IEEE Press Series on Computational Intelligence
David B. Fogel, Series Editor

Complex-Valued Neural Networks

Advances and Applications

Edited by
Akira Hirose

IEEE
IEEE PRESS

WILEY

Complex-Valued Neural Networks

IEEE Press
445 Hoes Lane
Piscataway, NJ 08854

Technical Reviewers

George M. Georgiou
Gouhei Tanaka

Complex-Valued Neural Networks

Advances and Applications

Edited by

Akira Hirose
The University of Tokyo

IEEE Press Series on Computational Intelligence
David B. Fogel, Series Editor

IEEE
IEEE PRESS

WILEY

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data is available.

ISBN 9781118344606

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

CONTRIBUTORS

CHAPTER 1

AKIRA HIROSE, The University of Tokyo, Tokyo, Japan

CHAPTER 2

SIMONE FIORI, Università Politecnica delle Marche, Ancona, Italy

CHAPTER 3

TOHRU NITTA, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan

CHAPTER 4

MD. FAIJUL AMIN, University of Fukui, Fukui, Japan; Khulna University of Engineering and Technology, Khulna, Bangladesh

KAZUYUKI MURASE, University of Fukui, Fukui, Japan

CHAPTER 5

TEIJIRO ISOKAWA, University of Hyogo, Hyogo, Japan

HARUHIKO NISHIMURA, University of Hyogo, Hyogo, Japan

NOBUYUKI MATSUI, University of Hyogo, Hyogo, Japan

CHAPTER 6

YASUAKI KUROE, Kyoto Institute of Technology, Kyoto, Japan

CHAPTER 7

RAMASAMY SAVITHA, Nanyang Technological University, Singapore

SUNDARAM SURESH, Nanyang Technological University, Singapore

NARASIMHAN SUNDARARAJAN, Sri Jaya Chamarajendra College of Engineering (SJCE), Mysore, India

CHAPTER 8

NIKOLAY MANYAKOV, KU Leuven, Leuven, Belgium

IGOR AIZENBERG, Texas A&M University–Texarkana, Texarkana, Texas, U.S.A.

NILOLAY CHUMERIN, KU Leuven, Leuven, Belgium

MARK M. VAN HULLE, KU Leuven, Leuven, Belgium

CHAPTER 9

XIA HONG, University of Reading, Reading, U.K.

SHENG CHEN, University of Southampton, Southampton, U.K.; King Abdulaziz University, Jeddah, Saudi Arabia

CHRIS J. HARRIS, University of Southampton, Southampton, U.K.

CHAPTER 10

WAI KIT WONG, Multimedia University, Melaka, Malaysia

GIN CHONG LEE, Multimedia University, Melaka, Malaysia

CHU KIONG LOO, University of Malaya, Kuala Lumpur, Malaysia

WAY SOONG LIM, Multimedia University, Melaka, Malaysia

RAYMOND LOCK, Multimedia University, Melaka, Malaysia

CONTENTS

Preface xv

1 **Application Fields and Fundamental Merits** 1
Akira Hirose

1.1 Introduction 1
1.2 Applications of Complex-Valued Neural Networks 2
1.2.1 Antenna Design 2
1.2.2 Estimation of Direction of Arrival and Beamforming 3
1.2.3 Radar Imaging 3
1.2.4 Acoustic Signal Processing and Ultrasonic Imaging 3
1.2.5 Communications Signal Processing 3
1.2.6 Image Processing 4
1.2.7 Social Systems Such as Traffic and Power Systems 4
1.2.8 Quantum Devices Such as Superconductive Devices 4
1.2.9 Optical/Lightwave Information Processing Including Carrier-Frequency Multiplexing 4

1.2.10 Hypercomplex-Valued Neural Networks 5

1.3 What is a complex number? 5

1.3.1 Geometric and Intuitive Definition 5

1.3.2 Definition as Ordered Pair of Real Numbers 6

1.3.3 Real 2×2 Matrix Representation 7

1.4 Complex numbers in feedforward neural networks 8

1.4.1 Synapse and Network Function in Layered Feedforward Neural Networks 9

1.4.2 Circularity 11

1.5 Metric in complex domain 12

1.5.1 Metric in Complex-Valued Self-Organizing Map 12

1.5.2 Euclidean Metric 12

1.5.3 Complex Inner-Product Metric 14

1.5.4 Comparison Between Complex Inner Product and Euclidean Distance 14

1.5.5 Metric in Correlation Learning 15

1.6 Experiments to elucidate the generalization characteristics 16

1.6.1 Forward Processing and Learning Dynamics 17

1.6.2 Experimental Setup 21

1.6.3 Results 24

1.7 Conclusions 26

References 27

2 Neural System Learning on Complex-Valued Manifolds 33

Simone Fiori

2.1 Introduction 34

2.2 Learning Averages over the Lie Group of Unitary Matrices 35

2.2.1 Differential-Geometric Setting 36

2.2.2 An Averaging Procedure over the Lie Group of Unitary Matrices 37

2.3 Riemannian-Gradient-Based Learning on the Complex Matrix-Hypersphere 41

2.3.1 Geometric Characterization of the Matrix Hypersphere 43

2.3.2 Geodesic-Stepping Optimization Method 45

2.3.3 Application to Optimal Precoding in MIMO Broadcast Channels 46

2.4 Complex ICA Applied to Telecommunications 49

2.4.1 Complex-Weighted Rigid-Body Learning Equations for ICA 51
2.4.2 Application to the Blind Separation of QAM/PSK Signals 53
2.5 Conclusion 53
References 55

