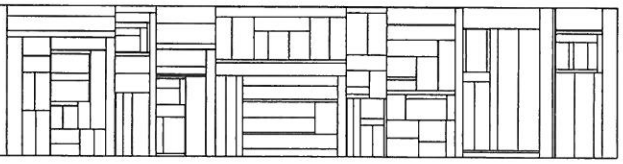
ISSN 0105-8517

Neural Networks With Complex Connection Matrices

Miklos Szilagyi

DAIMI PB – 260
August 1988

AARHUS UNIVERSITY
COMPUTER SCIENCE DEPARTMENT
Ny Munkegade 116 – DK 8000 Aarhus C – DENMARK
Telephone: + 45 6 12 71 88 *Telex: 64767 aausci dk*



NEURAL NETWORKS WITH COMPLEX CONNECTION MATRICES

Miklos Szilagyi

Department of Computer Science, Aarhus University, Aarhus, Denmark

(Permanent address: Department of Electrical and Computer Engineering, University of Arizona, Tucson, AZ 85721)

This work is supported by the National Science Foundation under Grant ECS 8700849 and by the Danish Research Council for Natural Sciences.

The author is indebted to Professors J. D. Garcia, G. A. Korn, B. Lautrup, B. Mayoh, R. H. Parmenter and T. Triffet for stimulating discussions.

Requests for reprints should be sent to Miklos Szilagyi, Department of Applied Physics, Delft Technological University, Delft, The Netherlands.

Running title: COMPLEX CONNECTION MATRICES.

NEURAL NETWORKS WITH COMPLEX CONNECTION MATRICES

Abstract-The concept of neural networks is generalized to include complex connections between the network's units. The similarity between the dynamics of some linear complex networks and the quantum mechanical behavior of atomic systems is shown.

Key words: Neural networks, Connection weights, Connection matrices, Quantum mechanics, Hamiltonian equations, Differential equations, Eigenvalues, Liapunov function, Energy.

The dynamic behavior of neural networks can be fully described by systems of differential equations (Hopfield, 1984; Grossberg, 1988). A large amount of information can be extracted from both the theoretical investigation and the numerical solution of such systems. It is also possible, however, to explore the behavior of neural networks by means of suitable physical models whose behavior is determined by similar systems of differential equations. The number of such physical systems is abundant. Let us just mention the oscillations of coupled mechanical and electrical systems such as coupled pendulums, delay-line circuits, analog amplifiers or ecological systems, e.g. the predator-pray model, etc. (Korn, 1988).

Classical harmonic oscillators represent a characteristic example of such systems. The states of such oscillators change in a continuous manner but the frequencies that define those states can only have discrete values. The oscillations of the system are superpositions of some monochromatic oscillations with normal-mode frequencies. These frequencies are the properties of the entire system and in the general case an infinite number of modes must be present, each having an appropriate amplitude and phase. Any linear oscillatory system is equivalent to a set of such harmonic oscillators. The natural frequencies of the system correspond to the normal modes of the harmonic oscillators.

For example, a set of N pendulums represents the sum of N oscillations with N different frequencies. The general motion is the superposition of these oscillations with a continuous energy transfer between the pendulums. The system is described by N coupled second-order linear differential equations. The speed at which the energy is swapped back and forth depends on the coup-

ling terms of the equations (Feynman, Leighton and Sands, 1965).

Complex functions are commonly used for the description of physical quantities. A well known example is the treatment of electric networks in terms of complex potentials, currents and impedances. The dynamics of such networks is represented by systems of differential equations in which the complex impedance matrix plays the role of the connections between the different elements of the network. One can even introduce the notion of complex power as the inner product of the complex voltage and current vectors. The real and imaginary parts of this power are the active (dissipative) and the reactive powers, respectively.

