Chapter 1

Introduction

An artificial neural network is designed to model how the human brain performs a particular task. The basic principle of the working mechanism is motivated by the function of biological neurons. In this thesis, we investigate the stability and synchronization problems of neural networks. The dynamics of neural networks are a crucial part that is needed in implementation in real world's problems such as associative memory, artificial intelligence, secure communication, signal processing, optimization problem [2, 3, 4]. For example, chaos synchronization of neural networks is applied in secure communication to increase the security of signals transmitting from a transmitter to a receiver.

In synchronization problems of neural networks, control theory is needed to stabilize the error systems. Many controllers have been designed so far, such as intermittent control [5], linear feedback control [6], integral sliding mode control[7], and impulsive control[8, 9, 10, 11]. Among these effective controllers, the intermittent and impulsive controls are discontinuous, which reduces the control cost; thus, they are more effective than continuous controllers. When we use the impulsive controller to control a nonlinear system, it makes the system hybrid, i.e., the system's state is continuous in the time-span except at a countable number of points. This type of system is called an impulsive system. This thesis's primary purpose is to study the dynamics of neural networks with different kinds of impulsive sequences. We focused on a hybrid impulsive sequence containing stabilizing and destabilizing impulses and tried to extend it to a more general class of impulsive sequences.

In the first section of this introductory chapter, we present the underlying motivating factor of artificial neural networks, biological neurons. We review a neuron's structure, the transmission of neural signals, and how the communication between these biological neurons can be modeled. In the following section, we define artificial neural networks and give introductory mathematical concepts applied in further chapters to investigate the dynamics of neural networks. In the final section of this chapter, we present the definitions of synchronization with its different forms.

1.1 Biological neurons

The human brain and the nervous system are a composition of a huge number of interconnected cellular units (nerve cells or neurons) and glial cells. The glial cells are supportive in providing physical and functional supports to neurons. Although neurons are found in a wide variety of shapes, sizes, and locations, most of them are of uniform structures and follow the same basic principles of transmitting neural signals [12, 13]. The essential features of the brain are the connectivity of neurons and the mechanism of transmitting electrochemical signals in neurons. Both of them help the brain to perform complex tasks.

A prototypical neuron is in Figure 1.1. The cell body (soma) is the central part of neuron which constitutes the metabolic center and contains the nucleus and other essential organelles. Many root-like extensions branch out from the cell body called dendrites that have receiving zones of synaptic signals or impulses or information from other neurons. A long fiber-like extension of the cell body is called an axon

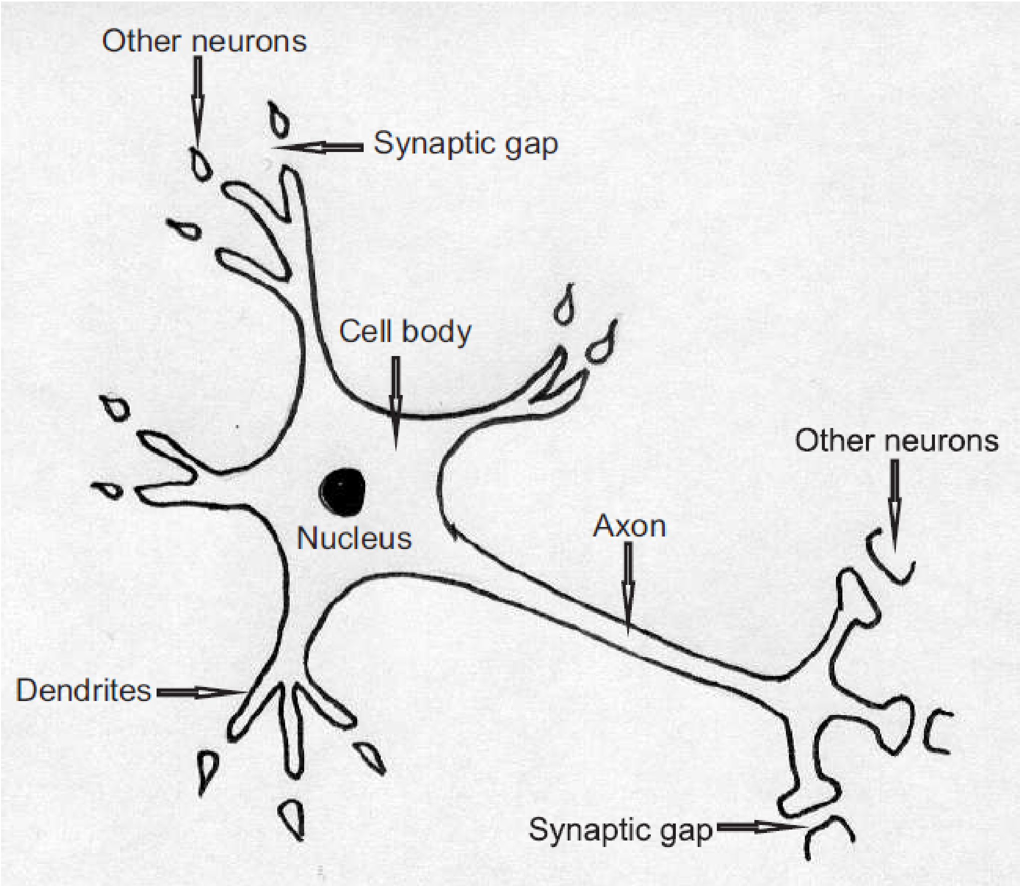

FIGURE 1.1: Schematic structure of a typical neuron [1].

that has specialized terminals (synaptic endings)-divided into many branches-to convey the signals to target neurons. The main purpose of axon is to propagate a self-generating electrical wave as electric signals from the point of initiation at the cell body to its terminals. The mechanism of transmitting electrochemical signals through axon is called the action potential. There are two types signaling mechanism in neurons one is electrically and another is chemically. Interior of neurons is prevailed with electrical signal, whereas chemical signals are operated at synaptic ends.

Now we describe the procedure through which an electric current is generated across the membrane of the axon, how it propagates, and the ways it can be modeled in the mathematical equation. The neuron membrane consists of two thin layers of the lipid molecules. The layers separate the axon cytoplasm (interior) of the neuron and the extracellular fluid. The membrane is selective to diffusion of particular molecules, and its diffusion selectivity varies with time and length along the axon. The selective permeability of membranes is due largely to ion channels which allow only certain kinds of ions to cross the membrane in the direction of their concentration and electro-chemical gradients.

There are ion concentrations of K^+, Na^+, Cl^-, and Ca^{2+}, and their differences across the membrane of the axon cause an electrical potential in the neurons. In the state of equilibrium the potential difference in the neuron is -70mV (resting potential) across the membrane of the axon. In an inactive state of neuron the ions distributions across the thin layers of the membrane makes the interior or cytoplasm of neuron negatively charged relative to the extra-cellular fluid. This is caused by a biological ion pump that works when the interior of the axon becomes more positive.

When an electric pulse injected in the axon of the neuron, the resting potential across the membrane is deviated. This deviation changes ions concentrations across the membrane that vary the potential difference. Thus, the membrane of the axon has the property of capacitance, i.e., separation of charges. The function of membrane is similar to the function of resistor-capacitor (RC)-type circuit that will be discussed in the next section [14, 15]. The membrane permeability allows Na^+ ions to enter in interior of the neuron through the ion channels. This increases the voltage across the membrane, as it exceeds the threshold, the injected current produces a single pulse which propagates through the axon terminals. The pulse signal traveling through the axon terminals stops at the synaptic ends due to the synaptic gap. From the synaptic ends, it is transferred to the target neuron by a special chemical mechanism called synaptic transmission. In synaptic transmission, special substances are released from the synaptic ends called neurotransmitters, and hits the target neurons passing through the synaptic gap.

In the target neurons, the signal pulse act as a current pulse that follows the same mechanism, to hit the other neurons, as it is followed in the pre-synaptic neuron.

1.1.1 Biological model

As we have discussed above that the neurons in human brain transmit an electrical pulse along their axons to other neurons. By transmitting signals from one neuron to other neurons in the network, our brain perform a particular task of interest. In order to understand the brain functioning, it is necessary to model the biological network of neurons into the mathematical equation. In this subsection we discuss about the general mathematical model of biological neural networks that was first formulated by Robert L. Harvey in his book *Neural Network Principles* published in 1994.

Assume that there are n neurons in the network shown in Figure 1.2, denote them by $N_1, N_2, ..., N_n$. Let a variable $x_i(t)$ describes ith-neuron's state and a variable $Z_{ij}(t)$ describes the coupling strength between the neurons N_i and N_j. More precisely,

$$x_i(t) = \text{the deviation of the } i\text{-th neuron from the resting potential.}$$

The activation level of the i-th neuron is described by the variable $x_i(t)$. It is called axon potential or short term memory(STM) trace. The strength of interaction between the neurons N_i and N_j is denoted by $Z_{ij}(t)$ which can be negative or positive depending on the signal that inhibits or excites the neurons to fire. If $Z_{ij}(t) > 0$, it means that i-th neuron is excited to send a signal to j-th neuron. If $Z_{ij}(t) < 0$,it

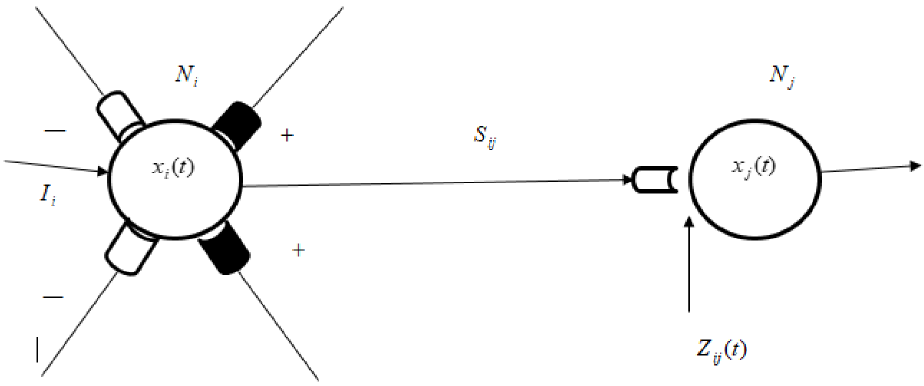

FIGURE 1.2: Schematic diagram of a network of neurons. Neuron N_i sends a signal pulse S_{ij} through its axon to hits the target neuron N_j with coupling strength $Z_{ij}(t)$. The external inputs I_i are stimuli for the neuron model.

means that i-th neuron is inhibited to send a signal to j-th neuron.