3 N-Dimensional Vector Neuron and Its Application to the N-Bit Parity Problem 59
Tohru Nitta

3.1 Introduction 59
3.2 Neuron Models with High-Dimensional Parameters 60
3.2.1 Complex-Valued Neuron 60
3.2.2 Hyperbolic Neuron 61
3.2.3 Three-Dimensional Vector Neuron 61
3.2.4 Three-Dimensional Vector Product Neuron 62
3.2.5 Quaternary Neuron 63
3.2.6 Clifford Neuron 63
3.3 N-Dimensional Vector Neuron 65
3.3.1 N-Dimensional Vector Neuron Model 65
3.3.2 Decision Boundary 65
3.3.3 N-Bit Parity Problem 67
3.3.4 A Solution 67
3.4 Discussion 69
3.5 Conclusion 70
References 71

4 Learning Algorithms in Complex-Valued Neural Networks using Wirtinger Calculus 75
Md. Faijul Amin and Kazuyuki Murase

4.1 Introduction 76
4.2 Derivatives in Wirtinger Calculus 78
4.3 Complex Gradient 80
4.4 Learning Algorithms for Feedforward CVNNs 82
4.4.1 Complex Gradient Descent Algorithm 82
4.4.2 Complex Levenberg–Marquardt Algorithm 86
4.4.3 Computer Simulations 89
4.5 Learning Algorithms for Recurrent CVNNs 91

4.5.1 Complex Real-Time Recurrent Learning Algorithm 92
4.5.2 Complex Extended Kalman Filter Algorithm 95
4.5.3 Computer Simulations 96
4.6 Conclusion 99
References 100

5 Quaternionic Neural Networks for Associative Memories 103
Teijiro Isokawa, Haruhiko Nishimura, and Nobuyuki Matsui

5.1 Introduction 104
5.2 Quaternionic Algebra 105
5.2.1 Definition of Quaternion 105
5.2.2 Phase Representation of Quaternion 106
5.2.3 Quaternionic Analyticity 106
5.3 Stability of Quaternionic Neural Networks 108
5.3.1 Network with Bipolar State Neurons 108
5.3.2 Network with Continuous State Neurons 111
5.3.3 Network with Continuous State Neurons Having Local Analytic Activation Function 115
5.3.4 Network with Multistate Neurons 118
5.4 Learning Schemes for Embedding Patterns 124
5.4.1 Hebbian Rule 124
5.4.2 Projection Rule 125
5.4.3 Iterative Learning for Quaternionic Multistate Neural Network 126
5.5 Conclusion 128
References 129

6 Models of Recurrent Clifford Neural Networks and Their Dynamics 133
Yasuaki Kuroe

6.1 Introduction 134
6.2 Clifford Algebra 134
6.2.1 Definition 134
6.2.2 Basic Properties and Algebraic Basis 136
6.3 Hopfield-Type Neural Networks and Their Energy Functions 137
6.4 Models of Hopfield-Type Clifford Neural Networks 139
6.5 Definition of Energy Functions 140

6.6 Existence Conditions of Energy Functions 142
6.6.1 Assumptions on Clifford Activation Functions 142
6.6.2 Existence Conditions for Clifford Neural Networks of Class $\mathbb{G}_{0,2,0}$ 143
6.6.3 Existence Conditions for Clifford Neural Networks of Class $\mathbb{G}_{0,1,0}$ 146
6.7 Conclusion 149
References 150

7 Meta-cognitive Complex-valued Relaxation Network and its Sequential Learning Algorithm 153
Ramasamy Savitha, Sundaram Suresh, and Narasimhan Sundararajan

7.1 Meta-cognition in Machine Learning 154
7.1.1 Models of Meta-cognition 154
7.1.2 Meta-cognitive Neural Networks 155
7.2 Meta-cognition in Complex-valued Neural Networks 156
7.2.1 Problem Definition 157
7.2.2 Meta-cognitive Fully Complex-valued Radial Basis Function Network 157
7.2.3 Complex-Valued Self-Regulatory Resource Allocation Network 160
7.2.4 Issues in Mc-FCRBF and CSRAN 164
7.3 Meta-cognitive Fully Complex-valued Relaxation Network 164
7.3.1 Cognitive Component: A Fully Complex-valued Relaxation Network (FCRN) 164
7.3.2 Meta-cognitive Component: A Self-regulatory Learning Mechanism 168
7.4 Performance Evaluation of McFCRN: Synthetic Complex-valued Function Approximation Problem 171
7.5 Performance Evaluation of McFCRN: Real-valued Classification Problems 172
7.5.1 Real-valued Classification Problem in the Complex Domain 173
7.5.2 Data Sets 174
7.5.3 Modifications in McFCRN Learning Algorithm to Solve Real-Valued Classification Problems 175
7.5.4 Performance Measures 176
7.5.5 Multi-category Benchmark Classification Problems 177