The Hamiltonian equations of Quantum Mechanics represent another well known example. Schrodinger's wave equation can be written in the form of the following system of N coupled differential equations (Feynman, Leighton and Sands, 1965):

$$i\hbar \frac{dC_j}{dt} = \sum_{k=1}^{N} H_{jk} C_k \qquad (j = 1, 2, \ldots, N) \qquad (1)$$

where N is the number of base states, $\hbar = h/(2\pi)$, h is Planck's constant, t is time, $i = (-1)^{1/2}$ is the imaginary unit, H_{jk} is an element of the Hamiltonian matrix and C_k is the probability amplitude to find the state of the system in the kth base state at time t. The system of differential equations (2) describes the time dependence of all these probability amplitudes, i.e. the global dynamic behavior of the entire atomic system.

The elements of the Hamiltonian matrix characterize the probability amplitudes of the generation of one base state from another one, i.e. the couplings between the different base states of the system.

In general, the probability amplitudes as well as the elements of the Hamiltonian matrix are complex. However, since the Hamiltonian matrix is hermitian, i.e. $H_{jk}^{*} = H_{kj}$ (where the asterisk represents the conjugate complex), its diagonal elements are all real.

Since the nature of information processing in biological nervous systems is totally different from that in digital computers (Kohonen, 1988), and it is still mostly unknown, one should not ex catedra exclude the existence of neural networks with complex connection weights between its individual units. Indeed, the response of biological neurons to stimuli can be represented as a measuring process and there is some evidence that the quantum mechanical theory of measurement can be applied to it (Jolesz and Szilagyi, 1978). Thus, although the connection weights in the anatomy of macroscopic neural networks may be real, there is no evidence that would exclude the existence of neural networks with complex connections at the functional level that can be anatomically realized e.g. between clusters of neurons.

Let us now consider a neural network with complex connections between its units. For simplicity, we shall restrict this treatment to linear networks but it can easily be extended to arbitrary nonlinear networks, too. The network dynamics is described by the system of differential equations written in the matrix form as

$$\underset{\sim}{dz}/dt = A \underset{\sim}{z} \qquad (2)$$

where $\underset{\sim}{z} = [z_1(t), z_2(t), \ldots, z_N(t)]$ is the complex state vector that describes the time dependence of the activation of each unit and A is the complex connection matrix (in this simplified treat-

ment we shall assume that the activation of each unit is equal to the output of the same unit). In case of an electric network $\underset{\sim}{z}$ may be the complex voltage vector and A is the complex impedance matrix. In Eq. (1) the components of the state vector $\underset{\sim}{z}$ are the probability amplitudes C_k. Then

$$A_{jk} = - (i / \hbar) H_{jk} \qquad (3)$$

and all diagonal elements A_{jj} are imaginary.

The solutions of Eq. (1) exist if and only if the condition

$$\text{Det} (H_{jk} - \delta_{jk} E) = 0 \qquad (4)$$

is satisfied where δ_{jk} is the Kronecker delta. This is the equation for the eigenvalues E_n of the Hamiltonian matrix. Equation (4), in general, has N solutions (some of them may be equal to each other). The eigenvalues are given by

$$E_{1,2} = (H_{11} + H_{22})/2 + [(H_{11} - H_{22})^2/4 + H_{12} H_{21}]^{1/2} \qquad (5)$$

Due to the hermitian nature of the Hamiltonian, its diagonal elements are always real and the product $H_{12} H_{21}$ is real and positive. Therefore, the eigenvalues are always real. They represent the possible energies of the system. The energy values are only dependent on the elements of the Hamiltonian matrix. The exact value of the energy is only defined for the stationary states. In the general case we can only talk about probabilities of being in one base state or another.

If the eigenvalues and eigenfunctions are known for a certain connection matrix, an approximate solution can be obtained for a slightly different matrix by using perturbation theory (Landau and Lifshitz, 1965). In case of a time-dependent Hamil-

tonian matrix the energy is not conserved, therefore the stationary states do not exist. In this case the perturbation is a function of time. Very little can be said about the energy in this case but we know that Heisenberg's uncertainty relation exists between the results of measurements of energy at two different moments of time and the time interval. As we have pointed out before (Jolesz and Szilagyi, 1978), the uncertainty relation is probably present in the measuring process of the nervous system, too.