$Z_{ij}(t) =$ the average release rate of neurotransmitter per unit axonal signal frequency.

This is called the synaptic coupling coefficient or long term memory(LTM) trace.

Assume a deviation in neuron's potential from equilibrium due to internal and external processes in neural network. The rate of change in neuron's potential is described by the following differential equation.

$$\frac{dx_i(t)}{dt} = \left(\frac{dx_i(t)}{dt}\right)_{external} + \left(\frac{dx_i(t)}{dt}\right)_{internal}, \quad \forall i. \tag{1.1}$$

Assume inputs from other neurons and stimuli are additive. Then, we have

$$\frac{dx_i(t)}{dt} = \left(\frac{dx_i(t)}{dt}\right)_{external} + \left(\frac{dx_i(t)}{dt}\right)_{excitatory} - \left(\frac{dx_i(t)}{dt}\right)_{inhibitory} + \left(\frac{dx_i(t)}{dt}\right)_{internal}, \forall i. \tag{1.2}$$

Assuming further that neuron's potential is decaying exponentially to equilibrium state without having external processes, we have

$$\left(\frac{dx_i(t)}{dt}\right)_{internal} = -\alpha_i(x_i(t))x_i(t), \text{where } \ \alpha_i(x_i(t)) > 0, \quad \forall i. \tag{1.3}$$

Assume additive synaptic excitation is proportional to the pulse train frequency

$$\left(\frac{dx_i(t)}{dt}\right)_{excitatory} \propto \sum_{other \ neurons} \text{(frequency of signal)(synaptic coupling strengths).} \tag{1.4}$$

It can be written as

$$\left(\frac{dx_i(t)}{dt}\right)_{excitatory} = \sum_{\substack{l=1 \\ l\neq i}}^{n} S_{li}(t)Z_{li}(t), \forall i, \tag{1.5}$$

where $S_{li}(t)$ is the average frequency of signal in the axon from neuron N_l to N_i, evaluated at N_i. The average signal frequency $S_{li}(t)$ depends on the propagation time delay τ_{li} taken by the signal to reach neuron N_i from N_l, and also depends on the threshold value Γ_l for firing of N_l in the following manner

$$S_{li}(t) = f_l(x_l(t - \tau_{li}) - \Gamma_l), \tag{1.6}$$

where $f_l : \mathbb{R} \to [0, \infty)$ is a given non-negative function called signal function. There are many different forms of signal functions which are commonly used in neural networks, we shall discuss it later in detail.

Substituting (1.6) in equation (1.5), we get

$$\left(\frac{dx_i(t)}{dt}\right)_{excitatory} = \sum_{\substack{l=1 \\ l\neq i}}^{n} Z_{li}(t) f_l(x_l(t-\tau_{li}) - \Gamma_l), \forall i. \tag{1.7}$$

Assume hardwiring of the inhibitory inputs from the other neurons,i.e., the coupling strengths between the inhibited neurons are constant. We have the following

$$\left(\frac{dx_i(t)}{dt}\right)_{inhibitory} = \sum_{\substack{l=1 \\ l\neq i}}^{n} C_{li}, \forall i, \tag{1.8}$$

where $C_{li} = b_{li}h_l(x_l(t-\tau_{li}) - \Gamma_l)$, and $b_{li} \geq 0$ is constant coupling strength. The function h_l is a signal function. Generally, the threshold value Γ_l is same for every neuron in the network.

The stimuli are other external sources to change the neuron's potential, so we have

$$\left(\frac{dx_i(t)}{dt}\right)_{external} = I_i, \forall i. \tag{1.9}$$

Now, substituting the equations (1.3), (1.7), (1.8), and (1.9) in the equation (1.2), we get the so-called additive STM trace equation for $i = 1, 2, ..., n$.

$$\frac{dx_i(t)}{dt} = -\alpha_i(x_i(t))x_i(t) + \sum_{\substack{l=1 \\ l\neq i}}^{n} Z_{li}(t) f_l(x_l(t-\tau_{li}) - \Gamma_l) - \sum_{\substack{l=1 \\ l\neq i}}^{n} b_{li}h_l(x_l(t-\tau_{li}) - \Gamma_l) + I_i. \tag{1.10}$$

As we have assumed that the excitatory synaptic coupling is varying with time, so we have the following equation based on Hebb's law.

$$\frac{dZ_{ij}(t)}{dt} = -A_{ij}(Z_{ij}(t))Z_{ij}(t) + P_{ij}(t)[x_j(t)]^+, \quad A_{ij}(Z_{ij}(t)) > 0, \forall i, j, \tag{1.11}$$

where $P_{ij}(t) = \beta_{ij} f_i(x_i(t - \tau_{ij}) - \Gamma_i)$, $\beta_{ij} \geq 0$, and

$$[x_j(t)]^+ = \begin{cases} x_j(t), & \text{if } x_j \geq 0, \\ 0, & \text{if } x_j < 0. \end{cases}$$

The second term of the equation (1.11) shows that to increase $Z_{ij}(t)$, neuron N_i must send a signal $P_{ij}(t)$ to neuron N_j, and at the same time N_j must be activated, i.e., $x_j(t) > 0$.

Note that the STM and LTM trace equations (1.10) and (1.11) respectively are not solvable until the coefficients $\alpha_i, A_{ij}, C_{li}, S_{li}, P_{ij}$ and the external stimuli I_i are given.

1.2 Artificial Neural Networks

An artificial neural network (or neural network) is a biologically motivated machine that is designed to model the way in which the biological neural network perform a particular task or function of interest. The network is usually implemented by using electronic components or is simulated in software on digital computer [16]. It consists of computational units that correspond to neurons, and of interconnections correspond to synapses. The architecture of artificial neurons and synaptic connections between them is inspired by the biological neuron model. The table 1.1 shows the comparison between an artificial neural network to a biological one.

The definition of an artificial neural network may be the following defined by Simon Haykin in his book [16] *Neural Network: A Comprehensive Foundation.*

Definition 1.2.1. A neural network is a massively parallel distributed processor consists of simple processing units, which has a natural propensity for experiential knowledge and making it available for use. It resembles the brain in two respects:

Biological neural network	Artificial neural network
Neuron	Processing unit
Dendrite	Input unit
Cell Body	Processing function
Axons	Output unit
Synapse	Weights
External stimulus	Bias

TABLE 1.1: Comparison between a human brain and an artificial neural network.

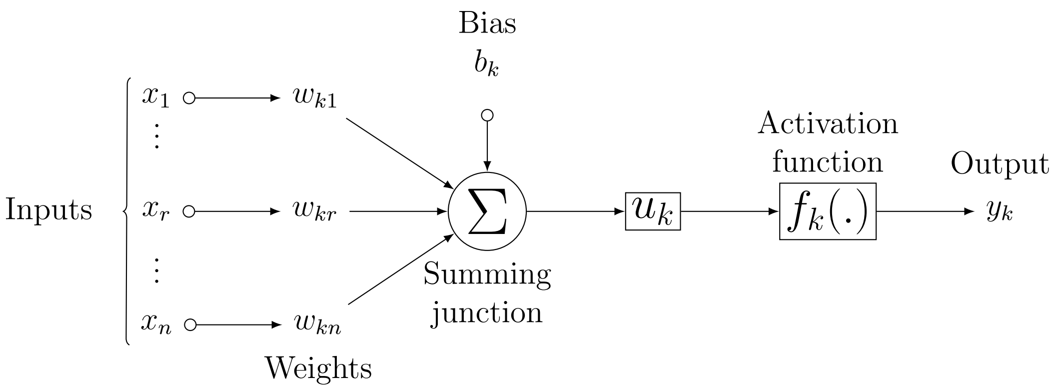

FIGURE 1.3: A diagram of an artificial neuron of the network.

(i) Neural network acquired knowledge from its environment through a learning process.

(ii) Interconnection strengths, known as synaptic weights, are used to store the acquired knowledge.

The specific features of a neural network that make it a powerful computing machine is its massively parallel distributed structure, and its ability to learn and therefore generalize. Generalization refers to its ability to produce reasonable outputs for the inputs which are not encountered during learning. A neural network performs to compute a complex problem in integrated manner. It cannot perform working individually.

A basic diagram block of a neuron of the network is shown in Figure 1.3. The input signals $x_1, ..., x_r, ..., x_n$ with synaptic weights $w_{k1}, ..., w_{kr}, ..., w_{kn}$ are connected to the

k-th neuron. The synaptic weights are the measurement of connections' strength that can be negative or positive depending on whether the signal is inhibitory or excitatory, respectively. The manner of writing the subscript in synaptic weights w_{kr} is important to note, the first subscript refers to the neuron that is receiver of the signals, and the second subscript refers to the input ends of the synapse to which the weight refers. Different from biological neuron, the synaptic weights in artificial neural networks may lie in intervals of real number.