7.5.6 Binary Classification Problems 178
7.6 Conclusion 178
References 181

8 Multilayer Feedforward Neural Network with Multi-Valued Neurons for Brain–Computer Interfacing 185
Nikolay V. Manyakov, Igor Aizenberg, Nikolay Chumerin, and Marc M. Van Hulle

8.1 Brain–Computer Interface (BCI) 185
8.1.1 Invasive BCI 187
8.1.2 Noninvasive BCI 188
8.2 BCI Based on Steady-State Visual Evoked Potentials 188
8.2.1 Frequency-Coded SSVEP BCI 190
8.2.2 Phase-Coded SSVEP BCI 191
8.3 EEG Signal Preprocessing 192
8.3.1 EEG Data Acquisition 192
8.3.2 Experiment Description 192
8.3.3 Feature Selection 194
8.4 Decoding Based on MLMVN for Phase-Coded SSVEP BCI 196
8.4.1 Multi-Valued Neuron 196
8.4.2 Multilayer Feedforward Neural Network with Multi-Valued Neurons (MLMVN) 198
8.4.3 MLMVN for Phase-Coded SSVEP BCI 200
8.5 System Validation 201
8.6 Discussion 203

Appendix: Decoding Methods 204
A.1 Method of Jia and Co-workers 204
A.2 Method of Lee and Co-workers 204
References 205

9 Complex-Valued B-Spline Neural Networks for Modeling and Inverse of Wiener Systems 209
Xia Hong, Sheng Chen and Chris J. Harris

9.1 Introduction 210
9.2 Identification and Inverse of Complex-Valued Wiener Systems 211
9.2.1 The Complex-Valued Wiener System 212

9.2.2 Complex-Valued B-Spline Neural Network 212
9.2.3 Wiener System Identification 215
9.2.4 Wiener System Inverse 219
9.3 Application to Digital Predistorter Design 222
9.3.1 High-Power Amplifier Model 222
9.3.2 A Novel Digital Predistorter Design 224
9.3.3 A Simulation Example 224
9.4 Conclusions 229
References 229

10 Quaternionic Fuzzy Neural Network for View-invariant Color Face Image Recognition 235
Wai Kit Wong, Gin Chong Lee, Chu Kiong Loo, Way Soong Lim, and Raymond Lock

10.1 Introduction 236
10.2 Face Recognition System 238
10.2.1 Principal Component Analysis (PCA) Method 239
10.2.2 Non-negative Matrix Factorization (NMF) Method 240
10.2.3 Block Diagonal Non-negative Matrix Factorization (BDNMF) Method 242
10.3 Quaternion-Based View-invariant Color Face Image Recognition 244
10.3.1 Quaternion 244
10.3.2 Quaternion Fourier Transform 247
10.3.3 Quaternion-Based View-Invariant Color Face Image Recognition System Model 253
10.4 Enrollment Stage and Recognition Stage for Quaternion-Based Color Face Image Correlator 255
10.4.1 Enrollment Stage 255
10.4.2 Recognition Stage 259
10.5 Max-Product Fuzzy Neural Network Classifier 260
10.5.1 Fuzzy Neural Network System 261
10.5.2 Max-Product Fuzzy Neural Network Classification 264
10.6 Experimental Results 266
10.6.1 Database of Reference Face Images for 200 Persons 266
10.6.2 Quaternion-Based Face Image Correlation Using Unconstrained Optimal Tradeoff Synthetic Discriminant Filter (UOTSDF) 267

10.6.3 Efficiency of the View-invariant Color Face Image Recognition System 269
10.6.4 Comparative Study with the Parallel Method 271
10.7 Conclusion and Future Research Directions 274
References 274
Index 279

PREFACE

Complex-valued neural networks (CVNNs) have continued to open doors to various new applications. The CVNNs are the neural networks that deal with complex amplitude, i.e. signal having phase and amplitude, which is one of the most core concepts in science and technology, in particular in electrical and electronic engineering. A CVNN is not equivalent to a double-dimensional real-valued neural network. It has different dynamics and characteristics such as generalization, which is significantly useful in treatment of complex-amplitude information and wave-related phenomena. This is a critical point in applications in engineering fields. It is also crucial for developing new devices in the future. That is, the CVNN framework will play an important role in introduction of learning and self-organization into future quantum devices dealing with electron waves and photonic waves.

We can further expect that broad-sense CVNNs such as quaternion neural networks break ground in unique directions respectively. Quaternion has been essential in computer graphics to render three-dimensional moving objects. When we introduce learning and self-organization in virtual realities and computer-aided amenities, quaternion neural networks will surely bring an important fundamental basis. CVNNs may be useful even in physiological analysis and modeling where the researchers suggest, for example, that the phase information of neuron firing timing against the theta wave in electroencephalography possesses a close relationship to short-term position memory in the brain.

