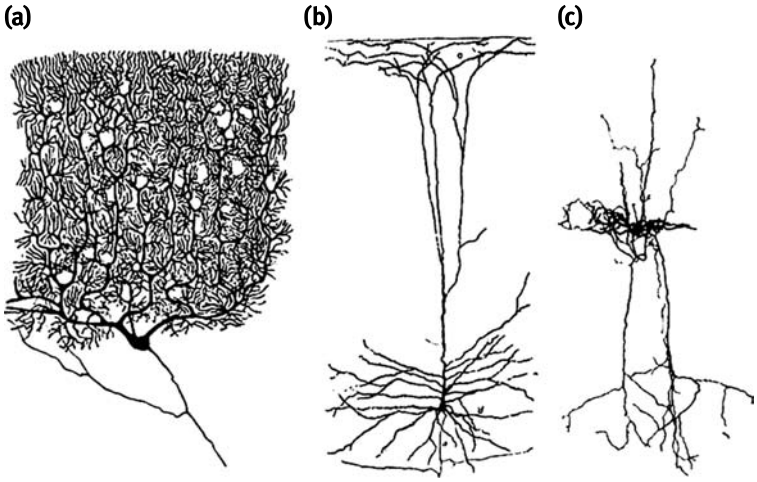

Fig. 1.2. Three examples of human brain neuron shapes: (a) Purkinje cell from cerebellum; (b) pyramidal cell from the visual cortex; (c) star cell from the cerebral cortex (drawings by Ramón y Cajal).

dition, the brain is energetically efficient: it consumes approximately 10^{-16} J per operation per second, whereas the corresponding value for the best modern computers is about 10^{-6} J per operation per second.

The brain possesses a number of striking features that make it beyond comparison with modern computers. From the information processing point of view, the following features are acquired by the brain nervous systems in the process of biological evolution during 600 millions of years: (a) self-organized style of functioning (with decentralized information processing); (b) massive parallel performance; (c) capability of adaptation and learning; (d) reliability and fault tolerance; (e) capability to deal with incomplete, noisy and redundant data.

Due to extreme brain complexity, the study of the brain functioning is one of the most difficult problems. Intensive interdisciplinary research efforts to understand how the brain works are being actively continued nowadays. Recently, international collaborative creation of large-scale computer models simulating functioning of real human brain structures has begun.

Two of the most impressive research cycles aimed at the real brain study via a detailed computer modeling are worth mentioning. The first one is the international "Blue Brain" Project (BBP) started in 2005 and directed by H. Markram at the Brain Mind Institute (EPFL, Switzerland). The project aims at computer simulation of the human brain neocortical column with great biological accuracy. In 2011 a morphologically detailed, electrically accurate and synaptically precise column model consisting of 10^6 neurons of 200 types and 10^9 synapses of 6 types was created. It is planned to further increase the model biological accuracy and to create finally a complete and accurate brain model capable of systematically reconstructing and eventually simu-

lating the human brain [15, 84, 91]. The second model is a detailed large-scale model of the brain thalamocortical system, created by Izhikevich and Edelman [52]. The model based on the experimental brain measurements of several mammalian species contains 10^9 spiking neurons and almost 10^{15} synapses. It incorporates multiple cortical regions, corticocortical connections, and synaptic plasticity of the mammalian thalamocortical system. The model exhibited spontaneous activity, emergence of collective waves and oscillations, alpha and gamma rhythms, functional connectivity at different scales and other interesting phenomena. The claimed final goal is to develop a large-scale biologically accurate computer model of the whole human brain. In addition to the created model of the thalamocortical system, later it is planned to add the other subcortical structures, including hippocampus, cerebellum, basal ganglia, etc.

In the opinion of researchers, working in the field of adaptive control, the brain modeling at microscopic level should be combined with modeling at the macroscopic level, when the brain is viewed as a hierarchically-organized system of modules [124]. It should also be added that we do not concern here such an extraordinary phenomenon as consciousness that is not yet completely studied and understood. There exist only separate preliminary attempts to elucidate the basic characteristic attributes of consciousness [3] that could permit us to formulate the problem in terms of neuroscience, information processing, cognition and control, and to start modeling artificial consciousness. Regular international conferences on machine consciousness have been organized since 2001.

From the informational viewpoint, the brain can be considered as a highly distributed self-organizing system, exploiting learning and adaptation as key principles of its functioning. All the impressive cognitive and executive functions of the brain are realized via modifiable neuron interconnections. One of the natural and effective ways to study parallel distributed information processing, inherent to the brain, consists in construction of simplified neural network models and their analysis via mathematical modeling tools.

Besides the dense interconnectivity structure of the brain neurons realized via dendrites and axons, there exists so-called dynamical binding–temporal coherence, realized in the form of synchronization of oscillatory neural activity. The dynamical binding serves as an important organizing principle in the brain work [105]. Selective extraction of certain neuron groups from spatially distributed neural networks via formation of dynamically coupled neural subensembles represents an additional quick, efficient and flexible tool, helpful in the brain performance. As was experimentally confirmed, synchronized oscillations of neural activity accompany information processing carried out by many brain neural structures. Various biologically motivated oscillatory network models – networks of interconnected nonlinear oscillators with synchronization-based performance – were designed to study the potential capabilities of dynamical binding used in olfactory, auditory and visual systems of the brain.

1.2 Artificial neural networks

One of the main of research direction in the neural network theory is the study of general capabilities of information systems realizing the parallel and distributive type of information processing. Construction and computer simulation of artificial neural networks, representing the simple versions of distributed information processing systems, can be regarded as one of the possible ways to study the capabilities of these systems.

ANN is a massively interconnected array of processing units (artificial neurons or others) capable of information storage, processing and retrieval. The ANN models should reflect the following features inherent to the brain neural structures: the information is stored in a distributed form, being encoded in terms of network interconnection architecture; new information is acquired by the brain neural networks by means of learning. Depending on the posed problem, either artificial neurons or neural oscillators usually figure as processing units of ANN.

1.2.1 Neuron models

An artificial neuron model most often figures as an information processing unit of a neural network. The artificial neurons are truly primitive in comparison with those found in the brain. However, the neuron model usually reflects three important characteristics of real neurons: the weighted values of synapses, the exploitation of a nonlinear activation function, limiting the permissible amplitude of the neuron output (see the block diagram of formal neuron, shown in Figure 1.3).

The model of an artificial neuron depicted in Figure 1.3 was suggested by McCulloch and Pitts in 1943 [85]. For a neuron with n inputs $x_1, \ldots, x_n$ and input synapse weights $w_1, \ldots, w_n$, the neuron output y is defined as

$$y(t+1) = G\left(\sum_{j=1}^{n} w_j x_j(t) - h_j\right), \tag{1.2-1}$$

where t is the discrete time, $G(x)$ is an activation function of the neuron, and h_j is a threshold. If $x_j \in \{0, 1\}$ are used, then the most simple $G(x)$ is the Heaviside step

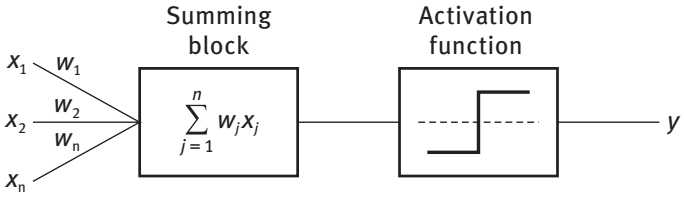

Fig. 1.3. Formal neuron by McCulloch–Pitts.

function, whereas at $x_j \in \{-1, 1\}$ the activation function $G(x) = \operatorname{sgn}(x)$ is naturally used. Continuous sigmoid functions are used as well. A simple example of a sigmoid activation function is the logistic function defined by

$$G(x) = \frac{1}{1 + e^{-\alpha x}}. \tag{1.2-2}$$

Many features of real neurons (pulse character of neural activity, existence of refractory period, nonlinear character of input information summing) are not reflected in the model of the formal neuron. However, the formal neuron was fruitfully used in a great variety of ANN models. Spiking neurons and oscillators were used for modeling of more complicated ANNs. In particular, networks of bursting oscillators were designed [50]. Two simplest types of ANN represent feedforward neural networks and fully connected recurrent neural networks (RNNs).

1.2.2 Feedforward networks

An open (feedforward) single-layer neural network of n identical formal neurons, not interconnected with each other, each possessing n inputs and a single output, is depicted in Figure 1.4 (a). Let $G(x)$ be the activation function of a single neuron, W_{jk} be the connection weight of the input j to the neuron k and $h_1, \ldots, h_n$ be neuron thresholds. Then the outputs $y_1, \ldots, y_m$ of the feedforward network (also called a single-layer perceptron) are defined as

$$y_j = G\left(\sum_{k=1}^{n} W_{jk} x_k - h_j\right), \quad j = 1, \ldots, m, \tag{1.2-3}$$

or, in the vector form

$$\mathbf{y} = G(\hat{W}\mathbf{x} - \mathbf{h}) \equiv \hat{T}\mathbf{x}, \tag{1.2-4}$$

where $\mathbf{x} = (x_1, \ldots, x_n)^T$, $\mathbf{y} = (y_1, \ldots, y_m)^T$, $\mathbf{h} = (h_1, \ldots, h_n)^T$ and $\hat{W} = [W_{jk}]$ is an $m \times n$-matrix.

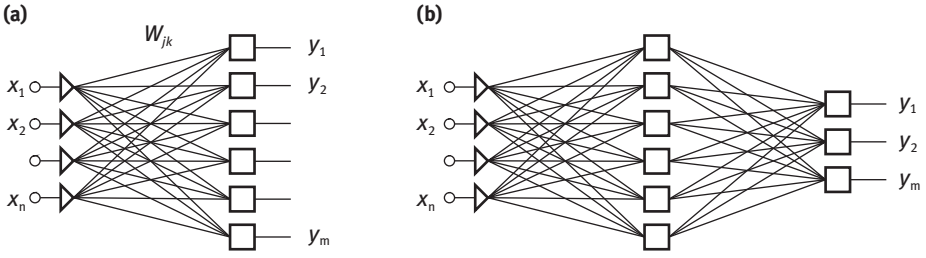

Fig. 1.4. Perceptrons: (a) single-layer and (b) two-layer feedforward networks.

Let a series of single-layer perceptrons be connected into a cascade – a chain of sequentially connected single-layer networks, so that the output of the previous network serves as the input of the subsequent network. Let $n^{(l)}$ be the number of neurons in the layer l. Then we have $\dim \mathbf{x}^{(l)} = n^{(l)}$ and $\hat{W}^{(l)}$ is an $n^{(l)} \times n^{(l-1)}$-matrix. The example of a two-layer perceptron (a cascade of two single-layer feedforward networks) is depicted in Figure 1.4 (b).

The perceptrons were the initial neural network models for which learning algorithms with a teacher (supervised learning) were applied [99]. At first, these simplest networks were used in pattern classification problems. The simplest learning algorithm (algorithm of the weights W_{jk} adjustment), the so-called least-mean-square algorithm, also known as the delta-rule, was initially implemented for perceptrons (see Section 1.5.1).

Nowadays perceptrons are widely exploited for the development of various parallel-performing computational algorithms, which often prove to be more effective than classical grid algorithms or finite element methods. Besides, multilayer perceptrons can be used as universal function approximators providing the exact calculation of complicated nonlinear functions of many variables. There exists the universal approximation theorem for a nonlinear input–output mapping, which is directly applicable to a two-layer perceptron of neurons with a sigmoid activation function [41]. The theorem is based on the generalization of function approximations by finite Fourier series. Therefore, in principle, two-layer perceptron is capable of approximating the arbitrary continuous function of many variables. The mentioned universal approximation theorem is an existence theorem: it provides the mathematical justification for the approximation, but provides no recipe for the construction of neuron activation functions. In practice, in the case of fixed perceptron architecture more flexible learning algorithms, admitting not only weight adjusting, but also activation function modification, are necessary. These algorithms have also been developed.

In training a multilayer perceptron by whatever method, it is necessary to create a nonlinear model of the physical phenomenon responsible for the generation of input–output examples used to train the feedforward network. Moreover, it is necessary to have an appropriate tradeoff between the reliability of training data for the network and goodness of the model (see Section 1.6 on neural network learning). That is, the network design is statistical in nature, and Tikhonov regularization of ill-posed problems is often necessary to be attracted. (The essence of the regularization consists in introduction of some additional statistical assumptions to get a stable solution.)

1.2.3 Recurrent neural networks

Recurrent (closed) neural networks (RNN) are characterized by the existence of feedback connection loops in the network connectivity structure. This essentially modifies network dynamical behavior and has a significant impact both on network per-

formance and on its learning capabilities. Recurrent neural networks should be considered as nonlinear dynamical systems.

The simplest type of a recurrent network can be obtained from a one-layer perceptron displayed in Figure 1.4 (a), if the network output $\mathbf{y}$ is fed back to the network input $\mathbf{x}$ (under the natural condition $\dim \mathbf{y} = \dim \mathbf{x}$). The network of such connectivity architecture can be called a one-layer fully recurrent network. When all network neurons work synchronously, the dynamics of the recurrent network is governed by the dynamical system which can be written either in the discrete time form

$$\mathbf{x}(t+1) = G(\hat{W}\mathbf{x} - \mathbf{h}), \tag{1.2-5}$$

or in the continuous time form (in the case of a continuous activation function):

$$\frac{d\mathbf{x}}{dt} = -\mathbf{x} + G(\hat{W}\mathbf{x} - \mathbf{h}). \tag{1.2-6}$$

The well-known Hopfield neural network model is an example of a fully connected RNN. One can analogously construct a recurrent multilayer perceptron (in the case when the dimension of perceptron output coincides with its input dimension). This network could be viewed as a multilayer fully recurrent network – a recurrent perceptron). The bidirectional associative memory (BAM) network model [60] can be viewed as a two-layer recurrent perceptron. Of course, partially recurrent network models can also be designed.

Many different types of RNN with various connectivity architectures were constructed and studied. They were used in various applications, such as sequence classification, time-series prediction, feature extraction, modeling of input–output behavior of dynamical systems and many others. We focus our attention mainly on single layer fully connected RNN of associative memory.

Dynamical systems theory can be used for the analysis of RNN dynamics in the deterministic approach (deterministic neurodynamics). In the presence of noise (statistical neurodynamics), the approaches based on the attraction of stochastic nonlinear dynamical systems are needed. In either case, the governing dynamical systems for RNNs belong to the class of multivariable multiparametrical nonlinear dissipative dynamical systems, which can have a complicated set of attractor structures. Learning algorithms can be formulated in terms of an additional dynamical system providing control of attractor locations in the phase space of the main dynamical system governing network dynamics.

1.3 Recurrent neural networks of associative memory

If a RNN with a rich set of stable attractors is designed, it can be used as an associative memory network for information storage. The following characteristics of recurrent associative memory are implied to be held: (a) distributed character of associative

memory (information is stored in the form of multicomponent network state vectors – stable attractors of network dynamics); (b) good "isolation" of memory vectors from each other (large enough and well-separated attractor basins); (c) appropriate degree of resilience to noise and network damage (robustness with respect to network parameter perturbation).

If the network with a set of stable fixed points is designed, the attractor basins of fixed points will play a crucially important role in associative memory network performance. The attraction basin of a fixed point $\mathbf{p}$ can be defined as the largest neighborhood of $\mathbf{p}$, such that any vector in this neighborhood is attracted by $\mathbf{p}$. For low-dimensional dynamical systems with continuous-time dynamics, the phase portrait of a dynamical system provides the complete picture of phase space separation into the collection attractor basins. In the case of large-scale recurrent networks with discrete-time dynamics the attractor basins usually have very complicated (fractal) form, and statistical methods are often needed for their analysis.

Various recurrent associative memory network models were developed since 1972 (see Section 2.1.1). The well-known Hopfield model, which can be regarded as a recurrent associative memory network with Hebbian learning, was suggested in 1982 (see Section 1.4). As was clarified, recurrent associative memory neural networks suffer from so-called spurious, or extraneous memory – the set of additional stable fixed points arising in the phase space of a dynamical system together with fixed points prescribed by network interconnection matrix construction. The Boltzmann machine model was introduced as an extension of the Hopfield network by incorporating stochastic neurons and using the so-called method of simulated annealing [1]. The Boltzmann machine with the learning algorithm of adaptive association type [53] does not suffer from extraneous stable fixed points. However, exponential complexity (with respect to the neuron number) is inherent to this model, restricting its applications.

1.4 Hopfield network model

The Hopfield model [47, 48] is a fully RNN with symmetric connections. If network connections are trained via Hebbian learning, then the Hopfield network can be viewed as an associative memory network with content-addressable memory. The network memory turned out to be robust with respect to connection perturbance.

The dynamical system governing the Hopfield model can be written both in discrete and in continuous time, depending on the choice of the neuron activation function. In the case of a continuous activation function, the network dynamics can be written as

$$\frac{d\mathbf{x}}{dt} = -\mathbf{x} + G(\hat{W}\mathbf{x} - \mathbf{h}). \tag{1.4-1}$$

The key advance for model analysis was introduction of energy

$$H = -\frac{1}{2}\sum_{j,k=1}^{n} W_{jk}x_j x_k. \tag{1.4-2}$$

The energy (1.4-2) can be interpreted as the Lyapunov function for the dynamical system (1.4-1), decreasing in the process of network dynamics. The energy H corresponds to a Hamiltonian in statistical mechanics, a cost function in optimization theory and a fitness function in evolution biology. The network memory vectors (stable attractors of dynamics) are located in local minima of the energy surface.

If the stochastic version of the Hopfield network of formal McCullock–Pitts neurons is considered, there is a direct analogy with the statistical mechanical Ising model of a magnetic spin system. The Hamiltonian (1.4-2) is then identical to that of infinite-range spin glasses – spatially distributed random spin systems with competing ferromagnetic and antiferromagnetic interaction. Randomness and competition (or frustration) are intrinsic features of spin glasses [20]. The Hopfield model corresponds to deterministic (zero temperature) dynamics of a spin system (where all spins are updated at the same time step). Asynchronous and block-sequential dynamics are also possible for RNN.

The statistical mean-field approach was used to analyze the Hopfield model with an infinite number of stable fixed points at nonzero temperature. The phase diagrams were obtained, indicating phase transitions from the paramagnetic phase to the phase of spin glass, and also from the retrieval phase (when the network is capable of performing as an associative memory) into the spin glass phase. The existence of a large number of metastable states was established, and the general bounds for storage capacity were found in the frames of the mean-field theory [28]. Much less information was obtained concerning the sizes and the shapes of attraction basins. These properties cannot be clarified without direct treatment of a nonequilibrium dynamical process. The Hopfield model with a small number of stored patterns was studied analytically via the master equation approach. In that way, the existence of limit cycles for the networks with asymmetric connections was shown.

1.5 Learning of neural networks

Learning algorithms are procedures that adapt the free parameters of a neural network to improve the network performance in a concrete task. Most learning algorithms adapt the network parameters in such a way that certain error measure (called the cost function) is minimized.

The capability of learning is one of the most significant features of biological neural networks. It ensures the modification of network functioning via learning from examples rather than any complicated externally given rules. As a rule, the modification of the connectivity matrix $\hat{W}$ is realized during a special "learning phase." However,

learning processes "online" are also possible. Then network dynamics is governed by a coupled system of equations for the network state $\mathbf{x}$ and the matrix $\hat{W}$ that can be written in the form

$$\begin{aligned} \frac{d\mathbf{x}}{dt} &= -\mathbf{x} + G(\hat{W}\mathbf{x} - \mathbf{h}), \\ \frac{d\hat{W}}{dt} &= \eta F(\mathbf{x}, \hat{W}), \end{aligned} \tag{1.5-1}$$

where $\eta \ll 1$ is the learning "intensity" (or the learning rate). So, the learning process causes slow (compared to the relaxation time of the network state $\mathbf{x}$) evolution of the operator $\hat{T}(\mathbf{x}, \hat{W}) = -\mathbf{x} + G(\hat{W}\mathbf{x} - \mathbf{h})$, and slow ("adiabatic") modification of $\hat{T}(\mathbf{x}, \hat{W})$ corresponds to the "online" learning.

The so-called supervised learning algorithms developed for feedforward multilayer neural networks turned out to be directly related to approaches developed for tasks of nonlinear global multidimensional optimization. Current components of matrix $\hat{W}$ are modified in terms of the proper cost function $E(\mathbf{y}, \mathbf{v}; \hat{W})$ specifying a measure of discrepancy between actual $\mathbf{y}$ and desirable $\mathbf{v}$ network outputs:

$$\Delta W_{jk} = -\eta \frac{\partial E(\mathbf{y}, \mathbf{v}; \hat{W})}{\partial W_{jk}}. \tag{1.5-2}$$

The unsupervised learning algorithms based on self-organized connectivity matrix modifications were developed mainly for recurrent networks (see Section 1.5.2). The well-known biologically inspired Hebbian learning algorithm belongs to the class of self-organized learning algorithms (Section 1.5.2 A). The reinforcement learning that can be viewed as learning via interacting with an environment also belongs to the class of unsupervised learning algorithms (Section 1.8.2). At last, so-called genetic learning algorithms based on modeling of the process of natural biological evolution can also be considered as self-organized optimization algorithms (Section 1.5.2 B).

1.5.1 The simplest learning algorithms for feedforward networks

The majority of learning algorithms for feedforward networks are based on gradient descent methods. We present two simplest supervised learning algorithms – the "delta-rule" algorithm for a single-layer perceptron and the backpropagation learning algorithm for a multilayer perceptron.

A. The learning algorithm "delta-rule"

Let us consider a single-layer perceptron with the input–output mapping $\mathbf{y} = G(\hat{W}\mathbf{x})$, where $\dim \mathbf{x} = m$, $\dim \mathbf{y} = n$, $\hat{W}$ is an $n \times m$-matrix and $G(x)$ is a continuous activation function (e.g. $G(x) = \tanh(\beta x)$). Let the "training" sample, that is, the set of known input–output pairs $\{\mathbf{u}^\mu, \mathbf{v}^\mu\}_{\mu=1}^{M}$ with $\dim \mathbf{u}^\mu = m$, $\dim \mathbf{v}^\mu = n$ be given.

Starting from an arbitrary $\hat{W}$, it is necessary to construct successive approximations to connectivity matrix, satisfying the condition: given the inputs $\mathbf{u}^1, \ldots, \mathbf{u}^M$, the corresponding outputs $\mathbf{y}^1, \ldots, \mathbf{y}^M$ should be asymptotically close to the target set $\mathbf{v}^1, \ldots, \mathbf{v}^M$. The solution can be obtained in the explicit form in the case of the simplest cost function

$$E(\mathbf{y}, \mathbf{v}; \hat{W}) = \frac{1}{2M} \sum_{\mu=1}^{M} |\mathbf{v}^\mu - \mathbf{y}^\mu|^2 = \frac{1}{2M} \sum_{\mu=1}^{M} \sum_{j=1}^{n} \left(v_j^\mu - y_j^\mu\right)^2. \tag{1.5-3}$$

The weight corrections are calculated as

$$\Delta W_{jk} = \eta \frac{\partial E}{\partial W_{jk}}, \tag{1.5-4}$$

where

$$\frac{\partial E}{\partial W_{jk}} = -\frac{1}{M} \sum_{\mu=1}^{M} \left(v_j^\mu - y_j^\mu\right) G'\left(\sum_{l=1}^{m} W_{jl} u_l^\mu\right) u_k^\mu, \quad G'(x) \equiv \frac{dG}{dx}. \tag{1.5-5}$$

Finally, we can rewrite the result in the form

$$\Delta W_{jk} = \frac{\eta}{M} \sum_{\mu=1}^{M} \delta_j^\mu u_k^\mu, \quad \delta_j^\mu = \left(v_j^\mu - y_j^\mu\right) G'\left(\sum_{k=1}^{m} W_{jk} u_k^\mu\right). \tag{1.5-6}$$

Formula (1.5-6) may also be written in the vector form

$$\Delta \hat{W} = \eta \langle \boldsymbol{\delta}^\mu (\mathbf{u}^\mu)^T \rangle, \quad \boldsymbol{\delta}^\mu = (\mathbf{v}^\mu - G(\hat{W}\mathbf{u}^\mu)) G'(\hat{W}\mathbf{u}^\mu), \tag{1.5-7}$$

where $\langle \mathbf{u} \rangle = M^{-1} \sum_{\mu=1}^{M} \mathbf{u}^\mu$.

The generalization of the gradient-descent delta-rule algorithm to the case of multilayer perceptron gives the known backpropagation algorithm.

B. Backpropagation learning algorithm

The backpropagation algorithm, the main working algorithm that was used for feedforward multilayer neural networks, was initiated in 1974 [11, 89, 100, 121]. It was independently proposed by Bryson and Ho [11]. Werbos [121] proposed the algorithm as a kind of nonlinear regression. Rummelhart with coauthors [100] used the algorithm for neural network optimization. However, the algorithm was developed earlier by researchers working in the field of optimal control in the early 1960s. Later it was described as a multistage optimization method for dynamical systems [11], and further summarized in the classical book on optimal control. The backpropagation algorithm, being rediscovered by the AI community, was supplemented by the concept of functions as dynamical objects that have to be evaluated by a network.

Let a multilayer perceptron of L layers be given with the numbers $n_1, \ldots, n_L$ of neurons in the layers, the activation functions $G^{(l)}$ and the connectivity matrices $W^{(l)}$, $l = 1, \ldots, L$. Let $\{\mathbf{u}^\mu, \mathbf{v}^\mu\}$, $\mu = 1, \ldots, M$ be the learning sample, and the cost

function $E(\mathbf{v}, \mathbf{y}; \hat{W})$ be defined by the same expression as in the delta-rule algorithm:

$$E(\mathbf{y}, \mathbf{v}; \hat{W}) = \frac{1}{2M} \sum_{\mu=1}^{M} \sum_{j=1}^{n} \left(v_j^{\mu} - y_j^{\mu}\right)^2. \tag{1.5-8}$$

Let us perform the following calculations:

(1) Calculate the outputs $\mathbf{y}^{(l),\mu}$, $l = 1, \ldots, L$, $\mu = 1, \ldots, M$ for all the layers:

$$\begin{aligned} \mathbf{y}^{(1),\mu} &= G^{(1)}\left(\hat{W}^{(1)}\mathbf{u}^{\mu}\right) \equiv T^{(1)}\mathbf{u}^{\mu}, \\ &\ldots \\ \mathbf{y}^{(l),\mu} &= G^{(l)}\left(\hat{W}^{(l)}\mathbf{y}^{(l-1),\mu}\right) \equiv T^{(l)}\mathbf{y}^{(l-1),\mu}, \\ &\ldots \\ \mathbf{y}^{(L),\mu} &= T^{(L)} \cdots T^{(1)}\mathbf{u}^{\mu}. \end{aligned} \tag{1.5-9}$$

(2) Calculate the errors $\boldsymbol{\delta}^{(L),\mu}$ for the last layer L according to the delta-rule, $\mathbf{y}^{(L-1),\mu}$ being the layer input:

$$\boldsymbol{\delta}^{(L),\mu} = \left(\mathbf{v}^{\mu} - \mathbf{y}^{(L),\mu}\right) G^{(L)\prime}(\hat{W}^{(L)}\mathbf{y}^{(L-1),\mu}). \tag{1.5-10}$$

(3) Find the errors $\delta^{(l),\mu}$ for layers $l = L-1, L-2, \ldots, 1$ (in the opposite order)

$$\boldsymbol{\delta}^{(L),\mu} = G^{(l)\prime}\left(\hat{W}^{(l)}\mathbf{y}^{(l-1),\mu}\right)[(\hat{W}^{(l+1)})^T\boldsymbol{\delta}^{(l+1),\mu}]. \tag{1.5-11}$$

(4) Using $\delta^{(l),\mu}$, find the corrections $\Delta\hat{W}^{(l)}$ for all the layers

$$\Delta\hat{W}^{(l)} = \frac{\eta}{M} \sum_{\mu=1}^{M} \boldsymbol{\delta}^{(l),\mu}\left(\mathbf{y}^{(l-1),\mu}\right)^T. \tag{1.5-12}$$

As one can see, it is possible to find the outputs for all the perceptron layers by carrying out the calculations in the direction from the first layer to the last one, and after that, to find the errors $\delta^{(l),\mu}$ and matrix corrections $\Delta\hat{W}^{(l)}$, carrying out the calculations in the opposite direction.

The backpropagation algorithm can also be extended to the case of recurrent networks. In this case, it is necessary to introduce a discrete time variable t. At the time t one should recompute all the network outputs and transmit them to the time $t + 1$. The procedure should be continued in a step-by-step manner. This strategy converts a recurrent network into a feedforward one, and this version of the backpropagation algorithm for RNN is called backpropagation through time.

The following limitations of backpropagation as an iterative process are well known: (a) the convergence of the algorithm is not guaranteed; (b) the convergence can be very slow; (c) the algorithm can also converge to the local minima on error surface.

The computational complexity of an algorithm is usually measured in terms of the number of arithmetical operations involved in its implementation. A learning algorithm is said to be computationally efficient if its computational complexity is polynomial in the number of adjustable parameters that are to be updated from one iteration to the next one. In the sense, the backpropagation algorithm is computationally efficient, and its computational complexity is linear in $\hat{W}$, i.e. $\mathcal{O}(\hat{W})$. Also it is noteworthy that backpropagation algorithm is equivalent to an optimal filter. The algorithm tends to converge slowly and in a zigzag way about the true direction to a minimum on the error surface. Another feature of the algorithm performance is related to the presence of local minima in addition to global ones: there is a risk of being trapped in a local minimum at every small change in $\hat{W}$ [41].

1.5.2 Learning algorithms for recurrent networks

A. Hebbian learning

Hebb in the book *The Organization of Behavior* [44] postulated the first rule for self-organized (unsupervised) learning motivated by neurobiological considerations. Hebb proposed that the brain connectivity is continually changing as an organism learns different functional tasks, and the brain neural assemblies are created by such connectivity changes. Hebb was the first who introduced the learning rule for synaptic modification. The Hebbian learning law is often shortly formulated as "neurons that fire together, wire together." The Hebbian learning law further inspired the development of computational models of learning and adaptation.

The use of the Shannon information concept was relevant to deeper understanding of self-organized learning and information processing in self-organized neural networks. The so-called maximum mutual information principle [79] provided the proper mathematical formalism permitting us to understand that information processing is, in some sense, analogous to information transmission through communication channels. The idea of a nonlinear adaptive filter proposed by Gabor [27] allowed us to consider learning as a stochastic process accomplished by feeding samples together with the target function into a neural network.

The simplest form of Hebbian learning can be written as a discrete-time evolution process for the connectivity matrix $\hat{W} = [W_{jk}]$:

$$W_{jk}(t) = W_{jk}(t-1) + \Delta W_{jk}(t), \quad \Delta W_{jk}(t) = \eta y_j(t) x_k(t), \tag{1.5-13}$$

where $x_k(t)$ is the state of an input neuron, $y_j(t)$ is the state of an output neuron, and η ($\eta \geq 0$) is the learning rate. The discrete-time process (1.5-13) can also be presented in the continuous-time version:

$$\frac{d\hat{W}}{dt} = \eta \mathbf{y}\mathbf{x}^T. \tag{1.5-14}$$

A series of natural modifications of the learning rule (1.5-13) was suggested to prevent the exponential growth of the postsynaptical state $y_j(t)$ due to the repeated learning application to the input state $x_k(t)$. For example, the continuous-time version of the learning algorithm by Oja [88] can be written as

$$\frac{d\hat{W}}{dt} = \eta\left(\mathbf{y}\mathbf{x}^T - \mathbf{y}^T\mathbf{y}\hat{W}\right). \tag{1.5-15}$$

The competitive learning is a special type of Hebbian learning where output neurons are said to be in competition for input neurons. During training, the output neuron that provides the highest activation level is declared the winner, whereas the rest neurons are left unchanged. This strategy is often called "winner-takes-all." The competitive learning algorithm suitable for multilayer recurrent networks can be expressed by

$$\Delta W_{jk}(t) = \begin{cases} \eta[x_j(t) - W_{jk}(t)] & \text{if the } j\text{th neuron wins the competition,} \\ 0 & \text{if the } j\text{th neuron loses the competition.} \end{cases} \tag{1.5-16}$$

There are many versions of the competitive learning [25, 26, 38, 39, 113]. A simple modification of competitive learning gave rise to a new powerful class of neural network models known as self-organizing maps (SOMs). These models were initiated by Kohonen [57] and are also known as Kohonen maps. The SOMs are trained using a special competitive learning algorithm (with neighborhood constraint for the output neurons, which are arranged into a special lattice). As a result, the SOM networks form a topology preserving map. The SOM networks operate in the two stages: training and mapping. The competitive learning algorithm for SOM is called a vector quantization. The mapping corresponding to the competitive learning algorithm is capable of automatically classifing new input states. In particular, as it turned out, this property can be used for visualization of high-dimensional input data. The vector quantization learning algorithm has a clear neurobiological inspiration based on the combination of three processes: cooperation, competition and adaptation.

Differential Hebbian learning is connectivity matrix modification according to the rule [60]:

$$\frac{dW_{jk}}{dt} = \eta x_j(t)\frac{dy_k(t)}{dt}. \tag{1.5-17}$$

The motivation for the differential Hebbian learning algorithm was that the concurrent change rather than concurrent activation more accurately captures synaptic modifications in the brain nervous structures. The algorithm moved into the focus of interest after 1997 in relation to the study of networks of spike-timing neurons.

The essential feature of any unsupervised learning (in particular, Hebbian learning and its natural generalizations) is an automatic, self-organizing character of network connectivity structure modification. Classical Hebbian learning continues to be popular and successfully applicable in a variety of problems. In particular, it adequately "works" in problems of extraction of regularities from statistical environment.

B. Another self-organized learning algorithms

Three main categories of learning algorithms for RNN can be selected: deterministic gradient-based algorithms, stochastic algorithms, and hybrid algorithms. The probabilistic principles, information-theoretical approach, Bayesian inference approach [10], structural risk minimization [41], and cross-entropy error measure are widely used for learning algorithm construction and for estimation of their performance quality. Second-order algorithms, much more efficient than those based on standard gradient descent, were also developed [10]. They require more calculations but far less iterations.

The stochastic learning algorithms based on random updates of connectivity matrices and providing optimization of network connectivity structure were developed as well. In particular, a powerful class of genetic algorithms belongs to the category of stochastic learning algorithms. In hybrid learning algorithms a gradient approach is combined with elements of stochastic algorithms. Sometimes hybrid algorithms use switching of a gradient algorithm to a stochastic one (when the gradient algorithm gets stuck in a local minimum) and switching back to gradient mode after some time.

Earlier Darwinian selective learning models [19] were relied on the presupposition that nervous systems often operate in a manner similar to natural selection in evolution. In the context of learning it means that the basic operational units of the nervous system are not single neurons but rather local groups of strongly interconnected neurons, and local competition and cooperation among the neurons are responsible for production of local order in the network.

Modern genetic algorithms belong to the wide research field of population-based evolutionary computations operating by models of inheritance, mutation, selection and crossover [22]. According to Neo-Darwinism (a universally accepted nowadays paradigm of biological evolution containing classical Darwinism combined with genetics and selection rules) there are only a very few statistical processes acting on and within populations and species. These processes are reproduction, mutation, competition and selection. Evolution is the result of these fundamental interacting stochastic processes as they act on population, generation after generation. The idea of a modeling of machine learning through simulated evolution has been proposed many times since 1932 [22]. Genetic algorithms (GA) arose as one of the alternatives of simulated evolution. These are the population-based optimization algorithms, in which the optimization process is capable of iteratively improving (via a selection model) the quality of the solution due to modeling the biological fitness (the ability to survive and reproduce). They are constructed so as to asymptotically converge to globally optimal solutions by properly chosen variation and selection operators and population size. The GA can be viewed as self-adaptive search procedures that are relatively insensitive to the specific choice of internal parameters. This robustness allows many applications.

One of the sources of difficulties of evolutionary algorithms is related to complicated drifting landscapes caused by the dynamics of evolving community.

1.5.3 Learning, function approximation, and support vector machines

Multivariable function approximation learning algorithms for feedforward networks are equivalent to finding a surface in a multidimensional space that provides a best fit to training data (with the fit criterion being understood in some statistical sense). Generalization is then equivalent to the interpolation (in traditional strict sense) of the test data with the help of a multidimensional surface. A multilayer feedforward network performs a nonlinear mapping from the input space to the "hidden" space, followed by a linear mapping from the hidden space to the output space. The construction of so-called radial-basis function networks (RBF-networks) is a special well-developed topic in the field of artificial neural networks. The radial-basis functions form a proper arbitrary basis for an input vector expansion in the "hidden" space.

The RBF networks usually have three layers. The input layer consists of units that connect the network to its environment. The second ("hidden") layer implements a nonlinear transformation from the input space to the hidden space. In most applications the hidden space is high dimensional. The output space realizes the response of the network to a signal applied to the input layer. The hidden space is directly related to the capacity of the network to approximate a smooth input–output mapping. Just like multilayer perceptrons, the RBF-networks are universal approximators, demonstrating good approximation properties for continuous functions defined on compact sets. However, realization of the interpolation procedure via training of RBF-networks (mapping reconstruction task) usually proves to be the ill-posed problem (when large data sets contain a small amount of information about the desired solution). In these cases the Tikhonov regularization approach for ill-posed problems should be applied. The corresponding RBF-networks are called regularization networks [41, 94].

New approaches in multivariable function estimation and statistical inference were inspired by supervised learning theory for multilayer perceptrons. In the way a new class of feedforward networks – the support vector machines (SVM) – were designed [110]. Based on the concept of algorithmic complexity, the essential concepts of visual cortex (VC) dimension and VC entropy (Vapnic–Chervonenkis dimension (VC dimension) and entropy [111]) were introduced and exploited in pattern recognition problems, solvable via training of feedforward networks. The characteristic features of the learning algorithm for SVM networks consist of the special choice of a proper subset of training samples and the use of a more appropriate minimization criterion than the mean-square error criterion used in the backpropagation algorithm. Usually the learning of SVM-networks can be reduced to a quadratic programming problems. This is attractive for two reasons: a global extremum of the error surface can always be found and the computations can be performed efficiently.

The important advantage of SVM-networks is that they provide control of network generalization ability while dealing with classification tasks. It was also recognized [41] that the performance of the learning algorithms for SVM-networks can be improved by incorporating prior knowledge about the task to be performed.

1.5.4 Statistical theory of learning

Statistical theory of learning elucidates the features of supervised learning algorithms for stochastic multilayer feedforward networks.

Consider the network with input–output mapping

$$\mathbf{y} = G(\hat{W}\mathbf{x} - \mathbf{h}) \equiv \hat{T}(\mathbf{x}; \mathbf{w}), \tag{1.5-18}$$

where vector $\mathbf{w}$ defines the full collection of internal network parameters (connectivity matrices $\hat{W}$ components and thresholds $\mathbf{h}$).

The components of the input vector $\mathbf{x}$ are considered as random variables specified by the conditional probability function $P(\mathbf{y}|\mathbf{x}; \mathbf{w})$. The teacher provides the desirable response $\mathbf{z}$ for every input $\mathbf{x}$ in accordance with the distribution function $F(\mathbf{x}|\mathbf{z})$ that is fixed and unknown. We can present the response as

$$\mathbf{z} = \hat{R}(\mathbf{x}; \xi), \tag{1.5-19}$$

where ξ is some noise. The problem of supervised learning consists in selecting $\hat{T}(\mathbf{x}; \mathbf{w})$ that optimally approximates (in some statistical sense) the desirable response $\mathbf{z}$. The selection of $\hat{T}(\mathbf{x}; \mathbf{w})$ is based on a training sample $\Omega = \{\mathbf{x}^{(k)}, \mathbf{z}^{(k)}\}_{k=1}^{M}$ – a set of M independent identically distributed training examples. That is, the leaning problem can be viewed as an approximation problem: to find the function $\hat{T}(\mathbf{x}; \mathbf{w})$ that would be the best possible approximation to the desired function $\hat{R}(\mathbf{x}; \xi)$. The important question is whether a training sample Ω contains sufficient information to construct a network capable of good generalization performance. The tools suggested in [111] permit us to answer this fundamental question.

To solve the approximation problem, a measure of discrepancy between the desired response $\mathbf{z}$ and the actual response $\mathbf{y} = \hat{T}(\mathbf{x}; \mathbf{w})$ should be introduced. Usually a quadratic loss function $L(\mathbf{z}, \hat{T}(\mathbf{x}; \mathbf{w}))$ defined as the squared distance between $\mathbf{z}$ and the approximation of $\mathbf{y}$ is introduced, and the ensemble-averaged of $L(\mathbf{z}, \hat{T}(\mathbf{x}; \mathbf{w}))$ (the risk functional) should be minimized.

Due to mathematical difficulties related to the computation of $L(\mathbf{z}, \hat{T}(\mathbf{x}; \mathbf{w}))$ in the case of a large number M of training samples, the minimization of the so-called empirical risk functional can be used instead of the minimization the actual risk functional [110]. The theory of uniform convergence of the empirical risk functional to the actual one is based on the concept of VC dimension. As turned out, the VC dimension is a measure of the capacity of a learning neural network. In some cases the VC dimension is defined by the set of free parameters of a neural network. However, in most of practical cases it is difficult to estimate the VC dimension analytically. The following two important results have been obtained:

(1) Let the activation function $G(x)$ of the feedforward network be the Heaviside function. Then the VC dimension of the neural network is $\mathcal{O}(W \log W)$ with W being the total number of network free parameters.

(2) If the network activation function $G(x)$ is a sigmoid continuous function, the VC dimension of the neural network is $\mathcal{O}(W^2)$.

It should also be noted that multilayer feedforward networks have a finite VC dimension. Thus, the VC dimension can be regarded as a central parameter of the statistical learning theory.

Another important characteristic of learning algorithms is their computational complexity. It concerns the computational effectiveness of a learning algorithm that can be estimated by the worst-case "running time" needed to train a neural network given a training sample of finite size M. The algorithm is said to be efficient if the running time is $\mathcal{O}(n)$, where $n = \dim \mathbf{x}$ (a polynomial time algorithm). Learning tasks performed by a polynomial time algorithm are usually regarded as "easy."

At last, one of the most informative characteristics of the learning algorithm is the so-called learning curve that shows how a neural network performance is improved as the number of training examples increases. The asymptotical properties of learning curves were analyzed in the frames of statistical approach via a stochastic descent method [4, 86]. Two learning curves were analyzed as functions of M (the size of the training sample): one concerning predictive (or generalization) loss and other the training loss. Besides, a natural definition of the complexity of the neural network was given. It was found that the asymptotical expression for the statistical mean of predictive loss can be written as

$$E[L(\mathbf{w})] = L(\mathbf{w}^*) + \frac{C_0}{M} + \eta C_1, \tag{1.5-20}$$

where the constant C_0, which can be directly calculated for a concrete network, may be interpreted as an effective complexity of the network, $\mathbf{w}^*$ is an optimal set of network internal parameters, M is the number of pairs in the training sample and η is the learning rate. The value

$$NIC(\mathbf{w}) = L(\mathbf{w}^*) + \frac{C_0}{M} \tag{1.5-21}$$

can be considered as the network informational criterion for the selection of an optimal network model (the network with smaller NIC can be regarded as a better model).

1.6 Oscillatory neural networks and synchronization

1.6.1 Synchronous oscillations of neural activity in the brain and neural oscillators

The brain is a highly distributed system in which numerous operations are executed in parallel. One of the important question is how the computations occurring simultaneously in spatially segregated processing areas are coordinated. Dynamical binding via synchronization of neuronal activity could play the role of a coordinating mechanism. The synchronization was experimentally found in various brain structures,

such as olfactory bulb and cortex, visual cortex, hippocampus, neocortex, thalamocortical system, and hypothesis about its functional significance in brain information processing was induced since 1980 [23, 114]. Experimental discovery of synchronous oscillations in the range of 40 Hz in the brain VC in 1988–1989 [18, 33] reinforced the attention to oscillatory aspects of visual information processing. As was clarified, the synchronization that accompanies the visual cortex performance in image processing possesses the following features: (a) the synchronization is a stimulus-dependent phenomenon; (b) the synchronization is achieved via neuronal interactions; (c) long-distance neuronal connections are responsible for binding via synchronization.

The desire to elucidate the role of synchronization and the dynamical binding capabilities in visual information processing stimulated the design of oscillatory network models with various types of oscillators as processing units and synchronization-based performance [30, 59, 98, 102–104, 106, 115, 117]. The discussions on dynamical binding via synchronization and elaboration of neuromorphic oscillatory network models, providing dynamical algorithms for image processing with synchronization-based performance, are actively continued nowadays [12, 13, 32, 34–37, 62, 66–77, 116] (see Sections 3.3–3.5).

First of all, a proper model of a neural oscillator was necessary for oscillatory network modeling. Neural oscillators in the brain are typically formed by a pair of excitatory and inhibitory neurons. More concretely, the following pairs of neurons can form an oscillator in various brain structures: pyramidal cell and interneuron in the visual cortex, mitral cell and granule cells in olfactory bulb and cortex, pyramidal cell and thalamic interneuron in thalamocortical system; pyramidal cell and basket cell in hippocampus, motoneuron, and Renshaw interneuron in the spinal cord. The first oscillator model was suggested by Freeman [23], when the prominent 40–60 Hz synchronous oscillations were reported in the rat and rabbit olfactory bulb and cortex. Following Freeman, Li, and Hopfield suggested the oscillator model closely imitating a real cortical neural oscillator, and further used it for modeling the olfactory brain system where oscillations and synchronization play a key role in odor recognition task [78].

We present three typical models of the neural oscillator that were successfully used in neuromorphic oscillatory network models. The first model used in series of papers by Wang et al. [12, 13, 116–118] is a relaxation oscillator representing a version of the Van der Pol oscillator. The oscillator dynamics is defined by two coupled units x and y and is governed by the two-dimensional dynamical system

$$\begin{aligned} \dot{x} &= f(x,y) + S + I + \eta, & f(x,y) &= 3x - x^3 + 2 - y, \\ \dot{y} &= \epsilon g(x,y), & g(x,y) &= \alpha(1 + \tanh(x/\beta)) - y, \end{aligned} \tag{1.6-1}$$

where S defines the total contribution from the network oscillator coupling, I is an external input (defined by local image characteristics), η is a Gaussian noise, and ϵ, α and β are positive parameters controlling the form of the oscillator limit cycle.

A biologically motivated model of a limit cycle oscillator formed by an excitatory pyramidal neuron and inhibitory interneuron of the visual cortex was designed by Li

and used in a series of oscillatory network models for visual image segmentation [74–78]. The governing dynamical system of the oscillator model reflecting orientation-selective response of simple cells of the brain visual cortex can be written as

$$\begin{aligned} \dot{u} &= -u + \alpha h(u) - g(v) + I^{+}, \\ \dot{v} &= -v + \beta h(u) + I^{-}, \end{aligned} \tag{1.6-2}$$

where u and v are the membrane potentials of the excitatory and inhibitory neurons, $h(u)$ is an activation function of the excitatory neuron (of sigmoid form), $g(v)$ is an activation function of the inhibitory neuron (monotonic nondecreasing function), and $I^{\pm}$ are external inputs on excitatory and inhibitory neurons, respectively. The dynamical system possesses stable limit cycle existing under the condition that the values $I^{\pm}$ belong to some finite interval. Otherwise the limit cycle bifurcates into a stable focus.

The features of the dynamical behavior of a biologically motivated neural oscillator designed by Li were accounted for in the oscillator model based on a modified version of the Ginzburg–Landau oscillator. The improved versions of the oscillator were further used in series of oscillatory network models providing a synchronization-based algorithm for image processing tasks [34–37, 66–68, 71–73]. The oscillator state is defined by a pair of real variables (u_1, u_2), and a two-dimensional governing dynamical system can be written in the form of a single equation for the complex-valued variable $u = u_1 + iu_2$

$$\dot{u} = f(u) + g(I), \quad f(u) = \left(\rho_0^2 + i\omega - |u - c|^2\right)(u - c), \tag{1.6-3}$$

where ρ_0 and $c = c_x + ic_y$ are constants, defining the parameters of the oscillator limit cycle at $g(I) = 0$ (the limit cycle is the circle of radius ρ_0 located at the point of the complex plane (u_1, u_2) with coordinates (c_x, c_y) at $g(I) = 0$), and ω is the natural oscillator frequency). The function $g(I)$, which dependends on local image characteristics, controls the limit cycle size and the cycle location in the complex plane (u_1, u_2). The cycle size is sufficiently large if the brightness I of the pixel that corresponds to the oscillator is greater than some predetermined threshold value I^* (in the case the oscillator state is considered as "active"). Otherwise the limit cycle bifurcates into a stable focus ("passive" oscillator state). A monotonic dependence of the limit cycle size on pixel brightness I at $I \geq I^*$ can be provided via special construction of the function $g(I)$ (see Section 4.2).

1.6.2 Synchronizable oscillatory networks

The first biologically motivated oscillatory network model was developed by Li [74–78]. It represents a spatially three-dimensional oscillatory network, modeling the columnar architecture of the primary VC. Although the network was composed of neural oscillators, the system of excitatory and inhibitory connections for network

oscillators was constructed separately, based on experimental neurobiological data on horizontal intracortical connections in the VC. The model was tested in problems of preattentive image processing including contour integration and texture segmentation tasks. It demonstrated successful synchronization-based performance.

The second remarkable oscillatory network model for visual image segmentation was the oscillatory network LEGION, whose initial version was designed in 1995 [117, 118]. The model was not intended for a direct imitation of the brain visual processing. Nevertheless, the most perfect model version [13] provided an effective dynamical image segmentation method based on the synchronization of an oscillatory network. The active network unit is a relaxation oscillator with internal dynamics dependent on image pixel brightness. Network oscillators are located in a two-dimensional spatial lattice being in one-to-one correspondence with the image pixel array. In addition to stationary local excitatory connections and global network inhibition, a kind of dynamical oscillator coupling in the network was designed. An algorithm of dynamical coupling adaptation that allowed us to significantly improve the network performance was developed. As a result, the final model version [13] provided successful segmentation of real gray-level images containing more than 400 000 pixels. The comparison of the dynamical oscillatory network algorithm with several modern traditional algorithms of image processing demonstrated real advantages of the former.

The first version of our oscillatory network model of VC [66–68, 71–73] was in fact inspired by the model by Li. However, in contrast with that model, the pure oscillatory network model with completely dynamical interaction (without any persistent oscillator connections) was designed. The attempt to imitate image processing typical to the low (preattentive) level of the brain vision system performance was undertaken via the network architecture and performance. Namely, only a single step of bottom-up VC performance in the task of image reconstruction (without recognition) was simulated. The network oscillators are located in a three-dimensional spatial lattice placed inside of a parallelepiped. The two-dimensional lattice of $M \times N$ nodes is located in the parallelepided base corresponding to the rectangle containing an $M \times N$ image pixel array. The three-dimensional lattice is considered as consisting of $M \times N$ one-dimensional columns, each of K oscillators, oriented normally to the parallelepiped base. The total number of the nodes of 3D lattice is $M \times N \times K$, and the lattice imitates the columnar architecture of the VC. There is the one-to-one correspondence between the image pixel array and oscillator columns: one oscillator column represents a single pixel. The limit-cycle oscillator (1.6-3) is chosen as a network processing unit. The proper oscillator response on pixel brightness was designed by taking into account the main features of cortical neural oscillator dynamics constructed in [74]. The parametrical dependence of oscillator dynamics on two image characteristics – a pixel brightness and an elementary bar orientation – was designed. The elementary bar orientation corresponding to a pixel characterizes the inhomogeneity of its brightness and is specified by a unit vector orthogonal to the direction of local brightness gradient. Nonlocal self-organized dynamical connections of the oscillatory net-

work are constructed based on the presumable properties of the dynamical binding. The dynamical connection of each oscillator pair depends on: (a) closeness of pixel brightness levels corresponding to the oscillator pair and (b) closeness of oscillator receptive field orientation (which is an internal oscillator parameter prescribed to the lattice node where the oscillator is located) to the corresponding pixel bar orientation of each oscillator. The reduced oscillatory network located in the two-dimensional lattice of $M \times N$ nodes being in one-to-one correspondence with the image pixel array was extracted from the initial three-dimensional network model. The reduced model turned out to be capable of brightness image segmentation and, in addition, of texture image segmentation in the cases of simplest texture types admitting a representation by a collection of oriented bars.

The initial version of the 2D oscillatory network model was further significantly improved via the construction of more convenient single oscillator dynamics and introduction of more flexible principles of oscillator coupling [34–37]. The improved model allowed us to include more complicated tasks of image processing such as (see Chapter 4) (a) brightness segmentation of real gray-level images, (b) selective image segmentation, (c) color image segmentation and (d) object selection in a visual scene.

On the whole, the dynamical image processing approach via the design of synchronizable oscillatory networks can provide new insights into features of information processing in spatially distributed oscillatory networks. In particular, it seems plausible that the effects of the versatile synchronization phenomenon could be exploited in information coding, different from dynamical binding [105]. Synchronization could be used as a mechanism of acceleration of signal propagation through spatially distributed networks, and the dynamical binding strategy could help to discover various flexible mechanisms of network processor coordination.

1.6.3 Networks of pulsed neurons

During the last decade a new generation of artificial neural network models – networks of spiking neurons communicating through pulses – has been designed [29, 51, 82]. The first model of spiking neuron was proposed by Hodgkin and Huxsley [46]. It was suggested that networks of spiking neurons can provide more powerful computational tools than traditional neural networks [81]. Due to their more realistic properties, the networks of pulsed neurons can be used to study so-called biological neural circuits. From the informational viewpoint, it would be interesting to propose models explaining how information can be encoded and decoded by a series of pulse trains. A number of complex computational models were created to simulate biologically realistic neural networks.

