

Chapter 1

Complex-Valued Neural Networks: An Introduction

Akira Hirose

Complex-valued neural networks deal with complex-valued data with complex-number weights and complex-valued neuron-activation functions. George M. Gerogiou describes clearly in the Foreword the necessity of the complex-valued networks. In this introductory short chapter, we discuss how they are or can be useful and effective. We begin with the role of $i \equiv \sqrt{-1}$ in the quantum mechanics.

According to the quantum mechanics, the motion of an electron is related to the Schrödinger equation:

$$i\hbar \frac{\partial \Psi(\mathbf{r}, t)}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \Psi(\mathbf{r}, t) + V(\mathbf{r}) \Psi(\mathbf{r}, t) \quad (1)$$

where $\Psi(\mathbf{r}, t)$ is the electron's wave function in terms of position \mathbf{r} and time t , and \hbar , m , $V(\mathbf{r}, t)$ and ∇ denote Plank constant divided by 2π , electron mass, potential function and spatial differential operator, respectively.

The probabilistic interpretation argues that the squared absolute value of the solution $|\Psi|^2$ is the probability density of electron existence. The probability is related to ensemble average. However, realistically, through repetitive or long-term experiment of electron observation in an ergodic condition, we find that the electrons obey the probability $|\Psi|^2$. The equation represents experimental results successfully.

The special feature of this equation lies in the fact that it contains the imaginary unit i in an ineliminable manner. Some physicists claim that fundamental equations in physics should not include i because they should consist of only really physical entities. Even if we regard $|\Psi|^2$ as a probability, probabilities in general are discussed by using real number, and again i is not desirable.

However, we can consider the equation as follows. "The probability possesses an amplitude entity $\log(A)$ and a phase entity ϕ . Its spatial and/or temporal evolution obeys

$$\Psi_k(\mathbf{r}, t) = e^{\log(A_k(\mathbf{r}, t)) + i\phi_k(\mathbf{r}, t)} \quad (2)$$

and the principle of superposition holds for plural solutions Ψ_k as

$$\Psi(\mathbf{r}, t) = \sum_k \Psi_k = \sum_k e^{\log(A_k(\mathbf{r}, t)) + i\phi_k(\mathbf{r}, t)} \quad (3)$$

That is to say, the imaginary unit i plays a role in the interaction of amplitude and phase entities as well as the superposition of probability functions." In this sense, i is an operator to connect entities rather than an existence itself. In the complex-valued neural networks, i has the same role to combine plural quantities consistently.

Then, what is the operation of i ? In the networks, we multiply input data by synaptic weights. If we pay attention to real and imaginary parts of the data, the multiplication of i , for example, converts the real quantity into imaginary one, and also does the imaginary one into real one with putting a negative sign. In this way the two quantities are exchanged. Multiplication of general complex-valued weight w mixes the real and imaginary quantities in a certain manner determined by the value w . In the network, the multiplication of w is executed for all the parallel input data elements $\mathbf{x} = [x_k]$. As a result, the output maintains a certain vector-direction relation in the complex plane. This is one of the properties that we can utilize effectively to treat two-dimensional information.

If we have a polar coordinate picture, we can regard the multiplication of $w = e^{\log |w| + i \arg(w)}$ as the magnification of the vector length by $|w|$ and the vector rotation of an angle $\arg(w)$. In the problems where the weight has the amplitude and phase entities $|w|$ and $\arg(w)$ in the real world, just like the solutions of Schrödinger equation, the neural operation can directly influence these entities rather than some other apparent phenomena. Actually, (quasi-)periodic signals can be processed in relation to phase because they are expressed as an integration of sinusoidal wave through the Fourier transform and the Fourier synthesis.

Furthermore, we know in the phasor treatment of signals with carrier wave that a temporal differentiation realized by a capacitor C , for example, is analytically equivalent to multiplication of $i\omega C$ where ω is angular frequency, while an integration is to division by $i\omega C$. Such relations are utilized in stabilization of dynamical or time-sequential behavior of neural networks. The phase topology is also related directly to a cyclic metric.

These properties are very important also in the device electronics. The amplitude and phase of an electron probability function is expressed by (1) and modulated by electrical potential, permittivity, magnetic field, and so on. Lightwave and electromagnetic wave also have a similar wave nature. The probability density of a photon is given by the squared absolute value of the wave function. The value multiplied by the energy of a single photon gives the energy of lightwave. It can be modulated by absorption and amplification of media in reality. On the other hand, the phase corresponds to time delay or advance and is modulated by permittivity, permeability and optical path length. In this way, the fundamental particles composing the world interact each other through the amplitude (energy) and the phase (time). (In a special case such as resonance, these two entities have a clear relation, the Kramers-Kronig relation, where the neural holomorphy will become significant.) The complex-valued neural networks are highly expected to reflect such a natural world.