This book includes recent advances and applications of CVNNs in the following ten chapters. Chapter 1 presents historical and latest advances in applications of CVNNs first. Then it illustrates one of the most important merits of CVNNs, namely, the suitability for adaptive processing of coherent signals. Chapter 2 deals with complex-valued parameter manifolds and with applications of CVNNs in which the connection parameters work in complex-valued manifolds. Successful applications are also shown, such as blind source separation of complex-valued sources, multichannel blind deconvolution of signals in telecommunications, nondestructive evaluation of materials in industrial metallic slab production, and a purely algorithmic problem of averaging the parameters of a pool of cooperative CVNNs. Chapter 3 describes the N-dimensional vector neuron, which can deal with N signals as one cluster, by extending the three-dimensional vector neuron to N dimensions. The N-bit parity problem is solved with a signal N-dimensional vector neuron with an orthogonal decision boundary. It is shown that the extension of the dimensionality of neural networks to N dimensions originates the enhancement of computational power in neural networks. Chapter 4 discusses the Wirtinger calculus and derives several algorithms for feedforward and recurrent CVNNs. A functional dependence diagram is shown for visual understanding of respective derivatives. For feedforward networks, two algorithms are considered, namely, the gradient descent (backpropagation) and the Levenberg–Marquardt (LM) algorithms. Simultaneously, for recurrent networks, the authors discuss the complex versions of the real-time recurrent learning (RTRL) and the extended Kalman filter (EKF) algorithms.

Chapter 5 presents quaternion associative memories. Quaternion is a four-dimensional hypercomplex number system and has been extensively employed in the fields of robotics, control of satellites, computer graphics, and so on. One of its benefits lies in the fact that affine transforms in three-dimensional space can be compactly and consistently represented. Thus neural networks based on quaternion are expected to process three-dimensional data with learning or self-organization more successfully. Several schemes to embed patterns into a network are presented. In addition to the quaternion version of the Hebbian learning scheme, the projection rule for embedding nonorthogonal patterns and local iterative learning are described. Chapter 6 extends neural networks into the Clifford algebraic domain. Since geometric product is non-commutative, some types of models are considered possible. In this chapter three models of fully connected recurrent networks are i nvestigated, in particular from the viewpoint of existence conditions of an energy function, for two classes of the Hopfield-type Clifford neural networks.

Chapter 7 presents a meta-cognitive learning algorithm for a single hidden layer CVNN called Meta-cognitive Fully Complex-valued Relaxation Network (McFCRN). McFCRN has two components, that is, cognitive and meta-cognitive components. The meta-cognitive component possesses a self-regulatory learning mechanism which controls the learning stability of FCRN by deciding *what to learn*, *when to learn*, and *how to learn* from a sequence of training data. They deal with the problem of explicit minimization of magnitude and phase errors in logarithmic error function. Chapter 8 describes a multilayer feedforward neural network equipped with multi-valued neurons and its application to the domain of brain–computer interface (BCI). A new methodology for electroencephalogram (EEG)-based BCI is developed with which subjects can issue commands by looking at corresponding targets that are flickering at the same frequency but with different initial phase. Chapter 9 develops a complex-valued (CV) B-spline (basis-spline) neural network approach for efficient identification of the CV Wiener system as well as the effective inverse of the estimated CV Wiener model. Specifically, the CV nonlinear static function in the Wiener system is represented using the tensor product from two univariate B-spline neural networks. The effectiveness is demonstrated using the application of digital predistorter for high-power amplifiers with memory. Chapter 10 presents an effective color image processing system for persons' face image recognition. The system carries out the recognition with a quaternion correlator and a max-product fuzzy neural network classifier. The performance is evaluated in terms of accuracy, calculation cost, and noise and/or scale tolerance.

This is the first book planned and published by the Complex-Valued Neural Networks Task Force (CVNN TF) of the IEEE Computational Intelligence Society (CIS) Neural Networks Technical Committee (NNTC). The CVNN TF has been established to promote research in this developing field. The authors expect readers to get more interested in this area, to send feedback in any form, and to join us. Please visit our website `http://www.eis.t.u-tokyo.ac.jp/news/NNTC_CVNN/`.

AKIRA HIROSE

Tokyo
January 2013

CHAPTER 1

APPLICATION FIELDS AND FUNDAMENTAL MERITS OF COMPLEX-VALUED NEURAL NETWORKS

AKIRA HIROSE

The University of Tokyo, Tokyo, Japan

This chapter presents historical and latest advances in applications of complex-valued neural networks (CVNNs) first. Then it also shows one of the most important merits of CVNNs, namely, the suitability for adaptive processing of coherent signals.

1.1 INTRODUCTION

This chapter presents historical and latest advances in applications of complex-valued neural networks (CVNNs) first. Then it also shows one of the most important merits of CVNNs, namely, the suitability for adaptive processing of coherent signals.

CVNNs are effective and powerful in particular to deal with wave phenomena such as electromagnetic and sonic waves, as well as to process wave-related information. Regarding the history of CVNNs, we can trace back to the middle of the 20th century. The first introduction of phase information in computation was made by Eiichi Goto in 1954 in his invention of "Parametron" [17, 18, 61]. He utilized the phase of a high-frequency carrier to represent binary or multivalued informa-

Complex-Valued Neural Networks: Advances and Applications. Edited by Akira Hirose

tion. However, the computational principle employed there was "logic" of Turing type, or von Neumann type, based on symbol processing, so that he could not make further extensive use of the phase. In the present CVNN researches, contrarily, the researchers extend the world of computation to pattern processing fields based on a novel use of the structure of complex-amplitude (phase and amplitude) information.