From the dynamical viewpoint, pulsed (spiking) neurons belong to a class of so-called excitable systems that possess coexisting limit cycles and stable equilibrium points as independent attractors of dynamics. Oscillators demonstrating bursting os-

cillations were used for modeling pulsed neuron dynamics. The bursting oscillations represent a dynamical regime of the periodically repeated transition from a rest state to a state of fast, almost periodical oscillations. A dynamical system demonstrating bursting oscillations can be written in the form

$$\begin{aligned} \dot{\mathbf{x}} &= f(\mathbf{x}, \mathbf{y}), \\ \dot{\mathbf{y}} &= \mu g(\mathbf{x}, \mathbf{y}), \end{aligned} \tag{1.6-4}$$

where $\mathbf{x}$ is a "fast" variable, $\mathbf{y}$ is a "slow" variable, and μ is a small parameter. The function $f(\mathbf{x}, \mathbf{y})$ specifies the fast oscillatory process, whereas $g(\mathbf{x}, \mathbf{y})$ describes a slow modulation of the fast oscillations. A detailed analysis of several classes of dynamical systems demonstrating bursting oscillations and classification of bursting oscillatory regimes is given in [50]. From the viewpoint of synchronizable systems, the important feature of networks of bursting oscillators is that a complete synchronization of fast oscillations contained in burster packets can be much harder achievable than the synchronization of burster packets themselves. Hence, the synchronization of the bursting oscillators is harder than the synchronization of oscillators demonstrating strongly periodical (even complicated) oscillations [50]. Nevertheless, as theoretical studies show, various types of complicated synchronized regimes that characterize networks of limit cycle oscillators (partial and generalized synchronization, rich variety of complex spatio-temporal types of dynamics) are inherent also to networks of pulsed neurons [120]. It means that attractors of different types and dynamic complexity can coexist for dynamical systems governing networks of spiking neurons.

1.6.4 Signal processing and communication via chaos synchronization

The discovery of chaotic dynamics in simple dynamical systems was one of most important events in the field of nonlinear dynamical systems. The idea that deterministic nonlinear systems can behave in a chaotic manner was first noticed by Henri Poincaré. At present, the phenomenon is investigated in detail, and chaos is now a multidisciplinary field of research ranging from pure mathematical topics to chaos control. In the phase space of a dynamical system demonstrating chaotic behavior there exists a complicated manifold named stochastic attractor. One of the simplest discrete-time two-dimensional dynamical systems that exhibits chaotic behavior is the Hénon map – the map of the plane $\mathbb{R}^2$ into itself [45]

$$\begin{aligned} x &\to x' = 1 - ax^2 + y, \\ y &\to y' = bx, \end{aligned} \tag{1.6-5}$$

where a and b are two real-valued parameters.

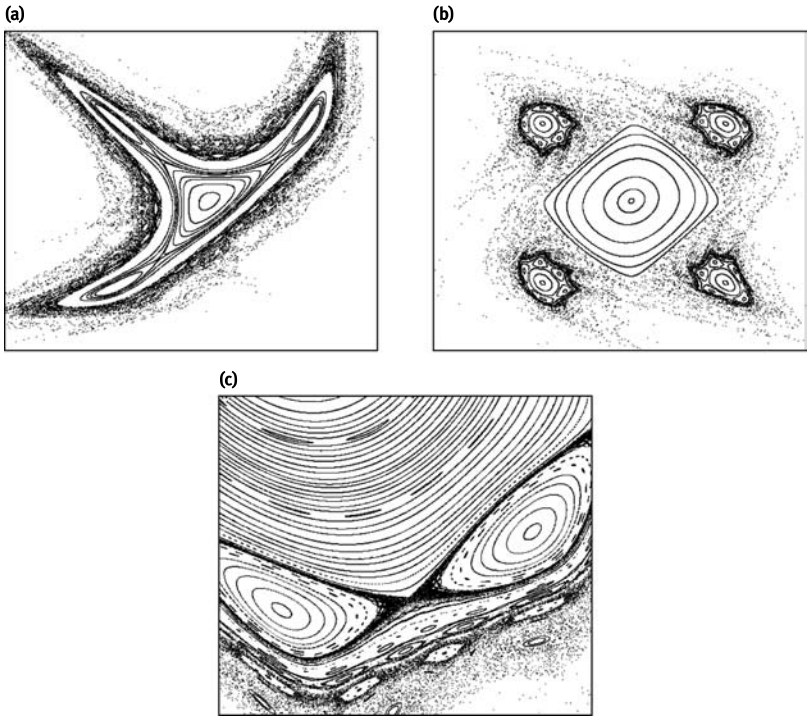

Fig. 1.5. The examples of fractal stochastic attractor in phase space of discrete map (1.6-6) (a version of Hénon map): (a) $a = 0.766$, $b = 0.4$; (b) $a = 0.8$, $b = -0.995$; (c) a fragment of fractal boundary of stochastic attractor at $a = 0.8$, $b = -0.9455$.

In Figure 1.5, the examples of chaotic dynamics that exhibit a version of Hénon map with cubic nonlinearity,

$$\begin{aligned} x &\to x' = a - x - bx^3 + y, \\ y &\to y' = -x, \end{aligned} \tag{1.6-6}$$

are presented.

The fractal stochastic attractors in the plane (x, y) (phase space of the discrete dynamical system) are shown in Figure 1.5 (a) and (b) at different values of parameters a and b; the fragment of fractal boundary of the stochastic attractor is depicted in Figure 1.5 (c).

The well-known Lorentz system [80] is an example of a three-dimensional continuous-time dynamical system possessing a stochastic attractor and demonstrating chaotic auto-oscillations. At appropriately chosen variables, the trajectory of the Lorentz system in the phase space looks like a blurred limit cycle, and temporal dependences of variables resemble periodic oscillations with variable amplitude and period [92]. The synchronization of chaotic oscillations may occur when two or more

systems demonstrating chaotic dynamics are coupled. Both phase and complete synchronization of chaotic oscillations are possible. This phenomenon has been studied for a long time, and is now well understood theoretically [2, 24, 90, 93]. Many of predicted theoretical results were further well established in computational experiments.

The capability of chaotic systems to be synchronized can be used in problems of signal processing and communication. Application of chaos synchronization in communication is aimed at providing the security of information transmission through telecommunication channels. Since 1990, when communication via chaotic system synchronization was first suggested [90], a great variety of different approaches to signal transmission via chaos synchronization have been proposed [6, 8, 9, 14, 16, 17, 40].

To implement chaos communication, two chaotic systems are required as a transmitter and a receiver. At the transmitter, a message signal is added to a chaotic signal, called the chaotic carrier. The undisturbed chaotic signal should be known to the response system (receiver). Via the synchronization of both chaotic systems (transmitter and receiver) the message can be decoded, and the message signal can be retrieved. This communication scheme is known as chaotic masking. There are a number of another communication schemes based on various ideas related to chaotic system synchronization: switching over chaotic regimes, using chaos pass filters, chaos frequency modulation, dual nonlinear transformation, advanced chaos control and many others. Chaos control is implied as stabilization by means of small system perturbations. The perturbations should be tiny, to avoid significant modifications of system natural dynamics [112].

Thus, the communication via chaos synchronization is a promising field of research with applications in cryptography. The isochronal synchronization of two chaotic systems provides information transmission with a small bit error rate by using chaos pass filters. The information transmission can be even error free if the chaos modulation approach is used. The introduction of the bidirectional coupling of transmitter and receiver chaotic systems allows additional promising advantages [55].

The following potential advantages of information transmission via chaos synchronization can be noted: chaos transmitter and receiver are easy to realize; chaos transmission provides lower interference with other kinds of transmission; privacy of the communication can be achieved via exploitation of chaos synchronization in information transmission.

The growing number of publications on the topic reflects permanent interest to the development of chaotic communication systems and their application in modern information technologies. New models provide realization of new methods of communication channel separation and a variety of facilities for transmitted information protection. Some of the research directions that are marked by experts as prospective are the following: elaboration of chaos generators, creation of fiber optic and cable networks, exploiting the features of chaotic systems, and development of computational methods of chaos generation.

1.7 Complicated networks

ANNs are currently used in a growing variety of research fields. The need to design the networks composed of processing units with more complicated internal dynamics than that of formal neurons often arises in applications. Interconnection architectures different from the extreme architecture types – completely regular or completely random – are also required. Such networks are regarded as complicated (or complex) ANN. There are the following sources of network complexity: complexity of network connectivity architecture (nonhomogeneity of network connections), dynamical complexity (complicated internal dynamics of network processing units and diversity of processing unit types) and gradual network evolution as a whole (under the learning procedure, a model of Darwinian evolution, or other prescribed evolution processes).

From the mathematical viewpoint it is interesting to understand the real evolution of a dynamical system governing network state dynamics. The subject of special interest concerns the types of network collective dynamics, the ways of information spread through the network and the influence of the network interconnectivity architecture on network functioning. Mathematical analysis of complicated networks started only recently, and a number of useful exactly solvable models were suggested.

As we saw above, RNNs tend to relax into stable static equilibria states if the dynamical system of single network processing unit has no other attractors besides stable fixed points. Many of such equilibrium network states can coexist. If the network dynamical system possesses stable attractors, different from fixed points, the attractor associative memory networks can be designed as well. For oscillatory networks, where processing units display stable attractors in the form of limit cycles, stable attractors of network dynamics are complicated stationary dynamical regimes. So, oscillatory networks can be referred to as a class of complicated networks. Depending on spatial dimension and symmetry properties, oscillatory networks can demonstrate synchronization and various inhomogeneous spatiotemporal dynamical regimes: traveling waves, rotating spiral waves, scroll waves, cluster states. Important strict results were obtained for oscillatory networks of special connectivity architectures: one-dimensional oscillatory chains, networks localized in spatial lattices of different dimensionality, networks with homogeneous all-to-all coupling. The features of collective dynamics types inherent to some of these oscillatory networks were studied analytically.

The phenomenon of the synchronization of large-scale oscillatory systems stands on the boundary between nonlinear dynamics and statistical physics. The analysis of synchronization as a phase transition was first given in the papers by Kuramoto with coauthors for oscillatory systems with global (all-to-all) homogeneous oscillator interaction [63–65, 101]. The macrovariable, admitting the interpretation of the self-consistent mean field, acting on each oscillator of an oscillatory system, was introduced, and dynamical equations governing microscopic dynamics of oscillatory system were rewritten in terms of the mean field variable. It was shown analytically that the transi-

tion into the synchronization state occurs at some critical value of interaction strength in the oscillatory network. The bifurcation corresponding to synchronization onset resembles the phase transition of second kind in statistical mechanics.

Another class of oscillatory systems is represented by networks of locally coupled oscillators localized in d-dimensional spatial lattices. In the case of homogeneously coupled oscillatory chains ($d = 1$), the critical value κ_c of interaction strength κ at which synchronization arises essentially depends on the form of the oscillator frequency distribution function. In the case of the Gaussian frequency distribution, the asymptotical dependence of κ_c on the number N of oscillators in the system was shown to be $\kappa_c = \mathcal{O}(\sqrt{N})$, whereas in the case of a deterministic frequency distribution with linear gradient (that is, when $\omega_{j+1} - \omega_j = \Delta$, $j = 1, \ldots, N-1$) the dependence is transformed into $\kappa_c = \mathcal{O}(N)$. Therefore, locally homogeneously coupled oscillatory chains with oscillator frequencies monotonically growing along the chain belong to oscillatory systems that are most difficult to synchronize [107].

Strict results concerning synchronization features of locally homogeneously connected oscillatory networks localized in d-dimensional cubic lattices were obtained for the case of random oscillator frequency distribution [107]. Formation and spatial dynamics of synchronized clusters were studied at the fixed interaction strength κ and a growing number N of oscillators in the system. It was shown that synchronized clusters of oscillators in a partially synchronized state possess a sponge (fractal) form.

The synchronization of oscillatory networks with special random graphs of connections was also studied. Namely, the random graphs with m connections, outgoing each graph node, were considered. It was found that at small m the network is in the state of partial synchronization, being fragmented into a collection of small synchronized clusters. As m grows, the bifurcation occurs at $m = N/2$: the set of separated clusters joins into a single giant cluster which size grows as $\mathcal{O}(N)$ at $N \to \infty$ [21].

The analytical results helped to construct a new class of networks admitting analytical study – networks with regular local and sparse global random connections. This class of networks was named "small-world" networks [87, 119]. The "small-world" networks allowed us to elucidate the role of long-range interaction for information processing in large-scale networks. As was found, the connectivity architecture of "small-world" networks is dynamically advantageous: the networks demonstrate better synchronization capabilities and facilities of fast information processing compared to analogous networks with pure local coupling. Besides, "small-world" networks possess better noise resistance. As a result, it is possible to say that long-range connections (even sparse) provide a global coordination of information stream in the system.

The "small-world" network models were developed and exploited in a variety of research areas ranging from neurodynamics and modeling of self-organization processes to the analysis of algorithms complexity. As it turned out, many graph topology types associated with some difficult problems are similar to the graph topology of "small-world" networks. The "small-world" connectivity architecture ensures fast sig-

nal propagation in the network, which cannot be achieved both for networks with regular pure deterministic connectivity and for those with pure random interconnections. Thus, "small-world" networks permitted us to make a step toward the interconnection architecture analysis of complicated large-scale networks and to study the influence of the network architecture on the set of possible dynamical regimes [122].

1.8 Multi-agent systems

1.8.1 Multi-agent modeling

Multi-agent modeling is a relatively new area of research dealing with modeling of interacting agent communities as dynamical systems. The multi-agent systems (MAS) can be viewed as natural extension of cellular automata models [123]. They represent a community of processors (finite automata) that function in common environment and work together at solving a common problem. Each processor may be a complicated system capable of changing its own state and internal dynamics to learn and to plan the strategy of cooperation with the other processors. The environment is also usually considered as dynamical. Cooperative solution is achieved due to various agent interactions, such as coordination, cooperation, learning, and planning of mutual problem solution. To obtain a cooperative solution, it is necessary to operate by several independent criterions: effectiveness of resource consumption, restriction of admissible time, safety demands, possibility of unforeseen situations, etc. The abilities to negotiate with each other, to cooperate and coordinate efforts are strongly necessary for successful work of MAS. Memory capabilities are needed to store knowledge about the environment and the information on members of agent society.

The MAS approach provides direct computational representation of the whole system through simulating dynamical behavior of individual agents and their interaction and studying the system evolution. Macroscopic dynamics of the whole system then will give information on emergent society behavior. The difference between mathematical modeling in terms of the system of ordinary differential equations and direct modeling via MAS can be outlined through the example of dynamics of a simple two-population predator–prey system. The two-dimensional Lotka–Volterra dynamical system describes an oscillatory behavior of the overall predator–prey population (undamped oscillations of constant amplitude for each subpopulation). Modeling the same system in the frames of MAS allows us to capture fluctuations: the corresponding subpopulations demonstrate oscillations with fluctuating amplitudes.

Nowadays highly flexible MAS models are successfully used in a diversity of research fields. MAS models are exploited in problems of artificial population evolution (artificial life), in modeling distributed intellectual systems (collectives of autonomous mobile robots), in monitoring problems (search for resources, study of resource distribution, modeling extreme situations), in bio-robotics (control of intelli-

gent robots, microrobotics), and in industry (planning, automation of plants, creation of control systems). In medicine MAS are used for tumor detection (redundant data are eliminated via genetic and immune algorithms). In mathematics MAS are attracted for the creation of self-organized adaptive parallel computational algorithms.

1.8.2 Reinforcement learning for multi-agent systems

Reinforcement learning [108] is the learning without a teacher. It is inspired by behavioral features of biological systems and can be viewed as a learning via interacting with an environment. Animals and artificial systems face similar problems in learning: how to optimize the behavior in the light of rewards or punishments. The performance of reinforcement learning algorithms is estimated with the help of the fitness (or reward) function. Reinforcement learning can be used in a variety of areas, such as optimal control theory, game theory, information theory, simulation-based optimization, and machine learning.

Being used in the frames of multi-agent systems, the learning deals with states, values of actions, etc. The agent performance is compared to that of an agent that acts optimally. There exist several ways to determine the optimal value function and the optimal strategy, which have their origins in the field of dynamic programming [7]. Reinforcement learning often focuses on an online performance and requires attraction of exploration mechanisms. Random selection strategy is known to give rise to poor performance. A deterministic stationary strategy exploits deterministic selection actions based on information about a current state. The strategy can be identified with a mapping from the set of states to the set of actions, the expectations being approximated by averaging over samples. The Monte Carlo methods are sometimes used to model strategy iterations, the recursive Bellman equation being attracted for the estimation of variance. The methods of stochastic optimization are used as an alternative method of good strategy search.

Currently the reinforcement learning is an actively studied topic in the area of artificial intelligence. This kind of learning is attractive because of its generality. It can be viewed as an extension of classical dynamic programming. Unlike supervised learning, the reinforcement learning does not require explicit train samples. By combining dynamic programming with the neural network approach many previously unsolvable problems can be solved.

1.9 Wireless networks and cognitive dynamical systems

During the last several years the Internet has evolved from a wired infrastructure to a hybrid of wired and wireless domains by spreading worldwide interaction for microwave access (WiMAX), Wi-Fi, and cellular networks. Therefore, there is a growing

need to facilitate a reliable content delivery over such heterogeneous networks. The following obvious advantages of wireless communication can be noted: (a) communication has been enhanced to convey the information quickly to the consumers; (b) professionals can access Internet anywhere without cables; (c) urgent situations can be alerted through wireless communication; (d) wireless networks are cheaper to install and maintain.

The proposed multiinput multioutput (MIMO) systems and networking technologies introduced a revolution in recent times, especially in the wireless and wired multicast (multipoint-to-multipoint) transmission field. MIMO networks and wired multicast transmission systems have a considerable influence on many applications such as teleconferencing and information dissemination services. Numerous research projects were carried out to explore how to support multicast in various networking environments. Design and analysis of wireless networks are on the agenda. In particular, such topics as the dependence of wireless network performance on network topology and multicast flow control using feedback mechanisms [83] are of interest.

Wireless sensor networks consisting of a number of distributed sensor nodes have received much attention in recent years as well. The nodes cooperate with each other to transmit the sensed data to an external base station. Due to their distributed character, a collaborative organization for robust communication is required. Innovative techniques are used for reliable communication organization. For example, an artificial bee colony algorithm, simulating the intelligent foraging behavior of honey bee swarms, was suggested for the routing protocol of wireless sensor networks [54].

The cognitive radio networks are complex multiuser wireless communication systems capable of emergent behavior. They are designed based on the concept of dynamic and intelligent network management and provide such capabilities as spectrum sensing, dynamic spectrum management, robust transmit power control self-configuration, self-learning, and emergent behavior. The primary goals of these wireless networks are: (a) to provide highly reliable communication for all users of the network; (b) to facilitate efficient utilization of the radio spectrum in a fair-minded way. The cognitive radio networks can be viewed as brain-inspired wireless communications. Many novel architectures, protocols and algorithms have been proposed and implemented. These networks can learn from the adaptations and allow the efficient utilization of radio spectrum and dynamic spectrum sharing [109]. Analyzing the properties of broadcasting in wireless networks, it was shown that asymptotically optimal broadcast capacity and latency can be achieved in mobile wireless networks [96]. Heterogeneous wireless systems are characterized by the coexistence of a variety of radio technologies, and a key aspect is then the implementation of efficient joint radio resource management mechanisms. The multimedia scenarios of politics on distribution decisions are developed [31].

The diversity of problems, arising in the field of wireless networking is related to such research areas as cognitive information processing, signal processing, communication and control. On the whole, the area of wireless networks stimulated the

emergence of a new discipline, called cognitive dynamic systems [43]. It is based on the ideas developed in statistical signal processing, stochastic control and information theory. It also uses the ideas drawn from neuroscience, statistical learning theory and game theory. This discipline promises to provide principal tools for further development of a new generation of wireless dynamic systems including cognitive radio and cognitive radar. The approaches elaborated in the fields of Kalman filtering and dynamic programming by Bellman for classical dynamical systems can be extended to problems of signal processing, communication and control. It is noteworthy that according to a neurobiological evidence, the Kalman-like filtering is used in the visual and motor brain cortices, and (plausible) also in the brain auditory system. On the other hand, as it is known, the reinforcement learning, which can be viewed as minimally supervised form of learning, is exploited by animals. Animals are not told what actions to perform, rather, they work out the actions themselves based on reinforcement received from the environment. A special form of reinforcement learning known as temporal-difference (TD) learning can be viewed as a combination of Monte Carlo methods with ideas contained in dynamic programming. Based on the animal brain features and on the ideas of Kalman-like filtering, a new approach was developed that led to the creation of new research area – cognitive dynamical systems [43].

Cognitive dynamical systems are built up based on the rules of behavior over time through learning from continuous experimental interactions with the environment. For instance, the cognitive radio system would be able to identify subbands of radio spectrum that are currently unemployed and assign them to unserviced users. A perfect example of cognitive radar can be constructed based the echolocation system of the bat. The bats store information concerning their habitat that was accumulated through their lifetime experience. With this information the bats are capable of locating their prey with accuracy and resolution that would be the envy for radar and sonar engineers. Although the intended applications of cognitive radio and cognitive radar are different, they do share two common features: scene analysis and operating by a feedback channel (connecting the receiver to the transmitter and providing the informational transmitter adaptation to the environment). Thus, both cognitive radio and cognitive radar represent examples of wireless control systems, and the feedback can be considered as the facilitator of intelligent information processing.

To complete the presentation of cognitive dynamical systems, it is worth adding that, in the expert opinion, the following research areas stand under the umbrella of cognitive dynamical systems: signal processing, communication theory, control theory, radar systems, neural networks and learning machines [42].

1.10 Quantum and classical quantum-inspired parallel computational algorithms

Nowadays there is a growing area of research dealing with the construction of quantum and quantum-inspired artificial neural networks and evolutionary algorithms. Qubit-inspired neuron models are used as processing units of these artificial neural networks, and concepts and principles of quantum computing are exploited in the design of network performance. Some quantum-inspired network models demonstrate more powerful computational capabilities in various complex tasks solvable with artificial neural networks, including data classification tasks, various problems of global search and others.

Quantum entanglement, a characteristic feature of quantum systems, is considered nowadays as the main physical resource of quantum information processing and quantum computing. Qualitative theory of quantum entanglement is one of modern directions in quantum information processing, and methods of entanglement measure calculation are currently under extensive development. Despite the ubiquitous nature of entangled states, many features of these states remain unusual and even mysterious from the viewpoint of classical physics.

1.10.1 Photonic one-way quantum computations

In photonic (optical) quantum computing photons figure as qubits (which are usually called polarization qubits). Photons are ideal carriers for quantum information, carrying it in the polarization degree of freedom. Photons are easily available, do not interact with the environment and therefore exhibit negligible decoherence. Their excellent manipulability allows the precise execution of single-qubit operations. Photons can be manipulated in an extremely fast and accurate manner and can easily be transmitted over long distances using optical glass fibers. Thus, a possibility of making a quantum computer that uses light seems very appealing.

However, the absence of photon-photon interactions makes two-qubit operations required for universal quantum computing very difficult to realize, and two-qubit gates for polarization qubits cannot be constructed. The construction of probabilistic optical two-qubit gates was suggested by Knill, Laflamme, and Milburn in 2000. This approach is now known as the KLM scheme [56]. Various experimental versions of the probabilistic gates were suggested later, and a number of experimental groups have already demonstrated all-optical probabilistic quantum gates [58].

An alternative approach to the usual circuit model of quantum computing was introduced by Raussendorf and Briegel [95]. It is the so-called one-way (or the cluster state) scheme of quantum computing, which is based on the creation of a cluster of entangled qubits as a computational resource. The computations are then proceed as a series of single-qubit measurements, the basis and order of measurements be-

ing defined by the computational algorithm itself. Since the entangled cluster can be prepared "off-line," the one-way scheme seems to be a promising quantum computing scheme. Currently all-optical implementation of the one-way computation scheme with four, six and ten qubit clusters has been experimentally realized. The underlying resource for the computations – entangled cluster states of different topological architectures – can be generated using the nonlinear optical process known as spontaneous parametric down-conversion.

1.10.2 Quantum-inspired neural networks and evolutionary algorithms

A number of novel classical parallel evolutionary algorithms, inspired by quantum computation algorithms, were developed based on the concepts and principles of quantum computing such as linear superposition of states. Quantum and quantum-inspired artificial neural networks (with qubit-inspired neurons as processing units) were designed and in many cases demonstrated more powerful computational capabilities than classical neural networks [5, 61, 97]. They are learnable and capable of solving complex tasks of data classification and problems of binary mapping. Attraction of quantum information processing principles into the problem of designing associative memory networks provided an exponential increase of network storage capacity [97]. Genetic quantum-inspired algorithms (stochastic, population-based optimization algorithms combining many principles in a global search) were also developed. The algorithms often incorporate operations of state superposition and reinforcement learning. In some problems they demonstrated rapid convergence and good global search capability.

Bibliography

[1] D. H. Ackley, G. E. Hinton and T. J. Sejnowski. A learning algorithm for Boltzmann machines. *Cogn. Sci.*, 9:147, 1985.
[2] V. S. Afraimovich, N. N. Verichev and M. I. Rabinovich. Stochastic synchronization of oscillation in dissipative systems. *Radiophys. Quantum El.*, 29:795, 1986.
[3] I. Aleksander. *The world in my mind, my mind in the world: Key mechanisms of consciousness in humans, animals and machines*. Imprint Academics, 2005.
[4] S. Amari and N. Murata. Statistical theory of learning curves under entropic loss criterion. *Neural Comput.*, 5:140, 1993.
[5] M. Andrecut and M. K. Ali. A quantum neural network model. *Int. J. Mod. Phys. C*, 13:75, 2002.
[6] Yu. V. Andreyev, Yu. L. Belsky and A. S. Dmitriev. Storing and retrieving information as stable cycles of 2-d and multi-dimensional maps. *Radiotekhnika i Elektronika*, 39:114, 1994. in Russian.
[7] R. E. Bellman. *Dynamic Programming*. Dover Publications, 2003.
[8] Yu. L. Belsky and A. S. Dmitriev. Communications using deterministic chaos. *Radiotekhnika i Elektronika*, 38:1310, 1993. in Russian.

[9] V. Belykh, I. Belykh and M. Hasler. Connection graph stability method for synchronized coupled chaotic systems. *Physica D*, 195:159, 2004.
[10] C. M. Bishop. *Neural networks for pattern recognition*. Clarendon Press, Oxford, 1995.
[11] A. E. Bryson and Yu-Chi Ho. *Applied optimal control: optimization, estimation, and control*. Taylor & Francis, 1975.
[12] E. Cesmeli and D. L. Wang. Texture segmentation using Gaussian–Markov random fields and neural oscillator networks. *IEEE T. Neural Networ.*, 12:394, 2001.
[13] K. Chen, D. Wang and X. Liu. Weight adaptation and oscillatory correlation for image segmentation. *IEEE T. Neural Networ.*, 11:1106, 2000.
[14] H. Dedieu, M. P. Kennedy and M. Hasler. Chaos shift keying: modulation and demodulation of a chaotic carrier using self-synchronizing Chua's circuits. *IEEE T. Circuits II*, 40:634, 1993.
[15] J. DeFelipe, H. Markram and K. S. Rockland. The neocortical column. *Front. Neuroanat.*, 6:22, 2012.
[16] A. S. Dmitriev, A. V. Kletsov, A. M. Laktushkin, A. I. Panas and S. O. Starkov. Ultrawideband wireless communications based on dynamic chaos. *Radiotekhnika i Elektronika*, 51:1193, 2006. in Russian.
[17] A. S. Dmitriev and S. O. Starkov. Message transmission using chaos and classical information theory. *Foreign Radioelectronics. Advances in Modern Radioelectronics*, 11:4, 1998. in Russian.
[18] R. Eckhorn, R. Bauer, W. Jordan, M. Brosch, W. Kruse, M. Munk and H. J. Reitboek. Coherent oscillations: a mechanism of feature linking in the visual cortex? multiple electrode and correlation analyses in the cat. *Biol. Cybern.*, 60:121, 1988.
[19] G. M. Edelman. *Neural Darwinism*. Basil Books, 1987.
[20] S. F. Edwards and P. W. Anderson. Theory of spin glasses. *J. Phys.: Metal Phys.*, 5:965, 1975.
[21] P. Erdős and R. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci.*, 5:17, 1960.
[22] D. B. Fogel. *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*. John Wiley & Sons, 2005.
[23] W. J. Freeman. Spatial properties of an EEG event in the olfactory bulb and cortex. *Electroen. Clin. Neuro.*, 44:586, 1978.
[24] H. Fujisaka and T. Yamada. Stability theory of synchronized motion in coupled oscillator systems. *Prog. Theor. Phys.*, 69:32, 1983.
[25] K. Fukushima. Cognitron: A self-organizing multilayered neural network. *Biol. Cybern.*, 20:121, 1975.
[26] K. Fukushima. Neocognitron: A hierarchical neural network capable of visual pattern recognition. *Neural Networks*, 1, 1988.
[27] D. Gabor. Communication theory and cybernetics. *T. IRE Prof. Group Circuit Theory*, CT-1:19, 1954.
[28] E. Gardner. Maximum storage capacity in neural networks. *Europhys. Lett.*, 4:481, 1987.
[29] W. Gerstner and W. M. Kistler. *Spiking neuron models*. Cambridge University Press, 2002.
[30] W. Gerstner, R. Ritz and J. L. van Hemmen. A biologically motivated and analytically soluble model of collective oscillations in the cortex: I. theory of weak locking. *Biol. Cybern.*, 68:363, 1993.
[31] J. Gonzalez, M. C. Lucas-Estañ and J. Sanches-Sortiano. Joint radio resource management for heterogeneous wireless systems. *Wirel. Netw.*, 18:443, 2012.
[32] C. M. Gray. The temporal correlation hypothesis of visual feature integration: Still alive and well. *Neuron*, 24:31, 1999.
[33] C. M. Gray and W. Singer. Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proc. Natl. Acad. Sci. USA*, 86:1698, 1989.

[34] E. Grichuk, M. Kuzmina and E. Manykin. Oscillatory network for synchronization-based adaptive image segmentation. In *Proc. IJCNN'06*, page 4529, 2006.
[35] E. Grichuk, M. Kuzmina and E. Manykin. Image processing via synchronization in self-organizing oscillatory network. In *Proc. ECC'07, Vol. 1*, volume 27 of *Lecture Notes in Electrical Engineering*, page 97, 2009.
[36] E. S. Grichuk, M. G. Kuzmina and E. A. Manykin. Oscillatory network for synchronization-based adaptive image segmentation. *Opt. Mem. Neural Networks*, 15:57, 2006.
[37] E. S. Grichuk, M. G. Kuzmina and E. A. Manykin. Object selection in visual scene via oscillatory network with controllable coupling and self-organized performance. *Opt. Mem. Neural Networks*, 20:113, 2011.
[38] S. Grossberg. Adaptive pattern classification and universal receding: I. Parallel development and coding of neural detectors. *Biol. Cybern.*, 23:121, 1976.
[39] S. Grossberg. Adaptive pattern classification and universal receding: II. Feedback, expectation, olfaction, illusions. *Biol. Cybern.*, 23:187, 1976.
[40] M. Hasler and T. Schimming. Optimal and suboptimal chaos receivers. *Proc. IEEE*, 90:733, 2002.
[41] S. Haykin. *Neural networks: A comprehensive foundation*. Prentice Hall, 1998.
[42] S. Haykin. Cognitive radio: brain-inspired wireless communications, 2008. Invited lecture.
[43] S. Haykin. *Cognitive dynamical systems*. McMaster University, 2012.
[44] D. Hebb. *The Organization of Behavior: A Neuropsychological Theory*. Taylor & Francis, 2002.
[45] M. Hénon. A two-dimensional mapping with a strange attractor. *Comm. Math. Phys.*, 50:69, 1976.
[46] A. L. Hodgkin and A. F. Huxsley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.*, 117:500, 1952.
[47] J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proc. Nat. Acad. Sci.*, 79:2554, 1982.
[48] J. J. Hopfield. Neurons with graded response have collective computational properties like those of two-state neurons. *Proc. Nat. Acad. Sci.*, 81:3088, 1984.
[49] D. H. Hubel. *Eye, brain, and vision*. W. H. Freeman, 1995.
[50] E. M. Izhikevich. Synchronization of elliptic bursters. *SIAM Rev.*, 43:315, 2001.
[51] E. M. Izhikevich. Simple model of spiking neurons. *IEEE T. Neural Networ.*, 14:1569, 2003.
[52] E. M. Izhikevich and G. M. Edelman. Large-scale model of mammalian thalamocortical systems. *P. Natl. Acad. Sci.*, 105:3593, 2008.
[53] M. Kam and R. S. Cheng. Convergence and pattern stabilization in the Boltzmann machine. In *Advances in neural information processing systems*, volume 1, page 511. Morgan Kaufmann, San Francisco, 1989.
[54] D. Karaboga, S. Okdem and C. Ozturk. Cluster based wireless sensor network routing using artificial bee colony algorithm. *Wirel. Netw.*, 18:847, 2012.
[55] W. Kinzel, A. Englert and I. Kanter. On chaos synchronization and secure communication. *Phil. Trans. R. Soc. A*, 368:379, 2010.
[56] E. Knill, R. Laflamme and G. J. Milburn. A scheme for efficient quantum computation with linear optics. *Nature*, 409:46, 2000.
[57] T. Kohonen. Self-organized formation of topologically correct feature maps. *Biol. Cybern.*, 43:59, 1982.
[58] Pieter Kok, W. J. Munro, Kae Nemoto, T. C. Ralph, Jonathan P. Dowling and G. J. Milburn. Linear optical quantum computing with photonic qubits. *Rev. Mod. Phys.*, 79:135, 2007.
[59] P. König and T. B. Schillen. Stimulus-dependent assembly formation of oscillatory responses: I. synchronization. *Neural Comput.*, 3:155, 1991.
[60] B. Kosco. Differential Hebbian learning. In *AIP Conference proceedings*, volume 151, 1986.

[61] N. Kouda, N. Matsui, H. Nishimura and F. Peper. Qubit neural network and its learning efficiency. *Neural Comput. Appl.*, 14:114, 2005.
[62] A. K. Kreiter and W. Singer. On the role of neural synchrony in primate visual cortex. In A. Aertsen and V. Braitenberg, editors, *Brain Theory*, page 201. Elsevier, 1996.
[63] Y. Kuramoto. Self-entrainment of a population of coupled nonlinear oscillators. In H. Araki, editor, *Int. Symp. on Mathematical Problems in Theoretical Physics*, volume 39, page 420, Berlin, 1975. Springer.
[64] Y. Kuramoto. *Chemical oscillations, Waves and Turbulence*. Dover Publications, 2003.
[65] Y. Kuramoto and I. Nishikawa. Statistical macrodynamics of large dynamical systems. case of a phase transition in oscillator communities. *J. Stat. Phys.*, 49:569, 1987.
[66] M. Kuzmina and E. Manykin. Biologically motivated oscillatory network model for dynamical image segmentation. In *Proc. BICS'04*, 2004.
[67] M. Kuzmina, E. Manykin and I. Surina. Oscillatory network with self-organized dynamical connections for synchronization-based image segmentation. *Biosystems*, 76:43, 2004.
[68] M. G. Kuzmina and E. A. Manykin. Oscillatory neural network for adaptive dynamical image processing. In *Proc. CIMCA'05*, page 301, 2005.
[69] M. G. Kuzmina, E. A. Manykin and I. I. Surina. Memory estimation of homogeneous closed chains of oscillators. In *Proc. Int. Conf. "Neuroinformatics-2000"*, page 37, 2000. in Russian.
[70] M. G. Kuzmina, E. A. Manykin and I. I. Surina. Oscillatory network model of the brain visual cortex with controlled synchronization. In *Proc. COC'00*, volume 3, page 461, 2000.
[71] M. G. Kuzmina, E. A. Manykin and I. I. Surina. Tunable oscillatory network for visual image segmentation. In *Proc. ICANN'01*, volume 2130 of *Lecture Notes in Computer Science*, page 1013, 2001.
[72] M. G. Kuzmina, E. A. Manykin and I. I. Surina. Oscillatory network with self-organized dynamical connections for image segmentation. *Neurocomputers*, 4:34, 2004. in Russian.
[73] M. G. Kuzmina and I. I. Surina. Spatially distributed oscillatory networks related to modelling of the brain visual cortex. In *Proc. NOLTA'00*, volume 1, page 336, 2000.
[74] Z. Li. A neural model of contour integration in the primary visual cortex. *Neural Comput.*, 10:903, 1998.
[75] Z. Li. Visual segmentation by contextual influences via intra-cortical interactions in the primary visual cortex. *Network: Comput. Neural Syst.*, 10:187, 1999.
[76] Z. Li. Pre-attentive segmentation in the primary visual cortex. *Spatial Vision*, 13:25, 2000.
[77] Z. Li. Computational design and nonlinear dynamics of a recurrent network model of the primary visual cortex. *Neural Comput.*, 13:1749, 2001.
[78] Z. Li and J. J. Hopfield. Modeling the olfactory bulb and its neural oscillatory processing. *Biol. Cybern.*, 61:379, 1989.
[79] R. Linsker. Self-organization in a perceptual network. *Computer*, 21:105, 1988.
[80] E. N. Lorenz. Deterministic nonperiodic flow. *J. Atmos. Sci.*, 20:130, 1963.
[81] W. Maass. Networks of spiking neurons: the third generation of neural network models. *Neural Networks*, 10:1659, 1997.
[82] W. Maass and C. M. Bishop, editors. *Pulsed Neural Networks*. The MIT Press, 1998.
[83] M. S. Mahmoud and M. M. Hassan Hamid. Self-tuning control for MIMO network systems. *J. Signal and Inf. Proc.*, 3:154, 2012.
[84] H. Markram. The blue brain project. *Nat. Rev. Neurosci.*, 7:153, 2006.
[85] Warren S. McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *B. Math. Biophys.*, 5:115, 1943.
[86] N. Murata, S. Yoshizawa and S. Amari. Learning curves, model selection and complexity of neural networks. In S. J. Hanson, J. D. Cowan and C. L. Giles, editors, *Advances in Neural Information Processing System 5*, page 607. Morgan Kaufmann Publishers, San Mateo, 1993.

[87] M. E. J. Newman, S. H. Strogatz and D. J. Watts. Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E*, 64:026118, 2001.
[88] E. Oja. A simplified neuron model as a principal component analyzer. *J. Math. Biol.*, 15:267, 1982.
[89] D. B. Parker. Learning-logic: Casting the cortex of the human brain in silicon. Technical Report TR-47, MIT Center for Computational Research in Economics and Management Science, 1985.
[90] Louis M. Pecora and Thomas L. Carroll. Synchronization in chaotic systems. *Phys. Rev. Lett.*, 64:821, 1990.
[91] R. Perin, T. K. Berger and H. Markram. A synaptic organizing principle for cortical neuronal groups. *Proc. Nat. Acad. Sci.*, 108:5419, 2011.
[92] A. Pikovsky, M. Rosenblum and J. Kurths. *Synchronization: A universal concept in nonlinear sciences*. Cambridge University Press, 2001.
[93] A. S. Pikovsky. On the interaction of strange attractors. *Z. Phys. B*, 55:149, 1984.
[94] T. Poggio and F. Girosi. Regularization algorithms for learning that are equivalent to multilayer networks. *Science*, 247:978, 1990.
[95] R. Raussendorf and H. J. Briegel. A one-way quantum computer. *Phys. Rev. Lett.*, 86:5188, 2001.
[96] G. Resta and P. Santi. The fundamental limits of broadcasting in dense wireless mobile networks. *Wirel. Netw.*, 18:679, 2012.
[97] G. G. Rigatos and S. G. Tzafestas. Quantum learning for neural associative memories. *Fuzzy Set. Syst.*, 157:1797, 2006.
[98] R. Ritz, W. Gerstner, U. Fuentes and J. L. van Hemmen. A biologically motivated and analytically soluble model of collective oscillations in the cortex: II. application to binding and pattern segmentation. *Biol. Cybern.*, 71:349, 1994.
[99] F. Rosenblatt. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychol. Rev.*, 65:386, 1958.
[100] D. E. Rumelhart, G. E. Hinton and R. J. Williams. Learning representations by back-propagating errors. *Nature*, 323:533, 1986.
[101] H. Sakaguchi, S. Shinomoto and Y. Kuramoto. Phase transitions and their bifurcation analysis in a large population of active rotators with mean-field coupling. *Progr. Theor. Phys.*, 79:600, 1988.
[102] T. B. Schillen and P. König. Stimulus-dependent assembly formation of oscillatory responses: II. desynchronization. *Neural Comput.*, 3:167, 1991.
[103] H. G. Schuster and P. Wagner. A model for neuronal oscillations in the visual cortex: 1. mean-field theory and derivation of the phase equations. *Biol. Cybern.*, 64:77, 1990.
[104] H. G. Schuster and P. Wagner. A model for neuronal oscillations in the visual cortex: 2. phase description of the feature dependent synchronization. *Biol. Cybern.*, 64:83, 1990.
[105] W. Singer. The brain, a complex self-organizing system. *Eur. Rev.*, 17:321, 2009.
[106] H. Sompolinsky, D. Golomb and D. Kleinfeld. Cooperative dynamics in visual processing. *Phys. Rev. A*, 43:6990, 1991.
[107] S. H. Strogatz and R. E. Mirollo. Phase-locking and critical phenomena in lattices of coupled nonlinear oscillators with random intrinsic frequencies. *Physica D*, 31:143, 1988.
[108] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. A Bradford Book, 1998.
[109] Y. Teng, F. R. Yu, Y. Wei, L. Wang and Y. Zhang. Behavior modeling for spectrum sharing in wireless cognitive networks. *Wirel. Netw.*, 18:929, 2012.
[110] V. N. Vapnik. *Statistical Learning Theory*. Wiley-Interscience, 1998.
[111] V. N. Vapnik and A. Ya. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theor. Probab. Appl.*, 17:264, 1971.
[112] T. L. Vincent. Chaotic control systems. *Nonlinear Dynam. Syst. Theor*, 1:205, 2001.

[113] C. Von der Malsburg. Self-organization of orientation sensitive cells in the striate cortex. *Kybernetik*, 14:85, 1973.
[114] C. Von der Malsburg. The correlation theory of brain function. Technical report, Max-Plank-Institute for Biophys. Chem., Göttingen, Germany, 1981.
[115] C. Von der Malsburg and J. Buhmann. Sensory segmentation with coupled neural oscillators. *Biol. Cybern.*, 67:233, 1992.
[116] D. L. Wang. Relaxation oscillators and networks. In J. G. Webster, editor, *Encyclopedia of Electrical Electronics Engineering*, volume 18, page 396. Wiley, 1999.
[117] D. L. Wang and D. Terman. Locally excitatory globally inhibitory oscillator networks. *IEEE T. Neural Networ.*, 6:283, 1995.
[118] D. L. Wang and D. Terman. Image segmentation based on oscillatory correlation. *Neural Comput.*, 9:805, 1997.
[119] D. J. Watts and S. H. Strogatz. Collective dynamics of "small-world" networks. *Nature*, 393:440, 1998.
[120] T. Wennekers and F. Pasemann. Generalized types of synchronization in networks of spiking neurons. *Neurocomputing*, 38-40:1037, 2001.
[121] P. J. Werbos. *Beyond regression: New tools for prediction and analysis in the behavioral sciences*. PhD thesis, Harvard University, Cambridge, MA, 1974.
[122] G. B. West, J. R. Brown and B. J. Enquist. The fourth dimension of life: Fractal geometry and allometric scaling of organisms. *Science*, 284:1677, 2000.
[123] S. Wolfram. *Cellular Automata and Complexity*. Westview Press, 1994.
[124] A. A. Zhdanov. *Autonomous artificial intellect (Adaptive and intellectual systems)*. Binom, Moscow, Russia, 2008. (in Russian).

2 Neural and oscillatory networks of associative memory

2.1 Neural networks of associative memory

2.1.1 Types of associative memory networks

Design of artificial neural network models with simple architectures of connections is a necessary step to the elucidation of basic principles of parallel information processing. It allows us to understand the capabilities and limitations of these network models and obtain well-established mathematical results.

Various neural network models of correlation-type associative memory have been developed since 1972. The first models were proposed in [12, 29, 46]. There is a number of different types of associative memory networks, among which are cross-correlation associative memory networks (or networks of heteroassociative memory), autocorrelation associative memory networks, sequence recalling associative memory networks, bilateral associative memory (BAM) networks and others. In this section, we present the most important results of a qualitative analysis of the main properties of associative memory networks: their memory storage capacity and dynamical behavior.

Networks of heteroassociative memory

Networks of heteroassociative memory can be designed based on multilayered perceptrons. The simplest type of an associative memory network is provided by one-layered perceptron. Let the set of M input–output vector pairs $(\mathbf{p}^1, \mathbf{q}^1), \ldots, (\mathbf{p}^M, \mathbf{q}^M)$ be given. The network is required to emit the outputs $\mathbf{q}^1, \ldots, \mathbf{q}^M$ provided $\mathbf{p}^1, \ldots, \mathbf{p}^M$ are given as the inputs. The problem is equivalent to the construction of the matrix $\hat{W}$ for one-layered perceptron, satisfying the condition

$$\hat{T}\mathbf{p}^m = \mathbf{q}^m, \quad m = 1, \ldots, M, \tag{2.1-1}$$

where

$$\hat{T}\mathbf{x} = G\left(\sum_{j=1}^{n} \hat{W}\mathbf{x} - \mathbf{h}\right), \tag{2.1-2}$$

and $\mathbf{h} = (h_1, \ldots, h_N)$ is the vector of thresholds, which we further put equal to zero, for simplicity.

Obviously, the matrix, defined in the form of the sum of outer products,

$$\hat{W} = \frac{1}{M}\sum_{m=1}^{M} \mathbf{q}^m (\mathbf{p}^m)^T, \quad W_{jk} = \frac{1}{M}\sum_{m=1}^{M} q_j^m p_k^m \tag{2.1-3}$$

provides the solution. The one-layered perceptron with the matrix of connections (2.1-3) can be interpreted as a network of heteroassociative memory that stores the "memory" vectors $\mathbf{q}^1, \ldots, \mathbf{q}^M$ and recalls them when vectors $\mathbf{p}^1, \ldots, \mathbf{p}^M$, associatively related to memory vectors $\mathbf{q}^1, \ldots, \mathbf{q}^M$, are provided as network inputs.

The associative memory problem for a multilayer perceptron (a cascade $N^{(1)}, \ldots, N^{(L)}$ of L feedforward networks) is formulated as a natural generalization of that for one-layer perceptron. Let $L + 1$ sequences $S^{(1)}, \ldots, S^{(L+1)}$ of network state vectors, each containing M vectors, be given:

$$S^{(k)} = \{\mathbf{p}^{1,(k)}, \ldots, \mathbf{p}^{M,(k)}\}, \quad k = 1, \ldots, L + 1. \tag{2.1-4}$$

The sequence of matrices $\hat{W}^{(l)}$, $l = 1, \ldots, L$, for the L-layer perceptron should be constructed so that the perceptron recalls the sequence $S^{(l)}$ by emitting $\mathbf{p}^{m,(l+1)}$ from the layer $N^{(l)}$, the $\mathbf{p}^{m,(1)}$ being the input of the first perceptron layer. That is, the following relations should be fulfilled:

$$\hat{T}^{(l)}\mathbf{p}^{m,(l)} = \mathbf{p}^{m,(l+1)}, \quad m = 1, \ldots, M, \quad l = 1, \ldots, L. \tag{2.1-5}$$

The sequence of matrices for perceptron layers

$$\hat{W}^{(l)} = \frac{1}{M}\sum_{m=1}^{M} \mathbf{p}^{m,(l+1)} \cdot (\mathbf{p}^{m,(l)})^T \tag{2.1-6}$$

provides the solution to the problem. It is natural to call this multilayered perceptron a network of cascade associative memory.

Networks of cyclic associative memory

If we introduce a feedback connection loop by closing the output of cascade associative memory network to its input, we obtain a network of cyclic associative memory. A network consisting of L layers connected into a ring is often called L-AM.

Let $S^{(1)}, \ldots, S^{(L)}$ be L sequences of vectors and let the connection matrix $\hat{W}^{(l)}$ for the layer $N^{(l)}$ of the ring be given by

$$\hat{W}^{(l)} = \frac{1}{M}\sum_{m=1}^{M} \mathbf{p}^{m,(l+1)}(\mathbf{p}^{m,(l)})^T. \tag{2.1-7}$$

Then relations

$$\mathbf{p}^{m,(l+1)} = \hat{T}^{(l)}\mathbf{p}^{m,(l)}, \quad m = 1, \ldots, M, \quad l = 1, \ldots, L, \quad L + 1 \equiv 1 \tag{2.1-8}$$

hold for the L-AM associative memory network. When $L = 2$, the 2-AM network is called a BAM [30]. As it turned out, the dynamical behavior of BAM is slightly different from that of L-AM networks with $L \geq 3$.

Networks of autoassociative memory

The one-layered recurrent network (of Hopfield type) can be regarded as a 1-AM. As we saw, its dynamics is governed by the dynamic equation

$$\mathbf{x}(t+1) = \hat{T}\mathbf{x}(t) \equiv G(\hat{W}\mathbf{x}(t) - \mathbf{h}), \tag{2.1-9}$$

where $\mathbf{x}(t)$ is the network state vector at time t.

If the vectors $\mathbf{p}^1, \ldots, \mathbf{p}^M$ are given, the recurrent network of associative memory is obtained if we put

$$\hat{W} = \frac{1}{M}\sum_{m=1}^{M} \mathbf{p}^m (\mathbf{p}^m)^T. \tag{2.1-10}$$

All the vectors $\mathbf{p}^1, \ldots, \mathbf{p}^M$ are the equilibria of network dynamics and satisfy the equation

$$\mathbf{p}^m = \hat{T}\mathbf{p}^m, \quad m = 1, \ldots, M. \tag{2.1-11}$$

The main question of interest for autoassociative memory networks is the estimation of the admissible number M of memory vectors that can be stored by an autoassociative memory network with N network neurons and then retrieved as stable equilibria with guaranteed attraction basins. It was analyzed by many researchers since the 1980s (see, for instance, [10, 35, 48]).

General analysis of dynamics of nonlinear, highly multicomponent and densely interconnected neural networks is a very challenging problem. Therefore, special asymptotical methods for elucidation of "typical" neural network dynamics were required to be derived. A statistical macrodynamics approach, developed by Amari and colleagues [6–8], provided one of the asymptotical methods. Key ideas of this approach are close to those used by Khinchin for the development of a statistical approach to the problems of classical statistical mechanics [27]. An important step in establishing the theoretical background of the statistical macrodynamics approach is also contained in the papers [50–52]. Rigorous mathematical foundations of statistical neurodynamics are given in [11]. We outline here the most essential results on the macrodynamical approach.

2.1.2 Random neural networks and the statistical macrodynamics approach

In order to elucidate a typical behavior of large-scale neural networks it is necessary to know some macroscopic characteristics of the network dynamics. To this end, macroscopic state equations derived from microscopic network dynamics are required. They can be obtained by the study of an ensemble of random networks – networks in which parameters (interconnection matrices elements and neuron thresholds) are randomly generated according to some probability law and then kept fixed. The presence of correlations among microscopic states introduces certain theoretical difficulties into the

derivation of macroscopic state equations. The situation is quite similar to that encountered in deriving the Boltzmann equation in statistical mechanics of gases. A kind of statistical postulate related to the restriction on possible correlations between neural network states (analogous to the assumption of "molecular chaos" [50–52]) must also be formulated for random neural networks in the macrodynamical consideration. It should be noted, however, that the correlation properties of ensembles of random neural networks are different from those of statistical physics.

The statistical macrodynamics approach is aimed at the analysis of ensembles of neural networks of randomly interconnected idealized neurons [6–8]. Let $\mathbf{x}(t) = [x_t(1), \ldots, x_N(t)]^T$ be the state vector of neural network of N neurons, and the dynamical system governing network microstate dynamics be written in the form

$$\mathbf{x}(t+1) = \hat{T}\mathbf{x}(t) \equiv G(\hat{W}\mathbf{x}(t) - \mathbf{h}) \tag{2.1-12}$$

Denote by $\boldsymbol{\omega}$ the set of network parameters: N^2 components of $N \times N$ matrix $\hat{W}$ and N components of threshold vector $\mathbf{h}$:

$$\boldsymbol{\omega} = \{W_{11}, \ldots, W_{NN}; h_1, \ldots, h_N\}. \tag{2.1-13}$$

Introduce the ensemble E_N of random networks where all W_{jk} and h_k are independently and identically distributed random variables with the probability distribution $P_N(\boldsymbol{\omega})$, statistical means $\overline{W}, \overline{h}$ and variances σ_W, σ_h. Then the dynamical system (2.1-12) can be rewritten as

$$\mathbf{x}(t+1) = \hat{T}_{\boldsymbol{\omega}}\mathbf{x}(t). \tag{2.1-14}$$

The goal is to determine the macroscopic variable $\mathbf{X}(\mathbf{x}) = [X_1(\mathbf{x}), \ldots, X_K(\mathbf{x})]^T$ that describes the macroscopic features of the microstate $\mathbf{x}$ and to find a macroscopic state transition equation in the form

$$\mathbf{X}(t+1) = \hat{F}(\mathbf{X}(t)). \tag{2.1-15}$$

The network mean activity level

$$X(\mathbf{x}) = \frac{1}{N}\sum_{j=1}^{N} x_j \tag{2.1-16}$$

is a typical example of a scalar macrovariable. Another typical example is the so-called overlap, which arises in the analysis of memory recall process in associative memory networks. The overlap, characterizing similarity of current state vector $\mathbf{x}$ and a network memory vector $\mathbf{p}^k$, is defined by the inner product

$$a^{(k)} = \frac{1}{N}(\mathbf{p}^k)^T\mathbf{x} = \frac{1}{N}\sum_{j=1}^{N} p_j^k x_j. \tag{2.1-17}$$

A detailed analysis of macrovariable properties and macrostate dynamics is given in [8, 11]. The following important results established should be mentioned:

(1) A dynamical variable $\mathbf{X}(\mathbf{x}) = (X_1(\mathbf{x}), \dots, X_K(\mathbf{x}))^T$ is a macrovariable if it satisfies the macroscopic equation (2.1-15) that holds in a weak sense, that is

$$\lim_{N\to\infty} E[|\mathbf{X}(\hat{T}_{\boldsymbol{\omega}}\mathbf{x}) - \hat{F}(\mathbf{X})|^2] = 0. \tag{2.1-18}$$

The weak convergence (2.1-18) is transformed (under appropriate exact conditions) into the strong convergence (uniform convergence in time) only in the case of macrovariables for the so-called sparsely connected networks (a special class of random networks).