The summing junction acts as a linear combiner of inputs ,i.e., input signal x_r is multiplied by its respective synaptic weight w_{kr}. The bias b_k is an external stimulus that has the effects of decreasing or increasing the input of activation function $f_k(.)$, it depends on whether the bias is negative or positive, respectively. Mathematical expression for the linear combiner of input signals is the following.

$$u_k = \sum_{r=1}^{n} x_r w_{kr} + b_k. \tag{1.12}$$

The output of the linear combiner u_k in equation (1.12) is an input value or an induced local field for the activation function. The main purpose of an activation function is to get the desired output of a neuron. This is also known as a squashing function because it squashes the amplitude of the output of a neuron to some finite value. In figure 1.3, the output of a neuron after employing the activation function is as follows:

$$y_k = f_k(u_k) = f_k(\sum_{r=1}^{n} x_r w_{kr} + b_k). \tag{1.13}$$

An activation function of a neuron can be linear or nonlinear in nature. Typically, there are three types activation functions that are usually considered in designing neural networks:

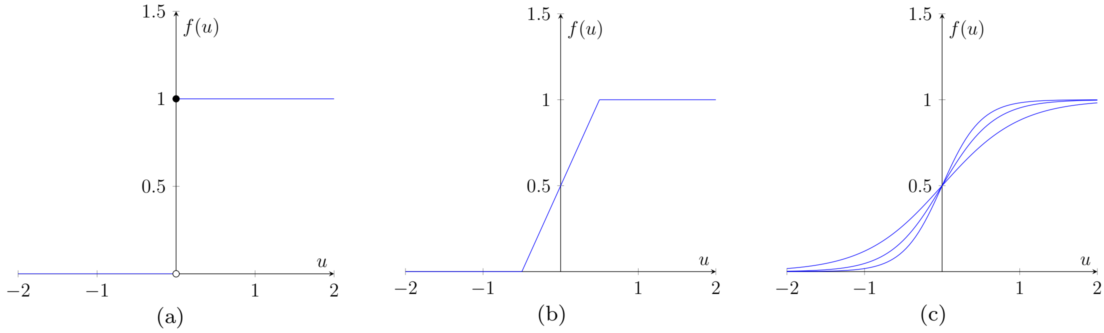

FIGURE 1.4: (a) A threshold function, (b)Piecewise linear function, and (c) Sigmoid activation function for $\beta = 2, 3$, and 4.

(a) *Threshold function (step function).* For this type of activation function shown in Figure 1.4(a), we shall have the following output for a neuron of a Figure 1.1.

$$y_k = f_k(u_k) = \begin{cases} 1 \ \text{ if } u_k \geq 0, \\ 0 \ \text{ if } u_k < 0. \end{cases} \tag{1.14}$$

The model involving threshold function as an activation function is known as McCulloch-Pitts model [17] in recognition of their pioneer work done by McCulloch and Pitts in 1943. The basic concept of this model is, a neuron fires a signal indicating "yes",i.e., "1" if the induced local field of that neuron is non-negative, and 0 otherwise. The activation function describes all-or-none property of a neuron in the McCulloch-Pitts model.

(b) *Piecewise linear function.* A piecewise linear function shown in Figure 1.4(b) is defined by the following function.

$$f(u) = \begin{cases} 0 \ \text{ if } u \leq -\frac{1}{\beta}, \\ u + \frac{1}{\beta} \ \text{ if } -\frac{1}{\beta} < u < \frac{1}{\beta}, \\ 1 \ \text{ if } \ u \geq \frac{1}{\beta}, \end{cases} \tag{1.15}$$

where β is called the amplification factor or the neural gain. Such type of activation function has been widely implemented in cellular neural network [18, 19]. The piecewise linear function plotted in figure 1.4(b) is for $\beta = 2$. It can be observed from the function 1.15 that if the amplification factor β approaches to infinity then the piecewise linear function reduces to a threshold function.

(c) *Sigmoid function.* The sigmoid function shown in figure 1.4(c) is the most commonly used activation function in the construction of artificial neural networks. It is a strictly increasing, smooth, and bounded function. A logistic function is one of the examples of a sigmoid function, defined by

$$f(u) = \frac{1}{1 + exp(-\beta u)}, \tag{1.16}$$

where β is an amplification factor, and it represents the slop parameter of the curve. By varying β, we can get different slopes of the curve as it is shown in Figure 1.4(c). It can be observed from the function 1.16 that the sigmoid function asymptotically approaches to 1 and 0 when $\beta \to \infty$ and $\beta \to -\infty$, respectively. That is the sigmoid function reduces to a threshold function if the slop parameter β approaches infinity. Whereas a threshold function has the range of only two elements 0 and 1, a sigmoid function has the range in open interval $(0, 1)$.

It is worth noting that a threshold function is not smooth whereas a sigmoid function is a smooth function. The smoothness of the sigmoid function is an important property for applications perspective: it allows analog signal processing and it makes many mathematical theories applicable [20]. If we consider random variables for the firing threshold of an all-or-non neuron with a Gaussian normal distribution function, then the expected output signal's value is a sigmoid function of activity [21]. The model of this type of neurons is called stochastic model. For this and other reasons, a sigmoid activation function has become increasingly popular in designing artificial neural network model. Other popular examples of sigmoid activation function are inverse tangent function and tangent hyperbolic function.

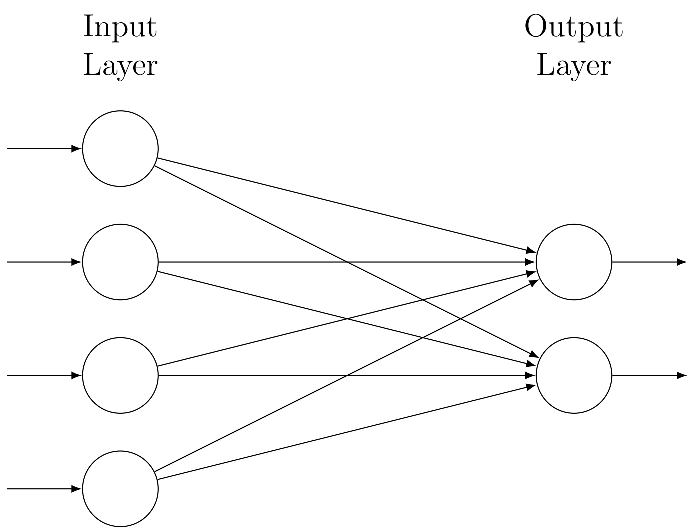

FIGURE 1.5: A feedforward network without hidden layer.

A network architecture of neurons in constructing the artificial neural networks is a very crucial point for its applications or learning point of view. Different structure of neurons in networks is designed for the different purpose of applications. In general, for layered neural networks, there are two types of network architectures:

(a) *Feedforward network.* A feedforward neural network is shown in Figure 1.5 which consists of input layer of 4 neurons and output layer of 2 neurons. Nodes in input layer are source of signals sending towards the neurons of output layer, and not vice versa. The layer of input neurons is not counted because there is no connections towards them to perform information processing. Hence, such type of neural networks is called a single layered neural network. More complicated feedforward neural network is a multilayer feedforward network

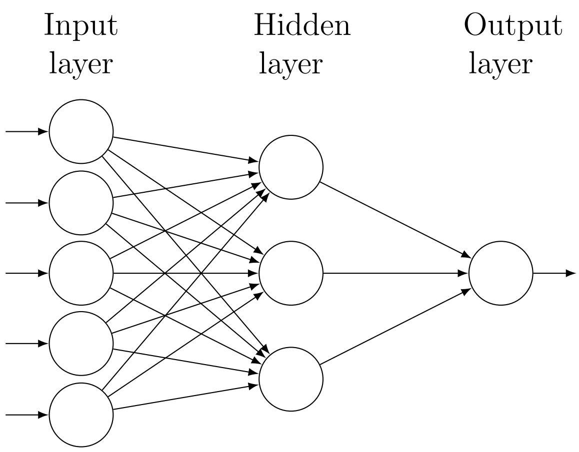

FIGURE 1.6: Multilayer feedforward neural network with one hidden layer.

which is shown in Figure 1.6. The network is consisted of a input layer of 5 neurons, a so-called hidden layer of 3 neurons, and a output layer of single neuron. This network is also known as a $5-3-1$ network. In general, a feedforward neural network consists of p neurons in input layer, q neurons in first layer, h neurons in second layer, and h_2 neurons in output layer is known as a $p-q-h_1-h_2$ network.

The main work of hidden layer's neurons is to intervene between the external inputs and the network output in a specific manner. If we add one or more hidden layers then the network becomes able to extract higher-order statistics. In other words, despite of having local connectivity in the network it acquires a global perspective due to the extra set of synaptic connections and the extra dimension of neural interactions [22]. The mechanism of forwarding the signals in a multilayer neural network is started from the neurons of input layer which supply the respective elements of the activation pattern (input vector) to the neurons of the first layer (i.e., the first hidden layer). Further, the output signals of the first hidden layer are supplied forward to the neurons of the second hidden layer, and so on for the rest of the network. Typically, the

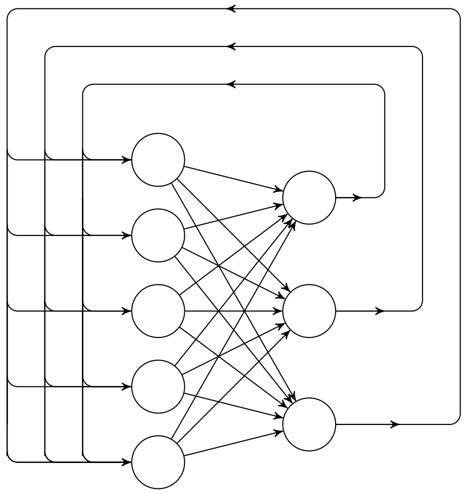

Figure 1.7: A recurrent network with one hidden layer of neurons.

input signals of the neurons in each layer of the network are the output signals of the preceding layers only. The overall response to the activation patterns of the neurons in the input layer is constituted by the set of output signals of the output (final) layer's neurons.