We notice that the above feature is significantly important when we give thought to the fact that various modern technologies centered on electronics orient toward coherent systems and devices rather than something incoherent. The feature will lead to future general probability statistics, stochastic methods, and statistical learning and self-organization framework in coherent signal processing and information analysis. The fundamental idea is applicable also to hypercomplex processing based on quaternion, octonion, and Clifford algebraic networks.

Some parts of the following contents of this chapter were published in detail in *the Journal of Society of Instrument and Control Engineers* [29], *the Frontiers in Electrical and Electronic Engineering in China* [28], and *IEEE Transactions in Neural Networks and Learning Systems* [35].

1.2 APPLICATIONS OF COMPLEX-VALUED NEURAL NETWORKS

Complex-valued neural networks (CVNNs) have become widely used in various fields. The basic ideas and fundamental principles have been published in several books in recent years [27, 22, 26, 41, 53, 2]. The following subsections present major application fields.

1.2.1 Antenna Design

The most notable feature of CVNNs is the compatibility with wave phenomena and wave information related to, for example, electromagnetic wave, lightwave, electron wave, and sonic wave [28]. Application fields include adaptive design of antennas such as patch antennas for microwave and millimeter wave. Many researches have been reported on how to determine patch-antenna shape and sub-element arrangement, as well as on the switching patterns of the sub-elements [46, 10, 47]. A designer assigns desired frequency-domain characteristics of complex amplitude, or simply amplitude, such as transmission characteristics, return loss, and radiation patterns. A CVNN mostly realizes a more suitable design than a real-valued network does even when he/she presents only simple amplitude. The reason lies in the elemental dynamics consisting of phase rotation (or time delay $\times$ carrier frequency) and amplitude increase or decrease, based on which dynamics the CVNN learning or self-organization works. As a result, the generalization characteristics (error magnitude at nonlearning points in supervised learning) and the classification manner often become quite different from those of real-valued neural networks [28, 35]. The feature plays the most important role also in other applications referred to below.

1.2.2 Estimation of Direction of Arrival and Beamforming

The estimation of direction of arrival (DoA) of electromagnetic wave using CVNNs has also been investigated for decades [67, 6]. A similar application field is the beamforming. When a signal has a narrow band, we can simply employ Huygens' principle. However, in an ultra-wideband (UWB) system, where the wavelength is distributed over a wide range, we cannot assume a single wavelength, resulting in unavailability of Huygens' principle. To overcome this difficulty, an adaptive method based on a CVNN has been proposed [60] where a unit module consists of a tapped-delay-line (TDL) network.

1.2.3 Radar Imaging

CVNNs are widely applied in coherent electromagnetic-wave signal processing. An area is adaptive processing of interferometric synthetic aperture radar (InSAR) images captured by satellite or airplane to observe land surface [59, 65]. There they aim at solving one of the most serious problems in InSAR imaging that there exist many rotational points (singular points) in the observed data so that the height cannot be determined in a straightforward way.

Ground penetrating radar (GPR) is another field [21, 66, 43, 44, 49, 34]. GPR systems usually suffer from serious clutter (scattering and reflection from non-target objects). Land surface as well as stones and clods generate such heavy clutter that we cannot observe what are underground if we pay attention only to the intensity. Complex-amplitude texture often provides us with highly informative features that can be processed adaptively in such a manner that we do in our early vision.

1.2.4 Acoustic Signal Processing and Ultrasonic Imaging

Another important application field is sonic and ultrasonic processing. Pioneering works were done into various directions [69, 58]. The problem of singular points exists also in ultrasonic imaging. They appear as speckles. A technique similar to that used in InSAR imaging was successfully applied to ultrasonic imaging [51].

1.2.5 Communications Signal Processing

In communication systems, we can regard CVNNs as an extension of adaptive complex filters, i.e., modular multiple-stage and nonlinear version. From this viewpoint, several groups work hard on time-sequential signal processing [15, 16], blind separation [68], channel prediction [12], equalization [63, 36, 55, 40, 33, 7, 8], and channel separation in multiple-input multiple-output (MIMO) systems [37]. Relevant circuit realization [13] is highly inspiring not only as working hardware but also for understanding of neural dynamics.

1.2.6 Image Processing

There are many ideas based on CVNNs also in image processing. An example is the adaptive processing for blur compensation by identifying point scattering function in the frequency domain [3]. In such a frequency-domain processing of images, we often utilize the fact that the phase information in frequency domain corresponds to position information in spatial domain. On the other hand, CVNN spatial-domain processing is also unique and powerful. A highly practical proposal was made for quick gesture recognition in smart phones by dealing with finger angle information adaptively by a CVNN [19]. Biological imaging is another expanding field. There we can find, for example, a classification of gene-expression stages in gene images [1], along with adaptive segmentation of magnetic resonance image (MRI) by placing a dynamic boundary curve (so-called "snake") in the obtained complex-amplitude MRI image for segmentation of blood vessels and other organs [20]. Since there are various types of active and coherent imaging systems in medicine, we can expect further applications of CVNNs to deal with complex-amplitude images.