(2) Let X_1 be a scalar macrovariable. Then there exist macrovariables $X_2, \dots, X_K$ such that $\mathbf{X}(\mathbf{x}) = (X_1(\mathbf{x}), \dots, X_K(\mathbf{x}))^T$ is the macrovariable satisfying a macrostate dynamical system

$$\mathbf{X}(t+1) = \hat{F}(\mathbf{X}(t)), \quad \text{or} \quad \dot{\mathbf{X}} = -\mathbf{X} + \hat{F}(\mathbf{X}). \tag{2.1-19}$$

It means that, as a rule, a scalar macrovariable is not a macrovariable by itself, but is contained as a component in some vector macrovariable.

The macrodynamical equation for scalar macrovariable X defined by (2.1-16) was obtained in [11]. This equation can be written as

$$X(t+1) = F(WX(t) - H), \quad \text{or} \quad \dot{X} = -X + F(WX - H), \tag{2.1-20}$$

where

$$F(u) = \frac{2}{\sqrt{2\pi}} \int_0^u e^{-t^2/2} dt = \operatorname{erf}\left(\frac{u}{\sqrt{2}}\right), \tag{2.1-21}$$

$$W = \frac{N\overline{W}}{\sigma}, \quad H = \frac{\overline{h}}{\sigma}, \quad \sigma = \sqrt{N\sigma_W^2 + \sigma_h^2}. \tag{2.1-22}$$

The structural portrait of the macrodynamical equation (2.1-20) was also obtained in [8]. In the parametrical plane (W, H) there exist three qualitatively different regimes: monostability regime *I*, when only single stable equilibrium exists; bistability regime *II*, when three equilibria exist (two stable and one unstable); the limit cycle regime *III*, when a limit cycle of period two is inherent to network dynamics. The bifurcation of "limit cycle birth" occurs when the curve that defines a change of network parameters crosses the boundary separating the regions *I* and *III*, whereas the bifurcation "the catastrophe gather" occurs when this curve crosses the boundary between regions *I* and *II*. Thus, the macrodynamical system provides full information about all the types of the network dynamics.

Another example of macrodynamical analysis can be demonstrated using the scalar variable of the Hamming distance between two network microstates

$$d = \frac{1}{2\pi} \sum_{j=1}^{N} |x_j - y_j|. \tag{2.1-23}$$

As was clarified in [8], the variable d is not a macrovariable by itself and should be included into some three-dimensional vector macrovariable $\mathbf{X} = (d, Z_1, Z_2)^T$. The macrodynamical system for $\mathbf{X}$ can be written in the form

$$\dot{\mathbf{X}} = -\mathbf{X} + \hat{\Phi}(\mathbf{X}; W, H, \alpha), \tag{2.1-24}$$

where

$$W = \frac{N\overline{W}}{\alpha}, \quad H = \frac{\overline{h}}{\alpha}, \quad \alpha = \left(1 + \frac{\sigma_h^2}{N\sigma_W^2}\right)^{-1}, \tag{2.1-25}$$

and the matrix-valued function $\hat{\Phi}$, defining the macrostate transition law, can be obtained via proper statistical averaging over random network parameters.

Only in the limit situation of $W = 0$, $H = 0$ the variable $X = d$ is the scalar macrovariable that satisfies the equation

$$\dot{X} = -X + f(X; \alpha), \quad f(X; \alpha) = \frac{2}{\pi} \arcsin \sqrt{\alpha X}. \tag{2.1-26}$$

This equation also admits simple structural analysis. The details can be found in [8].

Thus, the statistical neurodynamics approach is proved to be useful for the analysis of essential properties of associative memory networks. It has helped to elucidate the set of network dynamical regimes and has provided the estimation of the network storage capacity.

2.1.3 Macrodynamical approach for associative memory networks

Noise reduction by one-layer perceptron

The noise reduction property inherent to the simplest one-layered heteroassociative memory networks can be demonstrated based on the results obtained in [9].

Consider a one-layer perceptron with memory vectors $\mathbf{p}^1, \ldots, \mathbf{p}^M$, output vectors $\mathbf{q}^1, \ldots, \mathbf{q}^M$ and let us study the problem of recalling a vector $\mathbf{q}^1$ from a noisy version of $\mathbf{p}^1$. The similarity between an input $\mathbf{x}$ and $\mathbf{p}^1$ can be measured by the overlap

$$a = \frac{1}{N}(\mathbf{p}^1)^T\mathbf{x} = \frac{1}{N}\sum_{j=1}^{N} p_j^1 x_j. \tag{2.1-27}$$

We write the network dynamical system in the form

$$\mathbf{y}(t+1) = \operatorname{sgn}(\hat{W}\mathbf{x}) \tag{2.1-28}$$

(for simplicity we put $\mathbf{h} = 0$ and $G(\mathbf{x}) = \operatorname{sgn}(\mathbf{x})$), and calculate the overlap a^* of the output $\mathbf{y} = \hat{T}\mathbf{x}$ with one of the memory vectors, for instance, with $\mathbf{q} = \mathbf{q}^1$: $a^* = N^{-1}(\mathbf{q}^1)^T\mathbf{y}$.

The calculation of the jth component of the output $\mathbf{y}$ gives

$$y_j = \operatorname{sgn}\left(\frac{1}{N}\sum_{k=1}^{N} q_j^1 p_k^1 x_k\right) = q_j^1 \operatorname{sgn}\left(a + q_j^1 N_j\right), \tag{2.1-29}$$

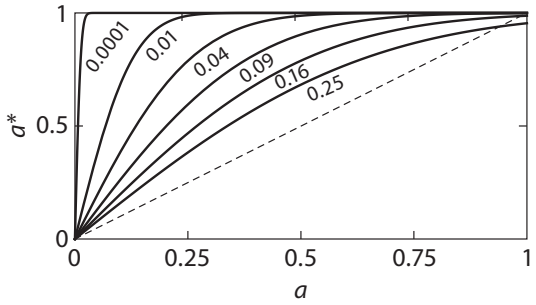

Fig. 2.1. Noise reduction by heteroassociative memory network: The curves $a^* = F(a/\sqrt{r})$ at different r.

where

$$N_j = \frac{1}{N} \sum_{m=1}^{M} \sum_{k=1}^{N} q_j^m p_k^m x_k. \tag{2.1-30}$$

The term N_j can be regarded as interference noise from superposed vector pairs in $\hat{W}$ other than $(\mathbf{p}^1, \mathbf{q}^1)$. The case $N_j = 0$ and $a > 0$ would correspond to the exact retrieval of $\mathbf{q}^1$, that is $\mathbf{y} = \mathbf{q}^1$. The term N_j is a sum of $N(M-1)$ random variables. According to the central limit theorem, at large N the random variables N_j are normally distributed with mean 0 and variance $\sigma^2 = (M-1)/N \approx M/N = r$, where r is the ratio of the number of memorized vectors to the number of neurons in the network (r is called the network storage capacity). Carrying out the calculation of the probability of $y_j = q_j^1$ using the law of large numbers, the following limit expression for a^* can be obtained [9]:

$$\lim_{N\to\infty} a^* = F\left(\frac{a}{\sqrt{r}}\right), \tag{2.1-31}$$

where $F(u)$ is defined by expression (2.1-21). The curves shown in Figure 2.1 demonstrate the noise reduction by one-layer heteroassociative memory for recalling of associative vector pairs.

Macrodynamical system for the retrieval process

We consider a network composed N formal neurons each taking two states ± 1, $\mathbf{x} = (x_1, \ldots, x_N)^T$ being a column vector of current network state, $\hat{W} = [W_{jk}]$ being an $N \times N$ matrix on network connections. We put vector of thresholds $\mathbf{h} = 0$ again and consider the network dynamical system in the form (2.1-28).

For a recurrent associative network of Hopfield-type M memory vectors $\{\mathbf{p}^m\}$, $m = 1, \ldots, M$, are constructed as equilibria of the network dynamics. It can be achieved by construction of the matrix $\hat{W}$ in the form of the sum of outer product over memory vectors,

$$\hat{W} = \frac{1}{M} \sum_{m=1}^{M} \mathbf{p}^m (\mathbf{p}^m)^T, \quad W_{jk} = \frac{1}{M} \sum_{m=1}^{M} p_j^m p_k^m. \tag{2.1-32}$$

In the case of recurrent heteroassociative memory network, arranged as a cascade of L layers with N neurons in each, the full set of memory vectors is

$$\{\mathbf{p}^{(l),m}\},\ m = 1,\dots,M,\ l = 1,\dots,L;\quad \mathbf{p}^{(l),m} = \left(p_1^{(l),m},\dots,p_N^{(l),m}\right)^T. \tag{2.1-33}$$

The memory construction can be achieved via defining the matrix of connections $\hat{W}^{(l)}$ of the layer l in the form

$$\hat{W}^{(l)} = \frac{1}{M}\sum_{m=1}^{M}\mathbf{p}^{(l+1),m}(\mathbf{p}^{(l),m})^T,\quad W_{jk}^{(l)} = \frac{1}{M}\sum_{m=1}^{M}p_j^{(l+1),m}p_k^{(l),m}. \tag{2.1-34}$$

If components of vectors $\mathbf{p}^m$ and $\mathbf{p}^{(l),m}$ are chosen as random identically distributed variables, taking values ± 1, the statistical neurodynamics method can be applied. The signal-to-noise analysis can be employed to obtain the following expression for the jth component of $\mathbf{x}(t+1)$:

$$x(t+1) = p_j^1\,\mathrm{sgn}(a(t) + p_j^1 N_j(t)), \tag{2.1-35}$$

where

$$N_j(t) = \frac{1}{N}\sum_{m=2}^{M}\sum_{k\neq j}^{N}p_j^m p_k^m x_k(t). \tag{2.1-36}$$

For $t = 0$ the interference noise term $N_j(t)$ is a sum of NM independent random variables $p_j^m p_k^m x_k(0)$ which take the values ± 1 with equal probability. Therefore, according to the central limit theorem $N_j(0)$ is the random variable obeying a Gaussian distribution with the mean zero and the variance a in the limit of large N. This is no longer the case for $N_j(t)$ at arbitrary t (since $x_k(t)$ depends on the set $\{\mathbf{p}^m\}$ through the dynamical system). The careful analysis, carried out in [10], shows that it is possible to consider $N_j(t)$ as random variables subject to Gaussian distribution with vanishing mean and time-dependent variance $\sigma^2(t)$. Moreover, as it turned out, the self-averaging property is inherent to the overlap macrovariable $a(t)$. As a result, the two-dimensional discrete time macrodynamical system was derived for the vector macrovariable with components X (the mean of overlap $a(t)$) and Y (the square of variance of interference noise $N_j(t)$):

$$\begin{aligned} &X = E[a(t)] \equiv \overline{a},\quad a(t) = \frac{1}{N}\mathbf{x}^T\mathbf{p}^1 = \frac{1}{N}\sum_{j=1}^{N}x_j(t)p_j^1,\\ &Y = E\left[\sum_{j=1}^{N}(N_j(t) - \overline{N_j})^2\right]. \end{aligned} \tag{2.1-37}$$

The macrodynamical system derived in [10] can be written in the form

$$\begin{aligned} &X(t+1) = F(u(t)),\quad u(t) = X(t)/\sqrt{Y(t)},\\ &Y(t+1) = r + 4f^2(u(t)) + 4ru(t)f(u(t))F(u(t)), \end{aligned} \tag{2.1-38}$$

where

$$f(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2}, \quad F(u) = \int_{-u}^{u} f(z)dz. \tag{2.1-39}$$

The extension of the statistical macrodynamical approach [10] to finite temperature dynamics was performed in [48].

2.1.4 Macrodynamical analysis of the memory retrieval process

In this section we present some results of macrodynamical analysis of the retrieval process for two different types of associative memory networks – recurrent autoassociative memory network and multilayered feedforward heteroassociative memory network. As we saw earlier, the macrodynamical approach allows us to describe the network dynamics asymptotically in the limit when the number of neurons N and the number of memory vectors M both tend to infinity under the condition that the ratio $r = M/N$ remains finite ($r < \infty$). The situation is similar to the so-called thermodynamical limit in the kinetic theory of gases, and the macrodynamical approach itself in fact is equivalent to accounting of correlations in the multicomponent system in frames of mean field approximation of statistical physics.

Continuous time analog of the macrodynamical system derived in [10] can be written as [35]

$$\begin{aligned} \dot{X} &= -X + F(u), \\ \dot{Y} &= -Y + r + f^2(u) + 2r\beta f(u)F(u), \end{aligned} \tag{2.1-40}$$

where the functions $f(u)$ and $F(u)$ are defined in (2.1-39), and the parameter β is equal to 1 in the case of a Hopfield network (a network of autoassociative memory) and to 0 in the case of a layered feedforward network (a network of heteroassociative memory).

Study of the phase portrait of the dynamical system (2.1-40) and its bifurcations at changing of the parameter r permits us to elucidate the features of the memory retrieval process and to find the critical value r_c of the network storage capacity, above which the memory retrieval is impossible. The phase portrait is defined by the number of fixed points of dynamics and their mutual locations in dependence of the parameter r. The analysis of fixed points of the dynamical system (2.1-40) can be reduced to studying the roots of the nonlinear transcendent equation [35]

$$R(u, r) \equiv F^2(u) - u^2\{4f^2(u) + r[1 + 4\beta u f(u)F(u)]\} = 0. \tag{2.1-41}$$

As a result, at $\beta = 1$ (Hopfield network) the phase portrait demonstrates the following behavior. At $r \in [0, 1]$ there exist exactly three fixed points of dynamics: stable nodes $\mathbf{z}^0 = (0, Y^0)$, $\mathbf{z}^* = (X^*, Y^*)$ and saddle point $\tilde{\mathbf{z}} = (\tilde{X}, \tilde{Y})$. The attraction basins of the nodes $\mathbf{z}^*$ and $\mathbf{z}^0$ are separated by the separatrix passing through the saddle $\tilde{\mathbf{z}}$. The node $\mathbf{z}^*$ corresponds to network capability of retrieving the memory vector $\mathbf{p}^1$ (since

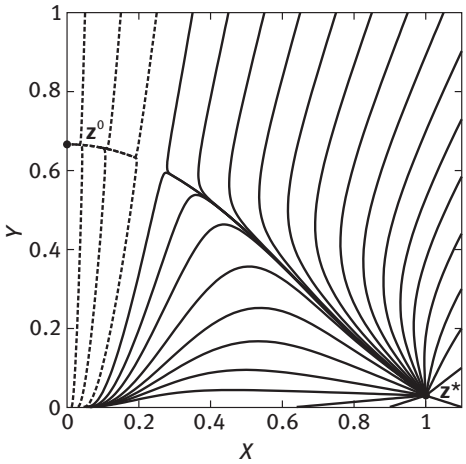

Fig. 2.2. Phase portrait of the dynamical system (2.1-40) at $r = 0.03$.

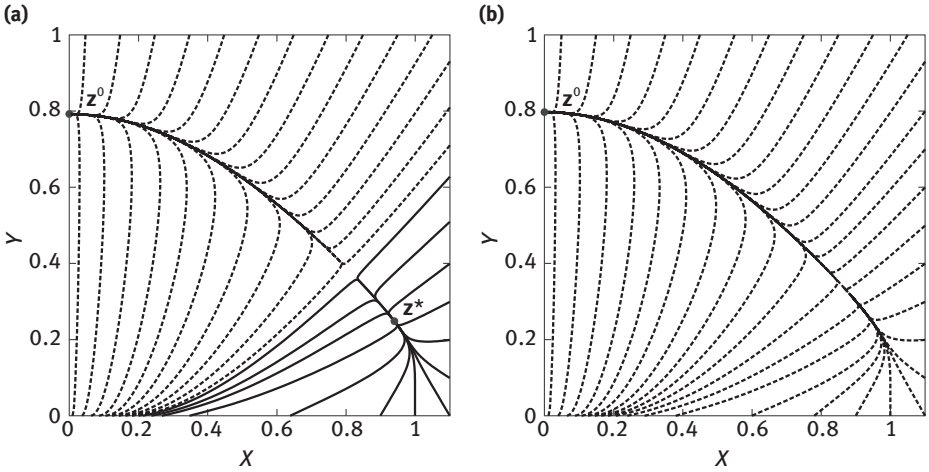

Fig. 2.3. Bifurcation of phase portrait of the dynamical system (2.1-40): (a) $r = 0.156$, (b) $r = 0.161$.

$0 < X^* < 1$), whereas the node $\mathbf{z}^0$ corresponds to the incapability of retrieving (since $X^0 = 0$).

At small r the node $\mathbf{z}^*$ is located in close vicinity of the point $\overline{\mathbf{z}} = (1, 0)$, and the attraction basin of the node $\mathbf{z}^0$ is very narrow (see Figure 2.2). When r grows, the basin size of the node $\mathbf{z}^*$ gradually decreases, and the node itself approaches to the basin boundary and to the saddle point $\tilde{\mathbf{z}}$. At $r = r_c$ the bifurcation saddle–node takes place: the point $\mathbf{z}^*$ merges with the approaching saddle point $\tilde{\mathbf{z}}$, and then both disappear (see Figure 2.3). The disappearance of the stable node $\mathbf{z}^*$ at $r > r_c$ means that the retrieval of memory vectors is impossible if $M > r_c N$.

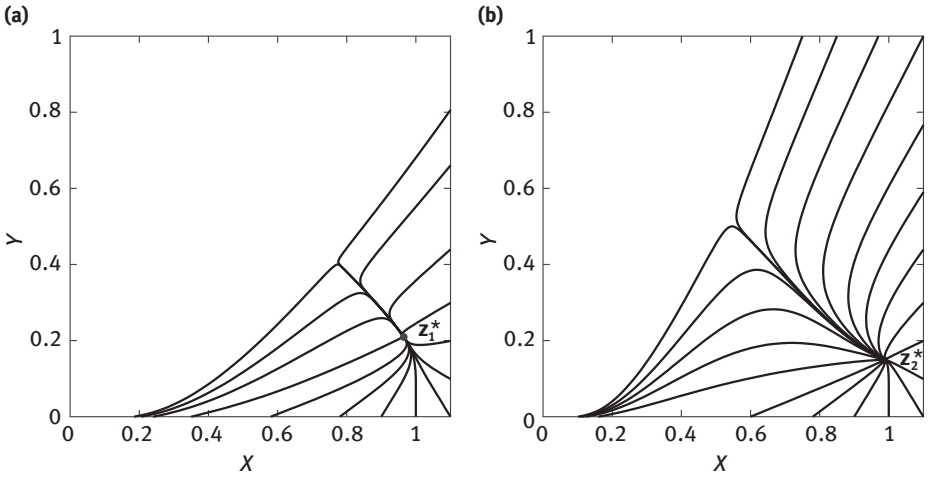

Fig. 2.4. The basins of attraction of stable nodes (a) $\mathbf{z}_1^*$ (at $\beta = 1$) and (b) $\mathbf{z}_2^*$ (at $\beta = 0$) at $r = 0.15$.

The behavior of the phase portrait of the dynamical system (2.1-40) at $\beta = 0$ is qualitatively similar to that at $\beta = 1$. The only difference is that now the saddle-node bifurcation takes place at larger value of r_c. For example, the comparison of attraction basins of the nodes $\mathbf{z}_1^*$ (at $\beta = 1$) and $\mathbf{z}_2^*$ (at $\beta = 0$) in the case of $r = 0.15$ is presented in Figure 2.4, where only the trajectories belonging the basin of care are shown.

The critical values r_c^1 and r_c^2 of the parameter r at the moment of bifurcation and the corresponding locations of the nodes $\mathbf{z}_1^*$ and $\mathbf{z}_2^*$ were obtained:

$$\begin{aligned} \beta = 1: \quad & r_c^1 = 0.16, \qquad \mathbf{z}_1^* = (X_1^*, Y_1^*) = (0.912, 0.286), \\ \beta = 0: \quad & r_c^2 = 0.269, \qquad \mathbf{z}_2^* = (X_2^*, Y_2^*) = (0.841, 0.357). \end{aligned} \tag{2.1-42}$$

The bifurcation diagrams for the dynamical system (2.1-40) at $\beta = 0$ and $\beta = 1$ are depicted in Figure 2.5. It should be noted that the bifurcation curve *2* and the bifurcation value r_c^2 are in excellent agreement with those obtained in [45], where the exact solution to the problem was found. This fact could be regarded as the evidence of applicability of the macrodynamical approach to the problem.

An example of time dependence of X and Y near bifurcation is shown in Figure 2.6. The feature of these curves $X(t)$ and $Y(t)$ is that at any fixed $Y(0)$ there exists a threshold $\overline{X}$ (dependent on $Y(0)$) so that $X(t) \to X^*$ at $t \to \infty$ for $X(0) \geq \overline{X}$ (that is, the retrieval is possible), whereas $X(t) \to 0$ at $t \to \infty$ for $X(0) < \overline{X}$ (the retrieval is impossible). This behavior is a direct consequence of the bounded size and the special form (triangle like) of the attraction basin of the stable node $\mathbf{z}^*$ in the vicinity of bifurcation. We recall that the existence of the critical value r_c for the network storage capacity is caused by network "overloading" – the appearance of a great number of extraneous ("spurious") fixed points of network dynamics, drastically growing as $N \to \infty$ and $M \to \infty$ [17].

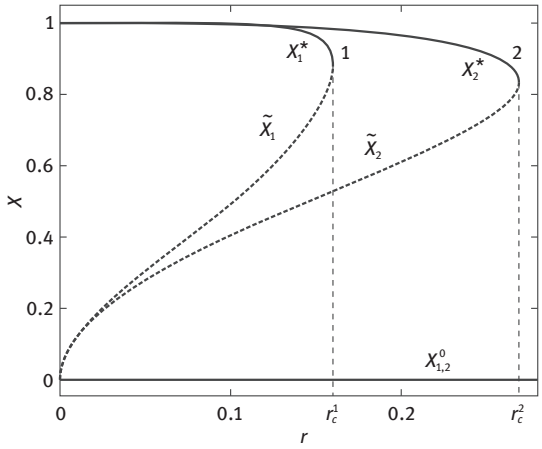

Fig. 2.5. The bifurcation diagram of the macrodynamical system (2.1-40): $\beta = 1$ (curve *1*) and $\beta = 0$ (curve *2*).

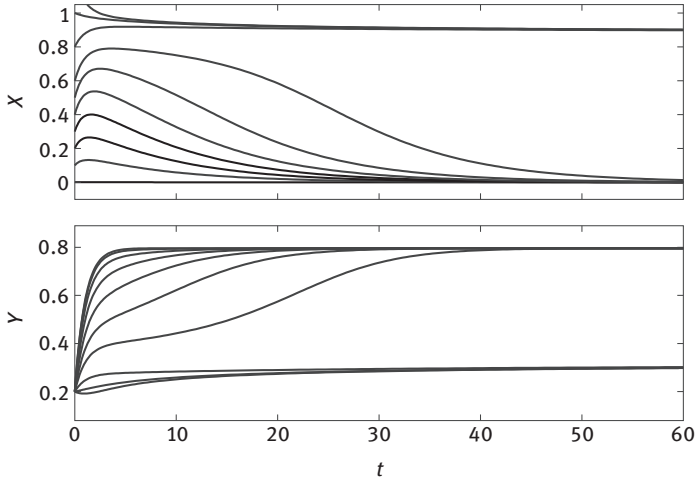

Fig. 2.6. The retrieval dynamics of autoassociative memory network in terms of $X(t)$ and $Y(t)$ near bifurcation ($r < r_c^{(1)}, r_c^{(1)} - r \ll 0.1$).

Concluding the macrodynamical analysis of associative memory networks, it is important to stress on the essential role of the dynamical equation for the variance Y (the second equation in the system (2.1-40)). Indeed, the macrodynamical system used in [28],

$$\dot{X} = -X + F\left(\frac{X}{\sqrt{\alpha}}\right), \quad Y(t) \equiv \alpha \tag{2.1-43}$$

provides more rough approximation for critical storage capacity: $r_c^1 = 2/\pi$.

2.2 Oscillatory networks of associative memory

2.2.1 Problem of the design of oscillatory associative memory networks

As discussed in Section 2.1, a low-memory storage capacity is inherent to associative memory networks of formal neurons. The attempts to design artificial neural-like networks of distributed associative memory consisting of other types of network processors seem natural. The network processing units, alternative to two-state formal neurons, were considered since 1980s. These were so-called phasors – the units in which state is defined by a continuous variable ranging over the unit circle. Recurrent associative memory phasor networks of both continuous-variable phasors and phasors with discrete q-state variable were designed and studied [19, 49]. The performance of associative memory phasor networks with continuous- and discrete-state phasor variable and sparse asymmetric network interconnection architecture admitted analytical analysis [49]. For fully connected phasor networks phase diagrams and estimations of network storage capacity were obtained in the frames of mean-field theory using a replica method [19].

The networks of continuous-variable phasors (so-called clock neural networks) are closely related to ensembles of limit cycle oscillators, arising in a wide variety of research areas, and being used to analyze different types of collective dynamics of various multicomponent systems. So, it seems reasonable to use oscillators (processors capable to perform stable undamped oscillations) as processing units for neural-like networks of the associative memory.

It should also be noted that the phasor networks are related to complex-valued neural networks (CVNN), which have been extensively developed over the last decade and have applications in various research areas. In the CVNN the artificial neurons with complex-valued states and complex-valued activation function are used as processing units, and interconnection matrices with complex-valued connection weights specify the neuron connections (a short outline is given in Section 2.2.4).

In the design of oscillatory neural-like networks of associative memory simple nonlinear limit cycle oscillators demonstrating sustained harmonic auto-oscillations are used as network processing units. Single oscillator dynamics is governed by a two-dimensional dynamical system that is traditionally written in the form of a single equation for the complex-valued variable $z = x + iy = re^{i\theta}$:

$$\dot{z} = (\rho^2 - |z|^2 + i\omega)z. \tag{2.2-1}$$

The stable limit cycle of the system (2.2-1) is the circle of radius ρ ($|z| = \rho$), and ω is the natural oscillator frequency on the limit cycle. The system (2.2-1) is the normal form of a dynamical system undergoing a supercritical Hopf bifurcation (the bifurcation of limit cycle birth or disappearance). The phase trajectory corresponding to the limit cycle and the curves of temporal behavior of the variables x and y are shown in Figure 2.7.

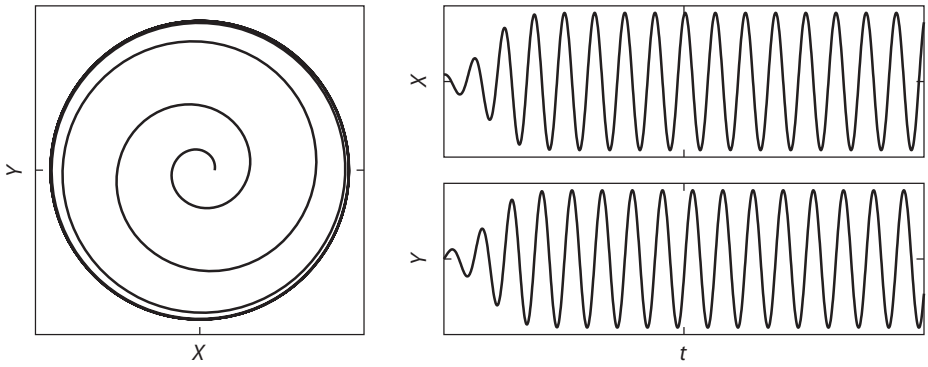

Fig. 2.7. A Ginzburg–Landau limit cycle oscillator.

The oscillator with dynamics (2.2-1) is known as a Ginzburg–Landau oscillator. In variables r, θ it reduces to the system of following two independent equations:

$$\begin{aligned} \dot{r} &= r(\rho^2 - r^2), \\ \dot{\theta} &= \omega. \end{aligned} \tag{2.2-2}$$

For modeling of recurrent associative memory networks composed of oscillators, we use a system of N limit cycle oscillators (2.2-1) with $\rho = 1$ and symmetric nonhomogeneous linear all-to-all oscillator coupling. The dynamical system of the oscillatory network can be written as

$$\dot{z}_j = \left(1 + i\omega_j - |z_j|^2\right) z_j + \kappa \sum_{k=1}^{N} W_{jk}(z_k - z_j). \tag{2.2-3}$$

Here $z_j = r_j e^{i\theta_j}$ is the state of the jth oscillator and ω_j is its natural frequency. The elements of the Hermitian matrix $\hat{W} = [W_{jk}]$ specify the coupling of jth and kth oscillators. This matrix satisfies the natural restrictions

$$|W_{jk}| \le 1 \quad \text{and} \quad \sum_{k=1}^{N} |W_{jk}| \le 1. \tag{2.2-4}$$

The real-valued parameter κ characterizes the absolute value and the sign of interaction strength in the system.

The system (2.2-3) can be written in the vector form

$$\dot{\mathbf{z}} = (\hat{D}(\mathbf{z}) + \kappa \hat{W})\mathbf{z}, \tag{2.2-5}$$

where $\mathbf{z} = (z_1, \ldots, z_N)^T$ is the state vector of the oscillatory system and $\hat{D}(\mathbf{z})$ is the diagonal matrix

$$\hat{D}(\mathbf{z}) = \operatorname{diag}(D_1(\mathbf{z}), \ldots, D_N(\mathbf{z})), \quad D_j(\mathbf{z}) = 1 + i\omega_j - |z_j|^2. \tag{2.2-6}$$

For arbitrary set of parameters $\{\omega_j\}$, κ, and $\hat{W}$ the system (2.2-3) demonstrates a great variety of complicated dynamical regimes, including synchronization and dynamical chaos (see Section 2.2.2). It is the synchronization regime that is used in associative memory modeling. If the condition $\sum_{j=1}^{N} \omega_j = 0$ is fulfilled (this restriction on the frequencies can be always satisfied by a proper rescaling of the dynamical system), the synchronization process in fact represents network relaxation into a stable equilibrium state. The problem of design of an oscillatory associative memory network can be formulated as a combination of two (independent) subproblems for governing dynamical system: (1) an inverse problem for the dynamical system: the design of a dynamical system of type (2.2-3) possessing the prescribed set of stable attractors with large enough attraction basins; (2) a control problem for the designed dynamical system: the selection of dynamical system parameters providing any prescribed subset of the system attractors from the full admissible attractor set. Here we mainly concentrate on the first subproblem.

Let $\{\mathbf{U}^m\}$, $m = 1, \ldots, M$ be a given set of state vectors of the oscillatory dynamical system (2.2-3). Our aim is to clarify the following questions:

(1) under what conditions parameters $\{\omega_j\}$, κ, and $\hat{W}$ the set $\{\mathbf{U}^m\}$ can belong to the set of stable attractors of the dynamical system (2.2-3);
(2) given N, what maximal number M of stable attractors from the set $\{\mathbf{U}^m\}$ can simultaneously exist (what the memory storage capacity of associative memory oscillatory network is);
(3) whether the dynamical system possesses another stable attractors of dynamics different from $\{\mathbf{U}^m\}$ ("extraneous memory" of the network).

Obviously, the study of general properties of the dynamical system (2.2-3) is required for the elucidation of the posed questions. The dynamical system with homogeneous all-to all coupling (when $W_{jk} = (1 - \delta_{jk})/N$) was extensively studied both analytically and via computer simulations. The most essential results are presented in the following section.

2.2.2 Systems of globally homogeneously coupled limit cycle oscillators

Types of collective dynamics of large-scaled systems of limit cycle oscillators

Here we consider systems of limit cycle oscillators with homogeneous all-to-all coupling and random frequency distribution governed by the equations

$$\dot{z}_j = \left(1 - |z_j|^2 + i\omega_j\right) z_j + \frac{\kappa}{N} \sum_{k=1}^{N} (z_k - z_j), \quad j = 1, \ldots, N. \tag{2.2-7}$$

Let the oscillator frequency distribution $\{\omega_j\}$ be characterized by a frequency distribution function $g(\omega)$ with the effective width Δ. The oscillatory system demonstrates a rich variety of dynamical regimes for various interaction strengths κ and widths Δ. A

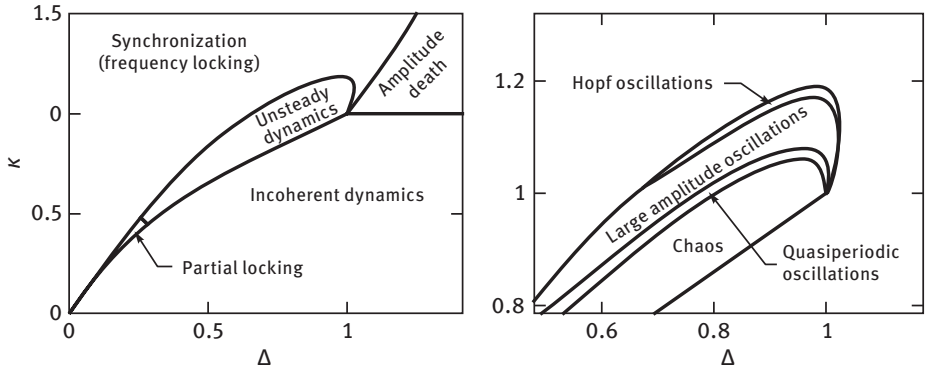

Fig. 2.8. The structural portrait of the dynamical system (2.2-7) in the parametrical plane (Δ, κ) [43].

kind of macrovariable is usually introduced to study collective dynamics. In the case of large-scale oscillatory system the macrovariable Z (known as an "order parameter," or an order function for oscillatory systems)

$$Z = \frac{1}{N}\sum_{j=1}^{N} z_j = Re^{i\phi} \tag{2.2-8}$$

turns out to be adequate. Different types of collective dynamics inherent to the dynamical system (2.2-7) were studied based on the analysis of a self-consistent equation for Z and computer modeling [43, 44]. The complete information about all types of dynamics is contained in the so-called structural portrait of dynamical system – the decomposition of parametrical space of the system into the set of domains, corresponding to qualitatively different dynamical regimes. The structural portrait of the dynamical system (2.2-7) in the parametrical plane (Δ, κ), was obtained in [43, 44] and is presented in Figure 2.8.

We qualitatively describe the regimes of collective dynamics in each parametrical domain. On the one hand, it should be noted that the variable Z defined in (2.2-8) can be viewed as a centroid coordinate of an oscillator "cloud" in the complex plane z. On the other hand, Z defines a measure of coherence degree of oscillatory dynamics. Both stationary and nonstationary regimes of collective dynamics are possible, manifesting themselves in different Z behavior. However, in the case of stationary dynamical regimes, the variable $R = |Z|$ either relaxes to some constant, or undergoes random fluctuations of order $1/\sqrt{N}$ about zero. However in the case of nonstationary states, Z oscillates. There are the following domains in the plane (Δ, κ) corresponding to different types of dynamics:

(1) The domain of incoherent collective dynamics.
The domain is located in the region $\kappa < C\Delta$ (C is a constant, $0 < C < 1$). In this domain the interaction is too small to synchronize the system. As a result, all oscillators perform independent oscillations with slightly perturbed amplitudes

and frequencies. At finite N the variable R undergoes statistical fluctuations of order $1/\sqrt{N}$ about zero.

(2) The domain of synchronization (frequency locking).
The domain is located in the region $\kappa \geq C\Delta$. The synchronization to a common frequency $\overline{\omega}$ emerges spontaneously. The variable R relaxes to a constant. Oscillations of each oscillator correspond to a limit cycle of radius $\rho = \sqrt{1-\kappa}$. The system is strongly ordered geometrically over the oscillation amplitudes. Oscillators form a stationary arc in the complex plane z, in which oscillators with the most extreme frequencies possess the lowest amplitudes.

(3) The domain of oscillation suppression ("amplitude death").
The domain is located in the region $\kappa \geq 1 + C(\Delta - 1)$. In this case oscillations are impossible. Oscillators pull each other off their limit cycles and collapse into the origin as $t \to \infty$. The origin $z_j = 0,\ j = 1,\ldots,N$ is a stable fixed point of the dynamical system (2.2-7). The function $R(t)$ approaches a constant value as $t \to \infty$. Since the amplitudes of all oscillators are zero, this state is often referred to as "amplitude death".

(4) The domain of unsteady collective dynamics.
This domain is located in the vicinity $\kappa = \Delta$ and is comparatively narrow. It contains a collection of different regimes (see [43, 44] for details). First, there exists a thin subdomain, located in the vicinity of the line $\kappa = \Delta$, corresponding to a partial synchronization. In this state the oscillator system splits into two subsystems with qualitatively different dynamics. The first subpopulation that contains oscillators having frequencies with absolute value less than some critical value is in a synchronization state. The second subsystem, formed by drifting oscillations, is characterized by oscillating $R(t)$. Both the states of partial synchronization and the incoherent state belong to statistically steady states. A thin domain bordering the synchronization domain corresponds to Hopf oscillations – small quasisinusoidal oscillations about the synchronization state. This nonstationary dynamical regime arises as a result of stability loss by synchronization state. The corresponding $R(t)$ is an asymptotically oscillating function with a small amplitude.
A wider domain adjacent to the domain of Hopf oscillations is a domain of so-called large amplitude oscillations. This regime is qualitatively different from Hopf oscillations. These are slow nonsinusoidal oscillations of a partially synchronized state. The frequency of large oscillations is typically lower than the width Δ of oscillator frequency distribution. The oscillations arise from the synchronization state as a result of a saddle-node bifurcation. In this case the function $R(t)$ undergoes slow nonsinusoidal oscillations of a large amplitude.
At last, there exists a regime of dynamical chaos, in which oscillators evolve in an irregular manner in the complex plane, and $R(t)$ undergoes chaotic oscillations. The "diagnostics" of a dynamic chaos state demands the estimation of Lyapunov exponents, which characterize an exponential divergence of dynamical system trajectories, and also the estimation of the spectrum of $R(t)$. The chaotic state

is characterized by "power spectrum" of $R(t)$, containing all the frequencies and demonstrating a typical decreasing at frequency growing.

A self-consistent method of the mean-field theory was used in [43] for calculation of all region boundaries in the plane (Δ, κ) and for stability analysis of all dynamical regimes. At small κ and $\Delta < 2/3$ there exist at least three formally permitted synchronization regimes, but only one of them is stable.

Analysis of critical phenomena for oscillatory systems in the frames of phase approximation of dynamics

The system of equations of phase approximation can be derived for small Δ and κ via a simple transformation of the of oscillator dynamical system (2.2-7). The phase system can be written in the form

$$\frac{d\theta_j}{d\tilde{t}} = \tilde{\omega}_j + \frac{\tilde{\kappa}}{N}\sum_{k=1}^{N}\sin(\theta_k - \theta_j), \tag{2.2-9}$$

where

$$\theta_j = \arg(z_j), \quad \tilde{\kappa} = \frac{\kappa}{\Delta}, \quad \tilde{\omega}_j = \frac{\omega_j}{\Delta}, \quad \tilde{t} = t\Delta. \tag{2.2-10}$$

The phase approximation is valid when all oscillator limit cycles are only slightly perturbed by the oscillator interaction. Hence, this phase approximation, which is attractive due to the simplicity of the system (2.2-9), is admissible when frequency distribution is narrow and oscillator interaction is sufficiently weak. In the frames of phase approximation a number of interesting analytical results, confirmed further by computer experiments, were obtained. These results mainly concern collective dynamics of oscillator systems with all-to-all homogeneous coupling. Sometimes it is convenient to rewrite the phase oscillatory dynamical system in more general form

$$\frac{d\theta_j}{d\tilde{t}} = \tilde{\omega}_j + \frac{\tilde{\kappa}}{N}\sum_{k=1}^{N}h(\theta_k - \theta_j). \tag{2.2-11}$$

where $h(\theta)$ is a periodic coupling function [20]. Then, besides the order function Z defined by (2.2-8), a more general order function H, related to the system of equations (2.2-11), can be used to study the synchronization as a type of the phase transition [20–22, 31, 32, 55]. Statistical macrodynamics approach for large-scale oscillator systems was developed in [33].

In terms of the order function $Z = Re^{i\psi}$ the dynamical system (2.2-9) can be presented in the form of the system of the uncoupled equations

$$\dot{\theta}_j = \omega_j + \kappa R\sum_{k=1}^{N}\sin(\theta_k - \psi). \tag{2.2-12}$$

The function $Z(t)$ can be interpreted as the mean field of the oscillatory ensemble. The appearance of the mean field can be described as self-consistent process that can

be studied via deriving and analysis of the evolution equation for Z. To derive the self-consistency equation, it is convenient to present the function Z in the form

$$Z(t) = \int_0^{2\pi} g(\psi, t) e^{i\psi} d\psi, \tag{2.2-13}$$

where $g(\psi, t)$ is the density function for oscillators with phase ψ at time t. The following self-consistency equation corresponding to dynamical system (2.2-12) can then be derived:

$$Z = 2 \int_0^1 \kappa Z g(\kappa |Z| x) \sqrt{1 - x^2} dx. \tag{2.2-14}$$

For the frequency distribution given by a Lorentzian,

$$g(\omega) = \frac{\gamma}{\pi} \frac{1}{(\omega - \omega_0)^2 + \gamma^2}, \tag{2.2-15}$$

the analytical expression for the function $R = R(\eta)$, $\eta = 2\gamma/\kappa$ was obtained [33]. The behavior of the function R in a small vicinity of the synchronization threshold κ_c characterizes the type of phase transition into the synchronization state. As it turned out, this phase transition is of the second type, demonstrating the behavior $R = \mathcal{O}(\sqrt{\kappa - \kappa_c})$ at small $\kappa - \kappa_c$. It was further named as Kuramoto phase transition.

In a more general situation, the order function H corresponding to the dynamical system (2.2-11) satisfies the functional equation dependent on κ, $g(\omega)$ and $h(\theta)$. In the following, the synchronization threshold (in the subcritical regime) the oscillatory system is in the incoherent state, and the order function is identically zero. In the supercritical regime, at $\kappa \geq \kappa_c$, the order function is essentially nonzero. The onset of synchronization can be viewed as a bifurcation of appearance of nontrivial order function in functional space. Computer simulations demonstrated excellent agreement with analytical results, obtained for function H in the vicinity of synchronization using the standard methods of bifurcation theory [20].

Another class of oscillatory systems studied in computer experiments was locally coupled spatially distributed oscillatory systems, localized in d-dimensional lattices ($d = 1, 2$) [53, 54, 56]. The main attention was focused on the cluster formation in the process of phase transition into the synchronization state (the cluster formation was studied in the frames of phase approximation of oscillator dynamics and the mean field theory). An interesting analytical result was obtained for one-dimensional oscillator chains: the critical value κ_c defining the synchronization onset behaves as $\kappa_c = \mathcal{O}(\sqrt{N})$ at large N [56]. Hence, a linear oscillator chain with any fixed coupling strength κ cannot be synchronized when the number of oscillators in the chain increases. For oscillator systems localized in lattices with $d \geq 2$ clusters of synchronized oscillators in an intermediate state of partial synchronization were demonstrated to possess the fractal structure.

Important results were also obtained for large-scale oscillator systems with homogeneous and nonhomogeneous global coupling in the presence of external noise. It required the consideration of infinite systems of stochastic differential equations coupled via mean-field interaction. The approach was based on the analysis of stable time-dependent solutions of a nonlinear Fokker–Planck equation [13–16]. It allowed us to give an analytical self-consistent description of the phase transition into the synchronization state. The phase transition was studied as a Hopf bifurcation for time-periodic probability density satisfying the Fokker–Planck equation. It was shown that the phase-transition-type depends on the shape of the frequency distribution function $g(\omega)$: both "soft" (subcritical) and "hard" (supercritical) bifurcation are possible. Oscillatory systems with a special type of random inhomogeneous oscillator coupling (the van Hemmen coupling) were also studied. In this case it was necessary to introduce three order functions instead of the single function Z. This allowed us to obtain a more detailed phase diagram for the oscillatory system. The phase diagram contains parametric domains corresponding to the "paramagnetic" state (the incoherent state of oscillatory system), "ferromagnetic" state (synchronized state), "spin glass" state (global synchronization is absent, but oscillators are synchronized with random disorder) and "mixed" state (a state of partial synchronization). The bifurcation analysis (transitions from the incoherent phase to the others) was performed. The results were confirmed by experiments on Brownian simulation [16].

2.2.3 Associative memory networks of limit-cycle oscillators

Oscillatory dynamical system, phase approximation of dynamics, and eigen-basis of coupling matrix.

In the regime of synchronization oscillatory systems have the capability to store information, just in the same manner as recurrent neural networks, and the dynamics of the oscillatory system has a resemblance with relaxation dynamics of recurrent neural networks into stable equilibrium states.

We consider oscillatory networks of limit cycle oscillators with dynamics governed by the dynamical system (2.2-3). It can be rewritten in the polar form

$$\begin{aligned} \dot{r}_j &= \left(1 - r_j^2 - \kappa a_j\right) r_j + \kappa \sum_{k=1}^{N} W_{jk} r_k \cos(\theta_k - \theta_j + \beta_{jk}), \\ \dot{\theta}_j &= \omega_j - \kappa b_j + \frac{\kappa}{r_j} \sum_{k=1}^{N} W_{jk} r_k \sin(\theta_k - \theta_j + \beta_{jk}), \end{aligned} \tag{2.2-16}$$

where

$$\begin{aligned} W_{jk} &= |W_{jk}| e^{i\beta_{jk}} = \mathrm{W}_{jk} e^{i\beta_{jk}}, \\ a_j &= \sum_{k=1}^{N} \mathrm{W}_{jk} \cos\beta_{jk}, \quad b_j = \sum_{k=1}^{N} \mathrm{W}_{jk} \sin\beta_{jk}. \end{aligned} \tag{2.2-17}$$

In the general case we have the system of coupled equations for r_j and θ_j, but in the special situation, when all r_j are close to the limit cycle $|r| = 1$, it is possible to separate the independent pure phase system from the complete dynamical system (2.2-16) (as a phase approximation of dynamics). Given the frequency distribution $\{\omega_j\}$ and the interaction strength κ, the parameter $\gamma = \Omega/\kappa$ with $\Omega = \max_j|\omega_j|$ can be introduced as the essential parameter of the oscillatory dynamics (the parameter Ω can be considered as a measure of the frequency distribution width). As the analysis shows, phase approximation is admissible for sufficiently small Ω and sufficiently small κ. Then, one can obtain the phase approximation of oscillatory dynamics (2.2-16) in the form

$$\dot{\theta}_j = \omega_j + \kappa \sum_{k=1}^{N} \mathrm{W}_{jk} r_k \sin(\theta_k - \theta_j + \beta_{jk}). \tag{2.2-18}$$

Initial attempts to design an oscillatory associative memory network were based on the dynamical system (2.2-18). It was clarified later that the exact dynamical system (2.2-1) is much more suitable for this task. Nevertheless, information on the properties of the phase system (2.2-18) was proved to be useful and instructive for elucidation of the behavior of attractors of full dynamical system (2.2-1) [36]. To present here some simple results on attractors of phase system (2.2-18), we introduce the eigenbasis of the matrix $\hat{W}$.

Let λ^m, $m = 1, \ldots, M$ ($M = \operatorname{rank}\hat{W}$, $M \le N$) be real nonzero eigenvalues of the Hermitian matrix $\hat{W}$ and $\mathbf{U}^m = (U_1^m, \ldots, U_N^m)^T$ be the corresponding mutually orthogonal N-dimensional eigenvectors vectors with complex-valued components, and $\mathbf{U}^\dagger$ be the Hermitian conjugate to $\mathbf{U}$. The one has

$$\hat{W}\mathbf{U}^m = \lambda^m \mathbf{U}^m, \quad (\mathbf{U}^k)^\dagger \mathbf{U}^m = \delta_{km}, \quad m = 1, \ldots, M. \tag{2.2-19}$$

Let the matrix $\hat{W}$ be presented in the form of the outer-product expansion

$$\hat{W} = \sum_{m=1}^{M} \lambda^m \mathbf{U}^m (\mathbf{U}^m)^\dagger. \tag{2.2-20}$$

Obviously, the expansion (2.2-20) is nothing but a representation of the linear operator $\hat{W}$ in the form of the sum of one-dimensional orthogonal projective operators $\mathbf{U}^m(\mathbf{U}^m)^\dagger$ onto the subspaces corresponding to the eigenvalues λ^m. The vector space $\mathcal{X}$ of state vectors $\mathbf{z} = (z_1, \ldots, z_N)^T$ of full amplitude-phase dynamical system can be expanded into the set sum

$$\mathcal{X} = \mathcal{X}^1 \cup \mathcal{X}^2, \tag{2.2-21}$$

where

$$\mathcal{X}^1 = \operatorname{im}\mathcal{X} = \{\mathbf{x} \,|\, \mathbf{x} = \hat{W}\mathbf{z}\}, \quad \mathcal{X}^2 = \ker\mathcal{X} = \{\mathbf{x} \,|\, \hat{W}\mathbf{x} = 0\}. \tag{2.2-22}$$

It is clear that $\mathcal{X} = \mathcal{X}^1$ when $M = \operatorname{rank}\hat{W} = N$. In the general case ($M < N$) any $\mathbf{z} \in \mathcal{X}^1$ can be presented by its expansion in the orthogonal basis $\{\mathbf{U}^m\}$:

$$\mathbf{z} = \sum_{m=1}^{M} Z^m \mathbf{U}^m, \quad Z^m = (\mathbf{U}^m)^\dagger \mathbf{z}. \tag{2.2-23}$$

It is easy to see that the state vectors of the phase approximation (2.2-18) belongs to the hypersphere S of radius $\sqrt{N}$, $S = \{\mathbf{z} \,|\, \mathbf{z} \in \mathcal{X}, \|\mathbf{z}\| = \sqrt{N}\}$. It is also convenient to introduce the set of vectors $\{\mathbf{V}^m\}$ related to $\{\mathbf{U}^m\}$ by

$$\mathbf{V}^m = \sqrt{N}\mathbf{U}^m, \quad V_j^m = e^{i\beta_j^m}, \quad \|\mathbf{V}^m\| = N. \tag{2.2-24}$$

Then, instead of (2.2-20) we obtain a more convenient decomposition

$$\hat{W} = \frac{1}{N}\sum_{m=1}^{M} \lambda^m \mathbf{V}^m (\mathbf{V}^m)^\dagger. \tag{2.2-25}$$

Now one can recognize the variables Z^m as the "overlaps" that were used for asymptotical analysis of the retrieval process in the case of neural networks of associative memory (see Section 2.1). When the oscillator number N is large, the variables Z^m are macrovariables. They can be considered as "order functions" governing the phase transition into the synchronization state of oscillatory networks with inhomogeneous all-to-all coupling.

In terms of the macrovariables Z^m, the phase system (2.2-18) can be rewritten in the form of N uncoupled equations

$$\dot{\theta}_j = \omega_j + \kappa \sum_{k=1}^{N} \lambda^m R^m \sin\left(\psi^m + \beta_j^m - \theta_j\right). \tag{2.2-26}$$

The advantage of representation of phase system (2.2-18) in the form (2.2-26) is that it provides the description of the phase dynamical system in terms of acting "mean field." In some simple situations it allows us to derive the self-consistency equations in a closed form and to clarify completely the question on the number of stable attractors of an oscillatory network and their bifurcations under network parameter variations [42]. The coupling matrix of homogeneously coupled oscillatory network can be defined as $\hat{W} = \mathbf{U}\mathbf{U}^\dagger$, $\mathbf{U} = N^{-1/2}(1, \ldots, 1)^T$, and the unique, already known to us, order function

$$Z = \frac{1}{N}\sum_{j=1}^{N} e^{i\theta_j} \tag{2.2-27}$$

exists in the case. The following self-consistency equation for $R = \operatorname{Re} Z$ can then be obtained:

$$\gamma = u\sum_{k=1}^{N} \sqrt{1 - \gamma_j^2 u^2}, \quad \gamma = \frac{\Omega}{\kappa}, \quad u = \frac{\gamma}{R}, \quad \gamma_j = \frac{\omega_j}{\Omega}. \tag{2.2-28}$$

Based on the self-consistency equation (2.2-28), the behavior of stable attractors of network dynamics under variation of the parameter γ for the network composed of three homogeneously all-to-all coupled oscillators was elucidated [42] (see Figure 2.9) with the following results. The value $\gamma = \overline{\gamma} = 0.588$ corresponds to synchronization threshold: the network is not synchronized at $\gamma > \overline{\gamma}$ (small interaction strength κ); synchronization arises at $\gamma = \overline{\gamma}$, and a pair node–saddle equilibria of the dynamical

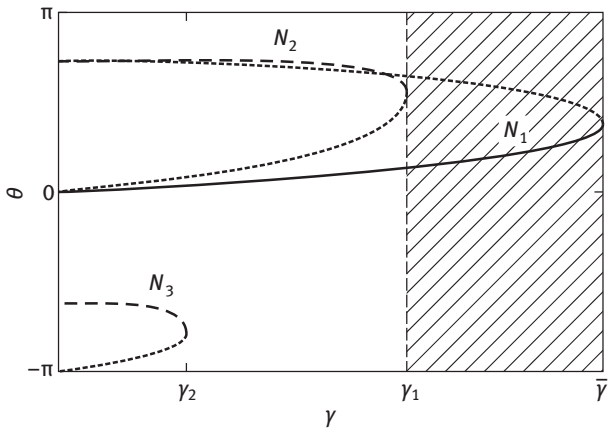

Fig. 2.9. Stable attractors of phase dynamical system in dependence on parameter $\gamma = \Omega/\kappa$.

system (2.2-26) appears. The single stable attractor (the stable node N_1) exists in the interval $\gamma_1 < \gamma \le \overline{\gamma}$ (with $0 < \gamma_1 < \overline{\gamma}$). Further, the second node–saddle pair appears at $\gamma = \gamma_2$ (with $0 < \gamma_2 < \gamma_1$), and the two stable attractors (nodes N_1 and N_2) coexist at $\gamma_2 < \gamma \le \gamma_1$. The third node–saddle pair appears at $\gamma = \gamma_2$, and three stable attractors (the nodes N_1, N_2 and N_3) coexist at $0 \le \gamma \le \gamma_2$. If, on the contrary, γ increases starting from $\gamma = 0$ (strong interaction), at first three stable nodes coexist at small γ ($0 \le \gamma \le \gamma_2$). Then, the stable nodes N_2 and N_3 disappear via a saddle–node bifurcation at $\gamma = \gamma_2$ and $\gamma = \gamma_1$, and the single stable node N_1 remains in the interval $\gamma_1 < \gamma \le \overline{\gamma}$. It should be noted that when the network with the coupling matrix $\hat{W} = \mathbf{U}\mathbf{U}^{\dagger}$, $\mathbf{U} = N^{-1/2}(1, \ldots, 1)^T$ is considered as an associative memory network, the stable attractors N_2 and N_3 should be referred to attractors belonging to the "extraneous" memory (that arise in addition to the "scheduled" attractors). The question on extraneous memory will be discussed in more detail in Section 2.2.3.