(b) *Feedback or Recurrent network.* The network of neurons containing the feedback of output signals as inputs is called a feedback or a recurrent neural network. A recurrent network may be of a single layer without a hidden layer as it is shown in Figure 1.8. The network in Figure 1.8 has four neurons, and each one of them feeds its output signal back to the inputs of all the other neurons. Note that each neuron in the network is not feeding back its output signal to the input of itself,i.e., network does not contain a self-feedback loop. Existing of self-feedback loops in the network has a profound impact on the learning capabilities, and its performance. Another class of recurrent neural networks is presented in Figure 1.7, where the network has a single hidden layer.

The feedforward neural network with a single hidden layer shown in Figure 1.6 is a fully connected network. The network is said to be fully connected if every neuron

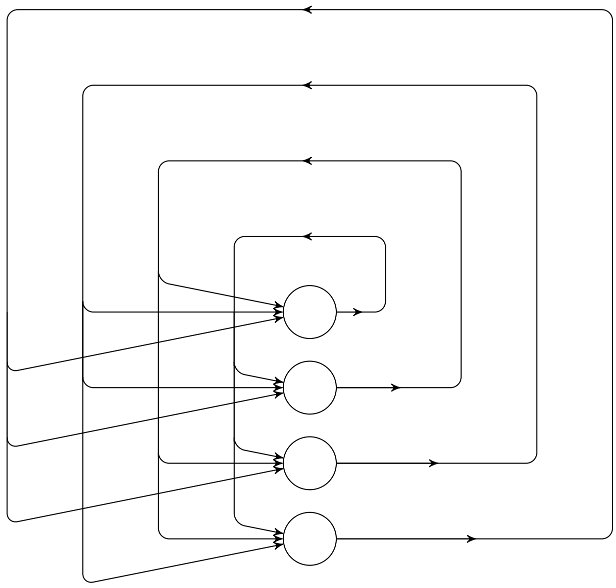

FIGURE 1.8: A recurrent neural network without a hidden layer and a self-feedback loop.

of each layer in the network is connected to every neuron in the adjacent forward layer. In other case, if some of the synaptic connections from the network is missing then it is called a partially connected network. An example of a partially connected network with one hidden layer is presented in Figure 1.9, where each neuron of the hidden layer has a local field of interconnections with the neurons of input layer.

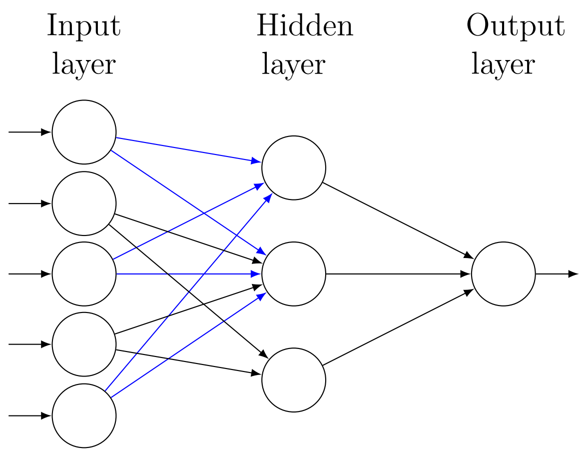

FIGURE 1.9: Partially connected neural network.

From the above discussions, it is concluded that the performance of a neural network is affected by the following important factors.

(1) External inputs.

(2) Internal decay rates.

(3) Synaptic weights.

(4) Activation functions.

(5) Propagation delays.

(6) Connections topology/Network architecture.

A deterministic model of a neural network is a dynamical system in the sense that the future activation levels and synaptic coupling coefficients can be calculated if their initial counterpart is given and all the other factors are assumed as parameters. This is called the joint activation-weight dynamics. However, there are many learning processes to train the network in which the weight dynamic is separated from the activation dynamic. In the learning processes, synaptic weights of a network are adaptively determined to achieve some particular kind of tasks such as pattern recognition or to obtain some desired network outputs from a given set of inputs. Such a scheme usually determines a discrete dynamical system (a system of difference equations) or a continuous dynamical system (a system of differential equations) in the space of matrices of synaptic coupling coefficients. This is called weight dynamics.

Now, we stop here the discussions about the basic properties of artificial neural networks, and proceed further to introduce the famous models of neural networks such as Hopfield model, Cohen-Grossberg model, and Bidirectional associative memory model. All are important for this thesis. In order to introduce Hopfield model, we have to understand the RC-circuit about which we have mentioned in the previous section.

1.2.1 Basics of electrical circuit

Electrical circuit is an important tool to understand the mechanism of signal processing among the neurons of a human brain. It has been discussed in the previous section that membrane in a biological neuron behaves like a capacitor that separates the ion concentrations. Due to the separation of charges between the layers of membrane, electric pulse is generated that travels through the axon of the neuron to hit the target neuron. This behavior can be modeled by a resistor-capacitor (RC) circuit shown in Figure 1.10. An electric circuit in Figure 1.10 is made of a resistor

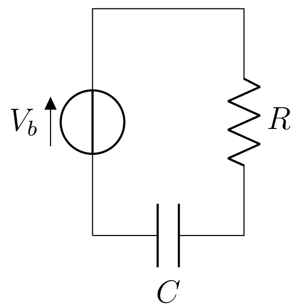

FIGURE 1.10: RC-circuit with a source of voltage.

(R), a capacitor (C), and a source of voltage(battery). All the electrical devices are added in a series. The battery in the circuit models the force due to the difference of ions' concentration inside and outside of the body cell. The resistor in the circuit behaves like ions channels that defines the permeability strength of membrane for the specific ions. The modeling of this circuit is based on the Kirchhoff's law of current and voltage.

(a) *Kirchhoff's law of current.* The sum of all currents entering into a junction is equal to the sum of all currents leaving the junction.

(b) *Kirchhoff's law of voltage.* The directed sum of the voltage differences around any closed loop is zero.

According to Kirchhoff's voltage law, the voltage drop across the capacitor is equal to the sum of the voltage drop across the battery and the voltage drop across the resistor.

$$V_c = V_b + I_R R, \tag{1.17}$$

where V_c is the voltage across the capacitor, V_b is the voltage across the batter, and I_R is the current passing the resistor that comes from Ohm's law $V_R = I_R R$. Thus, we have

$$I_R(t) = \frac{V_c(t) - V_b(t)}{R}. \tag{1.18}$$

The external current through the synapses is injecting into the neuron. Let it is denoted by I_{ext}. Then by Kirchhoff's current law, we have

$$I_{ext.} = I_C + I_R. \tag{1.19}$$

From equations 1.18 and 1.19, we obtain

$$C\frac{dV_c}{dt} + \frac{V_c(t) - V_b(t)}{R} = I_{ext.}. \tag{1.20}$$

1.2.2 Additive model

Motivated from the behavior of a biological neuron as an RC-circuit. The model (1.13) can be extended to the temporal nature of the input data,i.e., inputs to the neuron vary with time. One way to take into account the time-varying inputs is illustrated in Figure 1.11 where synaptic weights $w_{k1}, w_{k2}, ..., w_{kn}$ are represented by conductance (i.e.,reciprocal of resistance), and the respective inputs $x_1(t), x_2(t), ..., x_n(t)$

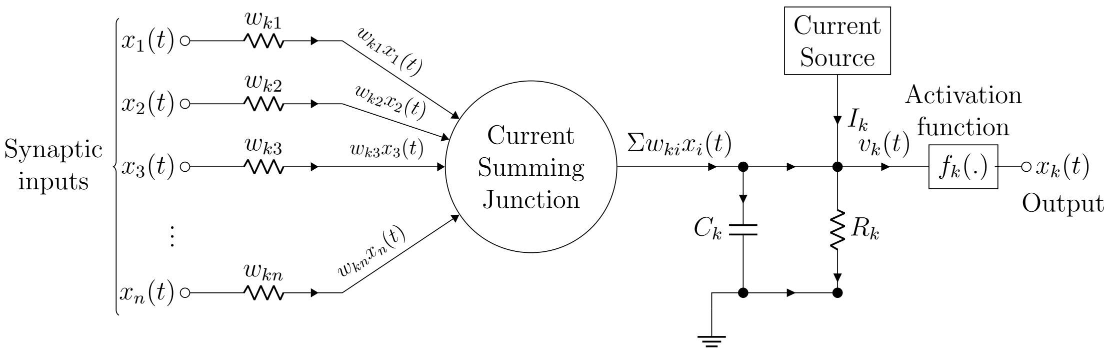

FIGURE 1.11: Additive model of a neuron.

are represented by potentials (i.e.,voltages). The inputs multiplied by their respective synaptic weights are summed up in the current summing junction characterized as low input resistance, unity current gain, and high output resistance. Total current flowing to the input node of the activation function $f_k(.)$ is

$$\sum_{i=1}^{n} w_{ki}x_i(t) + I_k, \tag{1.21}$$

where the first term is due to the input signals of k-th neuron and the second term is due to the external source of current applied as a bias to the neuron. Let $v_k(t)$ denotes the induced local field at the input node of the activation function, then the total current flowing away from the input node of the activation function is as follows:

$$\frac{v_k(t)}{R_k} + C_k\frac{dv_k(t)}{dt}, \tag{1.22}$$

where the first term presents the current across the resistor R_k and the second term presents the current due to potential drop across the capacitor C_k. According to Krichoff's current law, the total current flowing through the input node of the activation function in Figure 1.11 is zero. Thus, we have the following from the equations (1.21) and (1.22).

$$\frac{v_k(t)}{R_k} + C_k \frac{dv_k(t)}{dt} = \sum_{i=1}^{n} w_{ki} x_i(t) + I_k. \tag{1.23}$$

Given the induced input field $v_k(t)$, the output of the neuron k is determined by the activation function as

$$x_k(t) = f_k(v_k(t)). \tag{1.24}$$

The activation function in the additive model is generally chosen to be bounded and differentiable with asymptotic behavior like logistic function shown in Figure 1.4(c).

$$f_k(v_k) = \frac{1}{1 + \exp(-v_k)}. \tag{1.25}$$

The model described by the differential equation (1.23) is known as an additive model of a neuron, where the name 'additive' is used to discriminate it from the multiplicative (shunting) models which contain state dependent synaptic weights [23].