Eq. (4) is equivalent to the algebraic characteristic equation of the system of differential equations (2) and the eigenvalues E_n are proportional to the latent roots λ_n of the characteristic equation that are all imaginary in this case, according to Eq. (3). This fact represents the fundamental difference between real and complex networks. The complex solutions of Eq. (2) describe oscillatory behavior of the state vector in most cases.

The trivial case of N = 1 evidently yields that the only eigenvalue is $E_1 = H_{11}$ which is the energy of the single stationary state. The realization of a system with a very large value of N is, e. g., an electron in a crystal lattice. The solutions can still be sought in the form of normal modes (exponential functions of the eigenvalues).

In order to better understand the properties of neural networks with complex connection weights, one should consider the simplest nontrivial case of N = 2. The solutions of the system of equations (2) are simple enough but they depend on the relationships between the different elements of the connection matrix (Kamke, 1959). They can always be written in the form of sums of two complex exponential functions of λ_n.

Two special cases should be mentioned.

In the symmetric case of $H_{11} = H_{22}$ and $H_{12} = H_{21}$ all elements of the Hamiltonian are real. It means that all matrix elements of the corresponding neural network are imaginary. The solutions will oscillate like two identical coupled pendulums.

In case of $H_{11} = H_{22} = 0$ the integral curves in the phase plane can be hyperbolae or ellipses. If, in addition, we require that the matrix elements H_{12} and H_{21} be imaginary, then the corresponding neural network has real connection weights but they are antisymmetric. The integral curves are circles with their centers at the origin of the coordinate system which corresponds to oscillating solutions.

We can see that in most cases the basic difference between neural networks with complex and real connection weights is their oscillatory and dissipative dynamics, respectively. In case of complex connections it is generally difficult to construct a Liapunov function. It is, however, not impossible. In some cases asymptotic stability can be reached (La Salle and Lefschetz, 1961). We would like to mention that the Lyapunov function is a global map of the entire system while the eigenvalues of the energy discussed above represent only the stationary solutions. Evidently, the initial conditions determine which eigenvalue will be reached first.

In conclusion, we would like to emphasize that neural networks can be investigated on the basis of simple physical models. Since the functional behavior of biological neural networks is mostly unknown, their general models should include complex connection matrices. The dynamics of some linear networks is similar to the quantum mechanical behavior of atomic systems. Of course,

it does not mean that a classical model of quantum mechanics is possible but this approach may shed some light to the actual functioning of biological nervous systems. Further investigations of this problem should be based on the analysis of nonlinear neural networks with complex connection matrices.

REFERENCES

Feynman, R. P., Leighton, R. B., & Sands, M. (1965). The Feynman lectures on Physics, Vol. 3. Reading: Addison-Wesley.

Grossberg, S. (1988). Nonlinear neural networks: principles, mechanisms, and architectures. Neural Networks, 1, 17-61.

Hopfield, J. J. (1984). Neurons with graded response have collective computational properties like those of two-state neurons. Proceedings of the National Academy of Sciences, 81, 3088-3092.

Jolesz, F., & Szilagyi, M. (1978). A possible new model of neurons and neural processes based on the quantum mechanical theory of measurement. Acta Cybernetica, 4, 221-239.

Kamke, E. (1959). Differentialgleichungen. Leipzig: Akademische Verlagsgesellschaft.

Kohonen, T. (1988). An introduction to neural computing. Neural Networks, 1, 3-16.

Korn, G. A. (1988). Interactive dynamic system simulation and Mathematics on personal computers. New York: McGraw-Hill.

Landau, L. D., & Lifshitz, E. M. (1965). Quantum mechanics. Oxford & New York: Pergamon Press.

La Salle, J., & Lefschetz, S. (1961). Stability by Liapunov's direct method. New York: Academic Press.