Oscillatory associative memory networks, the admissible set of memory vectors, and phasor networks

We now return to the dynamical system (2.2-3) and to the problem of the oscillatory associative memory design. The dynamical system governing oscillatory network dynamics is given by (2.2-5) with the coupling matrix $\hat{W}$ satisfying the conditions (2.2-4). The stable equilibria are defined by the following system of linear equations with constant coefficients:

$$(\tilde{D}(\mathbf{z}) + \kappa\hat{W})\mathbf{z} = 0, \tag{2.2-29}$$

where

$$\tilde{D}(\mathbf{z}) = \mathrm{diag}(\tilde{D}_1(\mathbf{z}), \ldots, \tilde{D}_N(\mathbf{z})),$$
$$\tilde{D}_j(\mathbf{z}) = 1 + i\omega_j - |z_j|^2 - \kappa \sum_{k=1}^{N} W_{jk}. \tag{2.2-30}$$

The stability of equilibria is defined by eigenvalues of the Jacobian

$$\hat{J}(\mathbf{z}) = (\tilde{D}(\mathbf{z}) + \kappa \hat{W}). \tag{2.2-31}$$

As we saw in Section 2.2.1, the state vectors that can figure as stable attractors of the dynamical system (2.2-29) (the memory vectors that can be imposed into the oscillatory network memory) are not arbitrary but must be chosen from a special set $\{\mathbf{V}^m\}$ of mutually orthogonal vectors in $\mathbb{C}^N$ – the eigenvectors of matrix $\hat{W}$. This set can be naturally called the "phase" basis:

$$\mathcal{B} = \{\mathbf{V}^m \mid V_j^m = e^{i\beta_j^m},\ (\mathbf{V}^k)^\dagger \mathbf{V}^m = N\delta_{km},\ k, m = 1, \ldots, N\}. \tag{2.2-32}$$

The basis $\mathcal{B}$ is cyclical because equation (2.2-29) admits the cyclical N-polygonal pyramid group. Hence, all the vectors of $\mathcal{B}$ can be generated from the single unit vector $\mathbf{V}^0 = (1, \ldots, 1)^T$ by application of recurrent action of the group representation operator

$$T_g = \mathrm{diag}(1, e^{i\phi}, \ldots, e^{i(N-1)\phi}), \quad \phi = \frac{2\pi}{N}. \tag{2.2-33}$$

An arbitrary $N \times N$ matrix $\hat{W}$ satisfying the conditions (2.2-4) can be represented in the form

$$\hat{W} = \frac{1}{N} \sum_{m=1}^{N} \lambda^m \mathbf{V}^m (\mathbf{V}^m)^\dagger, \tag{2.2-34}$$

where λ^m are some real numbers. The matrix $\hat{W}^H$ of rank M,

$$\hat{W}^H = \frac{1}{N} \sum_{m=1}^{M} \lambda^m \mathbf{V}^m (\mathbf{V}^m)^\dagger, \quad M = \mathrm{rank}\, \hat{W}^H \tag{2.2-35}$$

is the matrix of the projection operator onto the M-dimensional subspace of $\mathbb{C}^N$ spanned by $\mathbf{V}^1, \ldots, \mathbf{V}^M$. Similarly to vectors of the basis $\mathcal{B}$, the matrices $\hat{W}$ and $\hat{W}^H$ are cyclic.

Let us denote oscillatory network with arbitrary frequency distribution $\{\omega_j\}$ satisfying the condition $\sum_{j=1}^{N} \omega_j = 0$, and coupling matrix $\hat{W}^H$ as $\mathcal{ON}(\{\omega_j\}, \kappa\hat{W})$. Introduce the special type of oscillatory network with zero frequencies: $\mathcal{PN}(\kappa\hat{W}) \equiv \mathcal{ON}(\{0\}, \kappa\hat{W})$. These networks, which can be regarded as phasor networks, are closely related to so--called clock neural networks (that, in tern, are related to systems of continuous-state magnetic spins in a plane) [19]. The phasor networks are also closely related to CVNN, which are now actively applied (see Section 2.2.5). On the other hand, the phasor networks $\mathcal{PN}(\kappa\hat{W})$ can be considered as the limit case of networks of coupled oscillators with extremely strong oscillator interaction (since $\gamma = \Omega/\kappa \to 0$ as $\kappa \to \infty$). In the

problem of oscillatory associative memory design the phasor networks $\mathcal{PN}(\{0\}, \kappa\hat{W})$ play a role of a key model, attaining the maximum storage capacity at special choice of network parameters.

Class of networks with guaranteed memory characteristics and poor extraneous memory

Stable attractors of dynamics of oscillatory network $\mathcal{ON}(\{\omega_j\}, \kappa\hat{W})$ and the corresponding $\mathcal{PN}(\kappa\hat{W})$ are proved to be closely related. Namely, the following proposition is valid.

Let $\mathbf{V}^1, \ldots, \mathbf{V}^M$ be the stable attractors of network dynamics of the phasor network $\mathcal{PN}(\kappa\hat{W})$, and the value of interaction strength $\tilde{\kappa}$ for the network $\mathcal{ON}(\{\omega_j\}, \tilde{\kappa}\hat{W})$ satisfies the condition $\gamma = \Omega/\tilde{\kappa} \ll 1$, where $\Omega = \max_j|\omega_j|$. Then, the the memory vectors $\tilde{\mathbf{V}}^1, \ldots, \tilde{\mathbf{V}}^M$ of the network $\mathcal{ON}(\{\omega_j\}, \kappa\hat{W})$ represent small perturbations of the corresponding $\mathbf{V}^1, \ldots, \mathbf{V}^M$.

The proof of this statement can be obtained by the estimation of the norm $\|\tilde{\mathbf{V}}^k - \mathbf{V}^k\|$ with the help of a perturbation method on the small parameter γ. The proximity of $\mathbf{V}^1, \ldots, \mathbf{V}^M$ and $\tilde{\mathbf{V}}^1, \ldots, \tilde{\mathbf{V}}^M$ was also confirmed in computer study of phase portraits of oscillatory dynamical systems at small N [37, 38].

The matrix $\hat{W}^H$, defined according to (2.2-35) is known as the Hebbian matrix of connections (due to its direct relation to the Hebbian learning). However, in contrast to neural networks of associative memory, memory vectors of oscillatory network cannot be arbitrarily chosen. Instead, they have to be chosen from the special set of mutually orthogonal vectors of $\mathbb{C}^N$ – the "phase" basis $\mathcal{B}$ defined according to (2.2-32). Besides, only the special choice of the number N of network processing units (oscillators or phasors) provides the networks with guaranteed (completely controllable) memory characteristics. Thus, a special class of oscillatory associative memory networks can be selected [37, 38]. Namely, if the number N of oscillators is a prime number, the memory storage capacity up to $r = M/N = 0.5$ can be achieved for the the phasor associative memory network. These networks are characterized by the following features: (a) any subset of M ($M < [N/2]$) vectors from the phase basis $\mathcal{B}$ can be chosen as the set of the network memory vectors; (b) the network possesses a poor and easily diagnosed "extraneous" memory.

If the number N is not prime, the network memory is not completely controllable: only special odd numbers M of memory vectors from $\mathcal{B}$ can be chosen as the sets of network memory vectors.

The extraneous memory – an additional set of stable equilibria of network dynamics that arises in the phase space of the governing dynamical system simultaneously with the scheduled memory vectors – crucially influences on the performance of associative memory network via decreasing the sizes of memory vector attraction basins. As a result, the network retrieval performance degrades. The extraneous memory of Hopfield networks was shown to grow exponentially under some conditions [17]. For

oscillatory networks with composite numbers N the extraneous memory definitely exists, but its features have not been completely studied. It is computationally confirmed that if some extraneous vector $\mathbf{P}$ belongs to the phasor network memory, the number $M' = N - 1$ of another extraneous vectors different from $\mathbf{P}$ also exists in the network memory. (This fact is a direct consequence of cyclicity of $\hat{W}$.) There is one more notable property of extraneous memory of the phasor networks: unlike the memory vectors $\mathbf{V}^1, \ldots, \mathbf{V}^M$ from the phase basis $\mathcal{B}$, the components P_j of extraneous vectors are characterized by different $|P_j|$ for different j. In addition to isolated extraneous memory vectors, a complicated degenerated stable equilibria can exist. These arise due to the peculiar phase space of the governing dynamical system (an N-dimensional torus). The detection of the degenerated extraneous equilibria is not a simple task. As was revealed in the computer study, a vast extraneous memory, which can exist for the oscillatory networks at composite numbers N, results in drastic degradation of network retrieval capabilities [37, 38].

Some analytical results on networks of limit cycle oscillators: Structural portrait of a two-oscillatory dynamical system

In the case of the system of two coupled oscillators

$$\begin{aligned} \dot{z}_1 &= (1 + i\omega_1 - |z_1|^2)z_1 + \kappa e^{i\beta}(z_2 - z_1), \\ \dot{z}_2 &= (1 + i\omega_2 - |z_2|^2)z_2 + \kappa e^{i\beta}(z_1 - z_2), \end{aligned} \tag{2.2-36}$$

the parametrical space of the dynamical system is $\{\omega, \kappa, \beta\}$ where $\omega = \omega_1 = -\omega_2$, $W_{12} = e^{i\beta}$ with the domains $-\infty < \omega < \infty$, $0 < \kappa < \infty$ and $-\pi/2 < \beta < \pi/2$. As one can see from the form of equation (2.2-5), the whole set of equilibria of the oscillatory dynamical system consists of two subsets: $\mathbf{z} = 0$ and $\{\mathbf{z} \neq 0, (\hat{D}(\mathbf{z}) + \kappa\hat{W})\mathbf{z} = 0\}$.

Let D denote the subdomain of the parametrical space where the equilibrium $\mathbf{z} = 0$ is stable and S denotes the subdomain where the equilibria $\{\mathbf{z} \neq 0, (\hat{D}(\mathbf{z}) + \kappa\hat{W})\mathbf{z} = 0\}$ are stable. The following results can be analytically obtained for the two-oscillatory system [39]:

(1) There exist from one up to four fixed points of (2.2-36) in different domains of the parametrical space, and only one of these points (the fixed point $\mathbf{U}^1 \in$ S) is a stable node, whereas the others are saddles.

(2) The structural portrait of the system (2.2-36) contains the following domains:

(a) the domain S:

$$\begin{aligned} \omega > 0: &\quad \kappa \geq f_1(\omega, \beta) \quad \text{at } 0 < \omega < \omega_1; &&\quad \kappa \geq f_2(\omega, \beta) \quad \text{at } \omega > \omega_1, \\ \omega < 0: &\quad \kappa \geq f_3(\omega, \beta) \quad \text{at } 0 < |\omega| < \omega_2; &&\quad \kappa \geq f_4(\omega, \beta) \quad \text{at } |\omega| > \omega_2, \end{aligned}$$

(b) the domain D:

$$\begin{aligned} \omega > 0: &\quad f_3(\omega, \beta) \leq \kappa \leq f_2(\omega, \beta) \quad \text{at } \omega \geq \omega_1, \\ \omega < 0: &\quad f_5(\omega, \beta) \leq \kappa \leq f_4(\omega, \beta) \quad \text{at } |\omega| \geq \omega_2, \end{aligned}$$

(c) the domain O:

$$\omega > 0: \quad \kappa < f_1(\omega,\beta) \quad \text{at } 0 < \omega < \omega_1; \qquad \kappa < f_2(\omega,\beta) \quad \text{at } \omega > \omega_1,$$
$$\omega < 0: \quad \kappa < f_3(\omega,\beta) \quad \text{at } 0 < |\omega| < \omega_2; \qquad \kappa < f_4(\omega,\beta) \quad \text{at } |\omega| > \omega_2,$$

where

$$f_1(\omega,\beta) = \frac{\omega}{1+\sin\beta}, \qquad f_2(\omega,\beta) = \frac{\omega^2+1}{2(\omega\sin\beta+\cos\beta)},$$
$$f_3(\omega,\beta) = \frac{\omega}{1-\sin\beta}, \qquad f_4(\omega,\beta) = f_2(-|\omega|,\beta),$$
$$\omega_1 = \frac{1+\sin\beta}{\cos\beta}, \qquad \omega_2 = \frac{1-\sin\beta}{\cos\beta}.$$

The projection of the structural portrait onto the quadrant $(\omega,\kappa,0)$, $\omega \geq 0$ is shown in Figure 2.10. The domains S (the synchronization domain), D (the domain of unsteady multifrequency oscillations) and O (the domain of oscillation suppression) can be written analytically in the following form:

$$\begin{aligned} \mathrm{S} &= \{a \geq \omega, \omega \leq 1\} \cup \{a \geq (\omega^2+1)/2, \omega > 1\}, \\ \mathrm{O} &= \{a < \omega, \omega \leq 1\} \cup \{a < 1, \omega > 1\}, \\ \mathrm{D} &= \{1 \leq a \leq (\omega^2+1)/2\}. \end{aligned} \tag{2.2-37}$$

The structural portrait is symmetrical with respect to the κ-axis at real-valued κ ($\mathrm{Im}\,\kappa = 0$) in the upper half-plane $\kappa > 0$. The symmetry is absent if $\mathrm{Im}\,\kappa \neq 0$. The simplicity of the structural portrait is a consequence of coupling symmetry. In the case of nonsymmetrical coupling, a narrow domain corresponding to more complicated oscillatory dynamics arises in the vicinity of the line $\kappa = \omega$, $\omega \in [0,1]$. However, in the

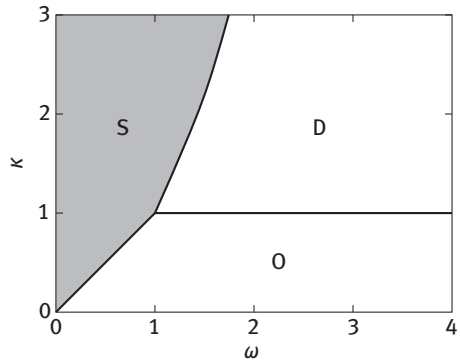

Fig. 2.10. Structural portrait of dynamical system of two coupled oscillators at $\mathrm{Im}\,\kappa = 0$ (the projection of full structural portrait of the system (2.2-36) onto the quadrant $(\omega,\kappa,0)$, $\omega \geq 0$).

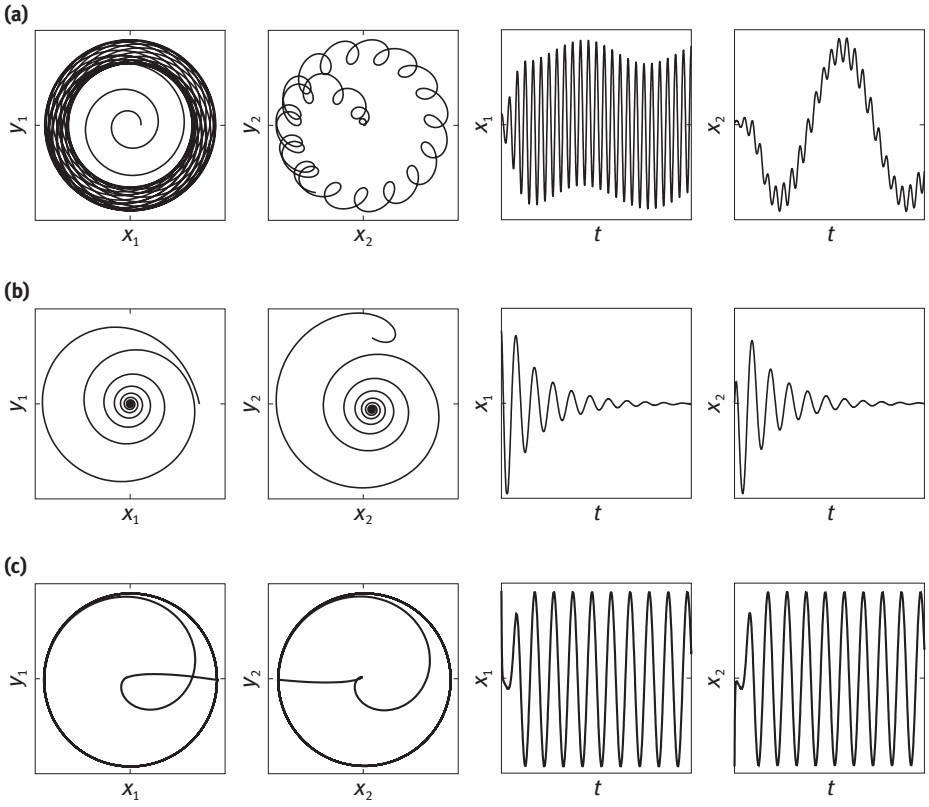

Fig. 2.11. Phase trajectories and temporal dynamics $x_1(t)$ and $x_2(t)$ of system of two coupled oscillators in parametric domains (a) O, (b) D, and (c) S.

case of the two-oscillatory system this narrow domain does not contain the subdomain corresponding to chaotic dynamics.

The examples of the dynamical regimes in the domains O, D, and S are given in Figure 2.11.

Some analytical results on networks of limit cycle oscillators: number of stable attractors for small oscillatory chains

A class of closed chains of limit cycle oscillators represents an important special example of globally coupled oscillatory systems and possesses some attractive features. These networks can be considered as building blocks in the problem of design of oscillatory associative memory networks with general interconnection architecture. Homogeneously coupled closed oscillatory chains are governed by the dynamical system (2.2-3) with $W_{jk} = w\delta_{j+1,k}$, where $w = b + ic \equiv ue^{i\beta}$. At small N ($N = 3, 4, 5, 6$) the

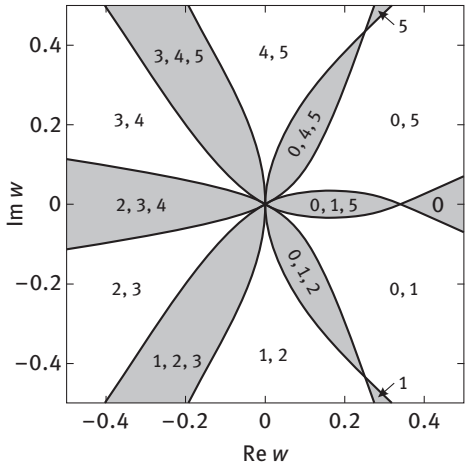

Fig. 2.12. The structural portrait of dynamical system governing the dynamics of the oscillatory chain of six oscillators. The numbers in the picture denote the numbers of stable attractors in dependence of coupling strength $w = b + ic$.

chains were analyzed analytically, and the obtained results were further confirmed in computer experiments.

The structural portrait of the dynamical system governing the dynamics of the oscillatory chain that comprises six oscillators in parametrical plane (b, c) is shown in Figure 2.12. It contains the domains of existence of one, two and three stable fixed points of the dynamical system. All fixed points correspond to memory vectors from the phase basis $\mathcal{B}$. The stability of the domains when N grows was estimated. For example, the domain of stability of the memory vector $\mathbf{V}^0 = (1, \ldots, 1)^T$ increases as N grows, whereas the domains of stability of the other vectors from $\mathcal{B}$ decreases [41]. An interesting feature of closed oscillator chains is that they provide an example of coupling architecture of oscillatory associative memory networks that lacks extraneous memory.

An asymptotical estimation of memory storage capacity of oscillatory chains gives the value $r_{as} = 0.27$.

2.2.4 Oscillatory media related to networks of limit cycle oscillators

In this section, we consider a class of oscillatory networks consisting of locally coupled limit cycle oscillators. In spatially continuous limit these networks can be viewed as oscillatory media governed by a nonlinear diffusion equation – the so-called reaction–diffusion equation (RDE). Spatiotemporal regimes in nonlinear media (wave trains, standing waves, spiral waves, stripe patterns, and cluster states) has been studied since 70s and are widely known as dissipative structures. Belousov–Zhabotin-

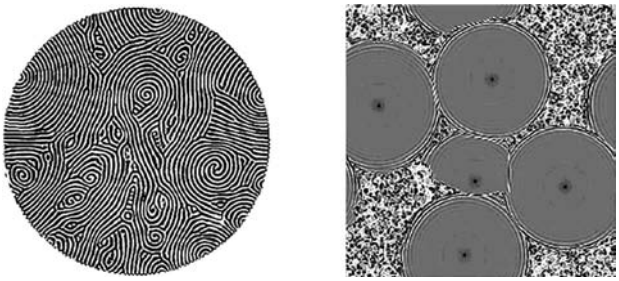

Fig. 2.13. The examples of spatiotemporal regimes (dissipative structures) in nonlinear media, governed by 2D CGL: spiral defect chaos near threshold of Rayleigh–Benard convection (left); coexistence of spirals and defect chaos (right).

sky oscillating chemical reaction in a thin layer of fluid and oscillatory media of Ginzburg–Landau oscillators belong to the most familiar examples of nonlinear active media. A considerable scope of oscillatory media studies exists, including strict mathematical results, physical level results and computer modeling [18, 23, 24]. The examples of dissipative structures arising in two-dimensional nonlinear media governed by the complex Ginzburg–Landau equation are shown in Figure 2.13.

Depending on the local coupling template (the number of closest neighbors coupled with each network unit), the networks can be considered as 1D, 2D, or nD spatially distributed arrays of processors. Similarly to locally connected neural networks known as cellular neural networks [57], the locally coupled oscillator arrays may be naturally regarded as cellular oscillatory networks. The two-dimensional oscillator arrays can be successfully used to model collective oscillatory neural activity. We present here some results of analytical study of possible dynamical regimes in one-dimensional oscillatory media corresponding to closed and open oscillator chains.

For locally coupled Ginzburg–Landau oscillator arrays it is convenient to rewrite the governing dynamical system (2.2-3) in the form

$$\dot{u}_j = (1 + i\omega_j - |u_j|^2)u_j + \kappa \sum_{k=1}^{N} W_{jk}(u_k - u_j), \tag{2.2-38}$$

where

$$u_j = r_j e^{i\theta_j}, \quad W_{jk} = W_{kj}^*. \tag{2.2-39}$$

In the case of homogeneously locally coupled oscillatory chains, the coupling matrix elements can be written as

$$W_{jk} = d\,\delta_{j-1,k}\delta_{j+1,k}, \quad d = \kappa e^{i\chi} \equiv d_1 + id_2, \tag{2.2-40}$$

where d is the coupling strength in the chain. To transform the dynamical system (2.2-38) into its spatially continuous analog, one should introduce the spatial coordinate x ($x \in [0, l]$), and the complex-valued function $u(x,t)$ instead of $u_j(t)$.

Then, the RDE can be easily derived:

$$\dot{u} = (1 + i\omega(x) - |u|^2)u + d\, u_{xx}, \tag{2.2-41}$$

where $u(x,t) = u_1(x,t) + iu_2(x,t)$, and the second derivative $u_{xx} \equiv \partial^2 u/\partial x^2$ plays the role of Laplacian Δu in 1D case. Equation (2.2-41) can be rewritten in terms of the two-component real-valued vector-function $\mathbf{u}(x,t) = (u_1, u_2)^T$:

$$\dot{\mathbf{u}} = \hat{F}(\mathbf{u})\mathbf{u} + \hat{D}\mathbf{u}_{xx} \tag{2.2-42}$$

where

$$\hat{F}(\mathbf{u}) = \begin{pmatrix} 1 - u_1^2 - u_2^2 & -\omega \\ \omega & 1 - u_1^2 - u_2^2 \end{pmatrix}, \quad \hat{D} = \begin{pmatrix} d_1 & -d_2 \\ d_2 & d_1 \end{pmatrix}. \tag{2.2-43}$$

Oscillatory media governed by the RDE (2.2-41) represent a special type of Ginzburg–Landau oscillatory media governed by the widely known equation

$$\dot{w} = [1 + i\omega - (1 + i\beta)|w|^2]w + (1 + i\epsilon)\Delta w. \tag{2.2-44}$$

However, the diffusion operator in (2.2-41) is of more general type than that in equation (2.2-44).

The following properties are inherent to the RDE (2.2-41).

(1) If $\omega(x) = \omega = \text{const}$, the function $w(x,t) = u(x,t)e^{-i\omega t}$ satisfies the RDE (2.2-41) with $\omega = 0$. So, it is sufficient to analyze only an RDE with $\omega = 0$.

(2) The function $u^0(x,t) = e^{i\theta_0}$ is a spatially homogeneous solution of the RDE (2.2-41) with $\omega = 0$.

(3) If $\omega = 0$, the complex-valued vector-function $\mathbf{w}(x,t) = (u,v)^T = 2^{-1/2}(u_1 + iu_2, u_1 - iu_2)^T$ satisfies the RDE with a diagonal diffusion operator:

$$\begin{bmatrix} \dot{u} \\ \dot{v} \end{bmatrix} = (1 - |u|^2 - |v|^2)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{bmatrix} u \\ v \end{bmatrix} + \kappa \begin{pmatrix} e^{i\chi} & 0 \\ 0 & e^{-i\chi} \end{pmatrix}\begin{bmatrix} u_{xx} \\ v_{xx} \end{bmatrix}. \tag{2.2-45}$$

(4) To further analyze the properties of the nonlinear RDE it is helpful to use an expansion of the solution into the series of orthonormalized system of eigenfunctions $\{X_m(x)\}$ of the corresponding linear scalar diffusion operator. For the RDE (2.2-41) with $\omega = 0$ these series can be written in the form

$$u_1(x,t) = \sum_{m=1}^{\infty} X_m(x)P_m(t), \quad u_2(x,t) = \sum_{m=1}^{\infty} X_m(x)Q_m(t), \tag{2.2-46}$$

where $X_m(x)$ are simple periodic functions of x (for open chain $X_m(x) = \cos(\pi m x/l)$), and $\{P_m(x), Q_m(x)\}$ can be obtained from the solution of the system of coupled ODE – the "moment system" [40].

By analyzing the series (2.2-46) it is possible to elucidate the main features of the RDE (2.2-41) solutions [40]. We mention of the most important of them.

(1) The diffusion instability of 1D oscillatory media corresponding to closed oscillatory chains occurs in some parametrical domain of (κ, χ).
(2) The regime of plane wave trains (the RDE solutions of the form $u(x, t) = U(\omega t - kx)$) can exist. It can arise as a result of a bifurcation from a spatially homogeneous dynamical state.
(3) In the case of oscillatory media corresponding to open oscillator chains, the regime of modulated standing waves (special RDE solutions with separated variables x and t) can also exist. Its existence was established by analyzing the series (2.2-46).
(4) Cluster states – the RDE solution with separated variables x and t of another type than standing waves – can exist as well. The cluster states correspond to the oscillatory medium decomposition into synchronously oscillating subdomains (clusters), where each cluster oscillates with its own frequency, amplitude and phase shift. The regimes of cluster synchronization are of special interest from the viewpoint of dynamical neuromorphic methods of image processing.

2.2.5 Associative memory networks of complex-valued neurons

The processing units of CVNN are neurons with complex-valued states and discrete-valued activation functions. Accordingly, network coupling is specified by matrices with complex-valued elements.

As was clarified, complex-valued neurons are more functional than real-valued ones. The multivalued activation function makes single network neuron more flexible and adaptive. As a result, the CVNN are capable to learn faster and generalize better. In particular, they permit us to overcome the Minsky–Papert limitation and to extend the solvable class of nonlinearly separable problems. The CVNN models with continuous-valued activation functions were successfully developed as well. It was possible to generalize the perceptron learning rule to construct the complex-valued back propagation algorithm. In addition, the learning algorithms for CVNN were constructed that do not suffer from the local minima phenomenon. A significant advantage is inherent to CVNN in signal and image processing problems due to the proper exploitation of phase information. Single-layer and multilayer CVNN of feedforward architecture are capable to analyze the phases of the Fourier spectral coefficients and to achieve successful recognition and restoration of blurred and corrupted images. The CVNN also permit us to naturally exploit a multivalued logic over the field of complex numbers [1–5, 25].

Recurrent complex-valued neural networks of associative memory were also designed and studied [26, 32]. They consist of N fully connected multistate neurons, the output of each neuron being determined by the equation

$$x'_j = \phi\left(\sum_{n=1}^{N} S_{jn} x_n\right), \tag{2.2-47}$$

where $\hat{S} = [S_{jn}]$ is the complex $N \times N$ matrix, and $\phi(x)$ is the complex multivalued sign function. The recurrent associative memory CVNN is capable to store complex-valued vectors

$$\mathbf{X}^k, \quad \mathbf{X}^k = \left(x_1^k, \ldots, x_N^k\right)^T, \quad \mathbf{X}^k \in \mathbb{C}^N \tag{2.2-48}$$

as a set of stable attractors of neural network dynamics. The components of $\mathbf{X}^k$ may be both continuous valued and quantized, that is, defined as

$$x_j^k \in \left\{e^{2\pi i n/K}\right\}_{n=1}^{K-1}, \quad j = 1, \ldots, N, \tag{2.2-49}$$

where the resolution factor K divides the complex unit circle into K quantization levels (separate sectors). Similarly to the case of real-valued neural networks, the energy function can be introduced for CVNN, providing the analysis of CVNN dynamics in terms of local energy minima. If the matrix $\hat{S}$ is chosen in the form of the sum of outer-products over the memory vectors (the generalized Hebb rule),

$$\hat{S} = \frac{1}{N} \sum_{m=1}^{M} \mathbf{X}^m (\mathbf{X}^m)^\dagger, \tag{2.2-50}$$

the set (2.2-48) of M memory vectors can be stored by a complex-valued autocorrelation associative memory network. In the case of CVNN each memory vector is not an isolated equilibrium point but an equilibrium subset of the network (a closed curve in the complex space). The asymptotic behavior of a dynamical system trajectory near memory vectors was analyzed, and conditions assuring the correct recall of a memory vector were discussed [34].

The temporal evolution of the overlap during the memory retrieval phase was also studied in computer experiments. It was found that in the case of unsuccessful memory vector recall, the variable overlap demonstrates a kind of oscillatory dynamics. It reflects the behavior of the trajectory of the network dynamical system in the vicinity of the stable attractor corresponding to the memory vector (which is similar to trajectory behavior near very slowly damping stable focus). The network storage capacity $r = M/N$ achieves the value $r_* = 0.5$ for CVNN of autoassociative memory [47].

On the whole, a great number of publications is devoted to various aspects of currently popular CVNN topic, including their numerous applications.

Concluding the section, we note a close resemblance between complex-valued associative memory neural networks and recurrent phasor networks of associative memory discussed in Section 2.2.3.

Bibliography

[1] I. Aizenberg. A periodic activation function and a modified learning algorithm for the multi-valued neuron. *IEEE T. Neural Networ.*, 21:1939, 2010.

[2] I. Aizenberg. *Complex-Valued Neural Networks with Multi-Valued Neuron*. Springer, 2011.

[3] I. Aizenberg and C. Moraga. Multilayer feedforward neural network based on multi-valued neurons and a backpropagation learning algorithm. *Soft Computing*, 11:169, 2007.
[4] I. Aizenberg, D. Paliy, J. Zurada and J. Astola. Blur identification by multilayer neural network based on multi-valued neurons. *IEEE T. Neural Networ.*, 19:883, 2008.
[5] I. N. Aizenberg, N. N. Aizenberg and J. Vandewalle. *Multi-Valued and Universal Binary Neurons: Theory, Learning and Applications*. Kluwer Academic Publishers, 2000.
[6] S. Amari. Characteristics of randomly connected threshold-element networks and network systems. In *Proc. IEEE*, volume 59, page 35, 1971.
[7] S. Amari. Characteristics of random nets of analog neuron-like elements. *IEEE Trans. Syst. Man Cybern.*, 2:643, 1972.
[8] S. Amari. A mathematical theory of nerve nets. In M Kotani, editor, *Advances in biophysics*, volume 6, page 75. Tokyo University Press, 1974.
[9] S. Amari. Neural theory of association and concept-formation. *Biol. Cybern.*, 26:175, 1977.
[10] S. Amari and K. Maginu. Statistical neurodynamics of associative mamory. *Neural Networks*, 1:63, 1988.
[11] S. Amari, K. Yoshida and K. Kanatani. A mathematical foundation for statistical neurodynamics. *SIAM J. Appl. Math.*, 33:95, 1977.
[12] J. A. Anderson. A simple neural network generating interactive memory. *Math. Biosci.*, 14:197, 1972.
[13] L. L. Bonilla. Nonequilibrium statistical mechanics model showing self-sustained oscillations. *Phys. Rev. Lett.*, 60:1398, 1988.
[14] L. L. Bonilla, J. M. Casado and M. Morillo. Self-synchronization of populations of nonlinear oscillators in thermodynamical limit. *J. Stat. Phys.*, 48:571, 1987.
[15] L. L. Bonilla, Pérez Vicente C. J., F. Ritort and J. Soler. Exactly solvable phase oscillator models with synchronization dynamics. *Phys. Rev. Lett.*, 81:3643, 1998.
[16] L. L. Bonilla, Pérez Vicente C. J. and J. M. Rubí. Glassy synchronization in a population of coupled oscillators. *J. Stat. Phys.*, 70:921, 1993.
[17] J. Bruck and V. P. Roychowdhury. On the number of spurious memories in the Hopfield model. *IEEE T. Inform. Theory*, 36:393, 1990.
[18] H. Chaté and P. Manneville. Phase diagram of the two-dimensional complex Ginzburg–Landau equation. *Physica A*, 224:348, 1996.
[19] J. Cook. The mean-field theory of a Q-state neural network model. *J. Phys. A: Math. Gen.*, 22:2057, 1989.
[20] H. Daido. Order functions and macroscopic mutual entrainment in uniformly coupled limit-cycle oscillators. *Progr. Theor. Phys.*, 88:1213, 1992.
[21] H. Daido. Critical conditions of macroscopical mutual entrainment in uniformly coupled limit--cycle oscillators. *Progr. Theor. Phys.*, 89:929, 1993.
[22] H. Daido. Onset of cooperative entrainment in limit-cycle oscillators with uniform all-to-all interactions: bifurcation of the order function. *Physica D*, 91:24, 1996.
[23] M. Falcke, H. Engel and M. Neufeld. Cluster formation, standibg waves, and stripe patterns in oscillatory active media with local and global coupling. *Phys. Rev. E*, 52:763, 1995.
[24] V. Hakim and W.-J. Rappel. Dynamics of the globally coupled complex Ginzburg–Landau equation. *Phys. Rev. A*, 46:R7347, 1992.
[25] A. Hirose, editor. *Complex-Valued Neural Networks: Theories and Applications*. World Scientific, 2003.
[26] S. Jankowski, A. Lozowski and J. M. Zurada. Complex-valued multistate neural associative memory. *IEEE T. Neural Networ.*, 7:1491, 1996.
[27] A. Ya. Khinchin. *Mathematical foundations of statistical mechanics*. Dover, 1949.

[28] W. Kinzel. Learning and pattern recognition in spin glass models. *Z. Phys. B Cond. Mat.*, 60:205, 1985.
[29] T. Kohonen. Correlation matrix memories. *IEEE T. Comput.*, 21:353, 1972.
[30] B. Kosko. Adaptive bidirectional associative memories. *Appl. Opt.*, 26:4947, 1987.
[31] Y. Kuramoto. Cooperative dynamics of oscillator community. *Progr. Theor. Phys. Suppl.*, 79, 1984.
[32] Y. Kuramoto. *Chemical oscillations, Waves and Turbulence*. Dover Publications, 2003.
[33] Y. Kuramoto and I. Nishikawa. Statistical macrodynamics of large dynamical systems. case of a phase transition in oscillator communities. *J. Stat. Phys.*, 49:569, 1987.
[34] Y. Kuroe. A model of complex-valued associative memories and its dynamics. In A. Hirose, editor, *Complex-Valued Neural Networks: Theories and Applications*, page 57. World Scientific, 2003.
[35] M. G. Kuzmina. The recalling process dynamics of associative memory networks in macrodynamical approach. *Neural Networks*, 8:1433, 1995.
[36] M. G. Kuzmina, E. A. Manykin and I. I. Surina. Oscillatory networks with Hebbian matrix of connections. In *Proc. IWANN'95*, volume 930 of *Lecture Notes in Computer Science*, page 246, 1995.
[37] M. G. Kuzmina, E. A. Manykin and I. I. Surina. Oscillatory networks of associative memory. *Opt. Mem. Neural Net.*, 5:91, 1996.
[38] M. G. Kuzmina, E. A. Manykin and I. I. Surina. Oscillatory networks with guaranteed memory characteristics. In *Proc. EUFIT'96*, volume 1, page 320, 1996.
[39] M. G. Kuzmina, E. A. Manykin and I. I. Surina. Exact solutions and modeling of associative memory in oscillatory networks. In *Proc. SPIE, Optical Information Science and Technology '97: Optical Memory and Neural Networks*, volume 3402, page 298, 1998.
[40] M. G. Kuzmina, E. A. Manykin and I. I. Surina. Pattern formation in locally connected oscillatory networks. In *Proc. ICANN'98*, volume 2, page 663, 1998.
[41] M. G. Kuzmina, E. A. Manykin and I. I. Surina. Memory estimation of homogeneous closed chains of oscillators. In *Proc. Int. Conf. "Neuroinformatics-2000"*, page 37, 2000. in Russian.
[42] M. G. Kuzmina and I. I. Surina. Macrodynamical approach for oscillatory networks. In *Proc. SPIE, Optical Memory & Neural Networks '94: Optical Neural Networks*, volume 2430, page 229, 1994.
[43] P. C. Matthews, R. E. Mirollo and S. H. Strogatz. Dynamics of a large system of coupled nonlinear oscillators. *Physica D*, 52:293, 1991.
[44] P. C. Matthews and S. H. Strogatz. Phase diagram for the collective behavior of limit-cycle oscillators. *Phys. Rev. Lett.*, 65:1701, 1990.
[45] R. Meir and E. Domany. Exact solution of a layered neural network memory. *Phys. Rev. Lett.*, 59:359, 1987.
[46] K. Nakano. Association – a model of associative memory. *IEEE T. Syst. Man Cyb.*, 2:381, 1972.
[47] I. Nemoto. Complex associative memory and complex single neuron model. In A. Hirose, editor, *Complex-Valued Neural Networks: Theories and Applications*, page 107. World Scientific, 2003.
[48] H. Nishimori and N. Ozeki. Retrieval dynamics of associative memory of the Hopfield type. *J. Phys. A: Math. Gen.*, 26:859, 1993.
[49] A. J. Noest. Discrete-state phasor neural networks. *Phys. Rev. A*, 38:2196, 1988.
[50] L. I. Rozonoér. Random logical nets I. *Automat. Rem. Contr.*, 5:773, 1969.
[51] L. I. Rozonoér. Random logical nets II. *Automat. Rem. Contr.*, 6:922, 1969.
[52] L. I. Rozonoér. Random logical nets III. *Automat. Rem. Contr.*, 7:1130, 1969.
[53] H. Sakaguchi, S. Shinomoto and Y. Kuramoto. Local and global self-entrainments in oscillator lattices. *Progr. Theor. Phys.*, 77:1005, 1987.

[54] H. Sakaguchi, S. Shinomoto and Y. Kuramoto. Mutual entrainment in oscillator lattices with nonvariational type interaction. *Progr. Theor. Phys.*, 79:1069, 1988.
[55] H. Sakaguchi, S. Shinomoto and Y. Kuramoto. Phase transitions and their bifurcation analysis in a large population of active rotators with mean-field coupling. *Progr. Theor. Phys.*, 79:600, 1988.
[56] S. H. Strogatz and R. E. Mirollo. Phase-locking and critical phenomena in lattices of coupled nonlinear oscillators with random intrinsic frequencies. *Physica D*, 31:143, 1988.
[57] P. Thiran, K. R. Crounse, L. O. Chua and M. Hasler. Pattern formation properties of autonomous cellular neural networks. *IEEE T. Circuits-I*, 42:757, 1995.

3 Oscillatory networks for modeling the brain structures performance

3.1 Motivations for oscillatory network modeling

Coherent oscillations and synchronization of the neural activity was observed in different areas of the brain of mammals, amphibians and insects. Synchronization was discovered in monkey's motor cortex and in the visual cortices of cats. Synchronous oscillations in the antennal lobe of insects appear to enhance the animal ability to distinguish between two closely related odors. The exhibition of synchronized neural activity in a wide range of brain areas of a diversity of organisms indicates that synchronization presumably plays a fundamental role in the brain information processing. In particular, synchronization and resonance are used in many human brain neural structures: olfactory and auditory systems, thalamocortical system, hippocampus, and neocortex.

All this induced large interest to theoretical studies of oscillatory networks and stimulated the development of oscillatory network models for various information processing tasks. As a result, a series of oscillatory neural network models, exploiting principle of dynamical binding via synchronization, was created. Besides the models developed for image processing [5, 6, 11–14, 17–19, 23–27, 31–34, 45, 46, 49, 50], oscillatory network model for processing of mixed sound streams [43, 47, 48] for olfactory information processing and odor recognition [30, 37, 38] was designed. It was clarified that the oscillatory dynamical methods of image processing demonstrate some advantages over traditional computational methods, mainly due to their self-organized, distributed and automatic style of performance.

3.2 Oscillatory network models for the olfactory and auditory brain systems

3.2.1 System of two coupled oscillatory networks for modeling the olfactory brain system functioning

The olfactory brain system consists of a pair of coupled modules – the olfactory bulb and the olfactory cortex. The main feature of this system is that the olfactory bulb transforms any input odor into the oscillatory neural activity. Further, all kinds of information processing are performed by cooperative work of the olfactory bulb and olfactory cortex using resonance, selective cortical response and odor-specific feedback suppression of olfactory bulb response. In addition, oscillatory associative memory is exploited by the olfactory system to successfully perform odor detection, recognition

and segmentation. Mechanism of odor-specific feedback provides adaptive functioning that can be viewed as the simplest example of cognitive neural computation.

The model of two coupled neural networks each consisting of excitatory and inhibitory neurons and capable of oscillatory dynamics was designed by Li, Hopfield, and Hertz [30, 37, 38] for modeling the olfactory bulb and olfactory cortex joint performance.

Before describing the model, it is necessary to say a few words about neuron types and interconnectivity architecture of the brain olfactory system. The main neuron types of the olfactory bulb are the excitatory mitral cells and the inhibitory granule cells. The mitral cells receive odor inputs and excite the granule cells, which in turn inhibit the mitral cells. The outputs of the bulb are carried to the olfactory cortex. Odors induce the oscillatory bulbar activity. These oscillations are an intrinsic property of the bulb, and upon repeated presentation of the odor they weaken. The primary olfactory cortex receives bulbar outputs. The signals are conveyed to the excitatory pyramidal cells of the cortex, both directly and via feedforward inhibitory cells of the cortex. The pyramidal cells are connected with each other and with feedback interneurons (which inhibit them). Thus, there is some excitatory–inhibitory circuitry both in the bulb and in the cortex. The cortex differs from the bulb in much greater spatial range of the excitatory connections and in presence of excitatory–excitatory connections. Cortical output (including the feedback to the bulb) starts from pyramidal cells. It is important to note that the oscillations in the cortex do not arise spontaneously and require an input from the bulb. This anatomical structure has led some researchers to model the olfactory cortex as an associative memory for odors.

In the model [38] only basic features of the olfactory system were retained to illustrate the basic operation of the system. The model consists of two modules, a bulb and a cortex, with forward and feedback connections between them. The bulb encodes odor inputs as patterns of oscillations entering into the cortex. The cortex, representing an oscillatory associative memory network, recognizes them via oscillatory resonance when the input from the bulb matches one of the stored odor memory patterns. The resonant activity pattern is transformed into a feedback control signal to the bulb, leading to approximate canceling the odor input that generated it. The system is then able to respond to a newly arrived odor (which can be superposed with the previous one). In this way, the system segments temporally different odors.

Damping limit cycle oscillators figure in the model of two coupled neural networks as functional units of both the bulb and the cortex networks, although local populations of excitatory and inhibitory neurons figure as the structural units of both networks. Single network oscillator is formed by these coupled local neuron subpopulations (of excitatory mitral cells and inhibitory granule cells). The first version of bulb oscillator model was created by Li and Hopfield [38]. Dynamical equations for the neural oscillator were written in terms of membrane potential of the mitral neuron (the variable x) and the granule cell (the variable y), averaged over local neuron

populations:

$$\begin{aligned} \dot{x}_j &= -\alpha x_j - \sum_k H_{jk} g_y(y_k) + I_j, \\ \dot{y}_j &= -\alpha y_j + \sum_k W_{jk} g_x(x_k) + I_j^c. \end{aligned} \tag{3.2-1}$$

Here I_j is the odor input to mitral neuron, I_j^c is the external input from the cortex (containing the feedback signal), H_{jk} is the (nonnegative) weight of inhibitory connection from the granule neuron k to the mitral neuron j, W_{jk} is the weight of corresponding mitral-to-granule connection, $g_y(y)$ is the sigmoid activation function of the granule neuron, $g_x(x)$ is the similar activation function of mitral neuron and $1/\alpha$ is the membrane time constant.

In fact, the network model represents a two-layered feedforward neural network with two external entries – the external input and the input from olfactory network (for realization of feedback control). When the odor intensity exceeds some threshold value, the bulb network demonstrates bursting oscillatory dynamics in the form of oscillatory packets of finite duration. There exists a finite discrete collection of oscillatory eigenpatterns of the bulb network defined by the network parameters. Nonlinear feedback from olfactory network permits to gradually change the parameters of bulb eigenoscillations and to control the response of the bulb to external input odor mixtures.

The model of the olfactory cortex is structurally similar to the bulb network. Its functional unit – a single oscillator – is formed by the (averaged) excitatory pyramidal neuron and the inhibitory interneuron. The difference between both networks mainly concerns the connection architecture. Namely, the cortex network receives oscillatory inputs from the bulb network and there exist internal excitatory-to-excitatory connections in the cortex network in contrast to the bulb network. Therefore, the cortex network can be represented as a two-layer recurrent network of associative memory. A set of stable focuses (that correspond to packets of damping oscillations of appropriate mean frequency) form a collection of network memory templates. The cortex network demonstrates a resonant response to the oscillatory input from the bulb network in the case when the frequency of the input is close to one of the mean eigenfrequencies of the cortex network (then it is in the corresponding eigenoscillatory state). Otherwise, the cortex network is in the "silence" state. The resonant oscillatory cortex network response is interpreted as "odor recognition." The cortex network feedback to the bulb is empirically designed but biologically motivated. It is not negligible only in the case when own network response does not vanish. The system of governing dynamical equations can be written as

$$\begin{aligned} \dot{u}_j &= -\alpha u_j - \beta g_u(u_j) + \sum_k J_{jk} g_v(v_k) - \sum_k H_{jk} g_v(v_k) + I_j^b, \\ \dot{v}_j &= -\alpha v_j + \gamma g_u(u_j) + \sum_k W_{jk} g_u(u_k). \end{aligned} \tag{3.2-2}$$

Here u_j is the average membrane potential of local population of pyramidal neurons and v_j is that one of interneuron population, $\hat{J} = [J_{jk}]$ is the matrix of excitatory-to-excitatory connections, $\hat{H} = [H_{jk}]$ is the matrix of inhibitory-to-excitatory connections and $\hat{W} = [W_{jk}]$ is the matrix of excitatory-to-inhibitory connections.

Detection and recognition of an input odor by the system is carried out in two steps: (1) if the odor of sufficient intensity has been recognized by cortex network as a familiar one, the bulb network response to the subsequent odor is "blocked" by the feedback from cortex network (the blocking is realized via shifting the location of bulb network attractor corresponding to the received odor); (2) at subsequent odor mixture the bulb network does not react to already diagnosed odor (as if it was excluded from the odor mixture), and so the bulb network is able to respond to the subsequent odor.

Computer simulations were performed with networks consisting of 50 excitatory and 50 inhibitory neurons. The Hebbian matrices $\hat{J}$ and $\hat{W}$ were designed so as to provide storing of three stable oscillatory attractors in the cortex network. Thus, the designed two-network model of olfactory brain system is capable to detect, recognize and segment (via exclusion from the unknown mixture) odors. The three tasks carried out by the system – detection, recognition and segmentation – are computationally linked. The designed system performs what might be called the simplest cognitive computation. It is natural to suppose that the flexible adaptation of the biological olfactory system was elaborated by biological evolution.

3.2.2 Oscillatory network approach to segmentation of mixed interfering acoustic streams

The acoustic streams reaching our ears are composed of sound energy from multiple environmental sources. Consequently, a fundamental task of auditory perception is to disentangle this acoustic mixture in order to retrieve its separate components. The problem can be categorized as auditory scene analysis. Human ear exhibits a remarkable ability to segregate the voice of a single speaker from a mixture of other interfering sounds. So, the creation of neuromorphic computational models capable to imitate neurobiological process of decomposition of mixed sound stream is always of interest. As usual, such models are also capable of robust and automatic performance. The auditory scene analysis is related to the problem of blind source separation, which was developed independently [43, 52].

The appeal to oscillatory approaches to auditory scene analysis was related to the fact that synchronized oscillations with frequency of about 40 Hz were experimentally discovered in the brain auditory cortex similarly and in the visual and olfactory brain systems. So, the application of the idea of temporal dynamical binding to auditory grouping problems was initiated in the early 1980s [44]. In particular, oscillatory binding via synchronization seemed to be relevant in problems of auditory scene analysis. Oscillatory network model with synchronization-based performance, capable to

segregate sound stream from a variety of interfering sound sources, was suggested by Wang and Brown [47]. It is a combination of two interactively performing networks of relaxation oscillators with local excitatory and global inhibitory network connections. The preliminary calculation of sound stream characteristics should be fulfilled to find the full set of frequency-temporal stream parameters ("segmentation"). In the frames of the model it was done via stream transmission through etalon filter system imitating the transfer function of cochlear filtering. At the first stage the full set of frequency-temporal stream characteristics including correlogram, integral correlogram and full cross-channel correlation map of sound stream is found. At the second stage the basic processing of auditory stream is carried out with the help of the two-layered oscillatory network. The first layer represents the network composed of relaxation oscillators located at a 2D grid with local excitatory and global inhibitory oscillator coupling. Single network oscillator is defined by excitatory variable x_{jk} and inhibitory variable y_{jk}. The dynamical equations are written in the form

$$\begin{aligned} \dot{x}_{jk} &= 3x_{jk} - x_{jk}^3 + 2 - y_{jk} + I_{jk} + S_{jk} + \rho, \\ \dot{y}_{jk} &= \epsilon \left[\gamma \left(1 + \tanh \frac{x_{jk}}{\beta} \right) \right] - y_{jk}. \end{aligned} \tag{3.2-3}$$

Here I_{jk} represents the external stimulation, S_{jk} denotes the overall coupling from other network oscillators, and ρ is the amplitude of Gaussian noise. The equations (3.2-3) define a relaxation oscillator with stable limit cycle and two time scales (the oscillator is similar to the Van der Pol oscillator). In the problem of sound stream processing, the weights of oscillator connections were derived from cross-correlation information found at previous stage. Under the network dynamics, the synchronized clusters of oscillators (segments) are formed in this oscillatory network layer in correspondence with connected regions of acoustic energy in the time–frequency plane. Different synchronized oscillatory clusters are desynchronized, and thus, the first layer provides the decomposition of acoustic mixture into a collection of sensory elements. It corresponds to auditory scene decomposition into the set of primary structure elements.

In the next stage, oscillatory network of the second layer realizes further processing of the auditory stream – grouping of different sensory elements into components. Two types of connections were designed in the network of the second layer: internal connections depending on correlation information of the stream and external (vertical) connections from the network of the first layer. The network of the second layer is capable to reconstruct the acoustic stream components from previously obtained sensory elements. The reconstruction includes: (1) the reconstruction of the basic acoustic component, (2) the reconstruction of peripheral components and (3) the reconstruction of the "middle" area. All these steps are achieved via cluster synchronization of oscillatory network. The acoustic stream is represented in the form of the network decomposition into the set of internally synchronized and mutually desynchronized oscillatory clusters encoding proximity in frequency and time (special methods of filtra-

tion were also used at the first stage). Therefore, it is possible to interpret the acoustic stream reconstruction realized by the oscillatory network method as auditory scene synthesis from its elementary constituents.

The two-layer oscillatory network performance was tested on several stream examples with voiced speech perturbed by interfering sounds of the telephone mixture. Good network performance was demonstrated. As was stressed by the authors, the designed model includes only bottom-up processing corresponding to primitive segregation. But it is known that auditory scene analysis is notably influenced by the prior knowledge (the so-called schema-based organization). The model was further elaborated based on statistical analysis of general signal properties (in particular, via accounting for the uncorrelatedness of speech and noise), but without using oscillatory networks for acoustic stream decomposition and reconstruction. In this way, the approach to solution of speech segregation problem based on the analysis of perceptual and speech properties and speech enhancement was developed [48].