1.2.3 Hopfield neural network

Let us consider a fully connected recurrent network consisting of n neurons, where each neuron has the structure like Figure 1.11. The basic diagram of interconnections is shown in Figure 1.12 in which each neuron feed back its output, via a unit delay element, to the inputs of the other neurons by making feedback loops. More precisely, there is no self feedback loop in the network. Thus, considering the instantaneous

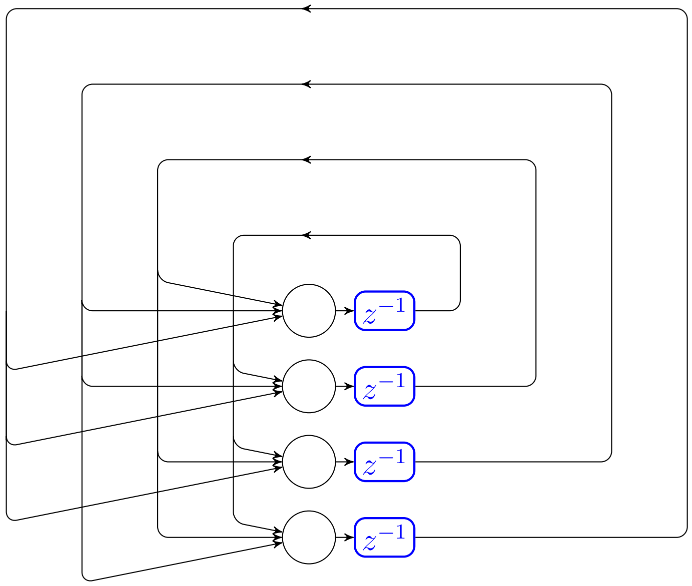

FIGURE 1.12: An architecture of Hopfield neural network consisting of $n = 4$ neurons.

propagation of signals between the neurons, and the equation (1.23). We have

$$C_k \frac{dv_k(t)}{dt} = -\frac{v_k(t)}{R_k} + \sum_{i=1}^{n} w_{ki} f_i(v_i(t)) + I_k, \text{for } k = 1, 2, ..., n, \tag{1.26}$$

where $x_i(t) = f_i(v_i(t))$, i.e., each neuron has its own activation function. The additive model (1.26) represents the Hopfield neural network without time delay [24]. Typically, the Hopfield model was designed into two forms of flow: one is discrete, and another is continuous flow. The solution of differential equation (1.26) is the continuous flow of the Hopfield neural network. The discrete flow of Hopfield network is based on the McCulloch-Pitts model where the state of each neuron has the values -1 and 1 depending on the signs of induced local field,i.e., $x_i(t) = -1$ if $v_i(t) > 0$, and $x_i(t) = 1$ if $v_i(t) < 0$, $\forall i$. This thesis is concerned with the continuous flow of neural networks rather than the discrete one.

The Hopfield neural network has attracted a great deal of attention in applications as a associative memory [25, 26]. The broad field of associative memory is a content

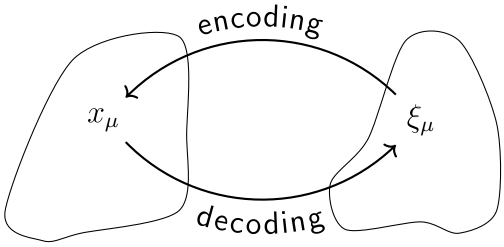

FIGURE 1.13: Encoding-Decoding illustration between the space of fundamental memories ξ_μ and the space of stored vectors x_μ.

addressable memory in that the fixed points of the network store the patterns in the memory. The primary function of the network is to retrieve a pattern stored in a memory on the basis of a given incomplete or noisy information about that pattern. Therefore, it is an important property of a content addressable memory to retrieve the stored pattern with the assistant of a given reasonable subpart or a subset of the information of that pattern.

Mathematically, we may define the essence of a content addressable memory as a mapping of a fundamental memory ξ_μ to a fixed (stable) point x_μ of the network. The arrow in Figure 1.13 from right to left describes encoding of a fundamental memory onto a fixed (stable) point, whereas left to right arrow describes decoding of the stored memory. Suppose now that the network has a pattern with partial but sufficient information about one of the fundamental memories. That partial information of the pattern is an initial state of the flow lies in the basin of attraction of the stable fixed point. In other words, from mathematical perspective it is the question of finding the conditions under which the flow of the network approach the fixed point in response to the arbitrary initial data. In content addressable memory, how many fixed points exist in the neural network is also an important question in applications perspective. More fixed points indicates greater storage capacity in the network.

The dynamics of Hopfield network was being studied by John Hopfield in his paper [24] published in 1984. He had defined an energy function for the network by assuming the following conditions.

(i) The matrix of synaptic weights is symmetric, i.e., $w_{ki} = w_{ik}$ for all i and k.

(ii) There exists an inverse of the activation function (1.25), so we may write $v = f_k^{-1}(x)$.

From the equation (1.25), we obtain

$$f_k^{-1}(x) = -\ln\left(\frac{1-x}{1+x}\right). \tag{1.27}$$

The energy or Lyapunov function of the Hopfield model (1.26) is defined as

$$E = -\frac{1}{2}\sum_{i=1}^{n}\sum_{k=1}^{n} w_{ik}x_i(t)x_k(t) + \sum_{k=1}^{n}\frac{1}{R_k}\int_0^{x_k} f_k^{-1}(x)dx - \sum_{k=1}^{n} I_k x_k. \tag{1.28}$$

The name 'Lyapunov function' will be discussed later so in this subsection we call it only energy function. The energy function defined in the equation (1.28) may have many minima points which represents the stable fixed points of the Hopfield model (1.26). The dynamics of the network is to find out those minima.

Hence, differentiating E with respect to time and using the equations (1.26) and (1.27), we obtain

$$\frac{dE}{dt} = -\sum_{k=1}^{n} C_k \left(\frac{dx_k}{dt}\right)^2 \left[\frac{df_k^{-1}(x_k)}{dx_k}\right]. \tag{1.29}$$

From the equation (1.27), we can see that the inverse activation function is monotonically non-decreasing function of the output x_k. Thus it follows that

$$\frac{df_k^{-1}(x_k)}{dx_k} \geq 0 \; \forall \; x_k. \tag{1.30}$$

Also note that

$$\left(\frac{dx_k}{dt}\right)^2 \geq 0 \; \forall \; x_k. \tag{1.31}$$

Hence, we conclude from the equation (1.29) that

$$\frac{dE}{dt} \leq 0. \tag{1.32}$$

According to the Lyapunov's method of stability, the inequality (1.32) implies that the time-evolution of the continuous Hopfield neural network described by the equation (1.26) represents a trajectory, which seeks out the minima of the energy function E and comes to a stop at such fixed points. It is also observable from the equation (1.29), using the inequalities (1.30) and (1.31), that the inequality (1.32) is zero only if

$$\frac{dx_k}{dt} = 0 \text{ for all } k. \tag{1.33}$$

We have, therefore, the following

$$\frac{dE}{dt} < 0 \text{ except at fixed points.} \tag{1.34}$$

Thus the trajectory of the continuous Hopfield neural network is converging asymptotically to the fixed points.

1.2.4 Cohen-Grossberg neural network

One of the most popular neural networks is a Cohen-Grossberg [27] neural network that can be considered as a generalized version of Hopfield network. Cohen and Grossberg in their paper published in 1983 described a general principle for designing content addressable memory networks by proving that models can be written in the form

$$\frac{dx_i(t)}{dt} = a_i(x_i(t))\big[b_i(x_i) - \sum_{k=1}^{n} w_{ik} f_k(x_k)\big], \text{ for all } i = 1, 2, ..., n, \tag{1.35}$$

where $x_i(t)$ is the activation state of the $i-$th neuron, $a_i(x_i(t))$ is the amplification function, $b_i(x_i)$ is the self-signal function, and $f_k(x_k)$ is the usual signal functions. From the equation (1.35), it can be observed that the rate of change in activity of the neuron decreases if and only if the net input to the neuron exceeds a certain intrinsic function b_i of its activity, and $a_i(x_i(t))$ is positive.

The nonlinear system (1.35) admits the Lyapunov function

$$V(x) = -\sum_{i=1}^{n} \int_0^{x_i} b_i(\xi_i) f_i^{'}(\xi_i) d\xi_i + \frac{1}{2} \sum_{k,j=1}^{n} w_{kj} f_k(x_k) f_j(x_j), \tag{1.36}$$

if the synaptic coefficients w_{ik}, and the other components a_i, b_i, f_k of the system satisfies the following conditions.

(G1) *symmetric*: $w_{ik} = w_{ki}$;

(G2) *continuity*: $a_i(\eta)$ is continuous for $\eta \geq 0$, and $b_i(\eta)$ is continuous for $\eta > 0$;

(G3) *positivity*: $a_i(\eta) > 0$ if $\eta > 0$;

(G4) *monotonicity*: $f_i(\eta)$ is continuously differential and $f_i^{'}(\eta) \geq 0$ for $\eta \geq 0$.

Integrating V along the trajectories implies that

$$\frac{dV}{dt} = -\sum_{i=1}^{n} a_i f_i' \big[b_i - \sum_{k=1}^{n} w_{ik} f_k\big]^2. \tag{1.37}$$

From (G3) and (G4), we have $\frac{dV}{dt} \leq 0$ along the trajectories. That is, the trajectories of the system (1.35) converge globally to the equilibrium points.