1.2.7 Social Systems Such as Traffic and Power Systems

Recent applications expand more multi-directionally even to social systems. In traffic systems, a CVNN will be effectively used for controlling mutual switching timing of traffic lights in complicatedly connected driving roads [50]. Since traffic lights have periodic operation, some CVNN dynamics is suitable for their adaptive control. Green energy and smart grid are also the fields. A CVNN-based prediction of wind strength and direction has been demonstrated for efficient electric power generation [14] in which amplitude and phase in the complex plane represent the strength and the direction, respectively.

1.2.8 Quantum Devices Such as Superconductive Devices

Applications to quantum computation using quantum devices such as superconductivity have also been investigated in many groups [57, 39, 48]. Their results suggest the future realization of intrinsically non-von Neumann computers including pattern-information representing devices. Conventional quantum computation is strictly limited in its treatable problems. Contrarily, CVNN-based quantum computation can deal with more general problems, which leads to wider applications of quantum computation.

1.2.9 Optical/Lightwave Information Processing Including Carrier-Frequency Multiplexing

Learning optical and lightwave computer is another field of CVNN applications. There are researches such as frequency-domain multiplexed learning [38] and real-time generation of a three-dimensional holographic movie for interactively controllable optical tweezers [32, 62]. In these networks, a signal has its carrier frequency, equivalent to a band signal in communications, and therefore the learning and processing dynamics is controllable by modulating the carrier frequency. The idea can be adapted to complex filters. It led to a novel developmental learning of motion control combined with reinforcement learning [30]. The success suggests further a possible influence of frequency modulation of brain wave on biological brain activity, indicating a new door to CVNN-related physiology.

1.2.10 Hypercomplex-Valued Neural Networks

Hypercomplex-valued neural networks have also been actively investigated [5]. An example is the adaptive learning in three-dimensional color space by using quaternion [45]. An adaptive super-high-sensitive color camera (so-called night vision) has been produced that realizes a compensation of nonlinear human color-vision characteristics in extremely dark environment. More generalized hypercomplex networks, namely, Clifford algebraic neural networks, are also discussed very actively in, e.g., special sessions in conferences [54].

1.3 WHAT IS A COMPLEX NUMBER?

In this section, we look back the history of complex numbers to extract the essence influential in neural dynamics.

1.3.1 Geometric and Intuitive Definition

Throughout history, the definition of the complex number has changed gradually [11]. In the 16th century, Cardano tried to work with imaginary roots in dealing with quadratic equations. Afterward, Euler used complex numbers in his calculations intuitively and correctly. It is said that by 1728 he knew the transcendental relationship $i \log i = -\pi/2$. The Euler formulae appear in his book as

$$\cos x = \frac{e^{ix} + e^{-ix}}{2} \quad \text{and} \quad \sin x = \frac{e^{ix} - e^{-ix}}{2i} \tag{1.1}$$

In 1798, Wessel described representation of the points of a plane by complex numbers to deal with directed line segments. Argand also interpreted $\sqrt{-1}$ as a rotation through a right angle in the plane, and he justified this idea on the ground that two $\sqrt{-1}$ rotations yields a reflection, i.e., -1. It is also believed that, in early 1749, Euler already had a visual concept of complex numbers as points of a plane. He described a number x on a unit circle as $x = \cos g + i \sin g$, where g is an arc of the circle. Gauss was in full possession of the geometrical theory by 1815. He proposed to refer to $+1$, -1, and $\sqrt{-1}$ as direct, inverse, and lateral unity, instead of positive, negative, and imaginary or "impossible" elements.

1.3.2 Definition as Ordered Pair of Real Numbers

The geometrical representation is intuitively simple and visually understandable, but may be weak in strictness. In 1835, Hamilton presented the formal definition of the complex number as an "ordered pair of real numbers," which also led to the discovery of quaternions, in his article entitled "Theory of conjugate functions, or algebra as the science of pure time." He defined addition and multiplication in such a manner that the distributive, associative, and commutative laws hold. The definition as the ordered pair of real numbers is algebraic, and it can be stricter than the intuitive rotation interpretation.

At the same time, the fact that a complex number is defined by two real numbers may lead present-day neural-network researchers to consider a complex network equivalent to just a doubled-dimension real-number network effectively. However, in this paper, the authors would like to clarify the merit by focusing on the rotational function even with this definition.

Based on the definition of the complex number as an ordered pair of real numbers, we represent a complex number z as

$$z \equiv (x, y) \tag{1.2}$$

where x and y are real numbers. Then the addition and multiplication of z_1 and z_2 are defined in *complex domain* as

$$\begin{aligned} (x_1, y_1) + (x_2, y_2) &\equiv (x_1 + x_2, y_1 + y_2) && (1.3) \\ (x_1, y_1) \cdot (x_2, y_2) &\equiv (x_1 x_2 - y_1 y_2, x_1 y_2 + y_1 x_2) && (1.4) \end{aligned}$$

As a reference, the addition and multiplication (as a step in correlation calculation, for example) of *two-dimensional real values* is expressed as

$$\begin{aligned} (x_1, y_1) + (x_2, y_2) &= (x_1 + x_2, y_1 + y_2) && (1.5) \\ (x_1, y_1) \cdot (x_2, y_2) &= (x_1 x_2, y_1 y_2) && (1.6) \end{aligned}$$

In the comparison, the addition process is identical. Contrarily, the complex multiplication seems quite artificial, but this definition (1.4) brings the complex number with its unique function, that is, the angle rotation, as well as amplitude amplification/attenuation, which are the result of the intermixture of the real and imaginary components.