3.3 Oscillatory network models for visual image processing tasks

Networks of nonlinear coupled oscillators with controllable connectivity architectures were successfully exploited in various problems of image processing, such as brightness image segmentation, image contour extraction and object selection in a visual scene. A great diversity of problems related to image processing stimulated the development of complicated traditional computational methods of image analysis and transformation. The most simple brightness image segmentation tasks are computationally complicated ones, being related to processing of great amount of information. These tasks still remain a crucial problem in machine vision. Although the human brain performs image analysis efficiently and with apparent ease, it is still a major challenge for computer vision systems. Dynamical approaches to image processing based on construction of oscillatory network models and exploiting principles of dynamical binding via synchronization represent a promising alternative to traditional computational methods of image processing. Most of such approaches were developed based on neuromorphic models imitating various aspects of functioning of the brain neural structures. They provide self-organized adaptive algorithms with automatic performance. Several of the most valuable oscillatory network models based on systems of limit cycle oscillators are presented and shortly described in the section.

3.3.1 Oscillatory model by Malsburg and Buhmann

The oscillatory network model suggested by Malsburg and Buhmann [45] was one of the first models for solving the problem of object separation in a visual scene. The network consists of a collection of identical 2D subnetworks. The oscillators of each sub-

network are localized in the 2D grid, so that one oscillator corresponds to each image pixel. The subnetworks correspond to different image objects that need to be selected. The internal oscillator connections in each layer are designed to provide the synchronization of oscillator groups corresponding to the same image segment, the desynchronization between oscillator groups corresponding to different image segments and the damping of activity of other oscillators that do not correspond to some image segment. To this end three types of oscillator connections were designed: (1) global (one-to-one) connections between the oscillators in each layer (horizontal connections); (2) local connections between the layers; (3) connections with inhibitor units preventing global correlations. It proved possible to achieve internal synchronization of the oscillatory ensembles corresponding to each object and mutual desynchronization of ensembles corresponding to different objects. The performance of the model was demonstrated in the problem of segmentation of synthetic (36×36 pixels) images. Further model development allowed us to extend it to the processing of real grayscaled multipixel images, however, at the price of appropriate image preprocessing.

3.3.2 Model LEGION by Wang and Terman

The remarkable oscillatory network model for visual image segmentation is the oscillatory network Locally Excitatory Globally Inhibitory Oscillator Network (LEGION) designed first in 1995 [5, 6, 46, 49, 50]. The model is not directly related to modeling the brain visual processing. But nevertheless, its most perfect version [6] delivers highly effective dynamical image segmentation algorithm based on the synchronization in oscillatory networks. Active network unit is a relaxation (limit cycle) oscillator (of the Van der Pol oscillator type). The internal dynamics of a single oscillator is dependent on the external input defined by image pixel brightness, the local stationary excitatory connections with closest neighbors, the action of global inhibitor and Gaussian noise. Network oscillators are located in 2D spatial lattice that is in one-to-one correspondence with the image pixel array. In addition to stationary connections and global inhibitor, a dynamical coupling was designed in the network. Besides, a method of dynamical coupling adaptation was developed that allowed us to essentially improve the quality of network performance.

The governing dynamical system for the LEGION network can be written as (cf. (3.2-3)):

$$\begin{aligned} \dot{x}_{jk} &= 3x_{jk} - x_{jk}^3 + 2 - y_{jk} + I_{jk}H(p_{jk} - \theta) + S_{jk} + \rho, \\ \dot{y}_{jk} &= \epsilon \left[\gamma \left(1 + \tanh \frac{x_{jk}}{\beta} \right) \right] - y_{jk}, \end{aligned} \tag{3.3-1}$$

where $H(x)$ is the Heaviside step function, p_{jk} is the lateral potential of oscillator (j, k), θ is the threshold defining a specific character of oscillator role in the network dynamics (whether the oscillator is a leader), ρ is the amplitude of the Gaussian noise and

the parameter ϵ ($0 < \epsilon \leq 1$) is chosen to ensure standard relaxation oscillations in the absence of noise. The oscillator is capable to demonstrate the active phase (stable steady oscillations) at $I > 0$ and the silent phase (the absence of oscillations) at $I < 0$. The function $H(x)$ is introduced to provide a mechanism to distinguish between object and noise fragments. The coupling term S_{jk} in (3.3-1) is defined by the total coupling term S^a_{jk} with active neighboring oscillators and the weight W_z of inhibition from the global inhibitor activity z:

$$S_{jk} = S^a_{jk} - W_z H(z - \theta_z), \tag{3.3-2}$$

where the threshold θ_z defines the condition of action of global inhibitor on each network oscillator (the action of inhibitor is triggered when the oscillator is in the active phase). If the activity of every network oscillator is below the threshold θ_z, the global inhibitor does not receive any input, and $z \to 0$. In this case all network oscillators do not receive any inhibition. Thus, the LEGION dynamics can be qualitatively presented as a competition between excitative action of network oscillator interaction causing synchronization, and desynchronizing action of global inhibitor.

The dynamical adaptation of network connection weights is aimed at noise removal and image features preservation. It just allowed us to extend the network approach to real gray-level image segmentation tasks. The fixed oscillator connections are directly determined by the image structure. Qualitatively, the dynamical contribution to the stationary connection weight has a small value if two oscillators correspond to two neighboring pixels in an image region with homogeneous brightness. The dynamical coupling design requires the detailed analysis of possible brightness discontinuities across the whole image brightness structure. The features of the constructed dynamical weight adaptation scheme were empirically analyzed, and it was found that it is similar to that used in some nonlinear smoothing algorithm [6].

The LEGION network was thoroughly tested in different problems of image segmentation: noisy synthetic images, real satellite images, and magnetic resonance images. It was illustrated that the LEGION with dynamical coupling adaptation definitely demonstrate much better performance, providing high quality in the reproduction of sharp boundaries between image fragments. The oscillatory network with dynamically adapted connections can also provide correct segmentation of images that are highly corrupted by noise. Segmentation of satellite images is a difficult task due to the additional inhomogeneity of image subregions that are actually homogeneous, and most of traditional approaches have had only limited success. The simulations were carried out at solving of the following tasks: (a) entire image segmentation; (b) extraction of hydrographic objects in the image; (c) extraction of the topographical map. Solution of the task of hydrographic objects extraction requires the grouping together the pixels corresponding to water body and putting other objects into the background. When LEGION was applied to the problem, the hydrographic objects were identified as leaders, and a special set of parameters controlling dynamical oscillator connections was used. The objects were separated precisely, including noisy bank of the compli-

cated form. The problem of extraction of topographical map was also qualitatively solved. The satisfactory results were also obtained in the problems of entire segmentation of images highly corrupted by noise and magnetic resonance images segmentation. There was a series of images of human brain containing 43 regions (including the cerebral cortex, the cerebellum, the corpus callosum and fornix area). All the significant regions were segregated. Hence, the effectiveness of the LEGION algorithm was demonstrated.

It is possible to say that the LEGION model version [6] with dynamical network coupling adaptation provides good segmentation of real gray-level images containing more than 400 000 pixels. The LEGION approach was compared with the adaptive multiscale method of image segmentation [4] and with the adaptive smoothing algorithms [40, 41] that are widely used in noisy image segmentation tasks. As computation results showed, the oscillatory network LEGION with adaptive dynamical oscillator coupling adjustment considerably more precise image segmentation.

3.3.3 Three-dimensional columnar neural network model by Li

Biologically motivated neural model of the brain visual cortex was designed and developed by Li in the series of papers [31–34]. It is a recurrent three-dimensional neural network model with oscillatory dynamics imitating information processing in the brain primary visual cortex (VC) at solving contour separation and contour integration tasks in visual images. Functional network unit is a neural oscillator formed by a pair of interconnected cortical neurons – an excitatory pyramidal neuron and an inhibitory interneuron. Similar model of neural oscillator was previously proposed by Freeman [8], when prominent synchronous oscillations of 40–60 Hz were discovered in the rat and rabbit olfactory bulb and cortex. Following Freeman, Li and Hopfield suggested oscillator model closely imitating a real cortical neural oscillator, and used it in modeling the olfactory brain system where oscillations and synchronization play a key role in odor recognition [38] (see Section 3.2.1). A cortical oscillator of the VC model, proposed further in [31], reflects orientation-selective response of simple cells of the primary visual cortex. The dynamical system governing single oscillator dynamics can be written in terms of membrane potentials of excitatory and inhibitory neurons, x and y, that form the neural oscillators:

$$\begin{aligned} \dot{x} &= -x - g_y(y) + J_0 g_x(x) + I, \\ \dot{y} &= -y + g_x(x) + I_c. \end{aligned} \tag{3.3-3}$$

Here $g_x(x)$ and $g_y(y)$ are the threshold sigmoid nondecreasing functions defining the outputs of excitatory and inhibitory neurons (neuron activation functions), J_0 is the weight of the back-loop connection to the excitatory neuron, and I and I_c are the external inputs to the excitatory and inhibitory neurons, respectively. The neural oscillator dynamical system possesses a limit cycle in which size decreases smoothly

under gradual decrease of input value I. It converts into a stable focus at some threshold value $I = I^*$.

The spatially 3D neural network model was designed in [31] based on known experimental data on VC interconnection architecture and functioning. The most important features of VC internal structure – orientation selective cells, local cortical circuits and horizontal intracortical connections – were reflected in the model construction. The orientation selective neurons respond to input edges only within their classical receptive fields (RFs), the local small regions in the visual field. The direction of the preferred orientation is presented by a bar in the RF center. The brain visual cortex transforms visual inputs received by the retina into VC responses in the form of dynamical regimes of neural activity. The simple edge detection mechanism probably "works" just in the primary visual cortex [39]. Contour integration is likely completed by higher visual centers. The VC neural network dynamics demonstrates synchronized oscillations in the process performance of contour enhancement tasks.

Structurally the VC is composed of a great number of excitatory and inhibitory neurons. Visual input is received mainly by the excitatory neurons. In addition to almost local connections in VC, there exist long-range horizontal connections from upper visual cortex layers. Probably, they are responsible for synchronization of the 40–60 Hz oscillations [7, 10]. Usually the degree of synchronization decreases as the distance between neurons increases. Both the synchronization and the enhancement of responses were postulated as mechanisms underlying feature linking [9, 42].

The designed three-dimensional neural network of columnar architecture is associated with a 3D spatial lattice consisting of M vertical columns, each containing K sites. The column bases are located at the sites of some two-dimensional lattice in a plane, the column direction being normal to the plane (M is the number of 2D lattice sites). The 2D lattice can be chosen either as square or as the hexagonal one. The pair (j,k), $j = 1,\dots,M$, $k = 1,\dots,K$, of interconnected excitatory and inhibitory neurons is prescribed to each site of the 3D lattice, and the angle $\theta = k\pi/K$ of the RF orientation is prescribed to the (j,k) neuron pair. It is suggested that the excitatory neuron of each pair receives the visual input I_{jk} modeling the response of orientation selective cells of VC hypercolumn. The input received by excitatory neuron depends on the difference between the direction θ of the RF-orientation corresponding to the site k and the direction β of preferred orientation of the incoming visual image input. The received input is maximal when the visual input orientation coincides with the RF orientation prescribed to the site (j,k). In the model it is realized via some function $\phi(\theta-\beta)$ quickly decreasing at increased values of difference $\theta-\beta$. The dynamical system, governing by the neural network dynamics, is written as

$$\begin{aligned}\dot{x}_{jk} &= -\alpha_x x_{jk} - \sum_{k'} c_{kk'} g_y(y_{jk'}) + J_0 g_x(x_{jk}) + \sum_{j'\neq j,k'} J_{jkj'k'} g_x(x_{j'k'}) + I_{jk} + I_0,\\ \dot{y}_{jk} &= -\alpha_y y_{jk} + g_x(x_{jk}) + \sum_{j'\neq j,k'} W_{jkj'k'} g_x(x_{j'k'}) + I_c,\end{aligned} \tag{3.3-4}$$

where $J_{jkj'k'}$ are the excitatory-to-excitatory connections, $W_{jkj'k'}$ are the inhibitory-to-excitatory connections, and α_x, α_y are the constants.

Stationary excitatory and inhibitory connections of neurons of cortical circuits forming neural oscillator were constructed based on experimental neurobiological data concerning horizontal intracortical connections in VC. As one can see from dynamical system (3.3-4), the network connections are organized so that each excitatory neuron of a column is connected with all excitatory and all inhibitory neurons of all the other columns (the connections are inhomogeneous, the connection weights decrease at increasing the spatial distance between the network oscillators) and each excitatory neuron is connected with all inhibitory neurons of the same column (the connections are also inhomogeneous). There also exist some additional controlling feedback coupling (from higher visual cortex areas) for all excitatory neurons.

The model was tested in problems of preattentive image processing including contour segmentation, contour integration, texture segmentation and figure-ground segmentation tasks. It demonstrated quite successful synchronization-based performance in these image processing tasks for synthetic images and for small number of real gray-level ones. Some further developments of the described network model version were suggested in [32–36].

3.4 Pure oscillatory network model for image processing tasks

An oscillatory network of columnar architecture located in 3D spatial lattice was designed by the authors as an oscillatory network model of the brain visual cortex. It was designed to simulate a synchronization-based performance of a single preattentive bottom-up step of visual image reconstruction. A pure oscillatory level of modeling was of our particular interest. The model is entirely formulated in terms of a system of oscillators and their mutual interaction. Self-organized dynamical network coupling was constructed, and the idea of dynamical binding on the proximity of oscillator activity levels and their RF orientations was reflected in the designed network connectivity principle. The network performance consisted in relaxation into a stable stationary state of clusterized synchronization. Internally synchronized network ensembles (clusters) corresponding to image brightness fragments were expected to be formed. The three-dimensional network model was considered further as a basic model for the creation of a number of spatially two-dimensional models for different image processing tasks.

A single network oscillator is a relaxation (limit cycle) oscillator with internal dynamics tunable by visual image characteristics – local brightness and local image elementary bar orientation. It is able to demonstrate either the active state (stable undamped oscillations) or the “silent” state (quickly damped oscillations). Nonlocal self-organized dynamical connections of network oscillators depend on oscillator activity levels and orientations of RFs prescribed to the sites of a spatial lattice. The network

performance consists in network state transfer into a state of clusterized synchronization. The reduced spatially two-dimensional oscillatory network was extracted as a simplified limiting version of the three-dimensional source network model. As it turned out, the initial version of the reduced network model is capable to perform a number of gray-level image segmentation tasks via supplemented gradual network coupling strength adjustment. As a result, satisfactory synchronization-based brightness image segmentation of synthetic low-pixel images was achieved. Simple tasks of texture segmentation and contour integration were also performed in the frames of the 2D oscillatory network approach in the case of low-pixel synthetic images. The 2D network model was developed later to provide solution of various segmentation problems for real multipixel gray-level and color images (see Chapter 4).

3.4.1 Known data on the brain visual cortex taken into account in the oscillatory network model

Designing the network model, we focused our attention on the performance of the primary visual cortex in tasks of preattentive image reconstruction. The initial step in visual perception consists in focusing of a visual image on the retina. The external input stimulus received by the retina is further transferred through the so-called visual path – the sequential cascade of the brain neural structures, each being sufficiently complicated both structurally and functionally. It includes the retina itself, the visual nerve, the chiasma, the lateral geniculate body, the visual radiation, the primary visual cortex and the higher cortex zones. The function of the primary visual cortex consists in extracting of all visual information (forms, color, texture, and movement). Although the VC is the most studied brain structure, only a small part of answers to the question how the VC realizes the visual information processing, is known nowadays. The main visual processing occurs in the primary visual cortex. The most important feature of neurons of the brain visual system is the existence of the so-called RF for each neuron. The RF can be viewed as the total set of neurons of the previous neural structure of the visual path, from which visual stimulation provides the response of the given neuron. The RFs of both the retina and the visual cortex possess the internal spatial structure – the internal and the peripheral spatial zones. Moreover, there exist so-called “on”- and “off”-centers. For “on”-centers the internal RF zone is excitatory, and the external RF zone is inhibitory, and vice versa for “off”-centers [15, 39].

Visual image processing is fulfilled by the VC via several stages and is carried out by neurons of three types, namely, simple, complex and hypercomplex cells. Receptive fields of complex cells respond only to moving stimulus, those of hypercomplex cells – only to area boundaries and line ends. Their RFs possess central symmetry: the internal circular zone is encompassed by a ring-shaped external zone. The existence of complex and hypercomplex cells in the VC is not accounted in our network model, and at first we restrict ourselves to the processing of still images. Vision stereoscopy,

for which so-called eye-dominance columns are responsible, was also ignored. Hence, only the subset of simple cells (about one-fourth of total number of VC cells), responsible for processing of motionless images, was taken into account in the model. The orientation selectivity of a simple cell response is the most important feature of these neurons. It is realized due to strong anisotropy of their RFs: the form of the excitatory RF zone is similar to a narrow bar intersecting the RF center. The simple cell response is maximal when visual stimulus also possesses the form of a narrow bar, and the stimulus bar orientation is close to the orientation of the RF internal zone. The response of a simple cell is also absent if the stimulus bar is located completely in the inhibitory zone of the RF, but the stimulus and RF bars orientations coincide [15]. Further, we always assume the orientation of the internal excitatory bar zone of the RF of a simple cell to be the RF orientation.

As for VC spatial architecture, so-called orientation hypercolumns (the neuron columns) represent functional modules of the VC. At simplified consideration, one can imagine the hypercolumns as oriented orthogonally to the visual cortex surface. So, a flat approximation – the plane retina and the VC in the form of parallelepiped – is acceptable. At last, the key feature is that the visual cortex is composed of excitatory and inhibitory neurons coupled by short internal connections. Excitatory neurons receive inputs from the retina. They "send" their outputs into higher zones of the visual cortex and receive feedback loop connections from there. In addition, there exists the coupling of VC neurons with close RF orientations. It is believed that these connections provide the experimentally discovered synchronized oscillations with frequency of 40–60 Hz in the VC. The existence of coupled pairs of excitatory and inhibitory neurons allows to say about neural oscillators in the VC. Therefore, the designed network models where oscillators figure as processing units represent a natural step in modeling the brain structures, which exploit oscillations and synchronization in their functioning.

3.4.2 Three-dimensional oscillatory network of columnar architecture

We designed oscillatory network consisting of neural oscillators localized at the nodes of three-dimensional spatial lattice being in one-to-one correspondence with a two-dimensional image pixel array located in a two-dimensional plane. The oscillatory network can be interpreted as an oscillatory model of the primary visual cortex (more exactly, the subset of VC simple cells). Let a 2D square lattice be related with the image pixel array of $M \times N$ pixels, which centers are located at the nodes of the lattice $G_{M\times N}^{(\mathrm{im})}$. We then design the oscillatory network consisting $M \times N$ oscillator columns each of K oscillators, so that the bases of the columns are located at the nodes of the 2D lattice $G_{M\times N}^{(\mathrm{net})}$ identical to $G_{M\times N}^{(\mathrm{im})}$, whereas oscillators of each column are located at the nodes of the 1D lattice $L_K^{(\mathrm{net})}$ oriented normally to the plane that contains $G_{M\times N}^{(\mathrm{net})}$. Thus, the oscillators of the whole network, containing $M\times N\times K$ oscillators, are located at the

nodes of the 3D spatial lattice $G^{(\mathrm{net})}_{M\times N} \times L^{(\mathrm{net})}_{K}$, and each oscillator column corresponds to some image pixel.

We prescribe to each network oscillator an internal parameter – the RF orientation specified by two-dimensional vector $\mathbf{n}^k_{jm} \equiv (\cos\psi^k_{jm}, \sin\psi^k_{jm})$ located in the plane orthogonal to the column direction. According to neurobiological data [15], vectors $\mathbf{n}^k_{jm}$ are uniformly distributed over each column, that is

$$\psi^k_{jm} = \psi^0_{jm} + \frac{k\pi}{K}, \quad k = 1, \ldots, K. \tag{3.4-1}$$

Further, the additional set of image characteristics is included into consideration – the set of unit vectors $\mathbf{s}_{jm}$ defining pixel elementary bar orientations. These parameters can be interpreted as a measure of pixel brightness I_{jm} inhomogeneity (the $\mathbf{s}_{jm}$ are orthogonal to the direction of pixel brightness gradient). We proposed that the set of $\mathbf{s}_{jm}$ can be extracted via proper image preprocessing, and so the array $[I_{jm}, \mathbf{s}_{jm}]$ of image characteristics can be defined. The array $[I_{jm}, \mathbf{s}_{jm}]$ figures in the model as the set of tuning parameters controlling oscillatory network dynamics.

3.4.3 Biologically inspired model of single network oscillator

In the design of network oscillator internal dynamics, we took into account and preserved the main features of dynamics of cortical neural oscillator (3.3-3) introduced in [31]. Preliminary analysis of dynamics of oscillator (3.3-3) shows that the dynamical system possesses a stable limit cycle when the value of external input I belongs to some finite interval. The limit cycle size monotonically depends on I: it decreases smoothly when the value I is gradually decreased. The limit cycle converts into a stable focus at some threshold value $I = I_{min} = I^*$.

After clarifying the dynamical features inherent to biologically motivated neural oscillator (3.3-3), we designed single network oscillator dynamics of our model based on proper modification of Ginzburg–Landau oscillator and preserving qualitatively the features of the oscillator dynamics (3.3-3), including the dependence of single oscillator dynamics on two image characteristics – pixel brightness and pixel elementary bar orientation.

Defining the oscillator state by a pair of variables (u_1, u_2), it is possible to write the system of two coupled ODE, governing the oscillator dynamics, in the form of a single equation for the complex-valued variable $u = u_1 + iu_2$:

$$\begin{aligned} &\dot{u} = f(u, \mu), \quad f(u, \mu) = (\rho_0^2 + i\omega - |u - c|^2)(u - c) + \mu, \\ &\mu = g(I, \mathbf{s}, \mathbf{n}) = p(I) + q(\mathbf{s}, \mathbf{n}). \end{aligned} \tag{3.4-2}$$

Here ρ_0, c and ω are the constants defining the parameters of the limit cycle of the the dynamical system (3.4-2): the limit cycle of the dynamical system (3.4-2) is the circle of radius ρ_0 with center located at the point (c_1, c_2), $c = c_1 + ic_2$, in the plane (u_1, u_2),

ω being cycle frequency. At $\mu > 0$ the size and the location of the limit cycle can be controlled by parameters I and $\mathbf{s}$ via properly designed functions $p(I)$ and $q(\mathbf{s}, \mathbf{n}) = q(|\beta - \psi|)$, where $\phi = \beta - \psi$ is the angle between the elementary bar orientation $\mathbf{s} = (\cos\beta, \sin\beta)$ and the RF orientation $\mathbf{n} = (\cos\psi, \sin\psi)$. The convenient functions for limit cycle control are:

$$\begin{aligned} p(I) &= 1 - H(I - h_0), \quad H(x) = \frac{1}{1 + e^{-\nu x}}, \qquad \nu \gg 1, \\ q(|\phi|) &= 1 - \Gamma(|\phi|), \quad \Gamma(|\phi|) = \frac{2e^{\sigma|\phi|}}{1 + e^{-2\sigma|\phi|}}, \qquad \sigma \gg 1. \end{aligned} \tag{3.4-3}$$

The parameter μ is a bifurcation parameter of dynamical system (3.4-2): the limit cycle radius $\rho = \rho(\mu)$ (the oscillation amplitude) is maximal at $\mu = 0$ ($\rho(0) = \rho_0$). It monotonically decreases at μ decreasing and bifurcates into a stable focus at some $\mu = \mu_*$, $\mu_* \in (0, 1)$. Due to the designed dependence μ on I and $\mathbf{s}$, the limit cycle size is sufficiently large if two following conditions are satisfied simultaneously: (a) I essentially exceeds the threshold value h_0; (b) the angle between $\mathbf{s}$ and $\mathbf{n}$ is sufficiently small. Otherwise, either the cycle size is very small or it degenerates into a stable focus (that corresponds to quickly damping oscillations). The bifurcation character of single oscillator dynamics is essentially exploited in image processing tasks performed by the oscillatory network. The response of a single network oscillator to variation of

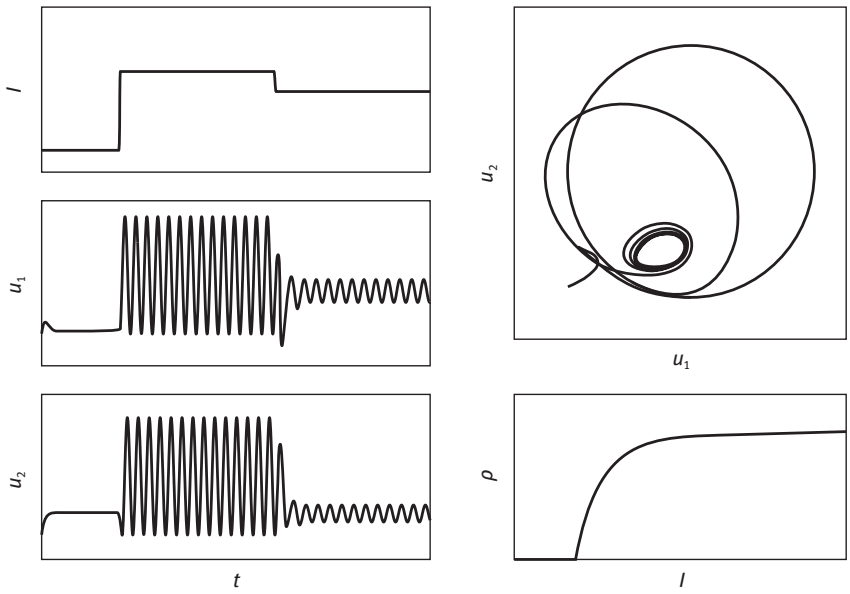

Fig. 3.1. The "response" of single network oscillator to variation of pixel brightness: temporal behavior of both oscillator state variables u_1 and u_2 (left); the trajectory of dynamical system (upper right); the limit cycle size dependence on I (lower right).

pixel brightness is shown in Figure 3.1. As one can see, the oscillator demonstrates almost instantaneous response to a sudden decrease of pixel brightness via oscillation amplitude reduction.

3.4.4 Self-organized dynamical network coupling

The principle of oscillatory network coupling has a key influence on the network dynamical behavior. We design self-organized dynamical coupling that nonlinearly depends on oscillatory activity levels (amplitudes of oscillation of network oscillators) and differences between **s** and RF orientations. The dynamical system governing the network dynamics can be written in the form

$$\begin{gathered}\dot{u}^k_{jm} = f\left(u^k_{jm}, \mu^k_{jm}\right) + S^k_{jm}, \\ j = 1, \ldots, M, \quad m = 1, \ldots, N, \quad k = 1, \ldots, K.\end{gathered} \tag{3.4-4}$$

Here $f(u, \mu)$ is defined according to (3.4-2) and $\mu^k_{jm} = p(I^k_{jm}) + q(\mathbf{s}^k_{jm}, \mathbf{n}^k_{jm})$. The term S^k_{jm}, specifying the interaction between network oscillators, is chosen in the form

$$S^k_{jm} = \sum_{j'm'k'} W^{kk'}_{jmj'm'}\left(u^k_{jm} - u^{k'}_{j'm'}\right). \tag{3.4-5}$$

The weights $W^{kk'}_{jmj'm'}$, defining the connection strength of network oscillators (j, m, k) and (j', m', k'), are specified by nonlinear functions that depend on state of the oscillator pair. The form of the functional dependence influences crucially both network dynamics and performance.

We constructed the functions $W^{kk'}_{jmj'm'}$ based on results of our previous mathematical study of synchronization in oscillatory networks, governed by more simple version of dynamical equations (3.4-4) and (3.4-5) (at $c = 0$, $\mu^k_{jm} = 0$, $W = \text{const}$) [20–22]. The following expressions for the matrix $\hat{W}$ elements were chosen:

$$W^{kk'}_{jmj'm'} = P\left(\rho^k_{jm}, \rho^{k'}_{j'm'}\right) Q\left(\mathbf{n}^k_{jm}, \mathbf{n}^{k'}_{j'm'}\right) D\left(\mathbf{r}^k_{jm}, \mathbf{r}^{k'}_{j'm'}\right), \tag{3.4-6}$$

where ρ^k_{jm} and $\rho^{k'}_{j'm'}$ are the limit cycle radii for oscillators defined by indices (j, m, k) and (j', m', k'), $\mathbf{n}^k_{jm}$ and $\mathbf{n}^{k'}_{j'm'}$ are the RF orientations for these oscillators, $\mathbf{r}^k_{jm}$ and $\mathbf{r}^{k'}_{j'm'}$ are the radius vectors defining their spatial locations in the network lattice. The cofactors $P(\rho, \rho')$, providing the connectivity principle dependence on oscillator activities, are chosen in the form

$$P(\rho, \rho') = w_0 H(\rho\rho' - h), \tag{3.4-7}$$

where $H(x)$ is the sigmoid function dependent on the threshold h, w_0 is the constant defining the total strength of network interaction. As it is clear from (3.4-7), the connection weight (3.4-6) is negligible if at least one of interacting oscillators is in the state of

low activity. The cofactors $Q(\mathbf{n}, \mathbf{n}')$, providing dependence on RF orientations, are defined in terms of a delta-shaped function Γ that depends on the orientation difference between $\mathbf{n}$ and $\mathbf{n}'$:

$$Q(\mathbf{n}, \mathbf{n}') = \Gamma(|\mathbf{n} - \mathbf{n}'|). \tag{3.4-8}$$

So, $Q(\mathbf{n}, \mathbf{n}')$ is nonzero only if the orientations $\mathbf{n}$ and $\mathbf{n}'$ are sufficiently close. The construction of $Q(\mathbf{n}, \mathbf{n}')$ reflects the neurobiologically confirmed fact of preferable connectivity of VC neurons with close RF orientations.

At last, the cofactors $D(\mathbf{r}, \mathbf{r}')$, permitting to control spatial radius of oscillator interaction, can be defined by any function vanishing at some finite spatial distance. For example, $D(\mathbf{r}, \mathbf{r}')$ can be chosen in the form

$$D(\mathbf{r}, \mathbf{r}') = 1 - H(|\mathbf{r} - \mathbf{r}'| - r_0), \tag{3.4-9}$$

where r_0 is the given radius of spatial interaction. In particular, one can define local interaction (with closest neighboring oscillators), or the global oscillator coupling (all-to-all).

As a result, according to the connectivity rule (3.4-6), any two network oscillators are proved to be sufficiently strongly dynamically coupled if they both are active, possess close RF orientations and are located at the distance not exceeding the prescribed radius of spatial interaction.

3.4.5 Reduced two-dimensional oscillatory network

The 3D oscillatory network can be naturally reduced to its limiting version – the 2D network in which oscillators are located in the nodes of the lattice $G_{M\times N}^{(\text{net})}$ and can be interpreted as idealized oscillator columns. The reduced network can be obtained in the following way. Let us fix parameters $I \geq h$ and $\mathbf{s}$ of some pixel of sufficient brightness and consider the response of the oscillator column corresponding to that pixel. Obviously, only several neighboring oscillators in the column will be active, i.e. those that possess RF orientations close to $\mathbf{s}$. Let the number of oscillators in the column be gradually increased and at the same time the width of the function Γ be reduced. Then the number of active oscillators in the column will be gradually decreased, and in the limit of an infinitely long column and an infinitely narrow Γ, we would get the single active oscillator in each column, namely, the one for which the RF orientation coincides with $\mathbf{s}$.

Hence the response of an infinitely long idealized column is reduced to the response of a single oscillator in this column. Its dynamics is governed by the system (3.4-2), where the function $g(I, \mathbf{s}, \mathbf{n})$ is reduced to $g(I, \mathbf{s}, \mathbf{n}) = 1 - H(I - h_0) = G(I)$. The 2D reduced network of these oscillator-columns is located in the $G_{M\times N}^{(\text{net})}$ lattice, which is in one-to-one correspondence with the retina lattice $G_{M\times N}^{(\text{im})}$. Therefore, one oscillator of the reduced network now corresponds to one image pixel (see Figure 3.2).

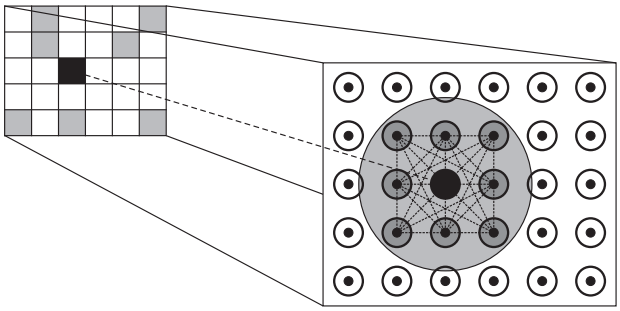

Fig. 3.2. Spatial architecture of the reduced two-dimensional oscillatory network. A fragment of image pixel array and the corresponding fragment of the oscillatory network are shown. The circle, defining the subset of network oscillators that will be dynamically coupled with a selected one, is marked by gray color.

The reduced network state is defined by the $M \times N$ matrix $\hat{u} = [u_{jm}]$, and network dynamical equations are:

$$\dot{u}_{jm} = \left(\rho_0^2 + i\omega_{jm} - |u_{jm} - c|^2\right)(u_{jm} - c) + G(I_{jm}) + \\ + \sum_{j'm'} W_{jmj'm'}(u_{j'm'} - u_{jm}), \tag{3.4-10}$$

where

$$W_{jmj'm'} = P(\rho_{jm}, \rho_{j'm'})Q(\mathbf{s}_{jm}, \mathbf{s}_{j'm'})D(\mathbf{r}_{jm}, \mathbf{r}_{j'm'}). \tag{3.4-11}$$

The cofactors P and Q in (3.4-11) are calculated in the same manner as in the case of 3D network, but the cofactor Q now depends only on image bar orientations. As it turned out, the presence of the cofactor Q in network connections provides the network with the capability to perform some texture segmentation tasks.

3.4.6 Brightness image segmentation via the reduced oscillatory network

The reduced network is capable to provide segmentation of pure brightness images, for which information on elementary bar orientations is absent. For simplicity we can suppose that for brightness images the bar orientations are the same for all the pixels and put $\mathbf{s}_{jm} = \text{const}$. Then, $Q(\mathbf{s}_{jm}, \mathbf{s}_{j'm'}) = 1$ in (3.4-11), and the dependence on bar orientations disappears. Due to the one-to-one correspondence between image pixels and the oscillators of the 2D network, in the initial network state the distribution of oscillator activities exactly corresponds to the pixel brightness distribution. The grouping of pixels into a whole image fragment will be achieved via synchronization of oscillators with close activities.

To improve the network performance in brightness image segmentation tasks (in particular, to achieve accurate detection of image fragment boundaries), we included

an additional procedure of synchronization control. It is a simple algorithm of gradual increasing of total network coupling strength, and it permits to realize successive selection of synchronized assemblies (clusters) corresponding to image fragments of different brightness levels. To this aim a matrix $\hat{N} = [\gamma_{jm}]$ of additional control parameters is introduced, and we put at first $\gamma_{jm} = 0$. A modified matrix $\hat{W}$ of network connections depending on γ_{jm} is further considered instead of $\hat{W}$: $\tilde{W}_{jmj'm'} = W_{jmj'm'}\Gamma(|\gamma_{jm} - \gamma_{j'm'}|)$. At the beginning of the interaction strength control stage we specify the initial interaction to be so weak that the network is completely desynchronized. This is realized by a choice of a sufficiently high initial threshold value h in equation (3.4-7) for the cofactor P. Further we gradually increase the $\tilde{W}$ elements via decreasing h up to the moment of synchronization of the first network cluster. It happens at some $h = h_1$. The first cluster, formed by oscillators of maximal activity, corresponds to an image fragment with maximal brightness. Under further decreasing h inside some interval (h_2, h_1), the first cluster remains as the single synchronized cluster of the network. Such oscillatory network behavior is a consequence of monotonic dependence of the oscillator limit cycle size on the pixel brightness I. When the first cluster is synchronized, whereas the rest network is desynchronized, we separate the cluster via excluding it from the interaction with all other network oscillators. It is achieved by means of the matrix $\hat{N}$ modification: we prescribe some nonzero value γ_1 to those components of $\hat{N}$ that correspond to spatial locations of oscillators belonging to the synchronized cluster. Further we continue the above process of interaction strengthening until the second cluster is synchronized, and after that it is excluded. Finally, all the clusters will be sequentially synchronized and separated. Thus, the network will be decomposed into a set of internally synchronized, but mutually desynchronized clusters, corresponding to image fragments of different brightness levels. Moreover, in the final state the desynchronized clusters oscillate with slightly different frequencies, which provides an additional tool for the analysis of the segmentation result. The described procedure of interaction control has been fulfilled manually. However, an automatic interaction control performance could be realized. In a series of computer experiments on segmentation of synthetic brightness images a satisfactory network performance was demonstrated. The example of processing of the synthetic image containing 2460 pixels is given in Figure 3.3.

The network state during image processing is represented in the form of the array of all the oscillator states (what exactly reproduces the current state of image pixel array), the brightness of each screen pixel being correspondent to $|u_{jm}(t)|$. Since different brightness fragments oscillate with slightly different frequencies, all image fragments are clearly distinguishable. A large number of different "versions" of segmented image is contained in the full set of oscillatory network states, which is quite helpful in the situations when some ambiguous image fragments exist (for instance, contours of low contrast). Several typical examples of such instantaneous image versions are presented in Figure 3.4. Some simple postprocessing is needed to extract the final image segmentation result from the collection of instantaneous images contained in the full set of the oscillatory network states arising in the oscillatory segmentation procedure.

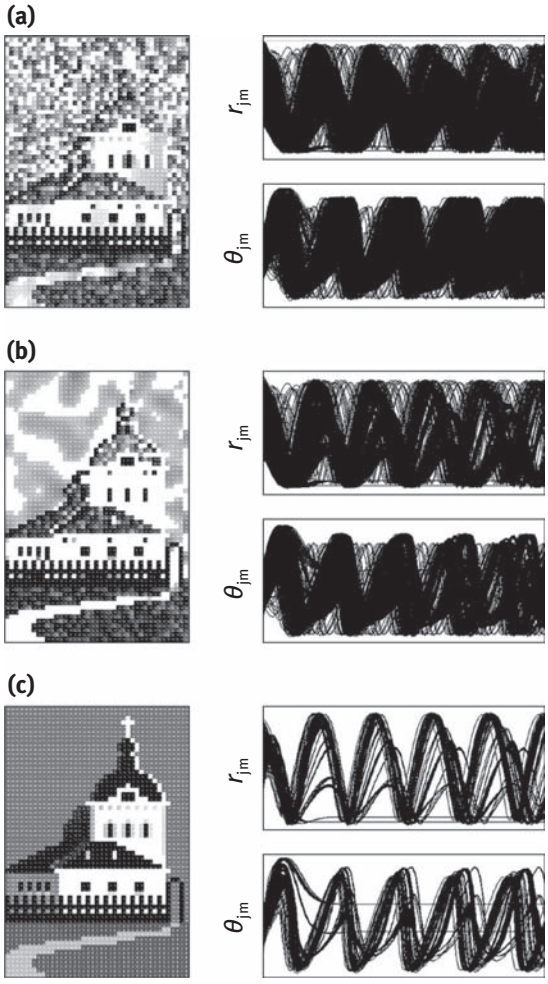

Fig. 3.3. Brightness segmentation of synthetic image of 2460 pixels. Three typical states of oscillatory network in the process of interaction strengthening: (a) total desynchronization ($h = 1.25$); (b) partial synchronization ($h = 0.5$); (c) almost complete synchronization ($h = 0.01$). The examples of instantaneous network states (left) and the temporal dependence of all oscillator variables (right) are shown ($r_{jm} = |u_{jm}|$, $\theta_{jm} = \arg u_{jm}$).

Note that there is an evidence of a sequential type of image processing performed by the brain visual cortex, i.e. image fragments of different brightness are not processed simultaneously. Instead, there is some time delay in fragment reproduction: the most bright fragments are reproduced faster than the less bright ones [51]. Probably such a behavior is achieved via additional image processing fulfilled in high cortical areas.

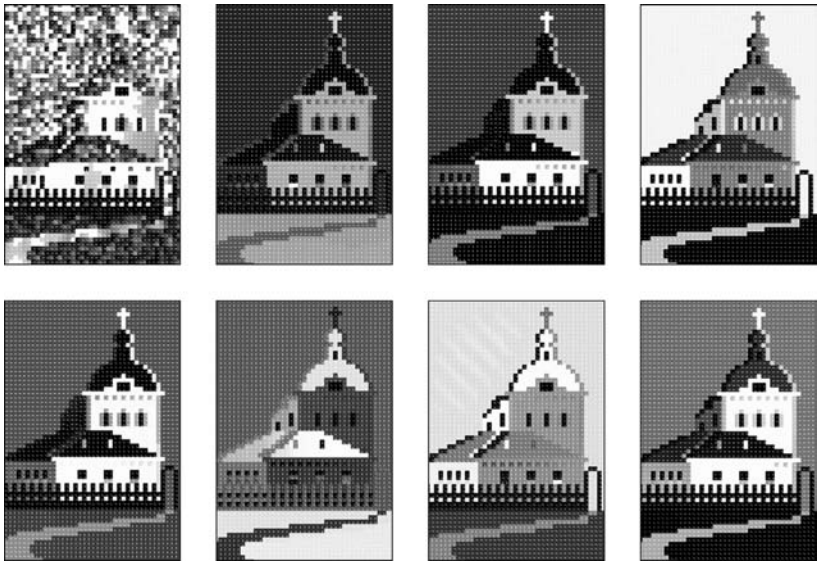

Fig. 3.4. The examples of instantaneous image "versions," arising during the oscillatory procedure of brightness image segmentation, carried out by the oscillatory network. The initial state of desynchronized network is shown in the left upper picture.

3.4.7 Texture segmentation and contour integration

The processing of texture visual images is usually regarded as a special class of problems in the field of traditional computer vision. In particular, there exist special methods of texture representation and synthesizing. In the frames of our approach it is possible to include into consideration only the simplest texture types that can be represented as collections of oriented bars.

The 2D reduced network is capable to process texture images due to the dependence on bar orientations preserved in the connectivity rule (3.4-11). In the general case of texture image segmentation problems, one has to deal with brightness–texture image fragments instead of pure brightness ones. In the first series of computer experiments on texture image segmentation we processed images with monodirected textures and concentrated our attention on texture images with homogeneous mean brightness. The network performance is based on desynchronization of clusters corresponding to different texture fragments. Unlike the case of pure brightness image segmentation, network coupling control is unnecessary here. Two examples of the texture image segmentation are shown in Figure 3.5. The source texture images are presented in the left panels, being defined in the form of the oscillator phase distribution at the initial network state. Two instantaneous states of synchronized network are shown in the middle and the right panels. Desynchronized clusters, corresponding to different

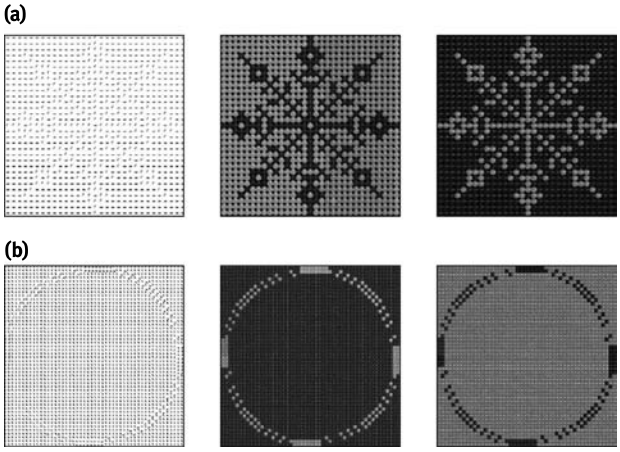

Fig. 3.5. Texture segmentation and contour integration via the 2D oscillatory network: (a) segmentation of a contour selected solely by monodirected texture; (b) segmentation of a double contour selected solely by a texture of continuously varied orientation.

texture fragments, are in different phases of oscillation and therefore are accurately segmented. For the processed texture-marked contour shown in Figure 3.5 (a) with the complicated, "fractal"-like form it was necessary to choose a sufficiently large spatial radius of network coupling. A solution of a simple contour integration task, fulfilled by the reduced oscillatory network, is demonstrated in Figure 3.5 (b). The image shown in the left panel contains two closed contours marked solely by texture, which is defined by oriented bars of continuously varied direction (approximating local contour tangent). As one can see, the network provides accurate segmentation of the double contour via its desynchronization with respect to the background.

3.4.8 Comparison of the model with other oscillatory network models

There are two oscillatory network models for image segmentation that are closely related to ours. The first one is the model by Li developed in the series of papers [31–34]. The following distinctions of our model from the model by Li should be marked:

(1) Two features of single oscillator dynamics designed in our model – the bifurcation character of oscillator dynamics and the monotonic dependence of oscillation amplitude on pixel brightness – were actively exploited in various tasks of image processing, demonstrating good performance of the reduced 2D oscillatory network extracted from the 3D oscillatory network model of the visual cortex.
(2) Self-organized dynamical network coupling is employed in our model: the network connections automatically emerge after tuning of single oscillator dynam-

ics by image characteristics (I_{jm}, $\mathbf{s}_{jm}$) (in contrast, stationary excitatory and inhibitory connections, designed in the model by Li, cannot be considered as self-organized ones).

(3) The reduced oscillatory network supplemented by the method of interaction adjustment demonstrated a capability of detailed brightness image segmentation (whereas the model by Li is designed mainly for contour integration and texture segmentation tasks).

The following remarks are also necessary concerning the relation of our 2D reduced network model to the oscillatory network model LEGION. The initial model described in Sections 3.4.5–3.4.7 required a considerable improvement to be used in real image segmentation problems. The LEGION was successfully used in various tasks of real gray-level image processing (see Section 3.3.2). Our 2D model was essentially improved and developed, including creation of new approaches, to provide successful processing of real gray-level and color images. However, some advantages of the described initial 2D model over the LEGION model should be mentioned:

(1) There exist only dynamical connections in our model (stationary connections are absent as unnecessary); this ensures the model simplicity and provides a flexibility of its performance.

(2) The network coupling principle constructed in our model leads to automatic origin of self-organized dynamical connections in oscillatory network and ensures flexibly controlled synchronization (in contrast, special sophisticated calculations are necessary for LEGION model to determine, for each network oscillator, the proper set of oscillators that should be coupled with it in a concrete image processing task; our method of interaction strength adjustment is definitely more simple as compared to algorithm of interaction adaptation developed in the last version of the LEGION model [6].

The described 2D oscillatory network model was further developed in the following directions:

(1) a new version of single oscillator dynamics was designed, providing more convenient and flexibly controlled oscillator response to pixel characteristics;

(2) several new principles of self-organized network coupling were developed and tested in various image processing tasks.

It allowed us to qualitatively solve a number of image processing tasks, including real (multipixel) gray level and color image segmentation, selective image segmentation and object selection in a visual scene (see Chapter 4).

3.5 Oscillatory network models including visual attention

The brain has to process large amounts of information. But it can carry out only a limited number of tasks at a time, and so it needs a capability to select the most relevant information. Attention is the cognitive process of selective concentration on one aspect of the environment while ignoring the others. Visual attention is believed to operate as a two-stage process. At the first stage attention is distributed uniformly over the visual scene, and information processing is performed in parallel. At the second stage attention is concentrated to a specific area of the visual scene, and information processing is performed in a serial manner. So, attention can be viewed as a kind of filtration. The attention is often related to the visual search process (active scan of a visual scene). In visual search attention is usually directed to the item with the highest priority.

In computation modeling of visual attention the prevailed approaches are related to reducing the amount of information to be processed. The approaches are concerned with the branches of active vision and selective attention. As a whole these areas stay out of our interests. So we present here only two oscillatory network models for visual image processing, supplemented by attention-like algorithm of visual information filtering. In the model by Labbi et al. [28, 29] a special network layer was introduced, named attention map, that realizes additional image processing via oscillator cluster desynchronization. It can be interpreted as "automatic attention focusing." In the series of oscillatory network models with central oscillator element [1, 2, 16] the attention is imitated via different states of partial synchronization for oscillatory network of peripheral oscillators (what is provided via synchronization of appropriate subensembles of the network of peripheral oscillators with the central oscillator). The central element, which consists of one or two neurons, reflects a hypothesis on existence of central element of the brain attention system, localized in the brain prefrontal cortex.

3.5.1 Oscillatory model by Labbi, Milanese, and Bosch

The oscillatory network model was designed for gray-level image segmentation based on the imitation of the visual cortex processing [28, 29]. It contains an array of FitzHugh–Nagumo oscillators governed by the dynamical system

$$\begin{aligned} \dot{u} &= u - \frac{u^3}{3} - v + I, \\ \dot{v} &= \epsilon(a + bu - v). \end{aligned} \tag{3.5-1}$$

Depending on the parameters a, b, ϵ, and I, the oscillator demonstrates either the active state (relaxation auto-oscillations) or the passive state (the absence of oscillations).

The oscillatory network of three-layered architecture was designed. The first (input) layer contains the pixel array of an image to be processed. The second layer consists of locally coupled FitzHugh–Nagumo oscillators being in one-to-one correspondence with the image pixel array. The external input for each oscillator of the second layer depends on the brightness of the corresponding pixel located in the first layer. The oscillation amplitude of the oscillators increases monotonically as a function of the input. The local connections designed in the second layer provide cluster synchronization of the oscillatory network (each cluster is supposed to correspond to the appropriate image segment). The third layer (named the attention map) is designed to introduce the desynchronization between the clusters corresponding to different spatially separated image objects. The connections of network oscillators of the third layer are organized in such a way as to provide detection of synchronized oscillator groups in which amplitudes strongly differ from oscillator amplitudes of the background.

The model demonstrated satisfactory performance for simple gray-level images (with 128×128 pixel array) containing several clearly separated objects. As was clarified in computer experiments, only large groups of neighboring oscillators that receive similar inputs are capable to synchronize.

3.5.2 Oscillatory neural network model by Borisyuk and Kazanovich

An oscillatory neural network model with imitation of selective visual attention was developed by Borisyuk and Kazanovich [1, 2, 16]. The model was designed as a network of locally coupled oscillators (peripheral oscillators) located in a two-dimensional spatial lattice. These oscillators are controlled by a central oscillator located outside the lattice and coupled with all the peripheral oscillators by feedforward and feedback connections. The peripheral oscillators imitate functioning of hypercolumns of the primary visual cortex, and the central oscillator plays a role of an attention control center, which is located in the brain septum–hippocampus neural structure. The state of a network oscillator is specified by the amplitude and the phase of oscillations, both dependent on the phase of the central oscillator. The dynamical system, governing network dynamics, is written in the form of three coupled ordinary differential equations. A single oscillator demonstrates periodical auto-oscillations. The pure phase interaction of network oscillators was designed. The attention is modeled via the synchronization of the central oscillator with the appropriate subensemble of peripheral network oscillators. In the frames of the model it was possible to model the phenomenon of the visual perception ambiguity [3].

As a further progress, a new version of the model was designed as a two-layer network of spiking peripheral neurons of the Hodgkin–Huxley type, containing excitatory and inhibitory connections, controlled by a central neuron [1]. The neurons of the first layer (peripheral neurons) are located in a 2D square lattice identical to that of the image pixel array, and coupled both via excitatory internal connections and with

the central element. The dynamics of a single neuron is written in the form

$$\begin{aligned} \dot{v} &= -I_{\text{ion}} + I_{\text{ext}} - I_{\text{syn}}, \\ \dot{x} &= A_x(v)(1-x) - B_x(v)x, \quad x \in \{m, h, n\}, \end{aligned} \tag{3.5-2}$$

where v is the membrane potential of the neuron, x is the variable defining the conductance of an ion channel (m, n, and h are the gating variables of the ionic channels, related to possible different colors of image objects), I_{ion} is the total ionic current, I_{ext} is the combination of constant current component and Gaussian noise, $A_x(v)$ and $B_x(v)$ characterize the dependence of membrane current on v. The value I_{syn} defines the synaptic current, provided by coupling with the neighboring network neurons. The conductance-based equations are used to determine the value of I_{syn}.

The central element consisting of two neurons is used for simplified imitation of the brain attention system. The image object in a visual scene is represented by an appropriate subensemble of peripheral neurons. The subensemble is capable to generate spikes synchronously with the central element, whereas the activity of the rest peripheral neurons is suppressed. Such regime of partial synchronization imitates the attention focusing. The network performs three different image processing tasks – selective attention, raw contour extraction and object separation. At the contour extraction stage, standard computational algorithms of contour extraction were attracted (the method of gradient calculation combined with Gabor filtering approach). At object separation stage, peripheral neurons "work" in the regime of partial synchronization with the central oscillator. The results of the model performance were demonstrated in the task of sequential object selection by a robot camera in a simple visual scene, containing four balls of different colors. As was found, local network excitatory connections facilitated synchronization with the central element. Pattern recognition considerations were not attracted in the frames of the model.

Bibliography

[1] R. Borisyuk, Y. Kazanovich, D. Chik, V. Tikhanoff and A. Cangelosi. A neural model of selective attention and objecy segmentation in the visual scene: An approach based on partial synchronization and star-like architecture of connections. *Neural Networks*, 22:707, 2009.

[2] R. M. Borisyuk and Y. B. Kazanovich. Oscillatory model of attention-guided object selection and novelty detection. *Neural Networks*, 17:899, 2004.

[3] R. M. Borisyuk and Y. B. Kazanovich. Oscillatory model of perception of dual images. In *Proc. Int. Conf. "Neuroinformatics-2009"*, volume 1, page 281, 2009. in Russian.

[4] T. Caelli and W. F. Bischof. *Machine Learning and Image Interpretation*. Plenum Press, 1997.