Cohen and Grossberg noted in his publication [27] that the differential equation (1.35) can represent the additive model (1.26) by using the coefficients of the standard electrical circuit interpretation as

$$a_i(x_i(t)) = \frac{1}{C_i}, \tag{1.38}$$

$$b_i(x_i(t)) = -\frac{1}{R_i} + I_i, \tag{1.39}$$

$$w_{ik}(1.26) = -w_{ik}(1.37), \tag{1.40}$$

where numbers in bracket is taken only for showing synaptic weights in respect of their model. Thus, in the additive case, the amplification function is positive constant, hence it satisfies positivity, and the self-signal function is linear. If we substitute the equations (1.38), (1.39), and (1.40) in the Lyapunov function (1.36) then we obtain

$$V = -\frac{1}{2}\sum_{j,k=1}^{n} w_{jk} f_j(x_j) f_k(x_k) + \sum_{i=1}^{n} \frac{1}{R_i} \int_0^{x_i} \xi_i f_i'(\xi_i) d\xi_i - \sum_{i=1}^{n} I_i f_i(x_i). \tag{1.41}$$

In Hopfield energy function, he has assumed the activation function should be invertible whereas Cohen and Grossberg has considered the activation function only to be non-decreasing. The work of Hopfield published in 1984 one year later than the work of Cohen and Grossberg. Despite of this, physicists and engineers named the additive model as Hopfield additive model in their literature. Therefore, Stephen

Grossberg had published a review paper [28] of his works and shown that how the additive model is a special case of the work published in collaboration with Michael A. Cohen [27].

1.2.5 Bidirectional Associative Memory

Bidirectional associative memory (BAM) is the extension of associative memory. The primary functions of associative memory is to recall the whole stored patterns in response to the given input patterns associated with the stored pattern. The classification of associative memory is such that while the memory in which the associated input and output patterns differ are called heteroassociative memory, it is called autoassociative memory if they are the same. The typical diagram of a bidirectional associative memory is shown in Figure 1.14 that is consisted of two layers of neurons. Let us denote the layers of the network by X-layer containing n neurons and Y-layer containing m neurons, where all the neurons of the layers have mutual interconnections. More precisely, there is not any interconnection among the neurons of the same layer. Information passes forward from one neuron layer to the other by passing through the connection matrix $W = [w_{ik}]_{n\times m}$. Information passes backward through the matrix transpose W^T. This backward-forward passes of information in search of the stored memory is the motivation behind naming the network BAM. In general, BAM is a symmetric neural network,i.e., $W \neq W^T$.

In the papers of Kosko [29, 30] published in 1987, he had extended the autoassociative neural networks Cohen and Grossberg [27] and Hopfield [24] to the heteroassociative neural network that performs bidirectional associative search. The Cohen-Grossberg theorem can not be applied in order to stabilize BAM neural network unless it is symmetric network. Bark Kosko in his paper [30] has devised the

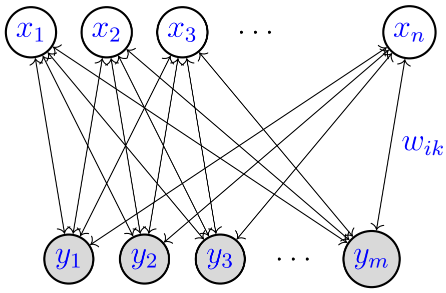

FIGURE 1.14: Bidirectional associative memory neural network.

procedure for dealing with systems having asymmetric coefficients. The interactions of the neurons from Y-layer to X-layer obey the following additive equation.

$$\dot{x}_j(t) = -A_j x_j + \sum_{i=1}^{m} w_{ij} f_i(y_i) + I_j, \text{ for all } j = 1, 2, ..., n. \tag{1.42}$$

Whereas the top-down interactions of the neurons obey the following additive equation.

$$\dot{y}_i(t) = -B_i y_i + \sum_{j=1}^{n} w_{ji} f_j(x_j) + I_i, \text{ for all } i = 1, 2, ..., m, \tag{1.43}$$

where the first term of the equation (1.42) represents a decay term with proportionality constant A_j, and I_j are external stimuli to the neurons having activation level x_j; similarly for the equation (1.43), B_i are proportionality constants, I_i are external stimuli, and the activation levels of the neurons are represented by y_i; the activation functions $f_k(.)$ is influencing the states x_j with connection strengths w_{ij} whereas $f_l(.)$ is influencing the states y_i with connection strengths w_{ji}; the activation functions $f(.)$ are bounded and monotonically increasing functions.

The connections matrix is $W = [w_{ij}]_{m \times n} \neq W^T$ in general model of BAM neural network. If we assume $W = W^T$ then by Cohen and Grossberg [27], B. Kosko in [30] has shown the Lyapunov global stability of the system governed by the equations

(1.42) and (1.43). The global Lyapunov or energy functional is defined as

$$E = -\sum_{j=1}^{n}\int_{0}^{x_j} f_j^{'}(a_j)a_j da_j - \sum_{i=1}^{m}\sum_{j=1}^{n} f_i(x_i)f_j(y_j)w_{ij} - \sum_{i=1}^{n} f_i(y_i)I_i - \sum_{j=1}^{m} f_j(x_j)I_j + \sum_{i=1}^{m}\int_{0}^{y_i} f_i^{'}(b_i)b_i db_i. \tag{1.44}$$

The total time derivative of E will result in the following equation after rearranging the terms.

$$\begin{aligned}\dot{E} &= -\sum_{j=1}^{n} f_j^{'}(x_j)\dot{x}_j\Big[-A_j x_j + \sum_{i=1}^{m} w_{ij} f_i(y_i) + I_j\Big] \\ &\quad -\sum_{i=1}^{m} f_i^{'}(y_i)\dot{y}_i\Big[-B_i y_i + \sum_{j=1}^{n} w_{ji} f_j(x_j) + I_i\Big] \\ &= -\sum_{j=1}^{n} f_j^{'}(x_j)\dot{x}_j^2 - \sum_{i=1}^{m} f_i^{'}(y_i)\dot{y}_i^2 \leq 0 \end{aligned} \tag{1.45}$$

From the inequality (1.45) it is concluded that the BAM neural network is globally stable. Since $f^{'}(.) > 0$ so the energy reaches its minimum if and only if $\dot{x}_j = \dot{y}_i = 0$. The BAM neural network can be implemented as an autoassociative neural network,i.e., Hopfield model (1.26), by defining the augmented vector $(x_1, x_2, ..., x_n, y_1, y_2, ..., y_m)$ with a symmetric connection matrix of order $m + n$.

1.3 An Overview of Mathematical Concepts

This section will summarize the mathematical concepts that are applied in the chapters of this thesis. We mainly focus on delay differential equations, impulsive differential equations, and matrix measure theory.

1.3.1 Delay Differential Equations

A delay differential equation (DDE) is a differential equation where the highest order derivative only occurs with one value of the argument, and this argument is not less than the argument of the unknown function and its lower order derivatives appearing in the equation. For example

$$\dot{x}(t) = x^2(t-3) + x(t), \tag{1.46}$$

is a delay differential equation but

$$\dot{x}(t-5) = x^2(t) + x^3(t), \tag{1.47}$$

is not.

Let us define $\mathcal{C} = C([-\tau, 0], \mathbb{R}^n)$ is a Banach space of continuous functions mapping interval $[-\tau, 0]$ into $\mathbb{R}^n$. The norm for the space $\mathcal{C}$ is designated as

$$\|\phi\|_\tau = \sup_{-\tau \leq s \leq 0} \|\phi(s)\|, \tag{1.48}$$

where $\|.\|$ is an Euclidean norm on $\mathbb{R}^n$. Suppose $x(t)$ is a function defined on at least $[t-\tau, t]$ then we can define a new function x_t as

$$x_t(s) = x(t+s), \ \forall \ s \in [-\tau, 0]. \tag{1.49}$$

For $D \subset \mathbb{R}^n$, we define $\mathcal{C}_D = C([-\tau, 0], D)$ is a Banach space of continuous functions mapping interval $[-\tau, 0]$ into D.

Definition 1.3.1. Let $I \subset \mathbb{R}$, $f : I \times C_D \to \mathbb{R}^n$ is a given function, and "." denotes the right hand time derivative then we say the following relation

$$\dot{x}(t) = f(t, x_t), \tag{1.50}$$

is a delay differential equation on $I \times C_D$.

Where right hand time derivative of a function $x(t) : \mathbb{R} \to \mathbb{R}$ is defined as

$$\lim_{h \to 0^+} \sup \frac{x(t+h) - x(t)}{h}. \tag{1.51}$$

For a given $t_0 \in I$ and $\phi_0 \in \mathcal{C}_D$, the initial value problem (IVP) associated with the DDE (1.50) is

$$\begin{cases} \dot{x}(t) = f(t, x_t), \ t > t_0, \\ x(s) = \phi_0(s - t_0), \ \forall \ s \in [t_0 - \tau, t_0]. \end{cases} \tag{1.52}$$

It means that the initial value at t_0 must specify the solution $x(t)$ for the whole past $[t_0 - \tau, t_0]$. Note that the initial function or history ϕ_0 is a continuous function but it is not necessarily compatible with the DDE (1.52). Thus the solution might not be differential at initial instant t_0. For example, take the simple model of DDE given as

$$\begin{cases} \dot{x}(t) = x(t-1), \ t > 0, \\ x(s) = 1, \ \forall \ s \in [-1, 0], \end{cases} \tag{1.53}$$

The solution of DDE (1.53) for interval $t \in (0, 1]$ is $x(t) = t + 1$ with initial condition $x(0) = 1$. It clear from the solution that $\dot{x}(0^+) = 1 \neq \dot{x}(0^-) = 0$,i.e., first derivative of $x(t))$ at $t_0 = 0$ is not continuous.