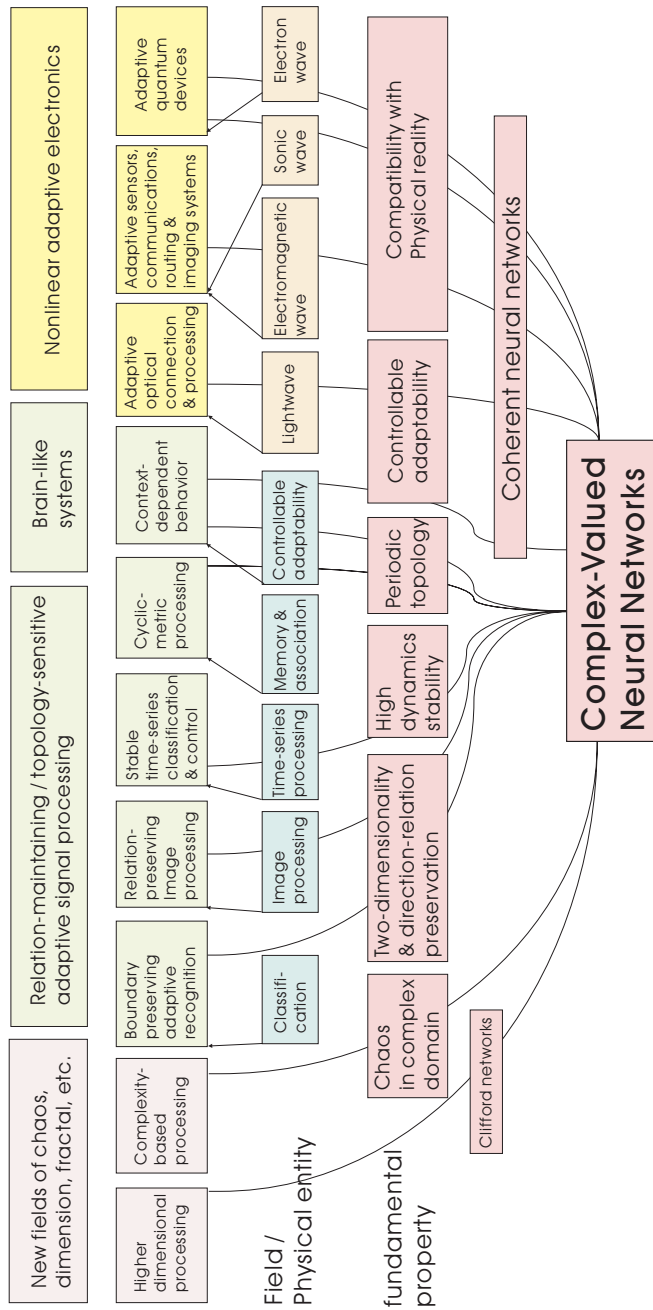


Figure 1. Fundamental properties and applications of complex-valued neural networks

Figure 1 presents the fundamental properties and application fields based on the discussion above. Most of the following chapters have a close relation to them. Recently the ideas and results have also been discussed in the Special Sessions in Conferences (KES 2001,2002) and (ICONIP 2002).

Further new ideas are also coming out to be presented in other Special Sessions in, for example, (ICANN / ICONIP 2003) and (KES 2003). The program includes quantum neural networks, radar imaging, array antenna signal processing, voice synthesis, spatiotemporal pattern processing, and so forth. These new results will be collected in a sequel. The complex-valued neural networks continue to extend the fields both in theories and applications.

References

Special Sessions on Complex-Valued Neural Networks, International Conference on Knowledge-Based Intelligent Information Engineering Systems (KES),

KES 2001 Osaka (Sept. 6-8, 2001) Proc., N.Baba, L.C.Jain, R.J.Howlett, Eds., IOS Press, Ohmsha (2001) Part 1, pp. 550-580/
KES 2002 Crema (Sept. 16-18, 2002) Proc., E.Damiani, R.J.Howlett, L.C.Jain, N.Ichalkaranje, Eds., IOS Press, Ohmsha (2002) Part 1, pp. 623-647.

Special Session on Complex-Valued Neural Networks, International Conference on Neural Information Processing (ICONIP) 2002 Singapore (Nov. 18-22, 2002) Proc., Lipo Wang, et al., Eds., Vol.3, pp. 1074-1103.

Special Session on Complex-Valued Neural Networks: Theories and Applications, International Conference on Artificial Neural Networks (ICANN) 2003 Istanbul/ International Conference on Neural Information Processing (ICONIP) 2003 Istanbul (June 26-29, 2003) to be held.

Special Session on Complex-Valued Neural Networks, International Conference on Knowledge-Based Intelligent Information Engineering Systems (KES) 2003 Oxford (Sept. 3-5, 2003) to be held.

Author's address

Akira Hirose: Department of Electrical and Electronic Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

ahirose@ee.t.u-tokyo.ac.jp