[5] E. Cesmeli and D. L. Wang. Texture segmentation using Gaussian–Markov random fields and neural oscillator networks. *IEEE T. Neural Networ.*, 12:394, 2001.

[6] K. Chen, D. Wang and X. Liu. Weight adaptation and oscillatory correlation for image segmentation. *IEEE T. Neural Networ.*, 11:1106, 2000.

[7] R. Eckhorn, R. Bauer, W. Jordan, M. Brosch, W. Kruse, M. Munk and H. J. Reitboek. Coherent oscillations: a mechanism of feature linking in the visual cortex? multiple electrode and corre-

lation analyses in the cat. *Biol. Cybern.*, 60:121, 1988.
[8] W. J. Freeman. Spatial properties of an EEG event in the olfactory bulb and cortex. *Electroen. Clin. Neuro.*, 44:586, 1978.
[9] C. M. Gray. Synchronous oscillations in neuronal systems: mechanisms and functions. *J. Comput. Neurosci.*, 1:11, 1994.
[10] C. M. Gray and W. Singer. Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proc. Natl. Acad. Sci. USA*, 86:1698, 1989.
[11] E. Grichuk, M. Kuzmina and E. Manykin. Oscillatory network for synchronization-based adaptive image segmentation. In *Proc. IJCNN'06*, page 4529, 2006.
[12] E. Grichuk, M. Kuzmina and E. Manykin. Image processing via synchronization in self-organizing oscillatory network. In *Proc. ECC'07, Vol. 1*, volume 27 of *Lecture Notes in Electrical Engineering*, page 97, 2009.
[13] E. S. Grichuk, M. G. Kuzmina and E. A. Manykin. Oscillatory network for synchronization-based adaptive image segmentation. *Opt. Mem. Neural Networks*, 15:57, 2006.
[14] E. S. Grichuk, M. G. Kuzmina and E. A. Manykin. Object selection in visual scene via oscillatory network with controllable coupling and self-organized performance. *Opt. Mem. Neural Networks*, 20:113, 2011.
[15] D. H. Hubel. *Eye, brain, and vision*. W. H. Freeman, 1995.
[16] Y. Kazanovich and R. Borisyuk. Object selection by an oscillatory neural network. *Biosystems*, 67:103, 2002.
[17] M. Kuzmina and E. Manykin. Biologically motivated oscillatory network model for dynamical image segmentation. In *Proc. BICS'04*, 2004.
[18] M. Kuzmina, E. Manykin and I. Surina. Oscillatory network with self-organized dynamical connections for synchronization-based image segmentation. *Biosystems*, 76:43, 2004.
[19] M. G. Kuzmina and E. A. Manykin. Oscillatory neural network for adaptive dynamical image processing. In *Proc. CIMCA'05*, page 301, 2005.
[20] M. G. Kuzmina, E. A. Manykin and I. I. Surina. Oscillatory networks with Hebbian matrix of connections. In *Proc. IWANN'95*, volume 930 of *Lecture Notes in Computer Science*, page 246, 1995.
[21] M. G. Kuzmina, E. A. Manykin and I. I. Surina. Oscillatory networks of associative memory. *Opt. Mem. Neural Net.*, 5:91, 1996.
[22] M. G. Kuzmina, E. A. Manykin and I. I. Surina. Exact solutions and modeling of associative memory in oscillatory networks. In *Proc. SPIE, Optical Information Science and Technology '97: Optical Memory and Neural Networks*, volume 3402, page 298, 1998.
[23] M. G. Kuzmina, E. A. Manykin and I. I. Surina. Memory estimation of homogeneous closed chains of oscillators. In *Proc. Int. Conf. "Neuroinformatics-2000"*, page 37, 2000. in Russian.
[24] M. G. Kuzmina, E. A. Manykin and I. I. Surina. Oscillatory network model of the brain visual cortex with controlled synchronization. In *Proc. COC'00*, volume 3, page 461, 2000.
[25] M. G. Kuzmina, E. A. Manykin and I. I. Surina. Tunable oscillatory network for visual image segmentation. In *Proc. ICANN'01*, volume 2130 of *Lecture Notes in Computer Science*, page 1013, 2001.
[26] M. G. Kuzmina, E. A. Manykin and I. I. Surina. Oscillatory network with self-organized dynamical connections for image segmentation. *Neurocomputers*, 4:34, 2004. in Russian.
[27] M. G. Kuzmina and I. I. Surina. Spatially distributed oscillatory networks related to modelling of the brain visual cortex. In *Proc. NOLTA'00*, volume 1, page 336, 2000.
[28] A. Labbi, R. Milanese and H. Bosch. Gray-level object segmentation with a network of FitzHugh–Nagumo oscillators. In *Proc. IWANN'97*, page 1075, 1997.
[29] A. Labbi, R. Milanese and H. Bosch. A network of FitzHugh–Nagumo oscillators for object segmentation. In *Proc. NOLTA'97*, page 581, 1997.

[30] Z. Li. A model of olfactory adaptation and sensitivity enhancement in the olfactory bulb. *Biol. Cybern.*, 62:349, 1990.
[31] Z. Li. A neural model of contour integration in the primary visual cortex. *Neural Comput.*, 10:903, 1998.
[32] Z. Li. Visual segmentation by contextual influences via intra-cortical interactions in the primary visual cortex. *Network: Comput. Neural Syst.*, 10:187, 1999.
[33] Z. Li. Pre-attentive segmentation in the primary visual cortex. *Spatial Vision*, 13:25, 2000.
[34] Z. Li. Computational design and nonlinear dynamics of a recurrent network model of the primary visual cortex. *Neural Comput.*, 13:1749, 2001.
[35] Z. Li and P. Dayan. Computational difference between asymmetrical and symmetrical networks. *Network: Comput. Neural Syst.*, 10:59, 1999.
[36] Z. Li and P. Dayan. Pre-attentive visual selection. *Neural Networks*, 19:1437, 2006.
[37] Z. Li and J. Hertz. Odor recognition and segmentation by a model olfactory bulb and cortex. *Network: Comp. Neural*, 11:83, 2000.
[38] Z. Li and J. J. Hopfield. Modeling the olfactory bulb and its neural oscillatory processing. *Biol. Cybern.*, 61:379, 1989.
[39] D. Marr. *Vision: A computational investigation into the human representation and processing of vision information*. MIT Press, 2010.
[40] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE T. Pattern Anal.*, 12:629, 1990.
[41] P. Saint-Marc, J.-S. Chen and G. Medioni. Adaptive smoothing: a general tool for early vision. *IEEE T. Pattern Anal.*, 13:514, 1991.
[42] W. Singer and C. M. Gray. Visual feature integration and the temporal correlation hypothesis. *Annu. Rev. Neurosci.*, 18:555, 1995.
[43] A. J. W. van der Kouwe, D. L. Wang and G. J. Brown. A comparison of auditory and blind separation techniques for speech segregation. *IEEE T. Speech Audi. P.*, 9:189, 2001.
[44] C. Von der Malsburg. The correlation theory of brain function. Technical report, Max-Plank-Institute for Biophys. Chem., Göttingen, Germany, 1981.
[45] C. Von der Malsburg and J. Buhmann. Sensory segmentation with coupled neural oscillators. *Biol. Cybern.*, 67:233, 1992.
[46] D. L. Wang. Relaxation oscillators and networks. In J. G. Webster, editor, *Encyclopedia of Electrical Electronics Engineering*, volume 18, page 396. Wiley, 1999.
[47] D. L. Wang and G. J. Brown. Separation of speech from interfering sounds based on oscillatory correlation. *IEEE T. Neural Networ.*, 10:684, 1999.
[48] D. L. Wang and G. J. Brown, editors. *Computational Auditory Scene Analysis*. IEEE Press / Wiley, 2006.
[49] D. L. Wang and D. Terman. Locally excitatory globally inhibitory oscillator networks. *IEEE T. Neural Networ.*, 6:283, 1995.
[50] D. L. Wang and D. Terman. Image segmentation based on oscillatory correlation. *Neural Comput.*, 9:805, 1997.
[51] F. Wörgötter. Bad design and good performance: strategies of the visual system for enhanced scene analysis. In *Proc. ICANN'01*, volume 2130 of *Lecture Notes in Computer Science*, page 13, 2001.
[52] H. H. Yang, S. Amari and A. Cichocki. Information-theoretic approach to blind separation of sources in nonlinear mixture. *Signal Process.*, 64:291, 1998.

4 Image processing based on the oscillatory network model

4.1 Problems of image segmentation and traditional methods of their solution

4.1.1 Digital images and methods of image analysis

Computational methods of image processing, image analysis, and machine vision belong to a research subarea of cognitive and computer science that exists about a half of the century and continues to be extensively developed nowadays [8, 23, 24]. Algorithms for a wide variety of sophisticated image processing applications required by software engineers and developers, advanced programmers, graphics programmers, scientists, and related specialists have been developed. Various tasks of image processing such as image segmentation (i.e. identification of homogeneous regions in the image), image reconstruction, spectral analysis, coding, compression, filtering, recognition, and quality enhancement are currently of great interest. Neural network architectures, which help to get the output in the real time due to their parallel processing ability, have also been used for image processing, demonstrating good results even when the noise level is very high.

A great variety and high complexity of computational problems related to image processing led to the development of theoretical background, which allowed us to investigate general properties of discretized images, their transformations and to construct effective computational algorithms for real color image processing. The area of computational image processing was significantly enlarged owing to computer technology modernization, and currently it involves science, technology, industry, agriculture, space monitoring, medicine, and arts. For example, in the field of cosmic research the tasks of automatic object revealing during information transfer from cosmic satellites require processing of great flows of numerical information in the real time. Powerful systems of image forming, transformation, visualization, and documentation are required in medicine. Automatic image processing systems are necessary in numerical problems of Earth state monitoring. Therefore, the necessity of automatic processing of great information flows looks like a characteristic feature of modern image processing tasks, and therefore it stimulates the development of new special approaches and alternative ways for sophisticated analysis of great amount of visual information.

In this section, we shortly outline the most important mathematical approaches and computational algorithms developed and mention some important results in the field of computational processing of digital images.

Digital image processing consists in the use of computational algorithms to perform processing of discretized images. An analog image is defined by a continuous

function $I(x, y)$ giving the intensity at position (x, y). The function $I(x, y)$, dependent on two spatial coordinates is defined in a rectangle. A discretized form of an image, or a digital image (2D array of numbers), is obtained by sampling and quantizing an analog image. The digital image can be considered as defined over a grid, each grid site being called a pixel. The image transformation into the digital form is the transfer from the continuous function $I(x, y)$ to some piecewise function. It is realized via combination of two operations: (a) discretization – replacement of continuous domain of the function $I(x, y)$ by some 2D discrete spatial grid; (b) quantization – replacement of continuous interval of the function $I(x, y)$ values by a discrete fixed number of function values $\{I^q\}$ – the given scale of intensity levels, or the scale of brightness levels. It is sometimes convenient to carry out the image discretization via multiplication of the continuous function $I(x, y)$ by a discretization function that is represented as a sum of δ-functions defined at the nodes of a 2D lattice with the cell $(\Delta x, \Delta y)$. In this case, the discretization operation and other function transformation can be fulfilled analytically.

The quantization error $I - I^q$ and its variance are important image characteristics, providing optimal choice of quantization for various image types. For instance, in the case of rather homogeneous image brightness distribution, a homogeneous quantization is optimal. The optimal quantization choice for noisy images was a subject of special investigations. As a result, a proper family of brightness scales was constructed for adequate choice of quantization type in dependence of noise character. The errors in image processing due to quantization type figure in the form of image spurious contours that appear in the image regions with smooth brightness variation.

The information on the frequency spectrum $F(\omega_x, \omega_y)$ is essentially used in image processing. The spectrum is obtained via direct two-dimensional Fourier transformation of the function $I(x, y)$. As is well known from the Fourier analysis, the spectrum of discretized image is obtained as a convolution of the spectrum of the continuous function $I(x, y)$ and the spectrum of the discretizing function. It gives the important result: if the initial continuous image possesses a finite frequency spectrum width (that is, the spectrum vanishes outside the rectangle $|\omega_x| \leq \Omega_x$, $|\omega_y| \leq \Omega_y$ with $\Omega_x = 2\pi/\Delta x$ and $\Omega_y = 2\pi/\Delta y$), then the spectrum of the discretized image can be obtained via infinite repetition of the continuous image spectrum by its shifting over $2\pi k/\Delta x$, $2\pi m/\Delta y$, $k, m = 1, \ldots, \infty$. Therefore, the neighboring spectrum fragments will be overlapped if Δx and Δy are chosen too large. This feature permits to understand the restrictions on conditions of reconstruction of the continuous image spectrum from that of the discretized image (via linear interpolation or spatial filtration of discretized image spectrum). In the case of filtration, the filter is required to reproduce the main spectrum component and to completely suppress all additional parasitic components. As the analysis shows, this condition can be fulfilled when discretization interval is not greater than spatial half-period corresponding to the smallest image fragments (Kotelnikov frequency). According to the Kotelnikov–Nyquist–Shannon theorem, only under this condition the initial continuous image can be reconstructed from the dis-

cretized one via spatial filtration with any given accuracy. The function that determines the spatial characteristics of the ideal reconstructing filter (called impulse response) and is defined in the infinite interval attains a maximum value at the origin and demonstrates quickly damping oscillations away from the origin. In practice, the special types of low-pass filters are used. They are constructed based on the principle of maximal similarity of filter frequency characteristics to those of the ideal low-pass filter.

For visualization quality enhancement elementwise transformations are often carried out. They can include linear contrast enhancement, transformation of brightness in a given range, and nonlinear brightness transformation. Besides, the elementwise transformations of the probability density function related to the image can also be carried out. The transformation is then realized based on preliminary obtained brightness histogram and the subsequent probability density function transformation (as a rule, the density function transformation corresponds to histogram equalization).

Probabilistic models are also widely used for image description and analysis. The brightness function $I(x, y)$ is considered as a stationary random function of coordinates. The first and the second moments, the autocorrelation function and its spectrum comprise the essential information on function $I(x, y)$ behavior, allowing us to estimate such image characteristics as quality, spatial distortions, color rendering, etc.

Image filtering is applied for the purpose of attenuation of various disturbances, such as noise and complicated background. The filtering is based on neighborhood operations (taking a filter mask from point to point and performing operations on pixels inside the mask), and can be realized both in spatial and in frequency domains. Filtering in frequency domain consists in transformation of the preliminary obtained frequency spectrum. Filtering with the help of filter masks is widely used in the tasks of digital image processing. Two-dimensional spatial filtering based on filters with finite spatial spectrum is one of popular filtering operations (a 3×3 square neighborhood and homogeneous mask being usually applied). The shortcoming of image processing via linear smoothing is that the noise suppression, inherent to the linear smoothing, is usually accompanied by boundary blurring. This is the reason why the methods of nonlinear filtering have been also developed. They are based on nonlinear brightness function transformation in a neighborhood specified by the filter mask. Good results are often provided by the so-called median filtering, in which action consists in reduction of both minimal and maximal local brightness "escapes." The median filtering better preserves image boundaries compared to linear filtering of any kind. The frequency filtering allows us to select filter frequency characteristics via discrete image Fourier transformation. As a result, high frequencies are suppressed in the filter frequency spectrum in the case of low-pass filter application, whereas low frequencies are reduces in the case of high-pass filter applications.

The necessity to process huge amount of information stimulated the development of approaches on image compressions. At the end of the 1960s the coding methods

for gray-level images were created, providing compression factors up to 5. In the 1980s the compression methods of the second generation were developed, including those based on wavelet approaches. These gave much more compression degree and ensured optimal compromise between spatial and frequency resolution. In the 1980s–1990s fractal coding methods were also developed, providing 50–2000 times compression. The compression methods can be divided into two classes of methods: compression without information loss and that with information loss. Some features of the brain visual systems performance related to the dependence of visual acuity on image brightness and contrast were attracted in the process of development of some image compression methods admitting the information loss.

4.1.2 Image segmentation problems and the examples of their solution via traditional methods of computer vision

Image segmentation is a partitioning of an image into a collection of homogeneous regions (image segments, or fragments). The homogeneity can be defined in terms of some feature characteristics (or a parameter, possibly multidimensional) such as gray-level value, color, shape, texture, and motion. The proper formulation of conditions for specification of the set of feature matching characteristics should naturally precede the segmentation problem solving. The following types of feature characteristics frequently figure in partitioning principles of image segmentation: brightness characteristics (such as brightness distribution over an image segment), geometrical characteristics (such as conditions of specification of fragment boundary form or appropriate information on proper reference points), and texture characteristics. The choice of parameters providing texture specification usually depends on concrete approach applied for texture description – structural, statistical or fractal.

It is worth to outline typical difficulties arising in the solution to the image segmentation problem via traditional computational methods. The extraction of the set of connected image regions is a typical task of image segmentation. The connected fragments of a digital image are the subregions of image pixel array in which for each pixel there exists at least one neighboring pixel belonging to the given subregion. Three main ways of segmentation of connected fragments via traditional approaches are reduced to iteration methods, recursive methods, and methods of boundary detection. The iteration methods are based on scanning of the whole image pixel array. The recursive methods of region-based segmentation are basically reduced to some algorithms of region growing, edge detection, region splitting, and merging. Edge detection is one of the important and difficult operations in image processing. It is an important step in the process of image partitioning into constituent objects. Mathematically the existence of edge is reflected by a sudden jump of intensity, which can be expressed in terms of first and second derivatives of $I(x, y)$: dI/dx is maximal and d^2I/dx^2 crosses zero at the edge point. In two-dimensional case the edge is indicated

by the maximum value of gradient, and gradient-based edge detection methods are applied, the Laplacian being used as the second derivative. A variety of boundary detection methods is based on versions of high-pass filtering. The main shortcoming of the iteration methods in the problems of real image segmentation is large running time and large volume of required computer memory. The slow running is also inherent to recursive methods, being related to the necessary enumeration of all segment pixels at the final stage of segment formation. The shortcomings of the methods based on boundary detection are related to insufficient accuracy of boundary identification and also to restrictions on segment boundary form. Various ways to overcome these problems were suggested. One of natural ways consists in the construction of parallel computational algorithms.

The image segmentation based on segment growing is usually accompanied by simultaneous detection of the boundaries between the growing segments. For instance, the task arises in image partitioning into a collection of macroscopic brightness fragments with a given brightness level between the fragments. In a number of situations, the fragment boundary detection can be successively fulfilled by histogram-based methods. These approaches are based on the use of functionals characterizing statistical properties of brightness field and often permit us to evaluate a probability of brightness variation at a given point of the processed image. An iteration histogram-based algorithm for spatial separation of image points can be then constructed. Such approaches can provide satisfactory boundary detection between image fragments in the case of image fragments with sharp brightness variation between the fragments, but fail in the cases when fragments with continuous brightness gradient exist.

For separation of contour boundaries in a segmented image the methods of contour filtering are used. They are based on the calculation of the first and the second derivatives of the brightness function $I(x, y)$. As it turned out, the linear filtration approaches lead to a number of systematic errors in segmentation task solution, for instance, to the appearance of spurious contours in the regions with continuous brightness gradient. Better results are provided by methods of quadratic filtering. And the best results demonstrate the improved methods of nonlinear filtering, permitting the construction of controllable adaptive oriented filters [2, 27]. In image segmentation problems where the exact calculations of fragment boundaries and thin contours are required, the so-called methods of contour tracing proved to be adequate.

At last, an interesting approach based on the application of a weighted graph associated with the image should be mentioned [26]. The graph vertexes were associated with image pixels, whereas the weights of graph edges specified the proximity degree of pixels. In dependence on concrete image segmentation task, the weights can be dependent on pixel brightness, pixel mutual spatial distance, color components, or proper filter parameters. Image segmentation problem is solved via the graph cutting into two components, the principle of maximizing internal connectivity of subgraphs and minimizing of mutual subgraph connectivity being exploited. The problem of the initial graph cutting was reduced to the problem of analysis of discrete spectrum of

a linear operator corresponding to the graph. The calculation of minimal (by module) eigenvalue of the operator was necessary to specify the first graph cutting. The graph cutting itself is related to image separation into two maximally independent partitions. Further the whole process was recurrently repeated with a complete segmentation of the processed image as a final result. This approach is in some sense similar to visual image processing that is realized by the brain visual system in the process of natural image perception.

It should be added that the principle of the graph weight construction used in [26] has a resemblance with the principle of construction of dynamical connections in the oscillatory network model designed for brightness image segmentation [10–13] that will be described below (see Section 4.2).

4.1.3 Neuromorphic methods of image processing

Neuromorphic methods of image processing imitating performance of the brain neural structures and representing a significant interest are considered as an alternative to traditional computational methods developed in the field of computer vision. In particular, the neuromorphic methods imitating the principle of dynamical binding via synchronization that is presumably exploited by the brain visual system are of special interest [7, 9, 28]. The brain visual system capability of self-organized functioning is especially attractive to use in the visual image processing tasks where it is necessary to exploit selective attention and an ability to exclude unessential information. The problems of image understanding, recognition, and selection of objects in a visual scene require using of these capabilities. The development of artificial visual systems of autonomous robots has led to clear understanding that robotic vision should be active, that is, be dynamically reconfigurable depending on a concrete visual processing task. So, the methods of flexible, adaptive visual information processing are necessary to be developed for artificial vision systems.

A dynamical oscillatory network method of visual information processing was developed based on the design of neural-like oscillatory network model of limit-cycle oscillators with controllable internal oscillator dynamics and controllable dynamical oscillator coupling [10–13]. The designed network coupling, being self-organized and tunable by a processed image, is capable to be quickly and adaptively reconfigurable. As a result, the neuromorphic dynamical synchronization-based oscillatory network method of image processing proved to be flexible and adaptive. The final oscillatory network model version possesses a number of useful features that are indispensable in the problems of active image processing. Besides the capability of qualitative brightness segmentation of real gray-level and color images, the oscillatory network performs selective image segmentation, and object separation in visual scenes in simple situations. The model is described in detail in Sections 4.2 and 4.3.

4.2 Oscillatory network model description

The model was developed based on the reduced 2D oscillatory network model that was obtained via reduction of previously designed biologically motivated 3D oscillatory network model (see Sections 3.4.2–3.4.5). New improvements of the 2D-reduced network model concern internal single oscillator dynamics and the principle of network oscillator coupling. Due to the advanced version of single oscillator dynamics, it is possible to specify the dependence of the oscillation amplitude on the pixel brightness via arbitrary continuous monotonic function. It gives a notable increase in segmentation quality. Several new versions of the network coupling principle were also constructed and tested. In particular, the filtration type version of the coupling principle was designed. It provided the network with the capability of selective image segmentation – the extraction of brightness fragment subset with brightness levels belonging to an a priori chosen interval of brightness. On the whole, the model improvements allowed us to significantly raise image segmentation accuracy and to control network noise suppression.

The oscillatory network model performance consists in network relaxation into the state of clusterized synchronization encoded by a segmented image. The set of internally synchronized but mutually desynchronized network ensembles (clusters) arising at the final synchronization state corresponds to the image decomposition into the set of brightness fragments specified by the choice of brightness levels.

4.2.1 Network architecture and governing dynamical system

Oscillators of a 2D-reduced network are located at the nodes of a 2D spatial lattice being in one-to-one correspondence with the pixel array of a segmented image (see Section 3.4.5 and Figure 3.2). Active network processing unit is the limit-cycle oscillator with the bifurcation dynamics. The stable attractor of single-oscillator dynamical system is either the stable limit cycle, of a stable focus. Image segmentation is performed by the oscillatory network via synchronization of network assemblies corresponding to image fragments of various brightness levels.

Let the image to be segmented be defined by the $M \times N$ matrix $[I_{jm}]$ of pixel brightness values. The state of a single oscillator is defined by a pair (u_1, u_2) of real-valued variables. The network state is defined by the $M \times N$ matrix $\hat{u} = [u_{jm}]$ of complex-valued variables $u_{jm} = u_{1jm} + iu_{2jm}$ defining the states of all network oscillators. System of ODE governing oscillatory network dynamics can be written as

$$\dot{u}_{jm} = f(u_{jm}, I_{jm}) + \sum_{j'=1}^{M} \sum_{m'=1}^{N} W_{jmj'm'}(u_{j'm'} - u_{jm}). \tag{4.2-1}$$

Here functions $f(u, I)$ define internal dynamics of isolated network oscillators, whereas the second term defines contribution into the dynamics via oscillator cou-

pling. In the first version of the model the values $W_{jmj'm'}$, defining coupling strength of network oscillators (j, m) and (j', m'), were designed so as to realize the principle of dynamical coupling based on brightness. The principle should provide self-organized emergence of synchronization in the oscillatory network, so that synchronized network ensembles correspond to image fragments of different brightness levels. The matrix elements $W_{jmj'm'}$ were designed in the form of the product of two nonlinear functions dependent on oscillation amplitudes (limit cycle radii) of oscillator pair and spatial distance between the oscillators in the network:

$$W_{jmj'm'} = P(\rho_{jm}, \rho_{j'm'})D(\mathbf{r}_{jm}, \mathbf{r}_{j'm'}). \tag{4.2-2}$$

The cofactors $P(\rho, \rho')$, providing the network coupling dependence on oscillation amplitudes, were constructed as

$$P(\rho, \rho') = w_0 H(\rho\rho' - h), \tag{4.2-3}$$

where $H(x)$ is a continuous step function, w_0 is a constant defining total strength of network interaction. The cofactors $D(\mathbf{r}, \mathbf{r}')$, providing "cut off" of spatial oscillator coupling, can be specified by any function of $|\mathbf{r} - \mathbf{r}'|$ vanishing at some finite distance. For instance, it is convenient to choose $D(\mathbf{r}, \mathbf{r}')$ in the form

$$D(\mathbf{r}, \mathbf{r}') = 1 - H(|\mathbf{r} - \mathbf{r}'| - r_0), \tag{4.2-4}$$

where r_0 is the chosen radius of spatial interaction. According to the coupling rule (4.2-2), any two network oscillators are coupled if they possess sufficiently great oscillation amplitudes and are separated by a distance not exceeding the prescribed radius of spatial interaction. Otherwise the coupling is absent.

4.2.2 Modified model of network oscillator

The new version of single oscillator dynamics was designed with preserving the main qualitative features of dynamics of biologically motivated model of neural oscillator designed in [18]. The modified version of dynamics allows us to specify the dependence of oscillation amplitude on pixel brightness by several types of simple monotone continuous functions, providing an additional tool of brightness segmentation accuracy rise. The new version of the dynamical system (3.4-2) can be written in the form [17]

$$\dot{u} = f(u, I), \quad f(u, I) = \left[\rho^2 + i\omega - |u - c|^2 + g(I)\right](u - c), \tag{4.2-5}$$

where $c = \rho(1 + i)$ and $g(I)$ is the continuous function of pixel brightness I.

If we choose $g(I)$ in the form of a step-function dependent on a threshold h_0,

$$g(I) = \alpha[1 - H(\rho)], \tag{4.2-6}$$

then the following features of isolated oscillator dynamics will be ensured: (1) at $I \leq h_0$ a stable focus is the only stable attractor of the dynamical system (4.2-5) (the oscillator demonstrates quickly damping oscillations); (2) at $I > h_0$ the limit cycle of radius $\rho(I)$ is the only stable attractor of the system (4.2-5) (the oscillator demonstrates stable auto-oscillations of the amplitude $\rho(I) < \rho_0$, where $\rho(I)$ can be specified by an arbitrary monotonically increasing function of I). The collection of limit cycles at different I (in the case of $\rho(I) = I$ and $h_0 = 0.25$) is shown in Figure 4.1 (a). The limit cycle centers are located at the points with coordinates (ρ, ρ) of the phase plane. The bifurcation of limit cycle into stable focus occurs at $I = h_0$. The oscillator "response" to pixel brightness variation is depicted in Figure 4.1 (b).

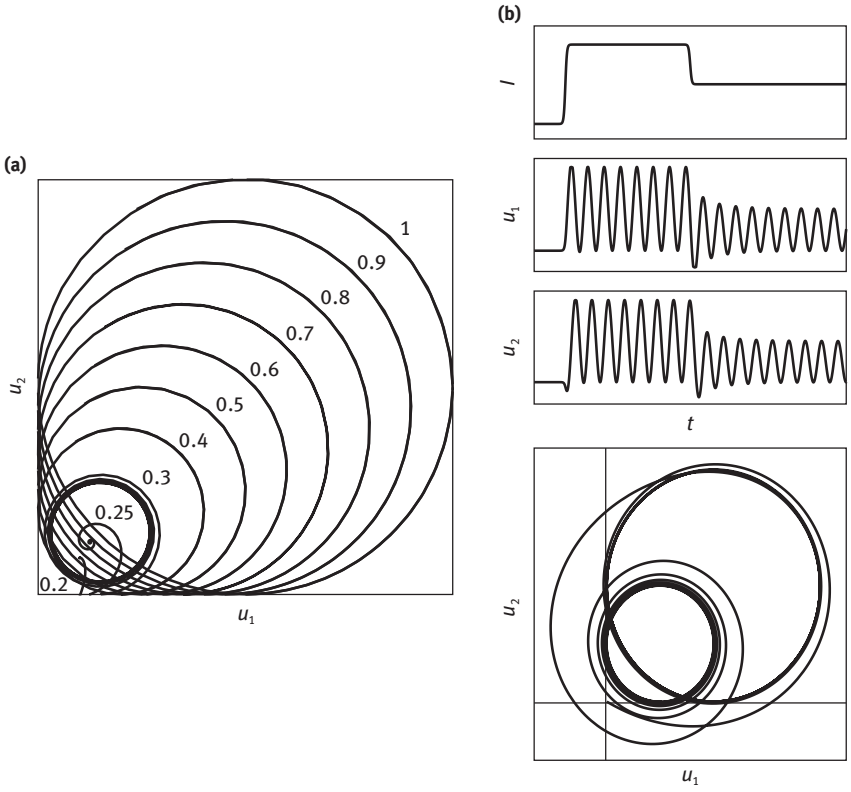

Fig. 4.1. New version of single oscillator dynamics: limit cycles and focuses of dynamical system (4.2-5) for various values of pixel brightness I and the oscillator "response" to the pixel brightness variation.

4.2.3 Modified principles of network coupling

In addition to the new version of single oscillator dynamics, providing an analog type of oscillator "response" on pixel brightness, the modified versions of network coupling principles were introduced. The aim was to provide acceptable brightness segmentation accuracy without application of the method of sequential segmentation, which was exploited previously in segmentation of synthetic images (see Section 3.4.6).

According to the coupling rule (4.2-2)–(4.2-4), any oscillator pair is coupled if both oscillators possess sufficiently large oscillation amplitude and the spatial distance between the oscillators does not exceed a given radius of spatial interaction. To raise the segmentation accuracy, more flexible connectivity rules were designed and tested. One of the coupling modifications, which demonstrated the segmentation accuracy improvement in processing of real gray-level images, was a coupling principle based on restriction for each oscillator of the set of other network oscillators to which this oscillator is coupled. The restriction can be realized via some "mask." Analytically the coupling principle can be defined by matrix (4.2-2) with the modified co-factor $\tilde{P}(\rho, \rho')$,

$$\tilde{P}(\rho, \rho') = H\left(\rho - \rho' + \Delta(\rho')\right) H\left(\rho' - \rho + \Delta(\rho)\right). \tag{4.2-7}$$

The function $\Delta(\rho)$ specifies the "mask," restricting the size of interaction vicinity of a network oscillator (see Figure 4.2).

If $\Delta(\rho) = 2\Delta_0$, then according to coupling rule defined by matrix (4.2-2) with $P(\rho, \rho') = \tilde{P}(\rho, \rho')$, any pair of network oscillators is coupled only in the case when the oscillator intervals of interaction $[\rho - \Delta_0, \rho + \Delta_0]$ and $[\rho' - \Delta_0, \rho' + \Delta_0]$ intersect. Computer experiments on brightness image segmentation demonstrate the segmentation accuracy improvement in the case of application of the coupling rule (4.2-2) with the modified cofactor $\tilde{P}$.

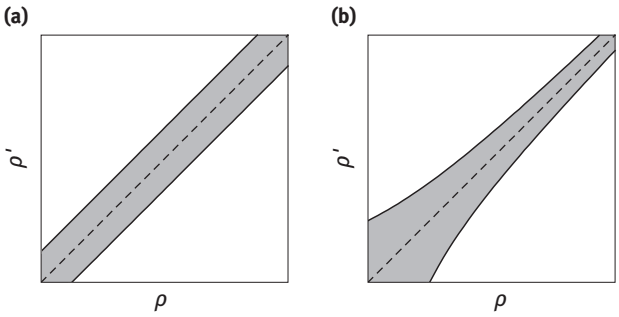

Fig. 4.2. The principle of network coupling with using of masks: (a) $\Delta(\rho) = \text{const}$, (b) $\Delta(\rho)$ is chosen as a monotonic function of ρ.

4.2.4 Network performance and model capabilities

Brightness image segmentation is realized by the oscillatory network via two steps: (1) preliminary tuning of oscillator dynamics by pixel brightness values (after the tuning operation the proper limit cycle size is specified for each network oscillator); (2) the network relaxation into the state of cluster synchronization (which results in oscillatory network decomposition into a set of internally synchronized and mutually desynchronized oscillator ensembles, so that each ensemble corresponds to a fragment of appropriate brightness). The combination of gradual (analog type) oscillator response to pixel brightness and the more flexible (filtration type) network coupling rule provides significantly higher segmentation quality of real brightness images.

Brightness segmentation of gray-level images

For computer experiments on real gray-level brightness image segmentation the interactive computer code was created that provided a large series of experiments on testing of model capabilities. Some results of real gray-level image segmentation are presented in Figures 4.3 and 4.4.

In the frames of the oscillatory network model, the segmented image is "reproduced" by the network in the oscillatory form: each ensemble of oscillators corresponding to a definite image fragment demonstrates internally synchronized oscillations, being at the same time desynchronized with the oscillators of all other oscillatory ensembles corresponding to different brightness fragments. To extract the correct stable version of reconstructed image from its oscillatory version provided by the network, some simple methods of postprocessing were used. One of the simplest postprocessing method was based on the extraction of the image via incorporation of its instantaneous versions at the points of the maximal values of all the $r_{jm}(t) = |u_{jm}(t)|$.

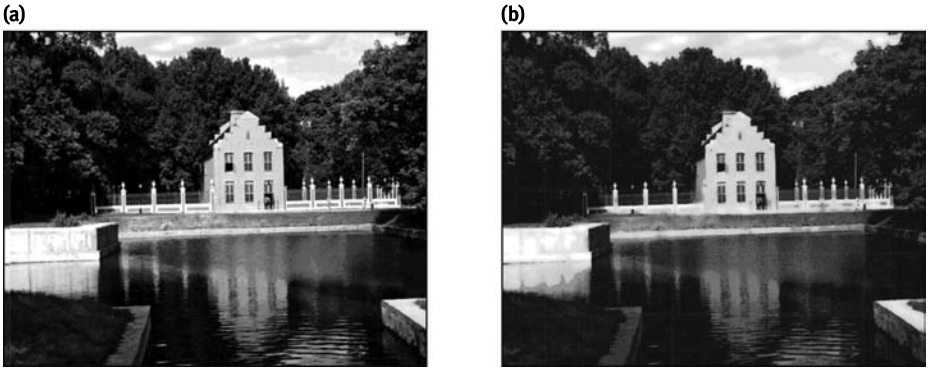

Fig. 4.3. Brightness segmentation of a gray-level photograph (657 × 432 pixels): (a) segmented image; (b) the final result of segmentation.

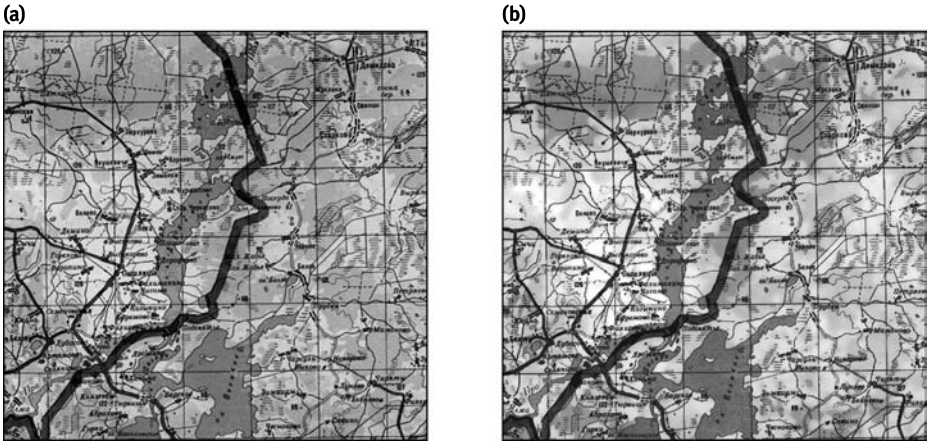

Fig. 4.4. Brightness segmentation of a map fragment (492 × 475 pixels): (a) segmented image; (b) the final result of segmentation.

Color image segmentation

Color image segmentation was realized via pixel array decomposition into three subarrays corresponding to red, blue, and green components of pixel colors. These three subarrays were processed by the oscillatory network independently. Final visualization of the segmentation result after accomplishment of network performance was realized via reverse joining up of three subarrays into the single array. The example of color image segmentation is presented in Figure 4.5.

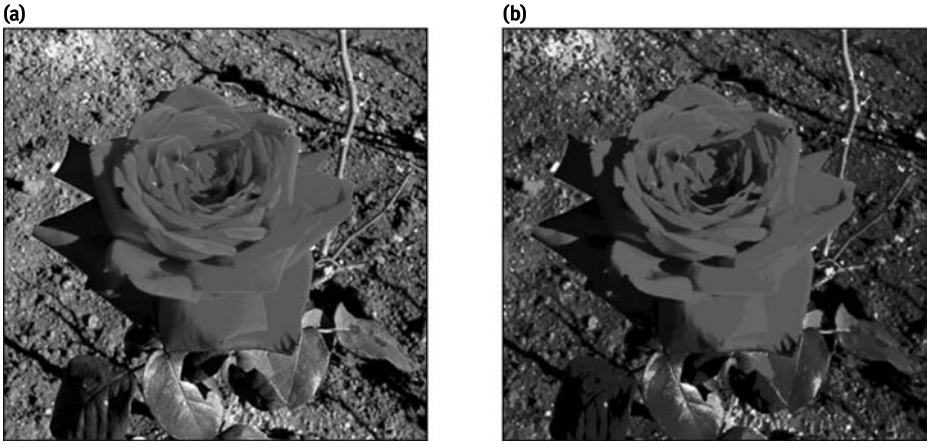

Fig. 4.5. Color image segmentation (524 × 374 pixels): (a) segmented image; (b) the final result of segmentation.

Another way of color image segmentation is also possible. Namely, one can use the vector parameter $\mathbf{I} = (I_1, I_2, I_3)$ with components equal to red, green, and blue pixel brightness components, respectively, and represent scalar pixel brightness I in the form

$$I = \sum_{k=1}^{3} c_k I_k, \quad c_k > 0, \ \sum_{k=1}^{3} c_k = 1. \tag{4.2-8}$$

Then the coupling principle (4.2-2) with $P(\rho, \rho') = P_c(\rho, \rho')$,

$$P_c = \prod_{k=1}^{3} H(2\Delta - |I_k - I_k'|), \tag{4.2-9}$$

can be used for segmentation of color images (Δ is a constant parameter).

Selective image segmentation

Selective image segmentation consists in the extraction of a desirable image fragment subset with brightness values contained in some a priori chosen interval of brightness levels. Selective segmentation can be viewed as a simplest type of active image processing. As it is intuitively clear, the selective segmentation can often be more informative compared to usual complete segmentation, because it allows us to exclude unnecessary information.

The modified version of oscillator dynamics (4.2-5) provides a natural way of selective segmentation realization. It is sufficient to introduce a new function $\tilde{\rho}(I)$ in equation (4.2-5) instead of $\rho(I)$ via introducing of a "mask" function $F(I)$ by putting

$$\tilde{\rho}(I) = \rho(I)F(I). \tag{4.2-10}$$

If it is required to select only image fragments of brightness values $I \in [I_1, I_2]$, we choose $F(I)$ to be equal to 1 inside the interval $[I_1, I_2]$ and vanishing outside the interval. It can be done analytically via the mask of a rectangle form

$$F^{(1)}(I) = H(I - I_1)H(I_2 - I). \tag{4.2-11}$$

In this case, obviously, only the oscillators corresponding to image fragments with brightness values $I \in [I_1, I_2]$ will possess nonzero oscillation amplitudes, whereas the rest network oscillators will drop out of network interaction because of vanishing limit cycle sizes. To select a collection of brightness fragments with a discrete brightness values $I^{(1)}, I^{(2)}, \ldots, I^{(m)}$, the mask in the form of the sum of delta-like functions $\Gamma(x)$ can be used:

$$F^{(2)} = \sum_{k=1}^{m} \Gamma(|I - I^{(k)}|). \tag{4.2-12}$$

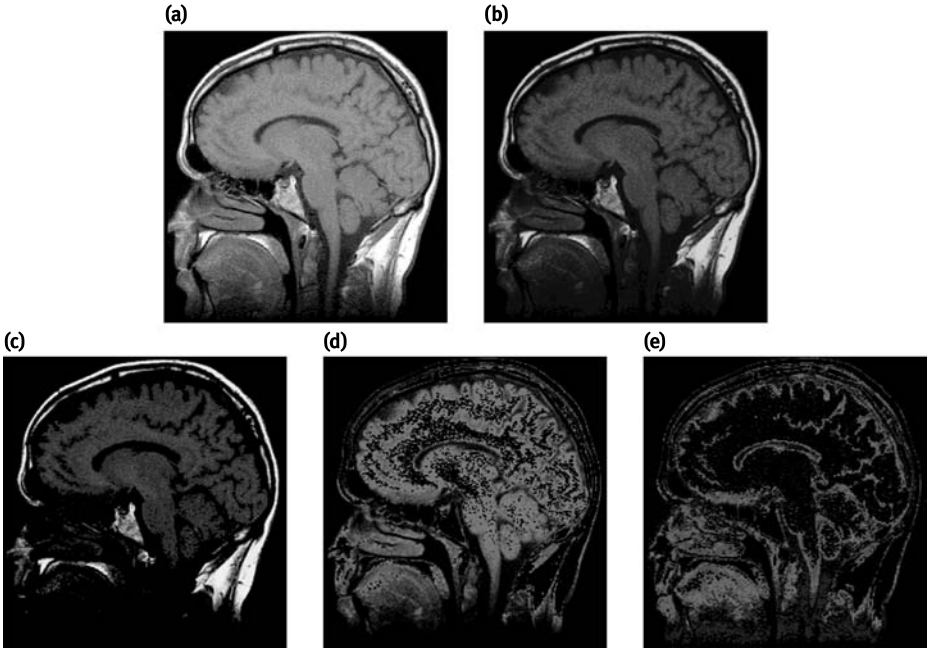

Fig. 4.6. Selective image segmentation 1: (a) original image; (b) complete image segmentation; (c) extraction of several the most bright image fragments; (d) extraction of a set of fragments of middle brightness; (e) extraction of several the least bright fragments.

The examples of selective brightness segmentation of a cross-sectional tomogram are given in Figures 4.6 and 4.7.

The capability to focus attention and exclude unnecessary information is the known significant feature inherent to the mammalian brain visual systems. In humanoid robotics the problems of active vision are particularly important, because the capability to concentrate attention on a single aspect is necessary for any animate vision system. It is of special importance in navigation of autonomous robots in unknown surroundings. The artificial neural network approaches, offering the analog with the brain neural systems working, represent the attempts to elucidate various principles of the brain functioning and to use them in robotics. In particular, modeling of principles of visual information reduction is helpful for developments of artificial vision systems. Thus, modeling of selective brightness image segmentation (simple exclusion of unnecessary visual information) could be considered as an initial step toward the creation of artificial active vision systems.

An example of the visual illusion shown in Figure 4.8 illustrates the informative character of selective image segmentation.

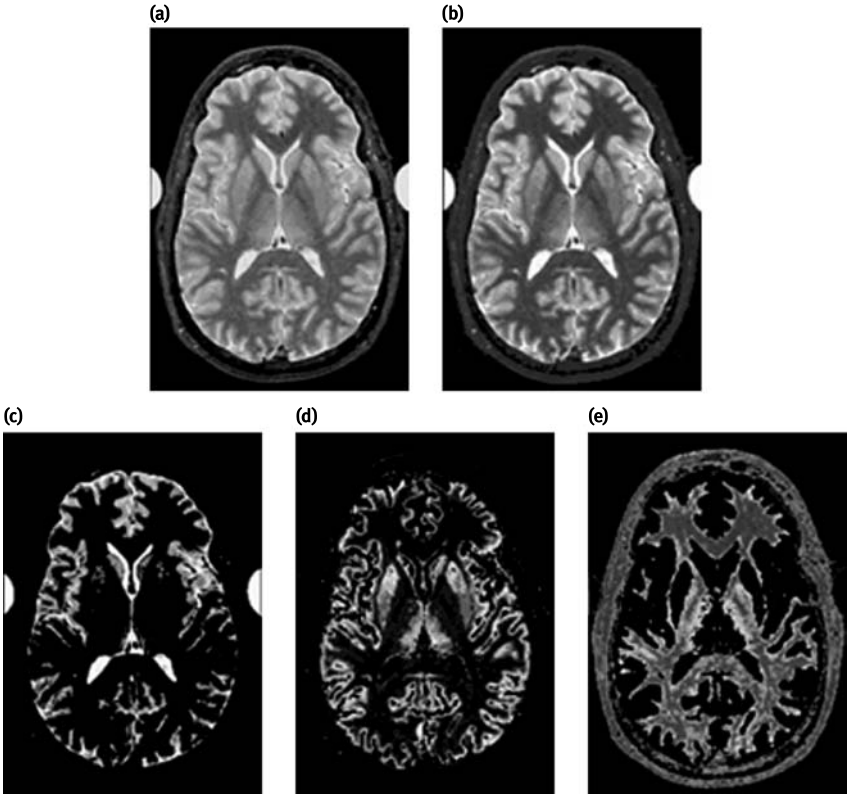

Fig. 4.7. Selective image segmentation 2: (a) original image; (b) complete image segmentation; (c) extraction of several the most bright image fragments; (d) extraction of a set of fragments of middle brightness; (e) extraction of several the least bright fragments.

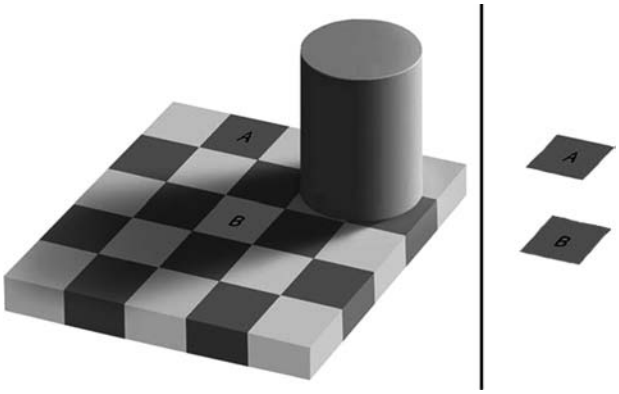

Fig. 4.8. The visual illusion (related to selective brightness image segmentation). The squares A and B are of the same gray level of brightness. It is demonstrated (right) via separation of the squares and the exclusion of the rest of the details from the picture. Picture by Edward H. Adelson.

4.2.5 Image fragment separation from a visual scene

Problems of visual scene analysis

The problem of object selection in a visual scene belongs to the class of higher level image processing tasks and demands additional methods of image analysis (such as filtering or pattern recognition), dependent on the problem statement. Visual scene interpretation consists in partitioning of the scene image into constituent subregions and in assigning of an appropriate label to each subregion.

In the mammalian visual systems the problems of scene analysis are solved via parallel distributed ways of visual information processing. Surprisingly, human visual system can recognize a gist of novel image in a single glance, independently on image complexity. So, it is natural to expect that in the frames of oscillatory network approach some methods of visual scene analysis based on temporal correlation can be suggested. Developing an example of such synchronization-based solution, we restricted ourselves at the first step to a simple problem of visual scene segmentation, which does not require additional computational tools (besides the developed network model). It is the problem of successive image fragment selection from some finite set of spatially separated image patches of almost equal, weakly inhomogeneous brightness. The condition of brightness homogeneity of the patches should be stressed, because in the case of patches of different brightness they could be easily separated by the improved version of the oscillatory network without any additional efforts. In the case of homogeneous patch brightness the problem can be solved in the frames of the oscillatory network approach by a proper model extension.

We now shortly outline the main traditional computational algorithms that were used in visual scene segmentation tasks.

Visual scene segmentation via traditional computational methods

Various traditional computational methods were proposed for visual scene analysis and segmentation, based on image representation in the form of a two-dimensional array where some brightness, color, and texture information is prescribed to each pixel. The first group of approaches contains region-based methods relied on generation of fixation point and finding the enclosing contour. A kind of attention mechanism is usually attracted in such approaches [21, 22]. This segmentation strategy, developed mainly for extraction of “simple” objects (presented by some connected image domains) from a visual scene, can be applied in problems for robots with cameras. The approach allows us to model a combination of top-down and bottom-up visual attention processes.

The second group of methods is based on the concept of graph cuts [26]. The graph-based algorithms exploit image representation in the form of hierarchical collection of salient segments. A graph is associated with an image so that each image pixel corresponds to the graph node, and the weight of graph edge, connecting two

nodes, is specified according to some pixels features. The contour-based segmentation strategy then involves two consecutive steps. At the first step a probabilistic boundary map is generated (in which the weight of each boundary pixel is proportional to the probability for the pixel to be at the region boundary). At the second step a fixation point and a closed boundary contour containing the fixation point are selected in the visual scene. This allows us to define a path that optimally cuts the boundary map into two halves [25]. Discrete layered models based on a sequence of ordered Markov random fields and on extended graph-cut methods have also been developed [29]. The approaches dealing with the normalized graph cuts admit estimation of image motion and can provide segmentation of moving scenes. Region-based segmentation approaches incorporating syntactic visual features (formulated in terms of shape and spatial configuration of image regions) can also sometimes provide satisfactory segmentation results.

Visual scene segmentation via oscillatory network approaches

The attempts to design neural network models reflecting the existence of so-called neural circuits and exploiting synchronization and desynchronization of neural assemblies were undertaken for scene segmentation problems [6]. A visual scene segmentation via synchronization in a two-layer network of Wilson–Cowan oscillators located in 50×50 spatial lattice was carried out with the help of image contours detection approach [30]. For solving the problem of texture patch selection in the frames of oscillatory network model it was necessary to use collections of filters and spectral histograms [19]. One more approach to the problem of sequential selection of objects from a visual scene can be formulated in terms of selective visual attention. Then the visual scene segmentation task can be solved based on partial synchronization in a neural network of spiking neurons using synchronization with a central element [1].

Image fragment separation based on in-phase synchronization of oscillatory network ensembles

The oscillatory network model described in Sections 4.2.1–4.2.4 was used for solving a simple visual scene segmentation task. A simple problem of visual scene segmentation, which does not require additional computational tools besides the developed oscillatory network model with controlled synchronization, was considered. It is the problem of successive fragment selection from some finite set of spatially separated image patches of weakly inhomogeneous (almost equal) brightness.

We introduce two identical independent (mutually uncoupled) oscillatory subnetworks denoted as the x-layer and the y-layer. The initial distributions of oscillator

phases of networks of the x-layer and the y-layer are defined as

$$\begin{aligned} \varphi_{jk}^{(x)}(0) &= \frac{2\pi j}{M}, \quad j = 1, \ldots, M, \\ \varphi_{jk}^{(y)}(0) &= \frac{2\pi k}{M}, \quad k = 1, \ldots, N, \end{aligned} \tag{4.2-13}$$

where M is the image width and N is its height. As one can see from (4.2-13), the initial phases of network oscillators of the x-layer are chosen to be proportional to the x-coordinates of the image pixels, whereas the initial phases of network oscillators of the y-layer are proportional to the pixel y-coordinates. After the relaxation of both subnetworks into the synchronization state, one will obtain the following relations for phase shifts: (1) in the x-layer the phase shift between two in-phase synchronized oscillator ensembles corresponding to two different image fragments is proportional to the difference of the x-coordinates of fragment "centers of mass" and (2) in the y-layer the phase shift between the same two oscillator ensembles is proportional to the difference of the y-coordinates of the fragment "centers of mass." So, the finite phase shifts arise between the oscillator ensembles, corresponding to different spatially separated fragments. Therefore, all spatially separated fragments (despite their equal brightness) can be clearly distinguished and successively selected. The difficulty in fragment separation can arise only in the rare situation when there exists a combination of two spatially separated fragments, located one inside another, so that the coordinates of both fragment centers coincide. The example of solvable problem of fragment separation is presented in Figure 4.9. A fragment of the Earth surface ob-

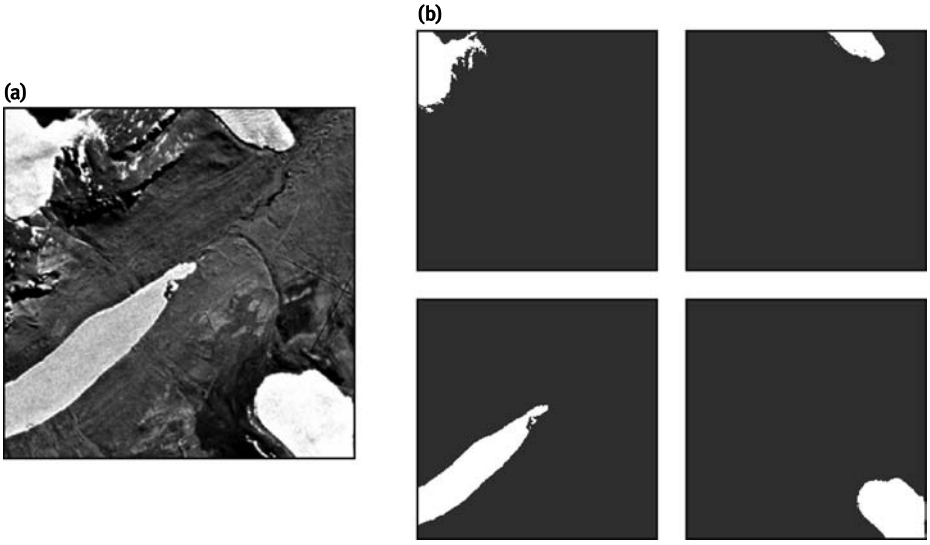

Fig. 4.9. Selection of spatially separated fragments of visual scene: (a) original image; (b) successive selection each of four fragments.