The equation (1.50) is a general equation for the following differential equations.

(i) If $\tau = 0$, then ordinary differential equation

$$\dot{x}(t) = f(t, x). \tag{1.54}$$

(ii) DDE with discrete delays as

$$\dot{x}(t) = f(t, x(t), x(t-\tau_1), ..., x(t-\tau_n)), \text{ where } \tau = \max_{1 \le i \le n} \tau_i. \tag{1.55}$$

(iii) DDE with distributed delay as

$$\dot{x}(t) = \int_{-\tau}^{0} f(t, s, x(t+s))ds. \tag{1.56}$$

In the third case, delay can be infinity as

$$\dot{x}(t) = \int_{-\infty}^{0} f(t, s, x(t+s))ds, \tag{1.57}$$

associated with the history function $x(s) = \phi_0(s - t_0)\ \forall\ s \in (-\infty, t_0]$.

Now we define the solutions of DDE (1.50).

Definition 1.3.2. A function $x(t)$ is a solution of equation (1.50) on $[t_0 - \tau, \beta)$ if there are $t_0, \beta \in \mathbb{R}$ with $\beta > t_0$ such that $x \in C([t_0 - \tau, \beta), D)$, $[t_0, \beta) \subset I$, and $x(t)$ satisfies equation (1.50) for $t \in [t_0, \beta)$.

As in an ordinary differential equation, the solution of IVP (1.52) is equivalent to finding the solution of the following integral equation.

$$\begin{cases} x(t) = \phi_0(t_0) + \int_{t_0}^{t} f(s, x_s)ds, \text{ for all } t \in [t_0, \beta) \\ x_{t_0}(s) = \phi_0(s), \text{ for all } s \in [t_0 - \tau, t_0]. \end{cases} \tag{1.58}$$

The concept of existence, uniqueness, continuation of solutions, and continuous dependence is very similar to the ordinary nonlinear differential equation [31]. We will introduce the theorems on existence and uniqueness of DDE with finite delay, the detailed proofs can be found in [32]. Before stating the theorems we need the following definition.

Definition 1.3.3. Let $f : I \times \mathcal{C}_D \to \mathbb{R}^n$ and let $S \subset I \times \mathcal{C}_D$. Then f is Lipschitz on S if there exists $L > 0$ such that

$$\|f(t, \phi_1) - f(t, \phi_2)\| \leq \|\phi_1 - \phi_2\|_\tau, \tag{1.59}$$

whenever $(t, \phi_1), (t, \phi_2) \in S$.

Theorem 1.1. *(Local Existence) Let Ω is an open subset in $\mathbb{R} \times \mathcal{C}$ and $f : \Omega \to \mathbb{R}^n$ be continuous on its domain. If $(t_0, \phi_0) \in \Omega$, then there is a solution of the IVP (1.52) passing through (t_0, ϕ_0) that exists on $[t_0 - \tau, t_0 + \delta]$ for some $\delta > 0$.*

Theorem 1.2. *(Uniqueness) Let Ω is an open subset in $\mathbb{R} \times \mathcal{C}$ and $f : \Omega \to \mathbb{R}^n$ be continuous with Lipschitz on each compact set in Ω. If $(t_0, \phi_0) \in \Omega$, then there is a unique solution of the IVP (1.52) passing through (t_0, ϕ_0).*

Time-delay in neural networks play an important role in dynamical behavior of the network. In previous section we have assumed that the transmission of signals in neurons is an instantaneous process but practically it is not always true. There are different shapes and sizes of axons due to which signal transmission from one neuron to others is not instantaneous [33]. To take this facts into account the additive model of Hopfield represented by the equation (1.26) can be transformed in the following DDE.

$$C_k \frac{dv_k(t)}{dt} = -\frac{v_k(t)}{R_k} + \sum_{i=1}^{n} w_{ki} f_i(v_i(t - \tau_{ki})) + I_k, \text{for } k = 1, 2, ..., n, \tag{1.60}$$

where τ_{ki} is the time taken by the signals to transmit from i-th to k-th neurons in the network. This type of delay is called discrete delay. Since neural networks have a spatial structure due to the presence of many parallel pathways connected among the neurons so it is not possible to model them with discrete delay. There will be a distribution of propagation time delay. Thus, it is desired to model them by introducing continuously distributed delays [34, 35, 36, 37, 38, 39]. The equation (1.26) can be represented in the form of DDE with distributive delays.

$$C_k \frac{dv_k(t)}{dt} = -\frac{v_k(t)}{R_k} + \sum_{i=1}^{n} w_{ki} f_i\left(\int_0^{\infty} v_i(t-u) g_{ki}(u) du\right) + I_k, \text{for } k = 1, 2, ..., n, \tag{1.61}$$

where u is a signal delay from i-th to j-th neuron occurs with probability distribution function $g_{ki}(u)$ with mean delay $\tau_{ki} = \int_0^{\infty} u g_{ki}(u) du$. A DDE consisting of discrete and distributed delays is called a DDE with mixed-time delays. The time derivative of state variable v_k depends on v_i for the past $(-\infty, t]$. There are literature involving results on delayed neural networks have been published in recent past [40, 41, 42, 43, 44, 45].

1.3.2 Impulsive Differential Equation

The theory of impulsive differential equations play an important role in modeling the real world problems that are the symbiosis of continuous and discontinuous systems. Many evolution process are characterized by the fact that at certain moments of time they experience a change of state abruptly [46, 47, 48]. This kind of sudden change in the states of the system is known as impulsive effects. There are many real world problems which can not be modeled only in continuous form, for example, bursting rhythm models in medicine and biology, frequency modulated systems, optimal control models in economics, and pharmacokinetics, do exhibit impulsive effects.

There are two different kinds of impulsive differential equations: impulses at fixed times, and impulses at variable times, the former one is described by the following equation.

$$\begin{cases} \dot{x}(t) = f(t,x), t \geq t_0,\ t \neq t_k,\ k = 1, 2, ... \\ \Delta x(t_k) = I_k(x),\ t = t_k, \end{cases} \tag{1.62}$$

where $I_k : \Omega \to \Omega,\ f : \mathbb{R}^+ \times \Omega \to \mathbb{R}^n$, $\Omega \subset \mathbb{R}^n$ is an open set, $\mathbb{R}^+$ is a non-negative real line, and $\mathbb{R}^n$ is a n- dimensional Euclidean space; t_k is a sequence of time in $\mathbb{R}^+$ at which the state $x(t)$ has a jump kind of discontinuity at $t = t_k$, more precisely, $x(t_k^-) = \lim_{h\to 0} x(t_k - h) = x(t_k)$ but $x(t_k^+) \neq x(t_k)$ so $\Delta x(t_k) = x(t_k^+) - x(t_k^-)$; impulsive sequence t_k is strictly increasing and unbounded above ,i.e., $t_0 < t_1 < t_2 < ... < t_k$ and $\lim_{k\to\infty} t_k = \infty$.

Let $x(t) = x(t, t_0, x_0)$ is a solution of the continuous part of the system (1.62) at initial point (t_0, x_0). Then evolution of the solution $x(t)$ of the system (1.62) is elaborated as follows: the point $P_t(t, x(t))$ starts its motion from the initial point $P_0(t_0, x_0)$ along the solution $\{(t, x(t)) : t \geq t_0\}$ until it reaches the time $t = t_1$ at which the first impulse effect is activated and transfer the point P_{t_1} to $P_{t_1^+}(t_1, x(t_1^+))$, where $x(t_1^+) = x(t_1) + I_1(x)$. Further the point P_t continues to move along the trajectory $x(t) = x(t, t_1^+, x(t_1^+))$ until it reaches the time $t_2 > t_1$ at which the second impulse effect is activated and transfer the point P_{t_2} to $P_{t_2^+}(t_2, x(t_2^+))$, where $x(t_2^+) = x(t_2) + I_2(x)$. As before, the evolution process continues as long as the solution of the system (1.26) exists.

The second form of the impulsive differential equation is described as follows:

$$\begin{cases} \dot{x}(t) = f(t, x), \;\; t \neq \tau_k(x), \;\; k = 1, 2, ..., \\ \Delta x(t_k) = I_k(x), \;\; t = \tau_k(x), \end{cases} \tag{1.63}$$

where $t = \tau_k(x)$ is a sequence of surfaces such that $\tau_k(x(t)) < \tau_{k+1}(x(t))$ and $\lim_{k\to\infty} \tau_k(x) = \infty$. The impulsive system with variable moment exhibits more difficult behavior than fixed moment. For example, it is clear from the system (1.63) that impulsive moment is depending on the solutions of $t_k = \tau_k(x(t))$, for each k. Thus, the trajectory started at different initial points will have different points of discontinuity. It may happen that a solution hits the same surface $t_k = \tau_k(x)$ several number of times, we call this behavior 'pulse phenomena'.

In this thesis, we studied impulsive neural network with fixed time impulses. The impulse phenomena was being considered in neural networks by K. Gopalsamy in his paper [49]. He has studied the stability problem of Hopfield network when states of the neurons have sudden changes at impulsive points. Generally, we study two types of neural network, discrete or continuous. An impulsive neural network is a symbiosis of both forms. Thus many researchers have drawn their attentions to the problems of stability and synchronization of neural networks with impulses [50, 51, 52, 53, 54, 55]. In particular, let us consider an impulsive system with fixed time impulses

$$\begin{cases} \dot{x}(t) = f(t,x), t \geq t_0,\ t \neq t_k,\ k = 1, 2, ... \\ \Delta x(t_k) = \mu_k x(t_k^+),\ t = t_k, \end{cases} \tag{1.64}$$

where μ_k is an impulsive strength or impulse gain at fixed time t_k. The impulse gain is characterized into three categories:(i) stabilizing impulses ($|\mu_k| < 1$), (ii) destabilizing impulses ($|\mu_k| > 1$), (iii) inactive impulses ($|\mu_k| = 1$). This thesis is concerned to investigate stability and synchronization problems of neural networks with generalized cases of impulses.