It is easily verified that the commutative, associative, and distributive laws hold. We have the unit element $(1, 0)$ and the inverse of z $(\neq 0)$, which is

$$\begin{aligned} z^{-1} &\equiv \left(\frac{x}{x^2 + y^2}, \frac{-y}{x^2 + y^2} \right) \\ &= \left(\frac{x}{|z|^2}, \frac{-y}{|z|^2} \right) && (1.7) \end{aligned}$$

where $|z| \equiv \sqrt{x^2 + y^2}$.

1.3.3 Real 2×2 Matrix Representation

We can also use real 2×2 matrices, instead of the ordered pairs of real numbers, to represent complex numbers [11, 9]. With every complex number $c = a + ib$, we associate the $\boldsymbol{C}$-linear transformation

$$T_c : \boldsymbol{C} \to \boldsymbol{C}, \quad z \mapsto cz = ax - by + i(bx + ay) \tag{1.8}$$

which includes a special case of $z \to iz$ that maps 1 into i, i into -1, ..., with a rotation with right angle each. In this sense, this definition is a more precise and general version of Argand's interpretation of complex numbers. If we identify $\boldsymbol{C}$ with $\boldsymbol{R}^2$ by

$$z = x + iy = \begin{pmatrix} x \\ y \end{pmatrix} \tag{1.9}$$

it follows that

$$\begin{aligned} T_c \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} ax - by \\ bx + ay \end{pmatrix} \\ &= \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \end{aligned} \tag{1.10}$$

In other words, the linear transformation T_c determined by $c = a + ib$ is described by the matrix $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$. Generally, a mapping represented by a 2×2 matrix is noncommutative. However, in the present case, it becomes *commutative*. By this real matrix representation, the imaginary unit i in $\boldsymbol{C}$ is given as

$$I \equiv \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad I^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -E \tag{1.11}$$

In the days of Hamilton, we did not have matrices yet. Even after the advent of matrices, it is very rare to define complex numbers in terms of real 2×2 matrices [11] (Chapter 3, §2, 5.), [9] . The introduction of complex numbers through 2×2 matrices has the advantage, over introducing them through ordered pairs of real numbers, that it is unnecessary to define an ad hoc multiplication. What is most important is that this matrix representation clearly expresses the function specific to the complex numbers—that is, the rotation and amplification or attenuation as

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} = r \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{1.12}$$

where r and θ denote amplification/attenuation of amplitude and rotation angle applied to signals, respectively, in the multiplication calculation. On the other hand, addition is rather plain. The complex addition function is identical to that in the case of doubled-dimension real numbers.

In summary, the phase rotation and amplitude amplification/attenuation are the most important features of complex numbers.

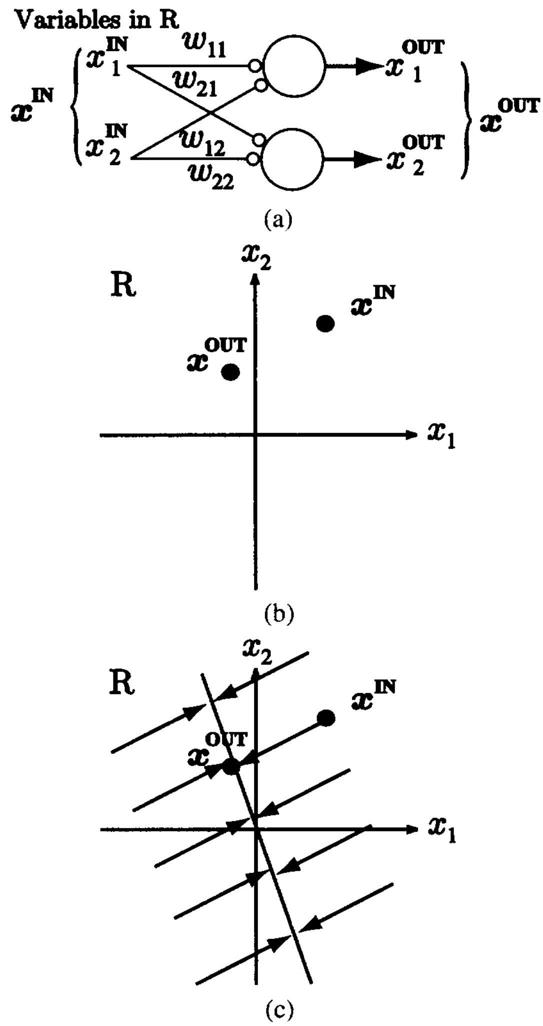

Figure 1.1 (a) A simple real-valued single-layered two-input two-output feedforward network to learn (b) a mapping that maps x^{IN} to x^{OUT} and (c) a possible but degenerate solution that is often unuseful [28].

1.4 COMPLEX NUMBERS IN FEEDFORWARD NEURAL NETWORKS

We consider intuitively what feature emerges in the dynamics of complex-valued neural networks. Here we first take a layered feedforward neural network. Then we consider metrics in correlation learning.