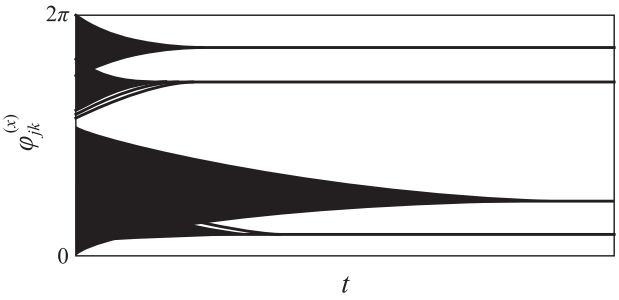

Fig. 4.10. Functions $\varphi_{jk}(t) = \theta_{jk}(t) - \omega t$ corresponding to patch selection, shown in Figure 4.9.

tained via satellite observation is presented in Figure 4.9 (a). The visual scene contains four spatially separated patches of almost equal brightness. The two-network approach allows us to successfully select all the fragments. The temporal behavior of the oscillator phases for oscillators of the x-layer is shown in Figure 4.10. Similar behavior of oscillator phases occurs for oscillators of the y-layer.

4.3 Relation of the oscillatory network method to other approaches

4.3.1 Relation of the model to controllable dynamical systems

The theory of dynamical system control is a well-developed research field existing during more than 30-year period. The purpose of the control theory is to design the control methods for complicated dynamical systems, that would allow to achieve the desirable dynamical system behavior without a detailed information on the system dynamics. A great variety of control methods has been developed: methods of a direct control, program control, control methods via feedback connections, adaptive control, methods of auto-tuning and self-tuning, optimization methods, dynamic programming methods, etc. The theory of adaptive control is related to the construction of strategies providing desirable dynamical behavior of the controlled system without the complete knowledge of the system. The control can be applied to systems of ordinary differential equations, stochastic differential equations, and stationary random processes. There exist close relations between adaptive control theory and appropriate aspects of signal processing theory, expert systems, fault-tolerant control, intelligent control, filtering algorithms, and learnable neural networks. In many situations the arising changes in controllable dynamics are so large that simple linear feedback controller will not work satisfactory (for instance, this is the case in problems of flight control). Adaptive control includes the possibility of control law modifications in situations when the parameters of controllable system are slowly time varying or uncer-

tain. Adaptive control does not need a priori information about the bounds of the uncertain or time-varying parameters. The design of an optimal strategy of autonomous robot behavior can serve an example of application of adaptive control methods. Self-organized (unsupervised) learning algorithms for recurrent neural networks can also be viewed as application adaptive control to network dynamical systems.

The oscillatory network model described in Section 4.2 is mathematically equivalent to a high-dimensional nonlinear controllable dynamical system. The control consists in preliminary tuning of internal parameters of the oscillatory dynamical system by the array of image pixel brightness. The tunable parameters are the oscillation amplitudes of single network oscillators (limit cycle radii) and the network coupling matrix, which depends on oscillation amplitudes. The type of control should be referred to as a kind of direct adaptive parametrical control. In fact, the control consists in self-consistent tuning of dynamics of the governing dynamical system. The controlling process itself is realized in discrete time and in a off-line manner.

4.3.2 Relation of the model to multi-agent systems

The principle of oscillatory network model design and the character of its performance is closely related to self-consistent performance of multi-agent systems (MAS) that are widely exploited in a great variety of research fields and applications (see Section 1.8). MAS applications cover such fields as computer vision, robotic control, computer-aided design, enterprise modeling, network monitoring, societies simulation, office and home automation, telecommunication, and traffic management.

MAS can be viewed as distributed controllable systems, providing modeling highly complicated nonlinear processes (more complicated than those formulated in terms of usual dynamical systems) admitting adaptive behavior. Robotic mobile intelligent MAS admit hierarchical programmable interconnection architectures. Distributed self-organizing algorithms for complicated program solving are currently under development. The application spheres include a great variety of new modern directions such as creation of distributed algorithms for planning and rational decision making, modeling of artificial social systems, argumentation and preferences modeling, mobile agent-based data mining for diagnosis support, remote control multi-agent systems for healthcare service, adaptive intelligent algorithms for various networks service, agent-based negotiation algorithms for cooperative decision making, agent-based dynamic data storage, intelligent solutions to problems of resource allocation, and agent-based WEB applications. Active development of reinforcement learning algorithms and the other adaptive algorithms for various multi-agent distributed systems is in progress. Also, it seems plausible that many complicated problems that are nowadays solved based on problem-oriented computation algorithms in future will be solved via MAS-based approaches.

Returning to the oscillatory network model discussed in Section 4.2, we note that network oscillators, located at the nodes of spatially distributed lattice and possessing image controllable internal dynamics, could be considered as a community of processors (agents). Cooperative solution to the problem – image processing – is performed by the community via a self-organized interaction of all processors. The interaction of the processors is easily and adaptively reconfigurable dependent on the concrete task to be solved.

Thus, the dynamical oscillatory method of image processing, developed on the base of neuromorphic oscillatory network model with synchronization-based performance, demonstrates flexible, inherently parallel, and automatic self-organized performance in various tasks of image processing.

4.4 Hardware implementation of oscillatory networks

The algorithms for image processing based on oscillatory network approach are highly parallel in essence. Hence, the desire to implement such networks as a standalone integrated circuit chip rather than to emulate them on a serial computer seems quite natural. Self-sustained oscillation systems either of Van der Pol type or Ginzburg–Landau type can be easily created using conventional discrete components such as bipolar and field effect transistors, resistors, and capacitors. Packing of these electronic oscillators into a single chip is well within current technological capabilities. We now briefly describe two implementations of the LEGION model (see Section 3.3.2).

The series of papers [3–5, 20] represents an ongoing work toward the designing the BIOSEG system – a bioinspired VLSI (very-large-scale integration) analog system for image segmentation. The prototype chips contain an array of 16×16 and 32×32 oscillators and are manufactures using 0.8 μm and 0.35 μm complementary metal-oxide semiconductor (CMOS) technologies, respectively. Experiments gave satisfactory results and demonstrated that the proposed network can properly segment small images, including noisy ones. The time required for the segmentation process is determined by the operating frequency of the device and in the present case is about 150 μs.

In the papers [14–16], the authors designed a VLSI chip using 0.35 μm CMOS technology. The chip contained about 100 000 transistors to create a matrix of 32×32 oscillators. It was demonstrated to be capable to perform fast segmentation of real biomedical binary images (magnetic resonance image of rat liver with pancreatic islets was used). Comparison with computer simulation of the same network revealed the advantage of its hardware implementation from the viewpoint of operation speed. The segmentation of 32×32 fragment took about 1 μs, and this should be compared with the simulation time of 0.24 s on a desktop computer operating at 4.3 GHz. The matrix size seems to be too small for real applications, but the authors claim that similar approach can be used to create much larger matrices containing, for example, 256×256 elements. This can be done using modern analog technology, like 65 nm CMOS process. The chip

requires further testing and much work still has to be done to significantly enhance its design. Nevertheless, the obtained results represent an important step toward the hardware implementation of neuromorphic image segmentation approach (with real-life applications) based on oscillatory networks.

It should be noted that in contrast to the LEGION model, our oscillatory network model described in Chapters 3 and 4 has only self-organized dynamical connections and no static ones. This feature could hamper the design of the corresponding integrated circuit, but the implementation still seems to be feasible. Finally we remark that it would be of particular interest to design an optoelectronic device with direct pure optical image input (CCD-like), that could be exploited for image processing in real time, for example, in video tracking tasks.

4.5 Code providing computer experiments on image processing

The collection of computer codes ONN was created for computer experiments on the oscillatory network method testing in different image segmentation tasks. The current version of the ONN code allows us

(1) to integrate the dynamical system governing the oscillatory network dynamics by specifying the image to be processed from the external graphical file and the main controlling parameters (the strength w_0 of oscillator interaction, the radius r_0 of spatial interaction in the oscillatory network, the "mask" Δ restricting the set of network oscillators that are coupled with each given oscillator and the level of network artificial noise);

(2) to display the functions $|u_{jm}(t)|$, $\arg u_{jm}(t)$, $\mathrm{Re}\, u_{jm}(t)$, $\mathrm{Im}\, u_{jm}(t)$, phase trajectories of individual oscillators, the oscillatory network state at any moment of time (in the form of an instantaneous state of the segmented image), the pixel brightness values being defined by either $|u_{jm}(t)|$ or $\arg u_{jm}(t)$.

The maximal size of pixel array of segmented image is restricted by the operating memory size and the computation procedure running time. A typical image segmentation task with image pixel array of 50 000 pixels takes about 10 min for complete problem solving using a desktop computer.

The ONN code provided a great variety of computer experiments on brightness segmentation of real gray level and color images with multipixel arrays and experiments on solving simple versions of vision scene segmentation task with the help of the oscillatory network of the two-layer architecture. The following directions of further model development are possible: the design of new principles of the oscillatory network coupling (in particular, for incorporation of more complicated visual scene segmentation problems), the development of new approaches for active image processing and the development of approaches for moving images processing.

Bibliography

[1] R. Borisyuk, Y. Kazanovich, D. Chik, V. Tikhanoff and A. Cangelosi. A neural model of selective attention and objecy segmentation in the visual scene: An approach based on partial synchronization and star-like architecture of connections. *Neural Networks*, 22:707, 2009.

[2] D. Comaniciu and P. Meer. Mean shift analysis and applications. In *Proc. ICCV'99*, volume 2, page 1197, 1999.

[3] J. Cosp and J. Madrenas. A microelectronic implementation of a bioinspired analog matrix for object segmentation of a visual scene. In *Proc. ESANN'01*, page 69, 2001.

[4] J. Cosp and J. Madrenas. Scene segmentation using neuromorphic oscillatory networks. *IEEE T. Neural Networ.*, 14:1278, 2003.

[5] J. Cosp, J. Madrenas and D. Fernández. Design and basic blocks of a neuromorphic VLSI analogue vision system. *Neurocomputing*, 69:1962, 2004.

[6] R. Eckhorn. Neural mechanisms of scene segmentation: recordings from the visual cortex suggest basic circuits for linking field models. *IEEE T. Neural Networ.*, 10:464, 1999.

[7] R. Eckhorn, R. Bauer, W. Jordan, M. Brosch, W. Kruse, M. Munk and H. J. Reitboek. Coherent oscillations: a mechanism of feature linking in the visual cortex? multiple electrode and correlation analyses in the cat. *Biol. Cybern.*, 60:121, 1988.

[8] R. C. Gonzalez and R. E. Woods. *Digital Image Processing Processing*. Prentice Hall, 2007.

[9] C. M. Gray and W. Singer. Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proc. Natl. Acad. Sci. USA*, 86:1698, 1989.

[10] E. Grichuk, M. Kuzmina and E. Manykin. Oscillatory network for synchronization-based adaptive image segmentation. In *Proc. IJCNN'06*, page 4529, 2006.

[11] E. Grichuk, M. Kuzmina and E. Manykin. Image processing via synchronization in self-organizing oscillatory network. In *Proc. ECC'07, Vol. 1*, volume 27 of *Lecture Notes in Electrical Engineering*, page 97, 2009.

[12] E. S. Grichuk, M. G. Kuzmina and E. A. Manykin. Oscillatory network for synchronization-based adaptive image segmentation. *Opt. Mem. Neural Networks*, 15:57, 2006.

[13] E. S. Grichuk, M. G. Kuzmina and E. A. Manykin. Object selection in visual scene via oscillatory network with controllable coupling and self-organized performance. *Opt. Mem. Neural Networks*, 20:113, 2011.

[14] J. Kowalski and M. Strzelecki. 32×32 oscillator network chip for binary image segmentation. In *Proc. ICSES'08*, page 227, 2008.

[15] J. Kowalski, M. Strzelecki and H. Kim. Implementation of a synchronized oscillator circuit for fast sensing and labeling of image objects. *Sensors*, 11:3401, 2011.

[16] J. Kowalski, M. Strzelecki and P. Majewski. CMOS VLSI chip of network of synchronised oscillators: Functional tests results. In *Proc. IEEE Workshop on Signal Processing 2006*, page 71, 2006.

[17] M. G. Kuzmina and E. A. Manykin. Oscillatory neural network for adaptive dynamical image processing. In *Proc. CIMCA'05*, page 301, 2005.

[18] Z. Li. A neural model of contour integration in the primary visual cortex. *Neural Comput.*, 10:903, 1998.

[19] X. Liu and D. L. Wang. Image and texture segmentation using local spectral histograms. *IEEE T. Image Process.*, 15:3066, 2006.

[20] J. Madrenas, J. Cosp, O. Lucas, E. Alarcón, E. Vidal and G. Villar. BIOSEG: A bioinspired VLSI analog system for image segmentation. In *Proc. ESANN'04*, page 411, 2004.

[21] A. K. Mishra and Y. Aloimonos. Active segmentation. *Int. J. Hum. Robot.*, 6:361, 2009.

[22] A. K. Mishra, Y. Aloimonos, L.-F. Cheong and A. A. Kassim. Active visual segmentation. *IEEE T. Pattern Anal.*, 34:639, 2012.

[23] W. K. Pratt. *Digital Image Processing*. Wiley-Interscience, 2007.
[24] P. Rosenfeld and A. C. Kak. *Digital picture processing*. Academic Press, 1982.
[25] E. Sharon, M. Galun, D. Sharon, R. Basri and A. Brandt. Hierarchy and adaptivity in segmenting visual scenes. *Nature*, 442:810, 2006.
[26] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE T. Pattern Anal.*, 22:888, 2000.
[27] E. P. Simoncelli and W. T. Freeman. The steerable pyramid: a flexible architecture for multi-scale derivative computation. In *Proc. ICIP'95*, volume 3, page 444, 1995.
[28] W. Singer and C. M. Gray. Visual feature integration and the temporal correlation hypothesis. *Annu. Rev. Neurosci.*, 18:555, 1995.
[29] D. Sun, E. Sudderth and M. Black. Layered image motion with explicit occlusions, temporal consistency, and depth ordering. In J. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. S. Zemel and Culotta. A., editors, *Advances in Neural Information Processing Systems*, volume 23, page 2226. MIT Press, 2010.
[30] M. Ursino, G.-E. La Cara and A. Sarti. Binding and segmentation of multiple objects through neural oscillators inhibited by contour information. *Biol. Cybern.*, 89:56, 2003.

5 Parallel information processing and photon echo

5.1 Properties of the photon echo effect

The photon echo effect (or time domain four wave mixing effect) was predicted in 1962 by Kopvillem and Nagibarov [6], and was experimentally observed two years later by Kurnit, Abella, and Hartmann in a ruby crystal [7]. This is a nonlinear optical phenomenon, analogous to a spin echo effect, that consists in the coherent response of optical inhomogeneously broadened mediums (gases, plasma, and crystals) to ultrashort optical pulses. It is widely utilized for studying, for example, processes of relaxation of elementary excitations. Investigations devoted to dynamic spatial and temporal properties of signals of coherent spontaneous emission have formed a separate direction in physics of photon echo. As a consequence, several practical applications of these properties in optical-memory and data-processing systems were proposed. Physically, various manifestations of the specific features of the formation of the temporal and spatial structure of optical echo responses are associated, on the one hand, with the sensitivity of quantum resonant systems to Fourier-transforming properties of incident light pulses and, on the other hand, with the sensitivity of such systems to spatial structure of the wave front of radiation. Optical processors that use photon echo may implement all the basis operations inherent in optical data-processing methods, such as spectral filtration of conventional and spatial frequencies and correlation comparison of optical signals and time-dependent images.

The physical essence of the photon echo phenomenon is as follows. Suppose we have a source of ultrashort laser pulses and some resonant medium. For example, we can use a dye laser that operates in a pulsed mode, and a ruby crystal (a corundum crystal Al_2O_3 with implanted chromium ions Cr^{3+}). The carrier frequency of laser radiation is adjusted to be resonant to some atomic transition of a chromium ion. In the first experiments the wavelength $\lambda = 0.635$ μm was used, which corresponds the photon energy $E = 1.9$ eV. An absorption of a photon excites a chromium ion. The typical duration of single laser pulse is 15–20 ns, whereas the lifetime of an ion excited state is about 20 μs, i.e. 1000 times longer than the pulse duration. Hence, an excited ion can interact with laser field a large number of times before it recombines.

If a ruby crystal is subjected to two laser pulses with an interval τ between them, a coherent excited state of chromium ions is formed. Upon its decay, an optical coherent pulse is emitted – a two-pulsed photon echo signal (see Figure 5.1).

The three-pulsed photon echo phenomena are much more important for scientific and practical applications. Suppose now a crystal is subjected to three pulses. Now an additional echo signal appears as a third pulse with a delay τ, which is equal to the interval τ separating first two pulses (see Figure 5.1).

The damping of an echo signal does not significantly depend on the moment T of the third pulse appearance. The value of T can be much larger than τ. More rigorously,

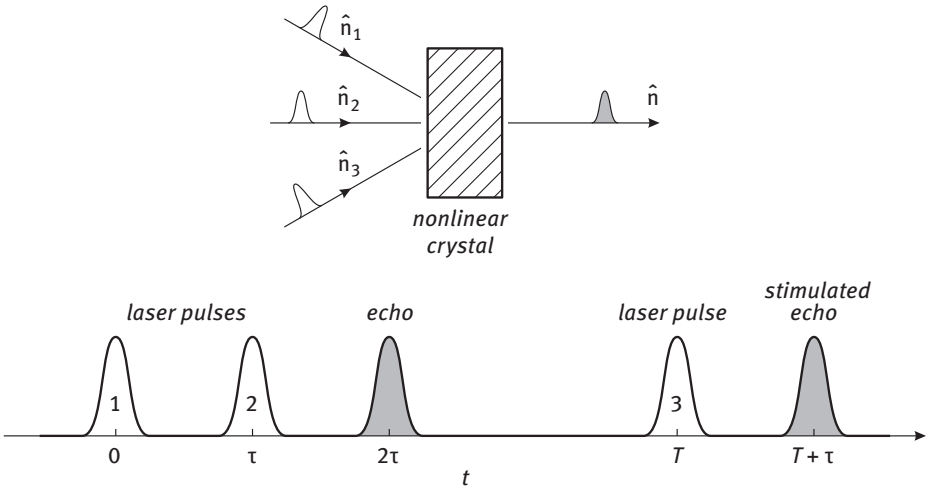

Fig. 5.1. A schematic of the two-pulse and three-pulse (stimulated) photon echo effect.

relaxation processes in resonant media can be characterized (in the simplest cases) by three main time scales: T_1 – longitudinal nonreversible relaxation time that determines the life-time of an excited state, T_2 – transverse nonreversible relaxation time, also known as phase memory time, and T_2^* – transverse reversible relaxation time that is determined by the width of inhomogeneously broadened line. For the photon echo it is important that the conditions $\delta_i < T_{1,2}$ and $\tau < T_{1,2}$ are fulfilled, where δ_i are the durations of excitation pulses. These conditions are natural: the photon echo is a coherent optical phenomenon and all its characteristic time scales should be smaller than nonreversible relaxation time of the medium. In three-pulsed photon echo the time interval T between second and third pulses also has an upper bound: $T < T_1$. There are no strict limiting relations between δ_i, τ, T and T_2^*. Different ratios of these quantities determine different regimes of the photon echo effect.

The typical wavelength of laser field (of the order of 1μm) is much smaller than the typical crystal size. Hence, in contrast to a spin echo, the photon echo is characterized by a spatial directionality: if all laser pulses are plain waves and $\mathbf{k}_1$, $\mathbf{k}_2$ and $\mathbf{k}_3$ ($\mathbf{k}_i = 2\pi\hat{\mathbf{n}}_i/\lambda$ with $\hat{\mathbf{n}}_i$ the direction of propagation of a laser beam) are the wavevectors of the first, the second and the third laser pulses, respectively, then the wavevector of the echo signal is $\mathbf{k} = \mathbf{k}_3 + \mathbf{k}_2 - \mathbf{k}_1$. This condition is known as a spatial wave-matching condition. If $\mathbf{k}_3 + \mathbf{k}_2 = 0$ then the echo signal propagates in the direction opposite to the first pulse direction. Moreover, if the first pulse is a divergent spherical wave, then the echo signal will be a convergent spherical wave. Therefore, the photon echo effect shares some spatial properties of the holography phenomenon.

As was demonstrated in numerous works, the photon echo effect can be used for various tasks of information processing. In such optoelectronic devices input informa-

tion is encoded in optical pulses, and the echo signal generated by a nonlinear crystal figures as the output. We will refer to these devices as echo processors.

5.2 Time processing of optical signals in photon echo processors

We now briefly review main methods to form various echo signals, and describe how these method can be used to implement some types of neural networks. Here we concentrate only on the final results, omitting their microscopic quantum mechanical justification. The details can be found, for example in [8].

Let $E_1(t)$ and $E_2(t)$ denote the envelopes of the electric field of excitation pulses in a two-pulsed photon echo. Then the echo signal $F(t)$ has the following form:

$$F(t) \sim \int G(t_1)E_2(t_2)E_2(t_3)E_1^*(t_1+t_2+t_3+t_1-t+\tau_e)dt_1dt_2dt_3, \tag{5.2-1}$$

where $\tau_e = 2\tau$ is the time of an echo signal emission and $G(t)$ is the Fourier transform of the frequency distribution function $g(\omega)$ of the resonant medium.

An analogous expression is obtained for a three-pulsed photon echo:

$$F(t) \sim \int G(t_1-t_2-t_3+t-\tau_e)E_3(t_2)E_2(t_3)E_1^*(t_1)dt_1dt_2dt_3, \tag{5.2-2}$$

where now $\tau_e = T + \tau$.

These expressions, involving triple integrals, can be significantly simplified under certain conditions. For example, if the excitation pulses have a broad spectrum, i.e. their duration δ_i is much lower than the relaxation time T_2^*, $\delta_i \ll T_2^*$, then the functions $E_i(t)$ can be approximated with delta-functions. The echo signal is then proportional to the $G(t-\tau_e)$, and its duration is characterized by T_2^*. On the other hand, if $\delta_i \gg T_2^*$ (this condition is typical for real experiments), then delta function approximation can be used for $G(t)$. Introducing the Fourier transform of the electric fields amplitudes

$$E_i(\epsilon) = \int E_i(t)e^{i\epsilon t}dt, \tag{5.2-3}$$

the output of the resonant medium for a stimulated echo is described by the expression

$$F(t) \sim \int E_3(\epsilon)E_2(\epsilon)E_1^*(\epsilon)e^{-i\epsilon(t-\tau_e)}\frac{d\epsilon}{2\pi}. \tag{5.2-4}$$

Suppose that one of the input signals $E_1(t)$ (which we will call the coding pulse) has an arbitrary temporal form with the duration δ_1, whereas other signals have much smaller duration, $\delta_i \ll \delta_1$ (and hence, much greater spectral width). Then all functions in (5.2-1), except $E_1(t)$, can be approximated with delta functions, and for the echo response we get

$$F(t) \sim E_1^*(\tau_e - t). \tag{5.2-5}$$

Hence, in this case the echo signal is just the time reversal of the input optical pulse. Similar situation is observed in a three-pulsed photon echo. Here, along with

time reversal of an optical pulse, it is also possible to introduce a time delay into its propagation without changing its form. It was shown [5, 9, 10] that the role of the coding pulse can be played by any of three excitation pulses. If the coding pulse comes first, the echo signal is time reversed, and if it comes second of third, only the time delay is introduced. For example, if $E_2(t)$ in (5.2-2) represents the coding pulse and two other pulses are much shorter, $\delta_{1,3} \ll \delta_2$, from (5.2-2) we obtain

$$F(t) \sim E_2(t - \tau_e). \tag{5.2-6}$$

The first experimental observation of this effect was performed in ruby crystal with implanted Cr^{3+} ions [12]. These experiments were later confirmed with numerous other nonlinear media, including crystal with paramagnetic impurities and gases. It is important to note, that the time delays in some experiments were as large as a few hours (see, e.g. [2]).

Similarly, if the spectral width of one of the functions involved in the product in expression (5.2-4) is considerably greater than the widths of other two functions, this function can be factored out outside the integral. Then, the envelope of the echo signal is proportional to

$$\begin{aligned} F_{\text{corr}}(t) &\sim \int E_2(t_1)E_1^*(t_1 - t + \tau_e)dt_1, \\ F_{\text{conv}}(t) &\sim \int E_3(t - \tau_e - t_1)E_2(t_1)dt_1, \end{aligned} \tag{5.2-7}$$

i.e. the correlation and convolution functions of the excitation pulses amplitudes.

5.3 Implementation of vector–matrix multiplier based on photon echo

Photon echo can be used for implementations of optical data-processing methods. In particular, a vector–matrix multiplier (VMM) can be built based on this effect [1, 11].

Since under certain conditions a photon echo signal can represent correlation and convolution functions of excitatory pulses, it can be employed for implementation of multiplication using an algorithm of digital multiplication by analog convolution (DMAC) [4]. Suppose, we want to multiply numbers $A = 22$ and $B = 9$. Their binary representations $A = 10110_2$ and $B = 01001_2$ should correspond to waveforms of light pulses (see Figure 5.2 (a)). The calculation of a product of two numbers A and B using the formula of convolution (5.2-7) yields waveforms shown in Figure 5.2 (b), which correspond to a number AB written in a mixed representation, $AB = 10120110_m$. An easy postprocessing is required to obtain a standard binary representation. It can be done using a shift-and-add algorithm [4]: the amplitude of the kth pulse of the output signal should be converted into binary representation, shifted by k digits, and then all the resulting values should be summed up. This process is demonstrated in Figure 5.2 (c). Its outcome is the result of multiplication: $AB = 11000110_2 = 198$.

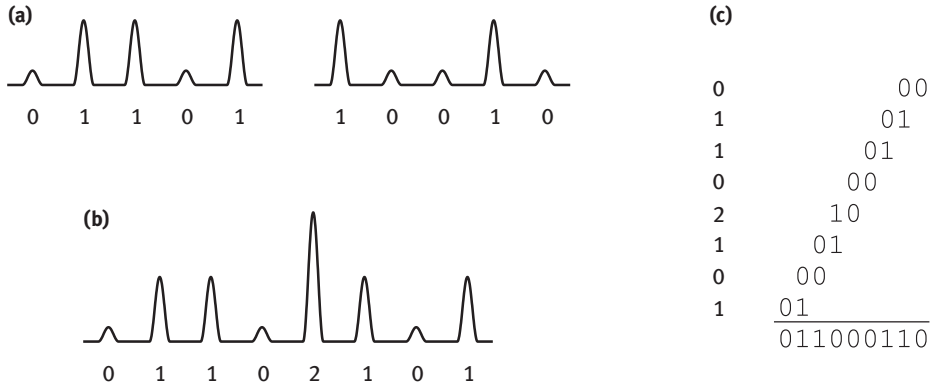

Fig. 5.2. DMAC algorithm in the time domain: (a) waveform of pumping light pulses (binary coding); (b) the result of the multiplication of two numbers in a mixed representation; (c) postprocessing with a shift-and-add algorithm.

A photon echo also makes it possible to produce optical images. In this case, a resonant medium plays the role of a spectral-selective hologram that is also sensitive to spatial frequencies. In particular, if a resonant medium coincides with the plane of spatial frequencies, photon-echo signals produce optical images that represent the correlation or the convolution of initial images in spatial coordinates,

$$F_{\text{corr}}(\mathbf{r}) \sim \int E_2(\mathbf{r}_1)E_1^*(\mathbf{r}+\mathbf{r}_1)d^3r_1,$$
$$F_{\text{conv}}(\mathbf{r}) \sim \int E_3(-\mathbf{r}-\mathbf{r}_1)E_2(\mathbf{r}_1)d^3r_1. \tag{5.3-1}$$

Thus, the DMAC algorithm can also be implemented for multiplied numbers in a parallel spatial coding on the basis of optical image processing. Generally, a photon echo provides an opportunity of spatial-temporal correlation analysis of time-dependent images. These opportunities are associated with the following fact: if a resonant medium is employed as a spectral-selective hologram that displays selective properties with respect to spatial frequencies, then the spatial and temporal frequencies enter into the total Fourier transform of electric field amplitude independently of each other. This allows us to perform correlation processing of time-dependent images.

Photon echo is a phenomenon that is determined by a coherent interaction of optical pulses with a resonant medium. Consequently, in implementing the DMAC algorithm with the use of the photon echo, we perform correlation processing or comparison of slowly varying amplitudes of optical pulses represented in the form of convolution integrals either in temporal or in spatial domain. At the same time, photodetectors can register a signal proportional to the intensity of incident light. In other words, the detected signal is quadratic in the amplitude. Therefore, a procedure of addition or integration in time, when a signal proportional to number digits in a mixed representation is accumulated in time, loses its meaning in echo processors. Integration or

amplitude addition in space is possible due to the use of coherent optical schemes. In this case, to produce optical images with a complex spatial structure of the wave front, one can employ schemes with the expansion of the laser beam and subsequent application of a spatial-temporal optical modulator (STOM). From the viewpoint of the DMAC algorithm implementation, two approaches are appropriate in echo processors with correlation comparison being implemented either in time or in space.

5.3.1 Photon echo vector–matrix multiplier with spatial integration

The pixel modification of an echo processor employs a binary representation of the elements of both a matrix $\hat{A}$ and a vector **B** in the form of a time-sequential code. Processing that corresponds to the calculation of the product of the two numbers $A_{ik}B_k$ is performed in independent spatially structured areas (pixels). The number of pixels in an STOM is equal to the number of pixels in a resonant medium (see Figure 5.3 (a)). Thus, each pixel of a resonant medium calculates the convolution or the correlation function of signals that correspond to a pair of light pulses in time domain (in the case of a two-pulse photon echo, the numbers being multiplied within various time intervals are related only to the second light pulse).

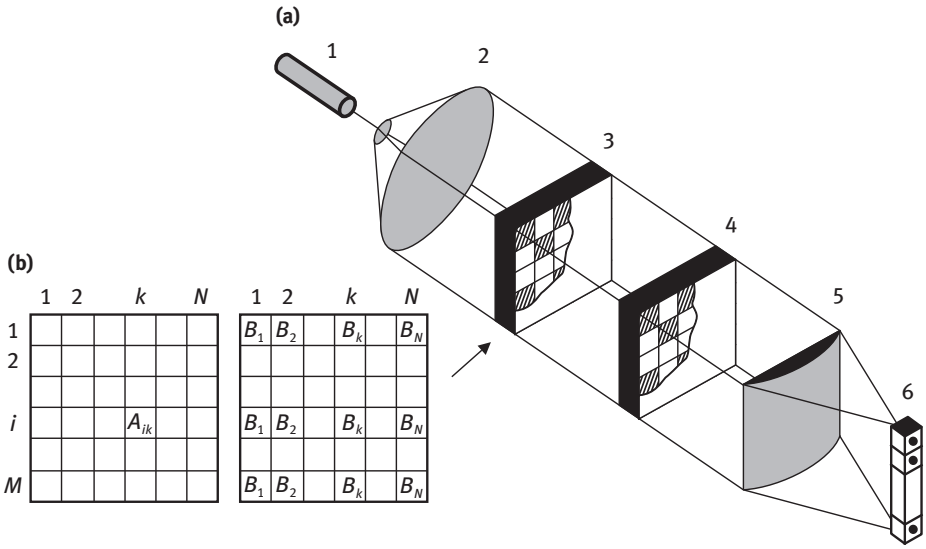

Fig. 5.3. A photon echo vector–matrix multiplier with spatial integration: (a) a simplified schematic: 1 – laser, 2 – former of light pulses and lenses that shape waves with a plane wave front, 3 – STOM, 4 – resonant medium, 5 – cylindrical lens that implements spatial integration, 6 – array of photodetectors; (b) images produced by a STOM in the form of two pumping light pulses for calculating a vector–matrix product.

The dimensions $M \times N$ of an image produced by the modulator correspond to the numbers of rows and columns of the matrix $\hat{A}$ or to the number M of identically arranged vectors **B** with elements $B_1, B_2, \ldots, B_N$. (see Figure 5.3 (b)). During the action of one of the pumping light pulses, the STOM should form different images. The number of such images is equal to the number of digits in the employed representation. Therefore, an STOM operating in this regime requires an electrically independent addressing of working elements (cells) of the modulator. Addition of different elements of the matrix $\hat{A}$ with the elements of the vector **B** is performed by a cylindrical lens. If coherent light is used, an STOM performs addition of the field amplitudes.

Eventually, an array of M photodetectors should register a signal proportional to $(\sum_k A_{ik}B_k)^2$ in a mixed representation in the form of a sequential code. Transformation of this code into a binary representation with subsequent element-by-element digitization and addition with a shift should involve conventional electronic means.

5.3.2 A holographic photon echo VMM scheme with spatial integration

In the situation considered in the previous section, the role of the STOM was reduced to the formation of a series of two-dimensional images within the time equal to the duration of light pulses encoded with the number of digits of the employed representation. Thus, the DMAC algorithm was implemented in the time domain. A scheme of the holographic type operates with numbers in a parallel code, and the DMAC algorithm is implemented in space. Similarly to the above-considered scheme, each of the images produced by an STOM in the case under study represents a space-pixel structure of domains that either transmit or do not transmit light (in accordance with encoding in the binary representation). At a definite moment in time, the image forms a vector **A** with N elements and the number of digits involved in the binary representation equal to R (see Figure 5.4 (b)). If a time-dependent image features change during the time interval corresponding to the action of one of the incident light pulses, then the above-specified procedure defines a matrix of elements consisting of M vectors (M rows). A complex optical system shown in Figure 5.4 (a) performs one-dimensional Fourier transformation in the x-coordinate, forms an optical image and compensates for the phase difference in the y-coordinate. Thus, a resonant medium represents a plane of a multichannel dynamic hologram with one-dimensional filtration of spatial frequencies.

The distribution of the field in the plane of the hologram can be written as

$$E_j(x', y') = \int E_j(x, -y') \exp\left(-\frac{ikxx'}{F}\right) dx, \quad j = 1, 2, \tag{5.3-2}$$

where k is the wave number, and F is the focal length of spherical and cylindrical lenses. Then, a signal arising in the observation plane (x_α, y_α) is proportional to the

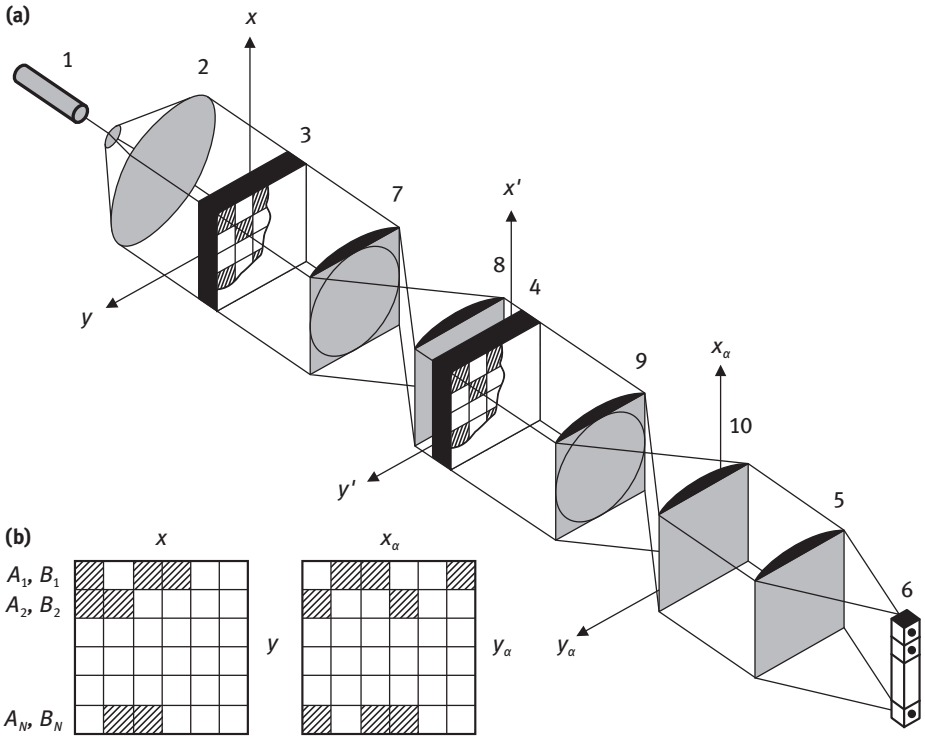

Fig. 5.4. Holographic VMM scheme based on photon echo: (a) 1–6 – components similar to those shown in Figure 5.3, 7, 9 – combinations of spherical and cylindrical thin lenses that implement image formation in the y-coordinate and one-dimensional Fourier transformation with respect to the x-coordinate, 8, 10 – cylindrical lenses that implement phase compensation; (b) data representation in parallel coding for a holographic device.

correlation function or the convolution-type function,

$$F_{\text{corr}}(x_\alpha, y_\alpha) \sim \int E_1^*(x', -y_\alpha) E_2(x', -y_\alpha) \exp\left(-\frac{ikxx'}{F}\right) dx',$$
$$F_{\text{corr}}(x_\alpha, y_\alpha) \sim \int E_2(x', -y_\alpha) E_3(x', -y_\alpha) \exp\left(-\frac{ikxx'}{F}\right) dx', \tag{5.3-3}$$

or

$$F_{\text{conv}}(x_\alpha, y_\alpha) \sim \int E_1^*(x, y_\alpha) E_2(x - x_\alpha, y_\alpha) dx,$$
$$F_{\text{conv}}(x_\alpha, y_\alpha) \sim \int E_2(x, y_\alpha) E_3(-x - x_\alpha, y_\alpha) dx. \tag{5.3-4}$$

Thus, a set of products A_iB_i, $i = 1, \ldots, N$ in a mixed representation is produced in the (x_α, y_α) plane, and a cylindrical lens placed in front of an array of photodetectors

performs the operation of addition

$$\sum_{i=1}^{N} A_i B_i.$$

The array of photodetectors produces a signal that is proportional to the square of the scalar product $(\mathbf{AB})^2$ of the vectors **A** and **B**.

Provided that the STOM forms a set of M images during the time interval corresponding to the action of one of the light pulses (which corresponds to the definition of M vectors or a data matrix), then we can obtain a set of M signals proportional to $(A_i\mathbf{B})^2$, $i = 1, 2, \ldots, M$ at the output of the resonant medium. The signal of a two-pulse photon echo can be considered as a particular case of degenerated stimulated photon echo when $E_2 = E_3$. Then, similarly to time-domain coding, the image that corresponds to the second pumping light pulse in two-pulse photon echo should contain various regions of the spatial coding of the numbers being multiplied. Thus, in contrast to an optical echo processor considered in the previous section (which represented a pixel structure with spectral filtration in conventional frequency domain), in the case analyzed in this section, a resonant medium plays the role of a dynamic multichannel hologram with spectral filtration in one-dimensional spatial frequency domain.

5.3.3 Estimates of the main technical characteristics of digital echo processors

The most important characteristics of echo processors are: (a) the speed, which determines, to a great extent, the performance of the system as a whole; (b) energy consumption per one data bit processed, and (c) the capability to operate in a conveyer (continuous) mode.

The speed of a processor is determined by the minimum time that is required to complete a single operating cycle of vector–matrix multiplication. If a stimulated photon echo is employed as a functional method of data processing, this time is determined by the durations δ_i of three incident light pulses and the duration of the photon echo signal, which can be assumed to be greater than all the δ_i and all the time intervals between the pumping light pulses and the photon echo signal itself. In the regime, when the system calculates the correlation function or the convolution, the duration of one of the pumping pulses should be less than the durations of other two pulses. Therefore, to estimate the minimum time of the working cycle, we can use the expression $t_{\min} = 3\delta + 3\tau + \tau^*$, where $\tau^* \geq T_1$ is the time required for the resonant medium to return to the ground state. The value $t_{\min} \sim 100$ ns can be attained experimentally.

Let us estimate the feasible performance of an echo processor. Since a VMM operates with MN numbers, the implementation of vector–matrix multiplication requires $2MN$ operations of multiplication and addition. The characteristic time that is re-

quired to perform these operations is mainly determined by the time intervals between pumping light pulses. Choosing $M \sim N \sim 20$, we can estimate the performance of the system as $2MN/t_{\min} \sim 10^{10}$ operations of multiplication and addition per second. To estimate the speed (in bits per second) of binary data processing, we should multiply the above result by the number of digits in the employed numbers.

5.4 Optical implementation of neural networks based on photon echo

Presently, optical implementation of associative memory network models (see Section 1.4) is an important subarea of optical computing and data processing. Successful realizations of associative memory networks, using the properties of Fourier transform holography, photorefractive crystals, and spatial light modulators have been reported. These elements are necessary to ensure the large number of interconnections of optical neural networks.

The photon echo effect allows the realization of an optically controlled matrix of interconnections. This matrix can operate during a time interval less than the population relaxation time of an optical medium. Moreover, all pulses in the photon echo effect are separated in time, and this can improve the architecture of optical neural networks in some cases.

5.4.1 Outer product scheme for optical neural networks

The properties of the photon echo effect can be used to realize an optical neural network [3]. Figure 5.5 presents the implementation principles of the Hopfield network based on the echo phenomenon. This system consists of input and output linear arrays (N elements) and an $N \times N$ storage matrix which is an inhomogeneously broadened medium interacting with a sequence of resonant optical pulses.

To store a network interconnection matrix, two 1D spatial light modulators (SLM) oriented along orthogonal directions (see Figure 5.5 (a)) can be used. These modulators allow the generation of the pulse train illustrated in Figure 5.5 (c), where $E_i^{(m)}$ is the light amplitude of the ith element of the input array, corresponding to the mth memory image. The optical pulses interact with the inhomogeneously broadened medium, which ensures the storage of the outer product matrix $E_i^{(m)*} E_j^{(m)}$. This implies that each pulse $E_i^{(m)}$ is the first pulse and each pulse $E_j^{(m)}$ is the second pulse of the excitation echo sequence in Figure 5.1 (c). These data pulses for each value of the index m are separated by a time interval T. The optics required to increase the dimensions of the input signals is omitted in the figures for simplicity.

The recognition stage is shown in Figure 5.5 (b). The image $E_i^{(in)}$ is supposed to be processed by the optical neural network. This 1D signal is generated with the help of a

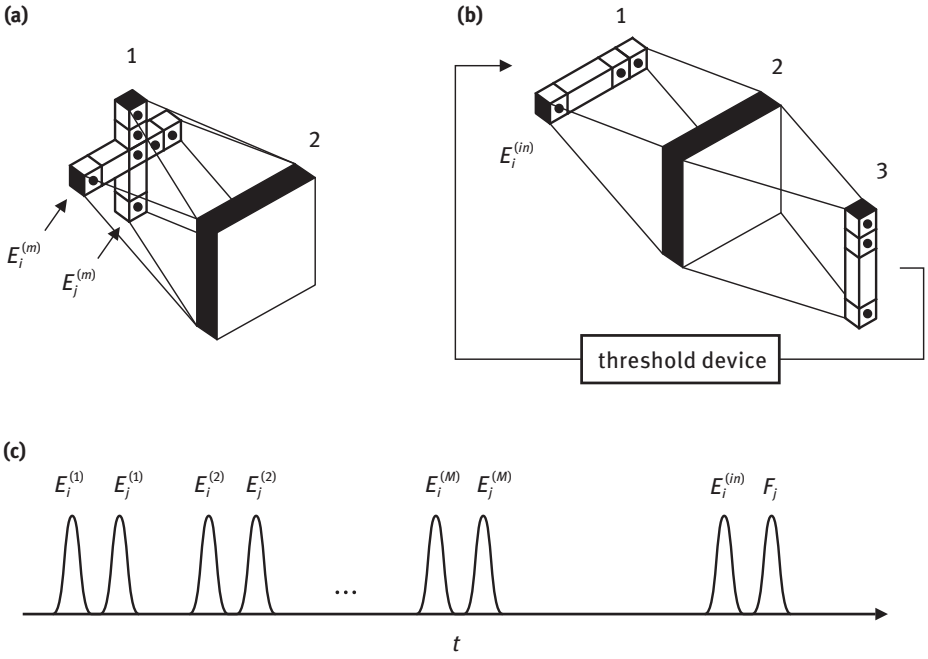

Fig. 5.5. Outer product scheme of the optical implementation of a Hopfield network based on the properties of the photon echo: (a) storage of the outer product scheme; (b) the recognition stage, 1 – the array of light emitters, 2 – resonant medium, 3 – the array of photodetectors.

horizontally oriented input SLM and then is spread in the vertical direction by means of the appropriate optics. A pulse with such an amplitude distribution causes a great number of echo responses corresponding to the pulse sequence that is preliminarily stored in the medium. It is seen that within a time interval τ after the recognition stage has begun the stimulated photon echo signal appears in the output plane,

$$F_j \sim \sum_{i=1}^{N} \left(\sum_{m=1}^{M} E_i^{(m)*} E_j^{(m)} \right) E_i^{(in)}, \tag{5.4-1}$$

where M is the number of stored images. The summation over the index i in (5.4-1) is carried out by anamorphic optics, which yields this 1D vertically oriented distribution. The output stimulated echo should be detected at the time $t = T + \tau$, where T corresponds to the beginning of the recognition stage. To separate the useful echo signal from the noise, an optical shutter can be placed before the photodetector array plane. The noise echoes arise at the times $T + t_0$, $T + t_0 + \tau$, $T + t_0 + 2\tau$ and so on. Consequently, if the time interval between adjacent data pulse pairs of the stored sequence in Figure 5.5 (c) is sufficiently longer than τ, the useful echo signal with the amplitude given by (5.4-1) can be selected by means of an optical shutter. If a threshold operation and

feedback are realized, as shown in Figure 5.5 (a), this system can be considered as an optical implementation of a Hopfield neural network.

5.4.2 Inner product scheme for optical neural networks

It is meaningful to investigate the optical implementation of an associative memory network with access to the correlation domain. This approach allows us to use nonlinearities in the correlation plane to considerably improve the network information capacity. The inner product scheme shown in Figure 5.6 is an example of an optical scheme with access to the correlation domain. The main idea is to change the order of the summations in (5.4-1). First, inner products of the input vector and memory vectors, which are stored in the $N \times M$ matrix columns, are calculated. Then these inner products weigh the corresponding memory vectors in the second matrix. Finally, the summation over the index m is carried out by means of anamorphic optics, yielding the output column. The realization of such optical scheme is possible using the photon echo effect. In this case, the storage matrix is a resonant inhomogeneously broadened optical medium. To store information, the first pulse should be a reference wave with amplitude $E_i = 1$ $(i = 1, 2, \ldots, N)$ for all M columns. The second pulse presents memory images that are stored in the columns of the first $N \times M$ matrix. The second matrix in Figure 5.6 is prepared in the same manner.

Upon arrival of the input image $E_i^{(in)}$, the stimulated photon echo signal appears in the output plane. The echo amplitude determines the output image and has the form

$$F_j \sim \sum_{m=1}^{M} \left(\sum_{i=1}^{N} E_i^{(m)} E_i^{(in)} \right) E_j^{(m)}. \tag{5.4-2}$$

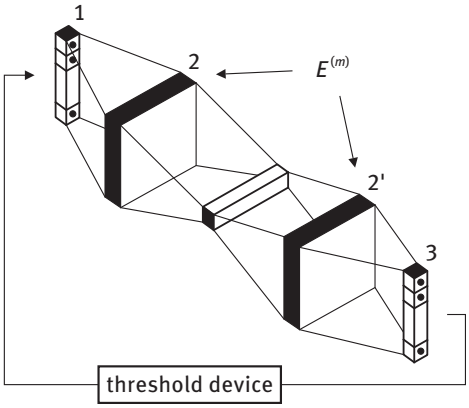

Fig. 5.6. Inner product scheme of the optical implementation of a Hopfield network based on the properties of the photon echo effect, 1 – array of light emitters, 2, 2' – resonant medium, 3 – array of photodetectors.

The considered system can store and retrieve only 1D optical arrays, because one transverse coordinate is needed for spatial separation of the memory vectors. In the next section we show how the properties of the photon echo effect allow us to process 2D optical arrays by replacing the spatial separation with a temporal one.

5.4.3 Inner product scheme for 2D array processing

A modified optical associative memory scheme, similar to that described in the previous section, can be introduced. The 2D optical arrays can be processed with conser-

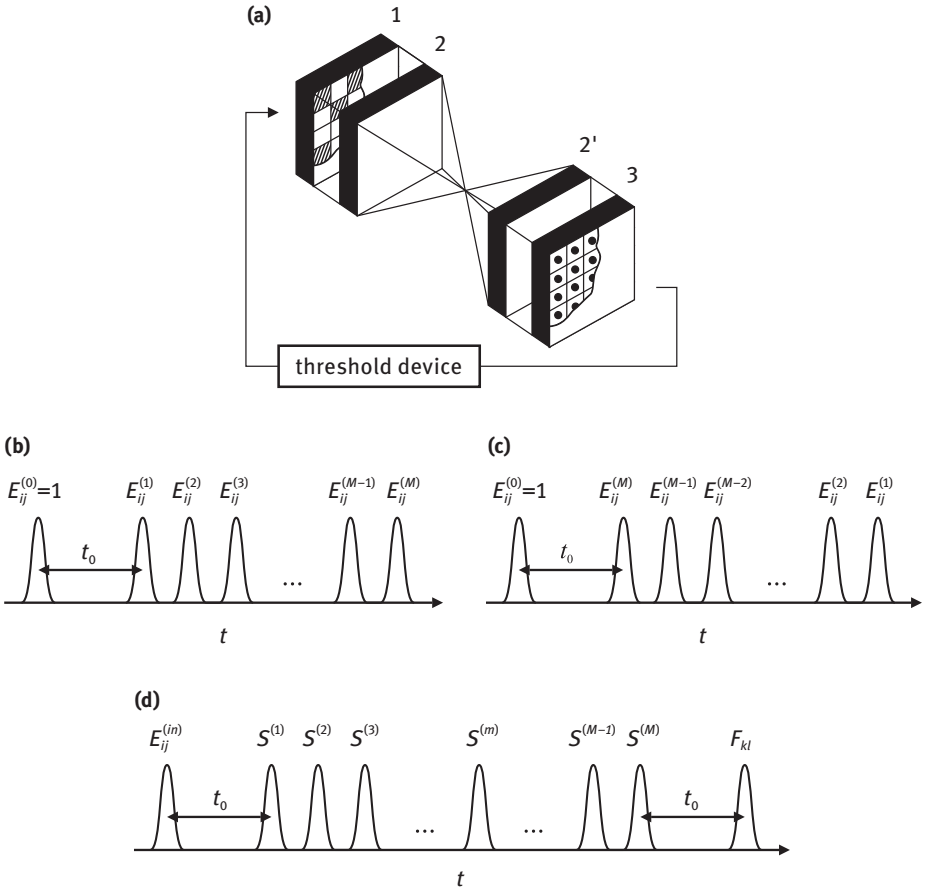

Fig. 5.7. a) Inner product scheme of the optical implementation of a Hopfield network for 2D array processing, 1 – the 2D array of light emitters, 2, 2' – resonant medium, 3 – the 2D array of photodetectors; (b) data pulse sequence for memory storage in the left matrix 2; (c) the same for the right matrix 2'; (d) optical pulse sequence during the recognition phase.

vation of parallelism, as shown in Figure 5.7 (a). Optical pulses with the 2D-amplitude distributions of the data can be generated with the help of a 2D SLM placed in the input plane. Then, these signals interact with the 2D storage matrix, which is an inhomogeneously broadened medium. Figure 5.7 (b) and (c) shows the data pulse sequence required to store memory images in the first and second matrices, respectively. The first pulse of each sequence is a reference wave and has an amplitude equal to unity for all 2D array elements. The time condition $t_0 > \tau(M-1)$ should be satisfied. It should be stressed that the overall durations of the data pulse train is limited by the dephasing time T_2.

Upon arrival of the input image $E_{ij}^{(in)}$, the temporal sequence of the stimulated echo signals appears in the correlation domain, as shown in Figure 5.7 (d). After summation in the transverse plane, their amplitudes are the inner products of the input and memory images,

$$S(m) \sim \sum_{ij} E_{ij}^{(m)} E_{ij}^{(in)}, \tag{5.4-3}$$

where $m = 1, 2, \ldots, M$ and the corresponding times are $t_m = T + t_0 + (m-1)\tau$. Then, these pulses weigh the memory images preliminarily stored in the second matrix and induce a set of stimulated echo responses that propagate into the output plane. The useful signal should be detected at the time $T + 2t_0 + (M-1)\tau$. The 2D amplitude distribution of the signal is

$$F_{kl} \sim \sum_m S(m) E_{kl}^{(m)} = \sum_m E_{kl}^{(m)} \sum_{ik} E_{ij}^{(m)} E_{ij}^{(in)}. \tag{5.4-4}$$

The nearest noise stimulated echoes are separated from this echo pulse by the time interval τ. Therefore, in this case an optical shutter can also be used to select the output signal with the amplitude given by (5.4-4). For this purpose, the delay τ in the data pulse sequence is chosen according to the operating time of the optical shutter.

The above-considered scheme, together with the threshold device and feedback, represent the optical implementation of a Hopfield neural network. One can see that the presence of the time coordinate in the echo phenomenon permits us to increase the dimensions of the optical array processed by the network. The temporal properties of the photon echo effect are also important for implementation of some other neural networks models. These are higher order neural networks and a neural network for storage and retrieval of a complex temporal sequence of images.