1.3.3 Matrix measure theory

There are many algebraic methods for stability analysis of neural networks, such as the methods based on the concept of Linear Matrix Inequality (LMI) [56, 57], M-Matrices [58], H-Matrices [59], and Matrix measure theory [60, 61, 62], etc. The concept of measuring the matrix is developed from the normed linear space or we can say it is an extension of the concept; norm of vectors, norm of matrices.

Let us consider a finite dimensional Euclidean linear space $\mathbb{R}^n$ over real field $\mathbb{R}$. The norm $\|(.)\|_p$ in $\mathbb{R}^n$ for $1 \leq p < \infty$ is defined as

$$\|x\|_p = \Big(\sum_{i=1}^{n} |x_i|^p\Big)^{\frac{1}{p}}. \tag{1.65}$$

The linear space $\mathbb{R}^n$ associated with the norm defined in (1.65) is called normed linear space. Thus for any $x \in \mathbb{R}^n$, we can have the following for $p = 1, 2, \infty$.

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|, \quad \|x\|_2 = \Big(\sum_{i=1}^{n} |x_i|^2\Big)^2, \quad \|x\|_\infty = \max_{1\leq i\leq n} |x_i|. \tag{1.66}$$

Similarly, we define the norm in linear space $\mathbb{R}^{n\times n}$ of matrices over the real field $\mathbb{R}$. For any matrix $A = [a_{ij}]_{n\times n} \in \mathbb{R}^{n\times n}$, we can get the following for $p = 1, 2, \infty$.

$$\|A\|_1 = \max_j \sum_{i=1}^{n} |a_{ij}| \text{ (column sum)}, \quad \|A\|_2 = (\max_i \lambda_i(A^T A))^{\frac{1}{2}},$$
$$\|A\|_\infty = \max_i \sum_{j=1}^{n} |a_{ij}| \text{ (row sum)}. \tag{1.67}$$

The function $\|(.)\| : \mathbb{R}^{n\times n} \to \mathbb{R}^+$ is uniformly continuous and convex in nature. Then at any point $X \in \mathbb{R}^{n\times n}$, the one-sided directional derivative of the function $\|(.)\|$ along the direction of matrix A is defined as

$$\lim_{h\to 0^+} \frac{\|X + hA\| - \|X\|}{h}. \tag{1.68}$$

Definition 1.3.4. The one-sided directional derivative of the norm function $\|(.)\|$ at point $I \in \mathbb{R}^{n\times n}$ in the direction of A is called the matrix measure of A and it is denoted by $\mu(A)$, i.e.,

$$\mu(A) = \lim_{h\to 0^+} \frac{\|I + hA\| - 1}{h}. \tag{1.69}$$

The existence of limit (1.69) can be shown by assuming function $f(h) = \frac{\|I+hA\|-1}{h}$. If $f(h)$ is decreasing w.r.t h and bounded below then the limit (1.69) must exists for all $A \in \mathbb{R}^{n\times n}$. Let $k \in (0,1)$, then we get

$$khf(kh) = \|I + khA\| - 1 = \|k(I + hA) + (1-k)I\| - 1 \leq k\|I + kA\| + 1 - k - 1$$
$$\leq k(\|I + hA\| - 1).$$

That is $f(kh) \leq f(h)$ implies $f(h)$ is decreasing w.r.t. h. Note that $f(h) \geq -\|A\|$. Therefore, the existence of the limit (1.69) is proved.

The matrix measure of A induced from the norm $\|(.)\|_p$ for $p = 1, 2, \infty$ is defined as

$$\mu_p(A) = \lim_{h\to 0^+} \frac{\|I + hA\|_p - 1}{h}. \tag{1.70}$$

Thus we have

$$\mu_1(A) = \max_j \left[a_{jj} + \sum_{\substack{i=1\\ i\neq j}}^{n} |a_{ij}|\right], \quad \mu_2(A) = \max_i [\lambda_i(\frac{A + A^T}{2})],$$
$$\mu_\infty(A) = \max_i \left[a_{ii} + \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}|\right]$$

Example 1.3.1. If matrix $A \in \mathbb{R}^{n\times n}$ is a skew-symmetric and $A \neq 0$ then $\mu_2(A) = 0$.

Important properties of matrix measure

The following properties of matrix measure are very useful for application purpose. Let $A, B \in \mathbb{R}^{n\times n}$, and $\mu(.)$ is defined in (1.69), then we have the following.

(i) $-\|A\| \leq -\mu(-A) \leq \mu(A) \leq \|A\|$.

(ii) $\mu(cA) = c\mu(A), \ \forall c \geq 0$.

(iii) $\mu(A + cI) = \mu(A) + c, \ \forall c \in \mathbb{R}$.

(iv) $\max[\mu(A) - \mu(-B), -\mu(-A) + \mu(B)] \le \mu(A + B) \le \mu(A) + \mu(B)$.

(v) $\mu : \mathbb{R}^{n\times n} \to \mathbb{R}$ is convex on $\mathbb{R}^{n\times n}$,i.e.,

$$\mu[\lambda A + (1 - \lambda)B] \le \lambda\mu(A) + (1 - \lambda)\mu(B),\ \forall\lambda \in (0, 1).$$

(vi) $|\mu(A) - \mu(B)| \le |\mu(A - B)| \le \|A - B\|$.

(vii) $-\mu(-A) \le \mathrm{Re}\lambda_i(A) \le \mu(A)$ for all $i \in \{1, 2, ..., n\}$.

(viii) If A is non-singular $-\mu(-A) \le (\|A\|)^{-1} \le \|A\|$.

Proof. Proof of the above listed properties of the matrix measure can be found in [62]. □

The benefits of using matrix measure method to investigate stability analysis of neural networks are as follows: (i) the measure of a matrix can be zero/negative/-positive whereas the norm is always positive; (ii) it is known well that constructing Lyapunov function for the system is a tough task due to unavailability of general method but in a matrix measure approach we can easily construct Lyapunov function. Due to these benefits of the matrix measure method, many authors have drawn their attentions to the problem of stability analysis of neural networks via matrix measure approach [60, 61, 62, 63, 64].

1.4 Synchronization

Chaos is an important behavior of nonlinear dynamical systems. If the trajectory of a nonlinear system is highly sensitive to the initial condition then it is called chaotic nature. Many researchers have demonstrated chaotic behavior of neural networks in their articles [65, 66, 67]. The synchronization of chaotic systems is a difficult problem owing to their extremely sensitive dependence on initial conditions. Any initial correlation present between identical systems, starting from very close initial conditions, exponentially decrease to zero with time. Thus, for all practical purposes, any initial synchronization between the systems is bound to disappear rapidly. In recent times, however, some methods of achieving synchronized behavior between chaotic systems have been proposed. Pioneering work in this respect has been done by Pecora and Carroll [68], who used the concept of a response system locking on to a driver system. So far, such studies have been limited to driving a response system by a single driver system. However, the knowledge gained from studying such simple systems may not be adequate to give us an idea as to how systems consisting of multiple independent driver systems, competing with each other to synchronize the same response system, will behave. The Pecora-Carroll driving mechanism can be seen as the ”strong-coupling” limit of a general scheme of directionally- oriented couplings in a network of chaotic elements.

There are many types of synchronization have been reported in literature [69, 70, 71]. Some of them are described below

(1) *Complete Synchronization*: This type of synchronization can be occurred when systems are identical and coupled unidirectionally or bidirectionally. Let us assume two coupled identical systems as follows

$$\dot{x}(t) = f(x(t)) \tag{1.71}$$

$$\dot{y}(t) = f(y(t)) + U(x(t), y(t)), \tag{1.72}$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$ is a continuous vector field and $U(x, y)$ is a coupling term. The systems (1.71) and (1.72) are said to be synchronized completely if

$\|y(t) - x(t)\| \to 0$ as $t \to \infty$.

(2) *Quasi Synchronization*: Let us consider the coupled systems as

$$\dot{x}(t) = f(x(t)) \tag{1.73}$$

$$\dot{y}(t) = g(y(t)) + U(x(t), y(t)), \tag{1.74}$$

where $g : \mathbb{R}^n \to \mathbb{R}^n$ is a continuous vector field. The systems (1.73) and (1.74) are said to be synchronized in quasi way if there exists $T > 0$ such that $e(t) \in D = \{e(t) : \|e(t)\| < \epsilon\}$ for all $t > T$,where $e(t) = y(t) - x(t)$ and ϵ is a synchronization error bound.

(3) *Projective synchronization*: In this type of synchronization response system is synchronized with drive system up to a scaling factor,i.e., error system is defined as

$$e(t) = x(t) - \alpha y(t), \tag{1.75}$$

where $\alpha \neq 0$ is a scaling factor and $\|e(t)\| \to 0$ as $t \to \infty$.

- In general, if α is replaced by a diagonal matrix $\mathit{\Omega} = \{\alpha_1, \alpha_2, ..., \alpha_n\}$ with constant elements $\alpha_i \neq 0 \forall i$ then it is called modified projective synchronization.
- If α is replaced by a diagonal matrix $\Omega(t) = \text{diag}\{\alpha_1(t), \alpha_2(t), ..., \alpha_n(t)\}$, where $\alpha_i(t) \neq 0$ is bounded and continuously differentiable function for all i, then it is called modified function projective synchronization (MFPS).
- If $\alpha_1(t) = \alpha_2(t) = ... = \alpha_n(t)$ then it is called function projective synchronization.

There are also other types of synchronization schemes have been reported in literature such as anticipated-synchronization, phase synchronization, generalized synchronization, see [72, 73, 74] and the references cited therein.

****