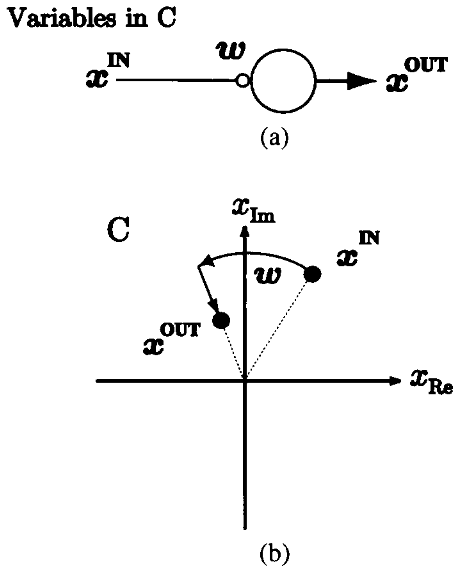

Figure 1.2 (a) A complex-valued neural network seemingly identical to Fig. 1.1(a) to learn the mapping shown in Fig. 1.1(b), and (b) a solution obtained in this small-degree-of-freedom case [28].

1.4.1 Synapse and Network Function in Layered Feedforward Neural Networks

In wave-related adaptive processing, we often obtain excellent performance with learning or self-organization based on the CVNNs. As already mentioned, the reason depends on situations. However, the discussion in Section 1.3 suggests that the origin lies in the complex rule of arithmetic. That is to say, the merit arises from the functions of the four fundamental rules of arithmetic of complex numbers, in particular the multiplication, rather than the representation of the complex numbers, which can be geometric, algebraic, or in matrices. Moreover, the essence of the complex numbers also lies in the characteristic multiplication function, the phase rotation, as overviewed in Section 1.3 [27].

Let us consider a very simple case shown in Fig. 1.1(a), where we have a single-layer 2-input 2-output feedforward neural network in real number. For simplicity, we omit the possible nonlinearity at the neurons, i.e., the activation function is the identity function, where the neurons have no threshold. We assume that the network should realize a mapping that transforms an input x^{IN} to an output x^{OUT} in Fig. 1.1(b) through supervised learning that adjusts the synaptic weights w_{ji}. Simply, we have only a single teacher pair of input and output signals. Then we can describe a general input–output relationship as

$$\begin{pmatrix} x_1^{OUT} \\ x_2^{OUT} \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1^{IN} \\ x_2^{IN} \end{pmatrix} \tag{1.13}$$

We have a variety of possible mapping obtained by the learning because the number of parameters to be determined is larger than the condition; i.e., the learning task is an ill-posed problem. The functional difference emerges as the difference in the generalization characteristics. For example, learning can result in a degenerate mapping shown in Fig. 1.1(c), which is often unuseful in practice.

Next, let us consider the mapping learning task in the one-dimensional complex domain, which transforms a complex value $\boldsymbol{x}^{IN} = (x_1^{IN}, x_2^{IN})$ to another complex value $\boldsymbol{x}^{OUT} = (x_1^{OUT}, x_2^{OUT})$. Figure 1.2(a) shows the complex-valued network, where the weight is a single complex value. The situation is expressed just like in (1.13) as

$$\begin{pmatrix} x_1^{OUT} \\ x_2^{OUT} \end{pmatrix} = \begin{pmatrix} |w|\cos\theta & -|w|\sin\theta \\ |w|\sin\theta & |w|\cos\theta \end{pmatrix} \begin{pmatrix} x_1^{IN} \\ x_2^{IN} \end{pmatrix} \tag{1.14}$$

where $\theta \equiv \arg(w)$. The degree of freedom is reduced, and the arbitrariness of the solution is also reduced. Figure 1.2(b) illustrates the result of the learning. The mapping is a combination of phase rotation and amplitude attenuation. This example is truly an extreme. The dynamics of a neural network is determined by various parameters such as network structure, input–output data dimensions, and teacher signal numbers. However, the above characteristics of phase rotation and amplitude modulation are embedded in the complex-valued network as a universal elemental process of weighting.

The essential merit of neural networks in general lies in the high degree of freedom in learning and self-organization. However, if we know *a priori* that the objective quantities include "phase" and/or "amplitude," we can reduce possibly harmful portion of the freedom by employing a complex-valued neural network, resulting in a more meaningful generalization characteristics. The "rotation" in the complex multiplication works as an elemental process at the synapse, and it realizes the advantageous reduction of the degree of freedom. This feature corresponds not only to the geometrical intuitive definition of complex numbers but also to the Hamilton's definition by ordered pairs of real numbers, or the real 2×2 matrix representation.

Though we considered a small feedforward network in this section, the conclusion is applicable also to other CVNNs such as complex-valued Hebbian-rule-based network and complex correlation learning networks, where the weight is updated by the multiplication results. The elemental process of phase rotation and amplitude modulation results in the network behavior consistent with phase rotation and amplitude modulation in total. The nature is a great advantage when we deal with not only waves such as electromagnetic wave and lightwave, but also arbitrary signals with the Fourier synthesis principle, or in the frequency domain through the Fourier transform.