Bibliography

[1] W. R. Babbitt, Y. S. Bai and T. W. Mossberg. Convolution, correlation, and storage of optical data in inhomogeneously broadened absorbing materials. In *Proc. SPIE, Optical Information Processing II*, volume 639, page 240, 1986.

[2] W. R. Babbitt and T. W. Mossberg. Time-domain frequency-selective optical data storage in a solid-state material. *Opt. Commun.*, 65:185, 1988.

[3] M. N. Belov and E. A. Manykin. Optical implementation of neural network models using the photon echo effect. *Opt. Commun.*, 84:1, 1991.
[4] L. N. Binh. *Photonic Signal Processing: Techniques and Applications*. CRC Press, 2007.
[5] S. O. Elyutin, S. M. Zakharov, V. A. Zuikov, E. A. Manykin and V. V. Samartsev. Spatial and temporal correlations of coherent optical pulses in the photon echo effect. *JETP*, 61:234, 1985.
[6] U. Kh. Kopvillem and V. R. Nagibarov. Photon echo in paramagnetic crystals. *Fiz. Met. Metalloved.*, 15:313, 1963. in Russian.
[7] N. A. Kurnit, I. D. Abella and S. R. Hartmann. Observation of a photon echo. *Phys. Rev. Lett.*, 13:567, 1963.
[8] E. A. Manykin and V. V. Samartsev. *Optical echo-spectroscopy*. Nauka, 1984. in Russian.
[9] T. W. Mossberg. Time-domain frequency-selective optical data storage. *Opt. Lett.*, 7:77, 1982.
[10] S. M. Zakharov and E. A. Manykin. Temporal and correlation properties of echo signals in two- and three-level systems under conditions of inhomogeneous broadening of resonant energy levels. *JETP*, 64:761, 1986.
[11] S. M. Zakharov and E. A. Manykin. Digital vector-matrix multiplication on the basis of a photon echo. *Laser Phys.*, 6:793, 1996.
[12] V. A. Zuikov, V. V. Samartsev and R. G. Usmanov. Correlation of the shape of light echo signals with the shape of the excitation pulses. *JETP Lett.*, 32:270, 1980.

6 Stochastic oscillators for modeling polarized light beams

6.1 Polarization analysis of quasimonochromatic light beams

6.1.1 Classical description of plane electromagnetic wave

Free classical electromagnetic field in the empty space satisfies homogeneous Maxwell equations [4]

$$\begin{aligned} \nabla \times \mathbf{E}(\mathbf{r},t) &= -\frac{1}{c}\frac{\partial \mathbf{H}(\mathbf{r},t)}{\partial t}, \qquad \nabla \cdot \mathbf{E}(\mathbf{r},t) = 0, \\ \nabla \times \mathbf{H}(\mathbf{r},t) &= \frac{1}{c}\frac{\partial \mathbf{E}(\mathbf{r},t)}{\partial t}, \qquad \nabla \cdot \mathbf{H}(\mathbf{r},t) = 0, \end{aligned} \tag{6.1-1}$$

where $\mathbf{E}(\mathbf{r},t)$ and $\mathbf{H}(\mathbf{r},t)$ are the electric and magnetic field vectors, respectively, at the space–time point $(\mathbf{r},t)$. Here and in what follows we use cgs units.

The vector potential $\mathbf{A}(\mathbf{r},t)$, related to $\mathbf{E}(\mathbf{r},t)$ and $\mathbf{H}(\mathbf{r},t)$ by the expressions

$$\mathbf{E}(\mathbf{r},t) = -\frac{1}{c}\frac{\partial \mathbf{A}(\mathbf{r},t)}{\partial t}, \quad \mathbf{H}(\mathbf{r},t) = \nabla \times \mathbf{A}(\mathbf{r},t), \tag{6.1-2}$$

satisfies the homogeneous wave equation, which can be obtained from (6.1-1),

$$\nabla^2 \mathbf{A}(\mathbf{r},t) - \frac{1}{c^2}\frac{\partial^2 \mathbf{A}(\mathbf{r},t)}{\partial t^2} = 0. \tag{6.1-3}$$

The solution of the wave equation specifies a plane transversal electromagnetic wave, for which both the vector potential $\mathbf{A}(\mathbf{r},t)$, and electric and magnetic fields $\mathbf{E}(\mathbf{r},t)$ and $\mathbf{H}(\mathbf{r},t)$ undergo oscillations in the plane orthogonal to the direction of the wave vector $\mathbf{k}$. Thus, in some real orthonormal basis $\mathcal{B} = \{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$, satisfying the conditions

$$\mathbf{e}_x \cdot \mathbf{e}_x = \mathbf{e}_y \cdot \mathbf{e}_y = 1, \quad \mathbf{e}_x \cdot \mathbf{e}_y = 0, \quad \mathbf{e}_z = \frac{\mathbf{k}}{|\mathbf{k}|} = \mathbf{e}_x \times \mathbf{e}_y, \tag{6.1-4}$$

the electric field of a plane monochromatic electromagnetic wave can be written in the form

$$\mathbf{E}(\mathbf{r},t) = \sum_{s=1}^{2} a_s \,\mathrm{Re}[e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t-\alpha_s)}]\mathbf{e}_s, \tag{6.1-5}$$

where the real amplitudes a_s and phases α_s determine the amplitude and polarization of the wave. For example, if $a_x = 0$, then the vector $\mathbf{E}(\mathbf{r},t)$ is always parallel or antiparallel to the fixed vector $\mathbf{e}_y$, i.e. the wave is linearly polarized. In contrast, if $a_x = a_y$ and $|\alpha_x - \alpha_y| = \pi/2$, then the end of the vector $\mathbf{E}(\mathbf{r},t)$ traces our a circle (in every fixed point $\mathbf{r}$), and the wave is circularly polarized.

In the general case, the end of vector $\mathbf{E}(\mathbf{r},t)$ traces out a spiral line in the space, which projection onto the plane $(\mathbf{e}_1, \mathbf{e}_2)$ is an ellipse, known as polarization ellipse (see Figure 6.1).

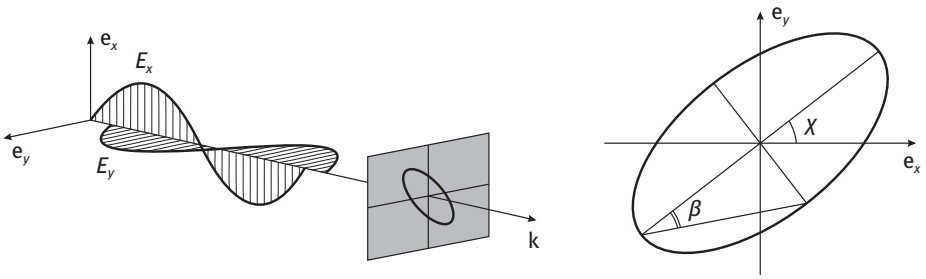

Fig. 6.1. Polarization ellipse.

Every monochromatic wave always has certain polarization. Real waves, however, are not monochromatic and contain components with frequencies in some narrow interval $\Delta\omega$:

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0(\mathbf{r}, t)e^{-i\overline{\omega}t}, \tag{6.1-6}$$

where $\mathbf{E}_0(\mathbf{r}, t)$ is a slow-varying function of time. Hence, polarization of the wave in every fixed point $\mathbf{r}$ varies in time. For adequate description of polarization plane electromagnetic wave the electric field vector $\mathbf{E}(\mathbf{r}, t)$ should be considered as a random vector function of time, or random process, because in practice we always deal with fluctuating electromagnetic fields. Then Figure 6.1 displays the behavior of $\overline{\mathbf{E}}(\mathbf{r}, t)$, the statistical mean of $\mathbf{E}(\mathbf{r}, t)$. The analysis of polarization of fluctuating electromagnetic waves can be carried out in the frames second-order correlation effects of two-dimensional vector random processes.

6.1.2 Polarization analysis in terms of coherence matrix

We consider well collimated, uniform beams of quasimonochromatic light of mean frequency $\overline{\omega}$. The complex amplitudes $E_x(t) = a_x(t)e^{-i\alpha_x(t)}$ and $E_y(t) = a_y(t)e^{-i\alpha_y(t)}$ in the expression (6.1-5) can be considered as slowly varying functions (in comparison with $\cos\overline{\omega}t$ and $\sin\overline{\omega}t$). We suppose them to be stationary random processes (random functions of t), so that $\mathbf{E}(t) = [E_x(t), E_y(t)]^T$ is a stationary two-dimensional vector random process. Let $\overline{\mathbf{E}}(t)$ be the mean of $\mathbf{E}(t)$. For stationary random processes the statistical mean coincides with the mean over time (the ergodicity [8, 9]), that is

$$\overline{\mathbf{E}}(t) = \langle \mathbf{E}(t) \rangle = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \mathbf{E}(t)dt. \tag{6.1-7}$$

It is known from the classical optics that second-order correlation characteristics of fluctuating electric field $\mathbf{E}(\mathbf{r}, t)$ are responsible for various interference phenomena. The properties of a partially polarized quasimonochromatic beam are determined by

the coherence matrix $\hat{J}$:

$$\hat{J} = \langle \mathbf{E}(t)\mathbf{E}^{\dagger}(t)\rangle = \begin{pmatrix} J_{xx} & J_{xy} \\ J_{yx} & J_{yy} \end{pmatrix} = \begin{pmatrix} \langle E_x E_x^*\rangle & \langle E_x E_y^*\rangle \\ \langle E_y E_x^*\rangle & \langle E_y E_y^*\rangle \end{pmatrix}. \tag{6.1-8}$$

The diagonal elements of $\hat{J}$ define the intensities (average electric energy densities) of electromagnetic wave components with electric field E_x and E_y, whereas the element J_{xy} defines correlation between the components. Other quadratic combinations of the electric field, such as $E_x(t)E_x(t)$ or $E_x(t)E_y(t)$, correspond to fast-oscillating factors $e^{\pm 2i\omega t}$ in (6.1-5) and vanish upon averaging.

There is the following relation between this matrix and covariation matrix $\hat{D}(\tau)$ of the random process $\mathbf{E}(t)$:

$$D_{ij}(\tau) = \langle E_i(t)E_j^*(t+\tau)\rangle, \quad J_{mn} = D_{mn}(0). \tag{6.1-9}$$

The trace of the coherence matrix is proportional to the average electric energy density W_e of the light beam, that is, to the light beam intensity I: $\operatorname{tr}\hat{J} = \langle E_x E_x^*\rangle + \langle E_y E_y^*\rangle = 4\pi W_e = I$. Factoring out the light beam intensity, we get the normalized coherence matrix $\hat{j}$,

$$\hat{j} = \frac{1}{\operatorname{tr}\hat{J}}\hat{J} = \frac{1}{I}\hat{J}. \tag{6.1-10}$$

The matrix $\hat{j}$ describes the polarization properties of the light beam, and its off-diagonal element

$$j_{xy} = |j_{xy}|e^{i\gamma_{xy}} = \frac{J_{xy}}{J_{xx}+J_{yy}} \tag{6.1-11}$$

defines statistical correlation between $E_x(t)$ and $E_y(t)$. The value $|j_{xy}|$ is a measure of correlation between the components $E_x(t)$ and $E_y(t)$, and the value γ_{xy} is the effective phase difference between them.

In optics the Stokes parameters I, Q, U and V, which represent the real-valued linear combinations of J_{mn}, are frequently used for characterization of polarization state of a light beam. The Stokes parameters are related with J_{mn} as follows:

$$\begin{aligned} I &= J_{xx} + J_{yy} = \langle a_x^2\rangle + \langle a_y^2\rangle, \\ Q &= J_{xx} - J_{yy} = \langle a_x^2\rangle - \langle a_y^2\rangle, \\ U &= J_{xy} + J_{yx} = 2\langle a_x a_y \cos(\alpha_x - \alpha_y)\rangle, \\ V &= i(J_{xy} - J_{yx}) = 2\langle a_x a_y \sin(\alpha_x - \alpha_y)\rangle. \end{aligned} \tag{6.1-12}$$

or, equivalently,

$$\hat{J} = \frac{1}{2}\begin{pmatrix} I+Q & U-iV \\ U+iV & I-Q \end{pmatrix} = \frac{1}{2}(I\sigma_0 + U\sigma_1 + V\sigma_2 + Q\sigma_3), \tag{6.1-13}$$

where σ_0 is the 2×2 identity matrix and σ_1, σ_2, σ_3 are the Pauli matrices.

Any coherence matrix can be uniquely expressed as the sum of two matrices, one of which corresponds to completely polarized light and the other – to completely unpolarized light:

$$\hat{J} = \hat{J}^{\text{pol}} + \hat{J}^{\text{unpol}}, \quad \hat{J}^{\text{pol}} = \begin{pmatrix} A & B \\ B^* & C \end{pmatrix}, \quad \hat{J}^{\text{unpol}} = \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$
$$A = \frac{1}{2}\left[(J_{xx} - J_{yy}) + \sqrt{I^2 - 4\det \hat{J}}\right], \quad B = J_{xy}, \tag{6.1-14}$$
$$C = \frac{1}{2}\left[(J_{yy} - J_{xx}) + \sqrt{I^2 - 4\det \hat{J}}\right], \quad \lambda = \frac{1}{2}\left[I \pm \sqrt{I^2 - 4\det \hat{J}}\right].$$

The value

$$P = \frac{\operatorname{tr} \hat{J}^{\text{pol}}}{I} \quad (0 \le P \le 1) \tag{6.1-15}$$

is called the degree of polarization. It defines the portion of completely polarized component of light in the total light beam. At $P = 1$ the light is completely polarized, at $P = 0$ – completely unpolarized, at $0 \le P \le 1$ – partially polarized. The degree of polarization can be expressed in terms of eigenvalues λ_1 and λ_2 of the coherence matrix $\hat{J}$:

$$P = \frac{|\lambda_1 - \lambda_2|}{\lambda_1 + \lambda_2} = \sqrt{1 - \frac{4\det \hat{J}}{I^2}} \tag{6.1-16}$$

and in terms of Stokes parameters is

$$P = \frac{\sqrt{Q^2 + U^2 + V^2}}{I}. \tag{6.1-17}$$

For completely polarized beam the Stokes parameters define the form and the orientation of polarization ellipse shown in Figure 6.1:

$$\tan(2\chi) = \frac{U}{Q}, \quad \sin(2\beta) = -\frac{V}{\sqrt{Q^2 + U^2 + V^2}}. \tag{6.1-18}$$

For unpolarized light beam the coherence matrix $\hat{J}$ is proportional to the unit matrix and does not depend on the choice of the polarization basis. This means that the x- and y-components of the electric field vector are uncorrelated for all pairs of directions, and the average $\langle E_i E_i^* \rangle$ has the same value for every direction $\mathbf{i}$ that is perpendicular to the direction $\mathbf{k}$ of beam propagation.

For completely polarized light beam the coherence matrix has the form

$$\hat{J}^{\text{pol}} = \begin{pmatrix} J_{xx} & \sqrt{J_{xx}J_{yy}}e^{i\alpha} \\ \sqrt{J_{xx}J_{yy}}e^{-i\alpha} & J_{yy} \end{pmatrix}. \tag{6.1-19}$$

It coincides with the coherence matrix of a plane monochromatic (deterministic) electromagnetic wave with complex amplitude $\mathbf{E}(t) = (\sqrt{J_{xx}}e^{i\beta_1}, \sqrt{J_{yy}}e^{i\beta_2})^T$, where β_1 and β_2 are arbitrary constants with $\beta_1 - \beta_2 = \alpha$. A quasimonochromatic wave with

the coherence matrix $\hat{J} = \hat{J}^{\mathrm{pol}}$ cannot be distinguished from a strictly deterministic monochromatic wave in canonical optical experiments, which include only polarizers and compensators.

6.1.3 Classical and quantum correlations in optics

Theory of classical electromagnetic field successfully describes all the phenomena related to wave propagation, interference, and diffraction. But in the situations with very few photons, the quantum optical description is necessary, because the very idea of an oscillatory field with definite phase becomes meaningless. Surprisingly, classical optics proceeds to correctly work even at extremely low light intensity levels. For instance, such nonclassical optical effects as photoelectric bunching and photoelectric counting statistics can be described in the frames of classical electromagnetic field theory. However, there are some optical phenomena for which electromagnetic field needs to be treated at quantum level of description.

Photons can be defined as quanta of the normal modes of the electromagnetic field and are associated with plane waves of definite wave vector and definite polarization. The concept of the photon as a localized particle traveling with the velocity c can be inappropriate and even misleading in some situations. There are several significant differences between the predictions of classical field theory and those of quantized field theory. An interesting nonclassical result can be obtained in the case of two correlated photon beams. Calculating the joint probability of detecting a photon at each of two output ports of a beam splitter with the help of general classical formula for joint probability, one can find that the probability can be equal to zero for special cases of one-photon states. This result has no analogy in the classical electromagnetic field. This is an example of quantum interference of the probability amplitude for a photon pair [5].

Another example is provided by the famous paradox, demonstrated by the so-called Bell state or, Einstein–Podolsky–Rosen state known as an EPR-pair [2]. In optics it can be illustrated by a two-photon state with zero total angular momentum J. Such state might be created in a cascade decay of an atom making a two-stage transition of the type $\Delta J = 0$ when two photons and, being in the states with orthogonal linear polarizations, leave the atom in the opposite directions. The results of the experiment with these two photons, using two polarizers and two photodetectors, were analyzed in detail. The joint probability that both photons are simultaneously detected by the photodetectors when both polarizers are oriented at angles θ_1 and θ_2, can be calculated both in frames quantum optics and as classical joint probability. A version of Bell inequalities can be written for classical joint probability, and it can be shown that the inequalities are violated for quantum mechanical expression. Several optical experiments confirming the Bell inequality violation have been carried out [1, 3, 6, 7].

6.2 Modeling superpositions of polarized light beams

In this section, we present a system of stochastic oscillators for modeling the electric fields of quasimonochromatic light beams and their superpositions. A single stochastic oscillator model imitates the electric field of a plane quasimonochromatic circularly polarized electromagnetic wave. A pair of properly parameterized stochastic oscillators simulates the electric field of a light beam at arbitrary polarization state. A superposition of $2N$ stochastic oscillators then simulates the electric field of the light beam obtained as a result of superposition of N light beams propagating in the same direction. The correlation matrix (or Stokes parameters) is used for estimation the state of polarization of the composed light beam.

6.2.1 Oscillatory model of polarized light beam

The model of stochastic oscillator is designed based on the simplest limit cycle oscillator named Ginzburg–Landau oscillator (see Section 2.2.1). The circle of radius ρ in the plane $\mathbb{R}^2$ is the limit cycle of the oscillator dynamics. The dynamical equation can be written as

$$\dot{z} = (\rho^2 + i\omega - |z|^2)z, \tag{6.2-1}$$

where $z = x + iy$, ρ is the limit cycle radius and ω is the oscillator frequency (see Figure 2.7). A pair of uncoupled Ginzgurg–Landau oscillators with chaotically modulated dynamics is necessary for modeling quasimonochromatic polarized light beams. In our model we use the oscillators with chaotically fluctuating limit cycle radii $\tilde{\rho}_1$, $\tilde{\rho}_2$ and oscillator frequencies $\tilde{\omega}_1$, $\tilde{\omega}_2$. The modulated dynamical equations are written in the form

$$\dot{z}_i = \left(\tilde{\rho}_i^2(t) + i\tilde{\omega}_i(t) - |z_i|^2\right) z_i, \quad i = 1, 2, \tag{6.2-2}$$

where

$$\begin{aligned} \tilde{\rho}_i(t) &= \rho_i + \xi_i(t), \\ \tilde{\omega}_i(t) &= \omega_i + \eta_i(t), \end{aligned} \tag{6.2-3}$$

and $\xi_i(t)$, $\eta_i(t)$ are the stationary random functions of time (stationary random processes) with zero mean: $\langle \xi_i(t) \rangle = 0$, $\langle \eta_i(t) \rangle = 0$. Due to ergodicity of stationary random processes all stochastic means are equal to the corresponding means over a large time period, that is $\langle \xi_i(t) \rangle = \overline{\xi_i(t)}$, $\langle \eta_i(t) \rangle = \overline{\eta_i(t)}$.

The oscillator pair with $\rho_1 = \rho_2 = \rho$, $\omega = -\omega_2 = -\omega$ corresponds to a pair of circularly polarized light beams in mutually orthogonal states of right and left circular polarization. To be convinced that dynamical equations (6.2-2) really specify a stochastic dynamics, one can rewrite each of two-dimensional equations (6.2-2) in the form of

two equivalent uncoupled equations for variables $r = |z|$ and $\theta = \arg z$:

$$\begin{aligned} \dot{r} &= [(\rho + \xi)^2 - r^2]r, \\ \dot{\theta} &= \omega + \eta. \end{aligned} \tag{6.2-4}$$

Then each of dynamic equation (6.2-2) can be represented in the form of two independent Langevin stochastic equations

$$\begin{aligned} \dot{r}_i &= [\rho_i^2 - r^2]r + \tilde{p}_i(t), \\ \dot{\theta}_i &= \omega_i + \tilde{q}_i(t), \end{aligned} \tag{6.2-5}$$

where

$$\begin{aligned} &\tilde{p}_i(t) = \xi(t)[\xi(t) + 2\rho_i]r, \qquad && \langle \tilde{p}_i(r,t)\rangle = 0, \\ &\tilde{q}_i = \eta_i(t), && \langle \tilde{q}_i(t)\rangle = 0. \end{aligned} \tag{6.2-6}$$

The Langevin equations (6.2-5) and (6.2-6) describe dynamics of a "particle" under the action of fluctuating external force (Brownian wandering).

If we assume the real-valued electric field components $E_x(\mathbf{r}, t)$ and $E_y(\mathbf{r}, t)$ to be proportional to $x = \operatorname{Re} z$ and $y = \operatorname{Im} z$, respectively, the trajectory of dynamical system (6.2-2) imitates the electric field behavior of plane quasimonochromatic circular polarized wave. As we saw earlier, a general fully polarized state of a plane electromagnetic wave is the state of elliptic polarization: the end of electric field vector $\mathbf{E}(\mathbf{r}, t)$ traces out a spiral line in the space, which projection onto a plane, orthogonal to the direction of wave propagation depicts an ellipse. To determine the form and orientation of the polarization ellipse, one should decompose the electric field into the sum of two mutually orthogonal components and find their amplitudes and a phase differences between them. This corresponds to the decomposition of the complex amplitude $\mathbf{E}$ of vector $\mathbf{E}(t)$ in some orthogonal basis. In particular, the states of right and left circular polarizations, denoted by the vectors $\mathbf{e}^{(\pm)} = (1/\sqrt{2})(\mathbf{e}_x \pm i\mathbf{e}_y)$, are often used as the basis states for representation of arbitrary elliptic polarization and polarization analysis. Obviously, the oscillator pair with parameters $\rho = \rho_0$, $\omega = \omega_0$ and $\rho = \rho_0$, $\omega = -\omega_0$ and vanishing random fluctuations $\xi_{1,2}(t) = 0$, $\eta_{1,2}(t) = 0$ corresponds to circular polarization basis states $\mathbf{e}^{(+)}$ and $\mathbf{e}^{(-)}$, respectively.

The examples of the electric field $\mathbf{E}(t)$ behavior corresponding to the light beam in the state of horizontal linear polarization are shown in Figure 6.2, where the variables $z_1 = x + iy$ and $z_2 = z + iu$ correspond to oscillators forming the total electric field $z_1 + z_2 = v + iw$. In terms of these variables the dynamical equations are

$$\begin{aligned} &\dot{x} = f_1(x,y), && f_1(x,y) = (\tilde{\rho}_1^2 - x^2 - y^2)x - \tilde{\omega}_1 y, \\ &\dot{y} = f_2(x,y), && f_2(x,y) = \tilde{\omega}_1 x + (\tilde{\rho}_1^2 - x^2 - y^2)y, \\ &\dot{z} = f_3(z,u), && f_3(z,u) = (\tilde{\rho}_2^2 - z^2 - u^2)z - \tilde{\omega}_2 u, \\ &\dot{u} = f_4(z,u), && f_4(z,u) = \tilde{\omega}_2 z + (\tilde{\rho}_2^2 - z^2 - u^2)u, \end{aligned} \tag{6.2-7}$$

and

$$\begin{aligned} \dot{v} &= f_1 + f_3, \\ \dot{w} &= f_2 + f_4. \end{aligned} \tag{6.2-8}$$

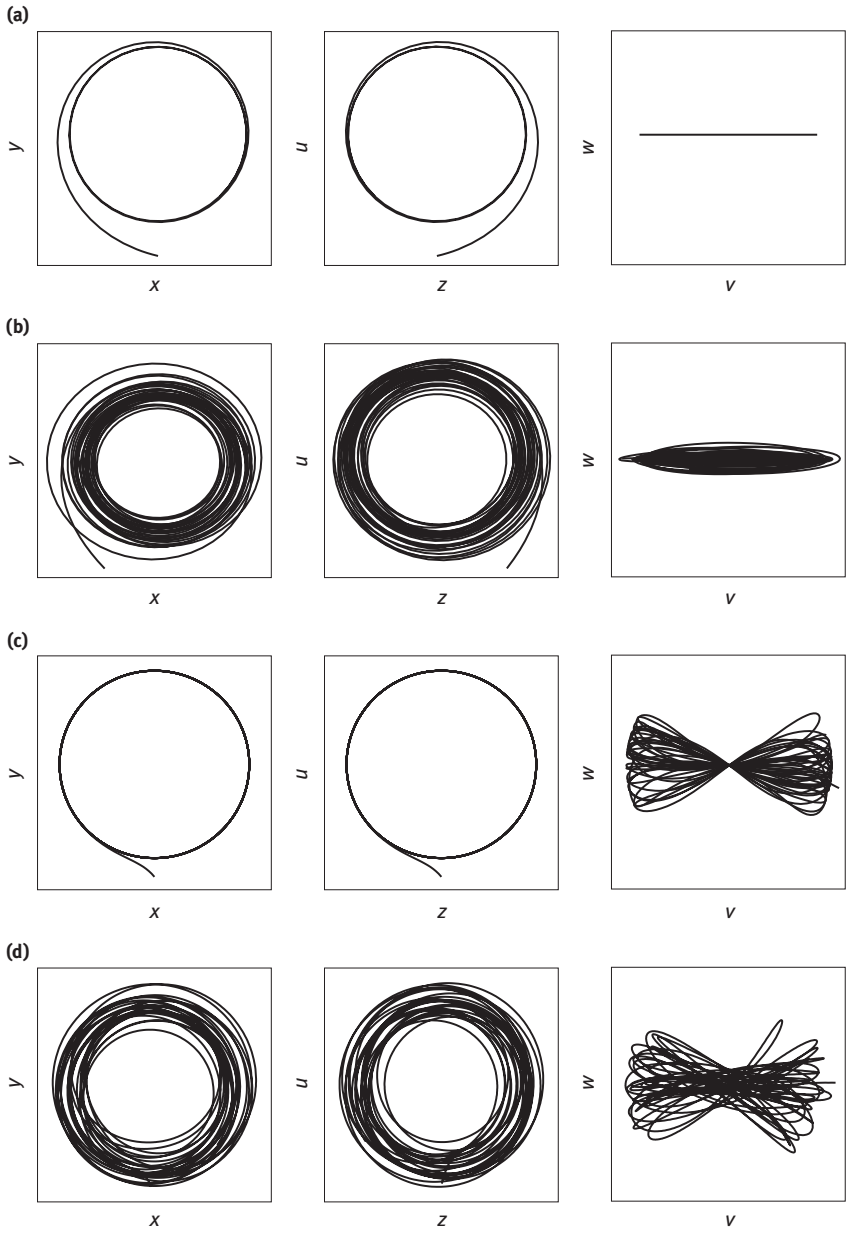

Fig. 6.2. Stochastic oscillator imitating electric field of light beam in the state of linear polarization: (a) $\xi_{1,2}(t) = 0$, $\eta_{1,2}(t) = 0$; (b) $\xi_{1,2}(t) \neq 0$, $\eta_{1,2}(t) = 0$; (c) $\xi_{1,2}(t) = 0$, $\eta_{1,2}(t) \neq 0$; (d) $\xi_{1,2}(t) \neq 0$, $\eta_{1,2}(t) \neq 0$.

The polarization state is controlled via the set of parameters: ρ_1, ρ_2, ω_1, ω_2, $x(0)$, $y(0)$, $z(0)$, $u(0)$, $\xi_{1,2}(t)$, and $\eta_{1,2}(t)$. The examples of the field of horizontally polarized wave are presented in the panel (v, w) of Figure 6.2. The field of deterministic horizontally linearly polarized electromagnetic wave (no amplitude and frequency chaotic fluctuations) is depicted in the panel (v, w) of Figure 6.2 (a). In this case the deterministic polarization ellipse is reduced to a horizontally oriented line segment. In Figure 6.2 (b) simulated electric field of fluctuating linearly polarized electromagnetic wave is shown in the case of chaotic modulation of the wave amplitude at constant frequency. As one can see, at solely amplitude chaotic modulation only polarization ellipse form fluctuates, whereas the polarization plane remains unperturbed (horizontal). Different situation occurs in the case of chaotic modulation of electric field frequency at constant amplitude: fluctuations of the polarization plane orientation arise (with zero mean) (see Figure 6.2 (c)). In general case of fluctuating both field amplitude and frequency, the electric field behavior is presented in Figure 6.2 (d). Both the polarization ellipse form and the polarization plane orientation now fluctuate.

In the above consideration, the random processes $\xi_{1,2}(t)$ and $\eta_{1,2}(t)$ are assumed to be slow-varying functions of time, i.e. their mean frequencies Ω_ξ and Ω_η are small compared with the mean frequency $\overline{\omega}$. Then, the state of the Ginzburg–Landau oscillator (6.2-2) never deviates from the limit cycle significantly, and its dynamics can be approximated with just a perturbed harmonic oscillator. In other words, equation (6.2-5) can be reduced to

$$\begin{aligned} \dot{r}_{1,2} &= \tilde{p}_{1,2}(t), \\ \dot{\theta}_{1,2} &= \omega_{1,2} + \tilde{q}_{1,2}(t). \end{aligned} \tag{6.2-9}$$

In the absence of fluctuations, that is when $\tilde{p}_{1,2}(t) = \tilde{q}_{1,2}(t) = 0$, these equations correspond to the exact harmonic solution (6.1-5) of the Maxwell equations (6.1-1).

6.2.2 Modeling the action of a polarizer

A polarizer is one of widely used optical devices that transforms light polarization, converting a beam of arbitrarily polarized light into the beam with well-defined polarization (for instance, linear polarization). Let $\mathbf{E}$ be the vector of complex amplitude of the electrical field of quasimonochromatic light wave of incident beam. The electric field vector $\mathbf{E}'$ of the light beam transmitted through optical device is defined as $\mathbf{E}' = \hat{A}\mathbf{E}$, where $\hat{A}$ is the so-called Jones operator (or a Jones matrix) of the device. In the Cartesian basis $\{\mathbf{e}_x, \mathbf{e}_y\}$ the Jones matrix of an ideal linear polarizer, can be written as

$$\hat{A}_P = C \begin{pmatrix} \cos^2\theta_0 & \cos\theta_0 \sin\theta_0 \\ \cos\theta_0 \sin\theta_0 & \sin^2\theta_0 \end{pmatrix}, \tag{6.2-10}$$

where θ_0 is the angle between $\mathbf{e}_x$ and the plane of transmitted polarization and the constant C depends on the thickness and refractive index of the polarizer. The Jones

matrix of an optical compensator is diagonal in the basis $\{\mathbf{e}^{(+)}, \mathbf{e}^{(-)}\}$. In the basis $\{\mathbf{e}_x, \mathbf{e}_y\}$ it has the form

$$\hat{A}_C = \begin{pmatrix} \cos(\delta/2) & \pm\sin(\delta/2) \\ \mp\sin(\delta/2) & \cos(\delta/2) \end{pmatrix}, \tag{6.2-11}$$

where δ is the additional phase difference between the field components, provided by the compensator.

In the frames of our model of stochastic oscillators, the action of any Jones operator on the beam is realized by adding the properly constructed free term $\mathbf{F} = (F_x, F_y)$ into the dynamical equations written in the basis $\{\mathbf{e}_x, \mathbf{e}_y\}$. The function $\mathbf{F}$ describes the "switching on" of the action of an optical device in the dynamical equations at desirable given moment of time. The result of the action of an ideal linear polarizer on the electric field with horizontal elliptical polarization is presented in Figure 6.3.

In the panel (v, w) of Figure 6.3, the large axis of stochastically modulated polarization ellipse was initially horizontally oriented. After the transmission through a linear polarizer with polarization angle $\theta_0 = 60^\circ$ the ellipse transforms into a line segment with orientation angle 60° with respect to $\mathbf{e}_x$.

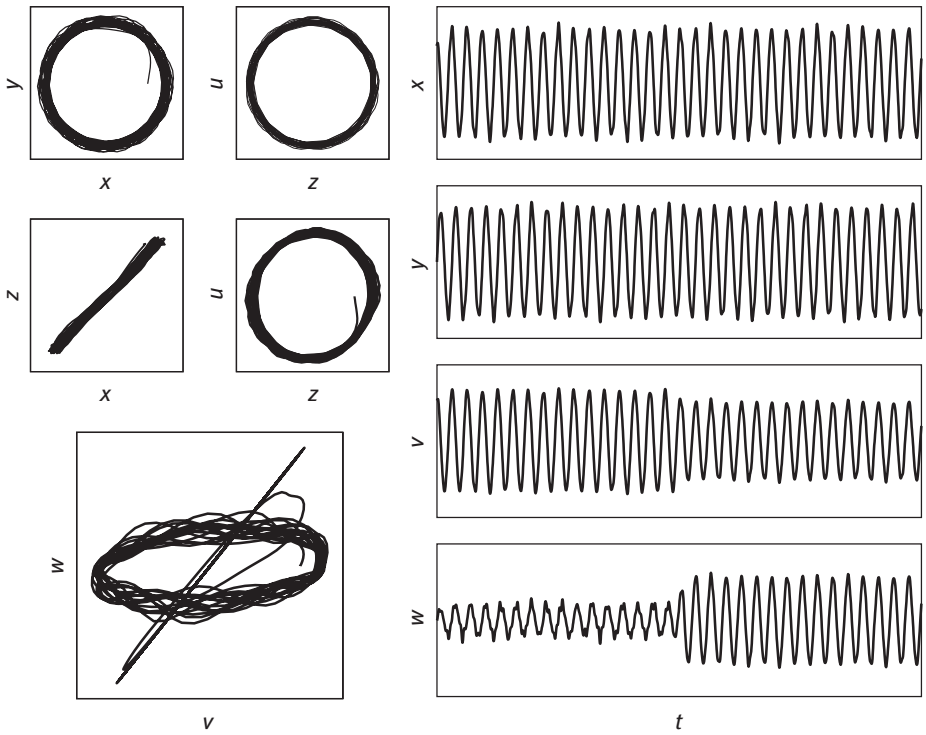

Fig. 6.3. Polarization state change of elliptically polarized qubit after transmission through ideal linear polarizer with polarization angle $\theta_0 = 60^\circ$.

6.2.3 Construction of superpositions of quasimonochromatic light beams

We now consider a classical superposition of two beams. First, we will give a simple analytical analysis of some general properties of this superposition. It is convenient to study the features of dynamical behavior of the electric field in terms of the dynamical system (6.2-5) written in variables $r = |z|$ and $\theta = \arg z$. If the mean frequencies Ω_ξ and Ω_η of processes $\xi_{1,2}(t)$ and $\eta_{1,2}(t)$ are small compared with wave frequency ω (this is always fulfilled for quasimonochromatic waves), the dynamical system (6.2-5) can be exactly integrated to give the explicit expressions for amplitude and phase of electrical field:

$$\begin{aligned} &\tilde{r}(t) \equiv \tilde{a}(t) = \frac{\tilde{\rho}}{\sqrt{1 + ce^{-2\tilde{\rho}^2 t}}}, \quad \tilde{\rho} = \rho + \xi(t), \\ &\tilde{\theta}(t) = \theta_0 + \tilde{\omega}t, \quad \tilde{\omega} = \omega + \eta(t), \end{aligned} \tag{6.2-12}$$

where $c = \xi(0)[2\rho + \xi(0)]/\rho^2$.

The expressions (6.2-12) allow us to clarify the dependence of the amplitude and phase fluctuations on properties of processes $\xi_{1,2}(t)$ and $\eta_{1,2}(t)$ and to analyze the properties of the correlation matrix. We can use the following expression for the electric field of the beam with arbitrary linear polarization:

$$\tilde{z} = \tilde{A}e^{i\tilde{\alpha}} = \tilde{a}^+ e^{i\tilde{\theta}^+} + \tilde{a}^- e^{i\tilde{\theta}^-}, \tag{6.2-13}$$

where

$$\begin{aligned} &A^2 = (\tilde{a}^+)^2 - 2\tilde{a}^+\tilde{a}^- + 4\tilde{a}^+\tilde{a}^- \cos^2\left(\tilde{\omega}t + \frac{\tilde{\delta}^+ + \tilde{\delta}^-}{2}\right), \\ &\tan\tilde{\alpha} = \frac{\tilde{a}^+ \sin\tilde{\theta}^+ - \tilde{a}^- \sin\tilde{\theta}^-}{\tilde{a}^+ \sin\tilde{\theta}^+ + \tilde{a}^- \sin\tilde{\theta}^-}. \end{aligned} \tag{6.2-14}$$

Due to ergodicity of $\xi_{1,2}(t)$ and $\eta_{1,2}(t)$, the means $\langle A^2\rangle$ and $\langle\tilde{\alpha}\rangle$ can be calculated as

$$\begin{aligned} &\langle A^2\rangle = \overline{A^2} = 4\rho^2 \cos^2\left(\tilde{\omega}t + \frac{\tilde{\delta}^+ + \tilde{\delta}^-}{2}\right) + \Delta^2_{\xi^+} + \Delta^2_{\xi^-}, \\ &\langle\tilde{\alpha}\rangle = \frac{\tilde{\delta}^+ + \tilde{\delta}^-}{2}, \quad \Delta^2_\xi \equiv D^2_\xi = \langle\xi^2\rangle. \end{aligned} \tag{6.2-15}$$

In particular, from (6.2-14) and (6.2-15), we obtain for the state with horizontal linear polarization ($\delta^+ = \delta^- = 0$):

$$\begin{aligned} &\tilde{A}^2_H = (\tilde{a}^+)^2 - 2\tilde{a}^+\tilde{a}^- + 4\tilde{a}^+\tilde{a}^- \cos^2(\tilde{\omega}t), \\ &\tan\tilde{\alpha}_H = \frac{\xi^+ - \xi^-}{2\rho + \xi^+ + \xi^-}\tan(\tilde{\omega}t), \\ &\langle A^2_H\rangle = 4\rho^2\cos^2(\tilde{\omega}t) + \Delta^2_{\xi^+} + \Delta^2_{\xi^-}, \\ &\langle\tilde{\alpha}_H\rangle = 0 \end{aligned} \tag{6.2-16}$$

and for the state with vertical linear polarization ($\delta^+ = \pi/2$, $\delta^- = -\pi/2$):

$$\tilde{A}_V = \tilde{A}_H, \quad \tan\tilde{\alpha}_V = -\frac{2\rho + \xi^+ + \xi^-}{\xi^+ - \xi^-}\cot(\tilde{\omega}t), \qquad \langle\tilde{\alpha}_V\rangle = \frac{\pi}{2}. \tag{6.2-17}$$

The correlation matrix $\hat{J}$ for the superposition of these states can be calculated using (6.2-16) and (6.2-17). The most essential fact is that in the case of equal mean frequencies $\langle\tilde{\omega}^H\rangle = \langle\tilde{\omega}^V\rangle$ of the waves, one finds for the matrix $\hat{J}$ the expression

$$\hat{J} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \tag{6.2-18}$$

which correspond to the completely linearly polarized wave with the angle of orientation of polarization plane $\chi = \pi/4$. Thus, we obtained the following important result: two quasimonochromatic waves in the states of mutually orthogonal polarization can be considered as completely noncorrelated (independent) only under the fulfillment of the condition $\langle\tilde{\omega}^H\rangle \neq \langle\tilde{\omega}^V\rangle$. If the condition is fulfilled, the superposition of two quasimonochromatic completely polarized electromagnetic waves gives unpolarized

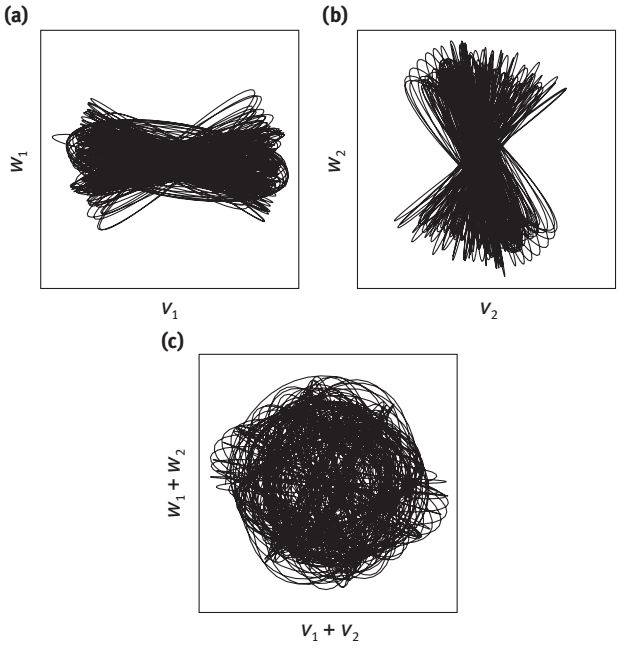

Fig. 6.4. Simulated electric field of light beam obtained as a result of a incoherent superposition of two light beams in the states of full horizontal (a) and vertical (b) polarization. The electric field of the incoherent two beam superposition is shown in (c).

quasimonochromatic wave, for which the coherence matrix $\hat{J}$ is proportional to the unit matrix.

The study of the simulated electric field confirmed the results of preliminary analysis, given earlier. The first example of electric field behavior corresponds the field of light beam obtained as a result of superposition of two light beams in the states of mutually orthogonal full linear polarization. The simulated electrical field is shown in Figure 6.4. Here the projection of phase trajectory of the dynamical system (6.2-7) onto the plane (x, y) displays the behavior of the random electric field vector $\mathbf{E}^H(t)$ in the state of horizontal linear polarization, the trajectory projection into the plane (z, u) – the behavior of the electric field vector $\mathbf{E}^V(t)$ in the state of vertical linear polarization, and the trajectory projection into the plane (v, w) – the behavior of the electric field

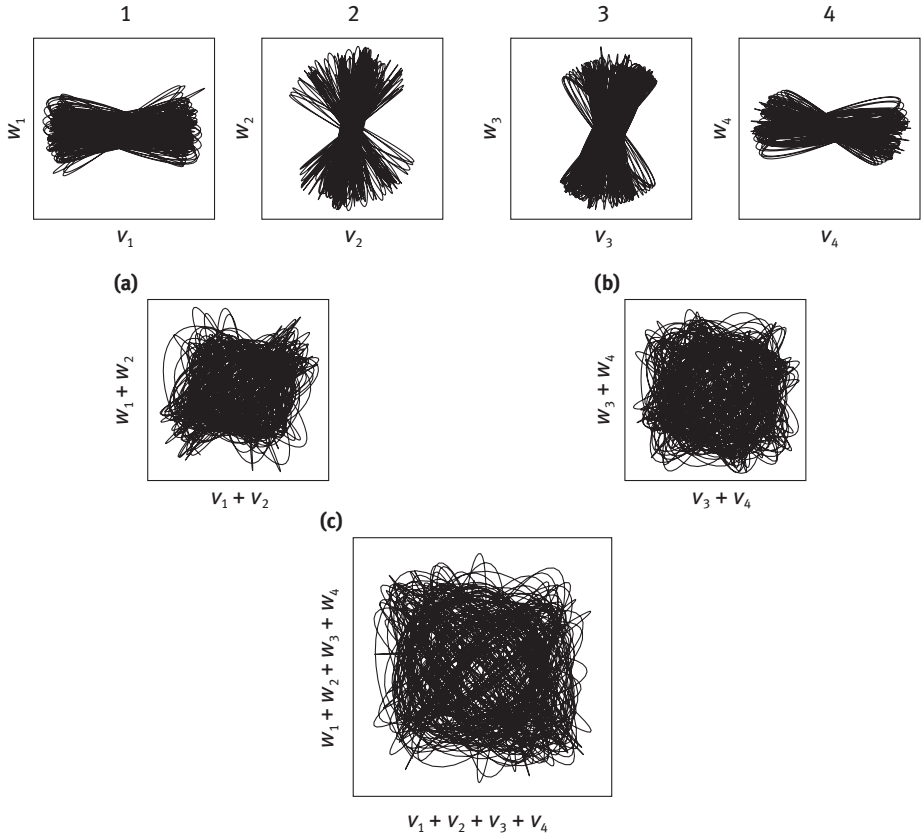

Fig. 6.5. Simulated electric field, corresponding to the superposition of four fully polarized light beams: (1), (4) the electric fields of fully horizontally polarized light beams; (2), (3) the electric fields of fully vertically polarized light beams; (a) the electric field of the superposition of light beams (1) and (2); (b) the electric field of the superposition of light beams (3) and (4); (c) the electric field of symmetrized superposition of four light beams.

vector $\mathbf{E}^B(t) = \mathbf{E}^H(t) + \mathbf{E}^V(t)$ that represents their superposition. The mean frequencies of $\mathbf{E}^H(t)$ and $\mathbf{E}^V(t)$ were chosen close, but different ($\tilde{\omega}^H \neq \tilde{\omega}^V$).

The characteristics of the random vector $\mathbf{E}^B(t)$ were estimated via calculation of the correlation matrix $\hat{J}^B$ according to (6.1-8). The correlation coefficient $|J_{xy}|/\sqrt{J_{xx}J_{yy}}$ for the components of the electric field $\mathbf{E}^B(t)$ was found to be less than 10^{-2}. Hence, the behavior of $\mathbf{E}^B(t)$ properly imitates the behavior of the electrical field vector of unpolarized quasimonochromatic electromagnetic wave.

The second example of simulated electric field is a four beam superposition. It is presented in Figure 6.5. The random electric fields of light beams in the states of full linear polarization, are depicted in the panels (1)–(4); the electric fields of light beams, representing the superpositions of two beams in the states of mutually orthogonal linear polarization, are displayed in Figure 6.5 (a) and (b); the electric field of light beam, obtained as a result of superposition of four fully polarized light beams, is presented in Figure 6.5 (c).

6.3 Outline of the design of feedforward network of stochastic oscillators

The model of system of stochastic oscillators, described in Section 6.2 is currently under development. It can be used for modeling transformations of coherent and incoherent superpositions polarized light beams that can be adequately described in the frames of classical stochastic electrodynamics. In the case of modeling the classical random electric field of a polarized light beam via the system of stochastic oscillators, the level of fluctuations of the simulated electric field, and the random electric field symmetry figure as the free parameters of the system. The beam electric field change caused by the beam transformations, modeled as optical measurements, can be used for organization of some computational process, similar to usual optical computing.

We design the feedforward oscillatory network model consisting of N processing units. The network processing unit modeling classical random electrical of a polarized light beam is formed by a superposition of two stochastic oscillators with internal dynamics defined by (6.2-2). A quasimonochromatic unpolarized light beam can be composed of N independent unpolarized sub-beams, in which mean frequencies belong to a collection of nonintersecting narrow frequency intervals. A network consisting of N uncoupled units can be related to the total composite light beam. The evolution of the initially uncoupled feedforward network can be studied. If the state of separate beam components are changed at some discrete time moments as a result of external optical device (e.g. a polarizer) actions modifying the polarization degree of light beam, the state of the whole network will be changed in a discrete manner in response to these external actions. The coherence degree change of the total light beam during the network evolution can be accurately estimated through the network correlation

(coherence) matrix calculation. The coherence degree of the network state will gradually increase, reflecting the incoherence degree destruction of the initial completely incoherent composite light beam. This increase in the total light beam coherence can be interpreted as gradual emergence of effective dynamical coupling in the stochastic oscillatory network.

It would be also interesting to analyze the nonlinear optical effect in the language of the oscillatory model. Nonlinear effects in optics arise when the intensity of light beams become sufficiently high (the electric field should be comparable to an atomic electric field). They cause the effective interaction of light beam, which do not interact in the linear case, and lead to a plethora of optical phenomena: harmonics generation, sum and difference frequency generation, self-focusing, photon echo and many others. Some of these phenomena can be described by the introduction of intraoscillator coupling in the network. Much work is still to be done in this area before some of the results could be stated.

Bibliography

[1] A. Aspect, J. Dalibard and G. Roger. Experimental test of Bell's inequalities using time-varying analyzers. *Phys. Rev. Lett.*, 49:1804, 1982.
[2] A. Einstein, B. Podolsky and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47:777, 1935.
[3] Stuart J. Freedman and John F. Clauser. Experimental test of local hidden-variable theories. *Phys. Rev. Lett.*, 28:938, 1972.
[4] L. D. Landau and E. M. Lifshitz. *The Classical Theory of Fields*. Butterworth-Heinemann, 1980.
[5] L. Mandel and E. Wolf. *Optic coherence and quantum optics*. Cambridge University Press, 1995.
[6] Z. Y. Ou and L. Mandel. Violation of Bell's inequality and classical probability in a two-photon correlation experiment. *Phys. Rev. Lett.*, 61:50, 1988.
[7] G. Weihs, T. Jennewein, C. Simon, H. Weinfurter and A. Zeilinger. Violation of Bell's inequality under strict Einstein locality conditions. *Phys. Rev. Lett.*, 81:5039, 1998.
[8] A. M. Yaglom. *Correlation Theory of Stationary and Related Random Functions: Volume I: Basic Results*. Springer-Verlag, 1987.
[9] A. M. Yaglom. *An Introduction to the Theory of Stationary Random Functions*. Dover Publications, 2004.

Summary and some perspectives

A series of oscillatory neural network models aimed at the elucidation of some aspects of the brain modules functioning and design of neuromorphic dynamical algorithms, realizing parallel distributed information processing, are presented in the book. Highly parallel, distributed, adaptive, self-organized manner of functioning is inherent to the brain. Oscillatory neural activity, synchronization, and resonance accompany the work of many brain neural structures, providing fast rearrangement of their work. Therefore, oscillatory network models, reflecting the features of the brain functioning, can be regarded as a relevant tool for construction of neuromorphic oscillatory computational algorithms with synchronization-based performance.

In Chapter 2, the recurrent associative memory networks of limit cycle oscillators and the related phasor networks, designed by the authors, are presented and discussed. The phasor networks appear to be closely related to artificial neural networks of complex-valued neurons. The choice of network parameters providing associative memory networks with guaranteed memory characteristic was specified. As was noticed earlier, symmetry properties of the phasor networks cardinally affect their associative memory structure. So, a more detailed analysis of these properties would be quite desirable. The macrodynamical analysis of a memory retrieval process for large scale feedforward and recurrent neural networks of associative memory as well as some their other features (such as noise reduction ability) are also presented in Chapter 2.

A series of oscillatory network models for modeling the brain structures performance is described in Chapter 3. Among them there are the famous oscillatory network model LEGION, the three-dimensional biologically inspired network model by Li imitating the brain visual cortex performance and a three-dimensional oscillatory network model with synchronization-based performance developed by the authors. In the latter two models the orientation selectivity, inherent to the subset of simple cells of the brain visual cortex, was reflected in model performance. It provided the network with the capability of segmentation of images containing the texture fragments and contour integration.

A spatially two-dimensional oscillatory network model with controllable synchronization and self-organized synchronization-based performance, developed by the authors for image processing tasks, is presented in Chapter 4. The network oscillators are localized in two-dimensional spatial lattice isomorphic with image pixel array. The internal dynamics of single network oscillator reflects the main features of dynamics of neural oscillator of the brain visual cortex. The amplitude of oscillation of single oscillator monotonically depends on the brightness of the image pixel corresponding to the oscillator. Several network coupling principles were constructed for different image processing tasks. As a result the model demonstrates a variety of capabilities, such as brightness segmentation of real gray-level and color images, selective image

segmentation and a simplest version of visual scene analysis – successive selection of spatially separated fragments of a visual scene. The further model development could allow us to encompass additional image processing tasks, including the selection of moving objects in a visual scene and other tasks of active vision.

The information on photon echo, a nonlinear optical phenomenon, that adjoins the phenomenon of dynamical binding in the brain structures, is given in Chapter 5. The photon echo is remarkable from the viewpoint of its capabilities of massive parallel information processing. In contrast to usual optical computing, which nowadays receives great attention of researchers, this effect allows us to use the temporal dimension in additional to the spatial one. In the chapter, the information on optical vector–matrix multipliers, and optical implementation of artificial neural networks is given.

Chapter 6 presents the initial results of using a model of stochastic oscillators for the simulation of the electric field of quasimonochromatic polarized light beams. Single stochastic oscillator is a Ginzburg–Landau limit cycle oscillator with chaotically modulated limit cycle size and frequency. Random electromagnetic field of a light beam in an arbitrary polarization state is modeled by a superposition of a pair of the stochastic oscillators. The electric field of a composite light beam consisting of N components is modeled by a proper superposition of $2N$ stochastic oscillators. The evolution of the electric field can be studied via evolution of the state of feedforward stochastic oscillatory network. The change of correlation characteristics of electric field of the total light beam can be estimated by calculating the coherence matrix of the composite light beam. It permits to observe the emergence of coherence in the composite system.

As we hope, the models of oscillatory neural networks, presented in the book and applied in various image processing tasks, will appear to be of interest for researchers dealing with the development of brain-inspired oscillatory network models with synchronization-based performance and related algorithms of parallel information processing.

Index